



Méthodes avancées d'optimisation par méta-modèles – Application à la performance des voiliers de compétition

Matthieu Sacher

► To cite this version:

Matthieu Sacher. Méthodes avancées d'optimisation par méta-modèles – Application à la performance des voiliers de compétition. Mécanique des matériaux [physics.class-ph]. Ecole nationale supérieure d'arts et métiers - ENSAM, 2018. Français. NNT : 2018ENAM0032 . tel-01924714

HAL Id: tel-01924714

<https://pastel.hal.science/tel-01924714>

Submitted on 16 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale n° 432 : Sciences des Métiers de l'Ingénieur

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École Nationale Supérieure des Arts et Métiers

Spécialité « Mécanique-matériaux »

présentée et soutenue publiquement par

Matthieu SACHER

le 10 septembre 2018

Méthodes Avancées d'Optimisation par Méta-Modèles – Application à la Performance des Voiliers de Compétition

Directeur de thèse : **Jacques-André ASTOLFI**

Co-Directeur de thèse : **Olivier LE MAÎTRE**

Encadrant de thèse : **Frédéric HAUVILLE**

Co-Encadrant de thèse : **Régis DUVIGNEAU**

Jury :

M^{me} Shabnam ARBAB CHIRANI,	Professeur – ENIB-IRDL	Présidente
M. Christophe CORRE,	Professeur – École Centrale de Lyon	Rapporteur
M. Rodolphe LE RICHE,	Directeur de Recherche – LIMOS-CNRS	Rapporteur
M. Jacques-André ASTOLFI,	Professeur – École-Navale	Examineur
M. Vincent CHAPIN,	Maître de Conférence – ISAE-SUPAERO	Examineur
M. Régis DUVIGNEAU,	Chargé de Recherche – INRIA	Examineur
M. Olivier LE MAÎTRE,	Directeur de Recherche – LIMSI-CNRS	Examineur
M^{me} Angela VINCENTI,	Maître de Conférence – Sorbonne Université	Examinatrice
M. Frédéric HAUVILLE,	Maître de Conférence – École-Navale	Invité

Vos enfants ne sont pas vos enfants. Ils sont les fils et les filles de l'appel de la Vie à elle-même. Ils viennent à travers vous mais non de vous. Et bien qu'ils soient avec vous, ils ne vous appartiennent pas.

KHALIL GIBRAN

Remerciements

Je tiens à profondément remercier mes encadrants, Jacques-André ASTOLFI, Olivier LE MAÎTRE, Frédéric HAUVILLE et Régis DUVIGNEAU, pour leur soutien et pour leur grande implication, tout au long de ces années de doctorat. Mes remerciements sont également pour Mathieu DURAND, inspirateur indispensable à la réalisation de ces travaux de thèse. De même, je remercie sincèrement les trois entités de cette thèse, qui sont l'École-Navale, le LIMSI-CNRS et l'INRIA, pour leur accueil et pour les importantes ressources mises en œuvre. Enfin, un grand merci à celles et ceux, et en particulier à ma famille, qui ont pu contribuer à l'aboutissement de ce projet de thèse.



À tous ceux qui y ont cru.

Sommaire

Introduction	1
I Contexte du voilier	7
1 Performance du voilier	9
1.1 Fonctionnement du voilier	10
1.1.1 Terminologie	10
1.1.2 Équilibre aéro-hydrodynamique	11
1.1.3 Vent apparent	12
1.2 Étude de la performance	15
1.2.1 Performance du voilier	15
1.2.2 Travaux scientifiques	17
2 Modéliser la physique du voilier	21
2.1 Physique du voilier	22
2.1.1 Mécanique des fluides	22
2.1.2 Interaction fluide-structure	29
2.2 Outils numériques	31
2.2.1 Mécanique des fluides numérique	31
2.2.2 Simulations des voiles en IFS	35
2.3 Essais expérimentaux	37
2.3.1 Hydrodynamique en bassin	37
2.3.2 Aérodynamique en soufflerie	40
2.3.3 Mesure <i>in situ</i>	41
II Méta-modèles pour l'optimisation	45
3 Généralités et contexte de l'optimisation	47
3.1 Motivations	47
3.2 Problème d'optimisation	48
3.3 Méthodes d'optimisation	49

3.3.1	Généralités	49
3.3.2	Algorithmes d'optimisation classiques	50
4	Méta-modèles Gaussiens	55
4.1	Définitions générales	56
4.1.1	Variables et vecteurs aléatoires	56
4.1.2	Processus Gaussiens	57
4.2	Méta-modèles Gaussiens	60
4.2.1	Prédiction	61
4.2.2	Fonction de covariance	63
4.2.3	Détermination des hyperparamètres	64
4.2.4	Plan d'expériences	65
4.3	Applications des méta-modèles Gaussiens	67
4.3.1	Analyse de sensibilité	67
4.3.2	Classification	71
4.3.3	Optimisation globale	73
5	Optimisation sous contraintes	81
5.1	Points non évaluables	82
5.2	Modélisation des contraintes	83
5.3	Classification des contraintes	85
5.3.1	Support Vector Machines, méthodes SVM et LS-SVM	85
5.3.2	Méthode EGO-LS-SVM	97
5.4	EGO sous contraintes évaluables et non évaluables	99
6	Optimisation multi-fidélité	105
6.1	Construction des méta-modèles	106
6.1.1	Méthode auto-régressive	106
6.1.2	Méthode récursive	108
6.2	EGO en multi-fidélité	110
6.2.1	Amélioration espérée augmentée en multi-fidélité	110
6.2.2	Points non évaluables	113
III	Contributions de la thèse	117
7	Réglages optimaux d'une grand-voile, applications expérimentales et numériques	119
7.1	Introduction	119
7.2	Publication dans « <i>Journal of Fluids and Structures</i> »	122
7.3	Conclusion	145

8	Optimisation en présence de points non évaluables	147
8.1	Introduction	147
8.2	Publication dans « <i>Structural and Multidisciplinary Optimization</i> »	149
8.3	Conclusion	170
9	Design optimal sous contraintes d'un hydrofoil à paramètres multiples	171
9.1	Introduction	171
9.2	Publication dans « <i>Ocean Engineering</i> »	174
9.3	Conclusion	185
10	Optimisation des réglages d'un gréement de voilier par méta-modèles multi-fidélité	187
10.1	Motivations	188
10.2	Publication soumise dans « <i>Computer Methods in Applied Mechanics and Engineering</i> »	191
10.2.1	Introduction	191
10.2.2	Publication	193
10.2.3	Conclusion	215
10.3	Optimisation à six paramètres	215
10.3.1	Problème d'optimisation	215
10.3.2	Validation du maillage type « overset »	217
10.3.3	Résultats	219
10.4	Conclusion	226
	Conclusion générale et perspectives	229
	Références bibliographiques	235

Liste des figures

1.1	La terminologie du voilier. Illustration provenant de (Paulet and Presles, 1998).	10
1.2	Les différentes allures d'un voilier. Illustration originale dans (Aubin, 2017).	11
1.3	Repère, attitudes et angles du voilier.	12
1.4	Équilibre aéro-hydrodynamique du voilier dans le plan (X,Y) . Illustration originale dans (Fossati, 2010).	13
1.5	Équilibre aéro-hydrodynamique du voilier dans le plan (Y,Z) . Illustration originale dans (Fossati, 2010).	13
1.6	Triangle des vitesses dans un cas de navigation au près ($\lambda = 0^\circ$).	14
1.7	Vrillage du vent apparent en fonction de l'altitude. Images originales dans (Hansen, 2006a).	14
1.8	Polaires de différents voiliers pour une vitesse du vent réel de 10 kts. Image originale dans (Garrett, 1996).	16
1.9	Programme de prévision de la vitesse du voilier.	16
2.1	Profils de vitesses au voisinage d'une paroi.	24
2.2	Volume de contrôle V contenu dans un domaine fluide D .	25
2.3	Efforts sur un corps portant soumis à un écoulement.	28
2.4	Effets de la cambrure d'une voile. Images tirées de (Fossati, 2010)	29
2.5	Effets des interactions entre une grand-voile et un génôis. Image issue de (Garrett, 1996)	29
2.6	Procédure d'un calcul IFS visqueux sur des voiles de Figaro.	37
2.7	Mesure de la résistance à l'avancement d'une carène. Image du « Wolfson Unit ».	38
2.8	Essais en tunnel de cavitation.	39
2.9	Soufflerie de l'université d'Auckland à veine ouverte générant un écoulement vrillé. Les dimensions de la veine d'essai sont de 7 m de large par 3.5 m de haut ; la vitesse de l'écoulement peut varier entre 1 et 5 m/s. Image provenant de (Aubin, 2017).	41
2.10	Mesures par photogrammétrie <i>in situ</i> d'un spinnaker. Images tirées de (Deparday, 2016).	42
3.1	Exemples d'évolutions de $f(x)$. Images issues de (Duvigneau, 2015).	49
3.2	Application à optimiser vue comme une boîte noire.	50
3.3	Distributions de lois normales multidimensionnelles où $d = 2$, $\mu = \mathbf{0}_2$ et $\sigma_{11}^2 = \sigma_{22}^2 = 1$, pour différentes valeurs de covariances σ_{12}^2 .	54
4.1	Densité de lois normales.	57

4.2	Réalisations de processus Gaussiens dans un cas unidimensionnel.	60
4.3	Prédictions à l'aide d'un méta-modèle Gaussien dans un cas unidimensionnel.	62
4.4	Influences du paramètre l sur les noyaux exponentiels et Gaussiens.	64
4.5	Plan factoriel complet sur $\Omega = [0,1]^2$	66
4.6	Plans d'expériences par hypercubes latins sur $\Omega = [0,1]^2$	67
4.7	Classification binaire non linéaire et non séparable.	72
4.8	Fonction Branin et plan d'expériences LHS.	78
4.9	Optimisation par méta-modèles Gaussiens de la fonction Branin, itérations 1, 2 et 3.	79
4.10	Optimisation par méta-modèles Gaussiens de la fonction Branin, itérations 15, 20 et 50.	80
5.1	Exemples de cas numériquement non évaluables.	83
5.2	Schéma d'un perceptron.	86
5.3	Classification binaire de données séparables avec un SVM linéaire.	87
5.4	Fonction séparatrice non linéaire.	90
5.5	Risque de surapprentissage.	96
5.6	Problème d'optimisation ($d = 2$) sous contraintes évaluables et non évaluables.	100
5.7	Itérations 5, 25, 50 et 100 de la fonction de mérite AEI_C et des probabilités P_C et $P_{A,B}$ du problème d'optimisation (5.4.2).	102
5.8	Convergences du problème d'optimisation (5.4.2).	103
6.1	Fonction unidimensionnelle (Forrester et al., 2007) illustrant un méta-modèle multi-fidélité.	109
6.2	Illustration de l'EGO en multi-fidélité sur le cas unidimensionnel défini par (6.1.14).	112
6.3	Illustration de la classification multi-fidélité dans un processus EGO. Cas du problème (6.2.9).	115
7.1	Illustration de la soufflerie du YRU.	120
7.2	Illustration d'un calcul IFS réalisé pendant la validation numérique de la méthode EGO.	121
8.1	Premiers essais de classifications des points non évaluables. Méta-modèles de la prédiction $\hat{y}(\mathbf{x})$ comprenant les zones classées comme non évaluables (zones blanches).	148
9.1	Illustration d'observations évaluables et non évaluables lors d'une optimisation à 4 paramètres d'un hydrofoil flexible.	172
10.1	Calculs IFS fluide parfait et visqueux pour différentes longueurs d'écoute de grand-voile. Résultats provenant de (Sacher et al., 2015).	189
10.2	Illustration du décollement fluide pour différents réglages de grand-voile. Résultats provenant de (Sacher et al., 2015).	190
10.3	Résultats de l'optimisation simple fidélité du problème (10.1.1). La construction d'un méta-modèle multi-fidélité utilisant cinq observations d'un niveau supérieur est également présentée.	190
10.4	Réglages du Figaro à optimiser.	216

10.5	Maillage fluide visqueux volumique en « overset ».	217
10.6	Validation IFS numérique globale du maillage overset.	218
10.7	Convergences de la performance et de la fonction de mérite par niveau de fidélité.	219
10.8	Effets de l'écoute de grand-voile, de l'écoute de génois et du D1 sur la performance.	220
10.9	Réglages optimums donnés par les méta-modèles $l = 1$ (voiles bleues) et $l = 2$ (voiles rouges). Les différences de pression sur les voiles, calculées par les deux modèles fluide, sont également présentées.	222
10.10	Effets de l'écoute de grand-voile, de l'écoute de génois et du D1 sur la performance maintenue optimale.	223
10.11	Effets de l'écoute de grand-voile, de l'écoute de génois et du D1 sur les cinq réglages, pour une performance maintenue optimale. Les courbes discontinues et continues correspondent respec- tivement aux réglages donnés par les méta-modèles simple et multi-fidélité. Les symboles $\triangle$ et $\square$ sont les optimums respectifs de $\mathbf{x}_{\text{opt}_1}$ et $\mathbf{x}_{\text{opt}_2}$.	225

Introduction

Le transport maritime est le plus important moyen de transport de marchandises. Aujourd'hui, une flotte gigantesque de près de 50000 navires sillonne les mers et assure le transport d'environ 90% du trafic mondial. Ces navires ne sont aujourd'hui propulsés que par des systèmes motorisés. En effet, la révolution industrielle a mis fin à la marine marchande à voile, au cours de la seconde moitié du XIXe siècle, en remplaçant la propulsion vélique par des hélices, entraînées initialement par des moteurs à vapeur. Aujourd'hui, l'augmentation du prix des énergies fossiles voit de nombreux acteurs anticiper un retour à une propulsion vélique, qu'elle soit auxiliaire ou principale. Certains grands groupes comme « Airbus » avec le projet « Airseas », « Beyond The Sea » porté par « Yves Parlier », ou la compagnie allemande « SkySails » parient sur la propulsion auxiliaire à l'aide d'un cerf-volant. Le projet « Neoline », porté par la société française du même nom, consiste en la mise au point d'un cargo conçu pour les navigations transocéaniques, dont la propulsion serait principalement vélique. A une plus petite échelle, mais déjà en action, la société « TOWT » effectue des liaisons transrégionales et transatlantiques sur des navires historiques à propulsion vélique pour transporter des produits alimentaires. Si le transport maritime à propulsion vélique reprend du service dans les années à venir, alors la période du transport maritime « tout motorisé » aura été d'une très courte durée, car en effet depuis les premiers peuplements de l'Indonésie, il y a plusieurs dizaines de milliers d'années, le vent a été utilisé pour se déplacer.

Si la voile a été délaissée pour le transport, elle est devenue parallèlement une source de compétitions sportives. La première compétition entre différents voiliers date de 1851 avec la « Coupe de l'America ». Aujourd'hui, il existe de nombreuses courses à voile, des plus locales proches des côtes terrestres, aux plus extrêmes en distances et durées comme le très connu « Vendée globe » ou encore la « Volvo Ocean Race ». Ces compétitions amènent la recherche et le développement au sein même des équipes de design, en particulier ces dernières années avec l'augmentation, d'une part des modèles disponibles en simulation numérique et d'autre part de la puissance de calcul.

Des équipes de scientifiques s'investissent également sur la thématique de la physique du voilier. Les chercheurs du « Yacht Research Unit (YRU) », de l'université d'Auckland, réalisent des mesures dans une soufflerie dédiée à l'étude des voiles, depuis maintenant quelques années (Flay, 1996; Viola and Flay, 2011a; Bot et al., 2014). Viola and Flay (2010) ont également effectué des mesures au réel, dites « *in situ* », de pression sur les voiles d'un voilier de 24'. Des comparaisons de ces mesures *in situ* ont été réalisées avec des essais en soufflerie et des simulations numériques dans (Viola and Flay, 2011b). Un voilier spécifiquement dédié aux mesures *in*

situ a aussi été développé et construit par l'équipe italienne du « Politecnico di Milano » (Fossati et al., 2015, 2016). Numériquement, les voiles sont aussi étudiées afin de mieux comprendre la complexité de leurs écoulements (Viola et al., 2014), d'analyser la stabilité de spinnakers (Durand et al., 2014), ou encore dans un but d'économie d'énergie propulsive de cargos (Viola et al., 2015). Enfin, très récemment, Viola and Arredondo-Galeana (2017) se sont intéressés au phénomène du « Leading Edge Vortex (LEV) » sur les spinnakers par des mesures hydrodynamiques. L'institut de recherche de l'École-Navale, l'« IRENav », travaille également sur ces problématiques de recherche en collaboration avec des industriels et universitaires. Un « J80 », voilier de 8 m, a été spécifiquement instrumenté pendant la thèse de Augier (2012) afin de réaliser des mesures expérimentales en mer pour des allures proches du vent. Les formes des voiles et les efforts au niveau des points de tire ont pu être mesurés dans différents cas de navigation, incluant des phases d'étude de la dynamique du voilier. En parallèle, un code numérique simulant les interactions fluide-structure sur les voiles a été développé par Durand (2012), dans le cadre d'une thèse financée par la compagnie « K-Epsilon » et menée en partenariat avec l'IRENav et l'École Centrale de Nantes. Des comparaisons avancées de mesures expérimentales à des simulations numériques, prenant en compte les interactions fluide-structure des voiles, ont ainsi été effectuées au cours de ces deux thèses. Deparday (2016) et Aubin (2017) se sont ensuite intéressés aux allures portantes dans leurs thèses, en étudiant les efforts et les formes de spinnakers du J80 par l'intermédiaire de mesures expérimentales *in situ* et en soufflerie. Un système de mesure par photogrammétrie a été développé dans (Deparday, 2016) pour reconstruire la forme du spinnaker soumis à des écoulements lors d'essais en mer. Dans (Aubin, 2017), des mesures ont été effectuées en soufflerie afin d'identifier les effets de « flapping » du spinnaker sur ses efforts aérodynamiques. Ces quatre dernières thèses ont permis de faire avancer la recherche sur la compréhension des phénomènes physiques rencontrés par le voilier. L'IRENav a donc acquis une expertise sur les effets influençant la performance d'un voilier, abordée de façon expérimentale et/ou numérique. L'optimisation de cette performance n'a quant à elle pas ou peu été étudiée jusqu'à aujourd'hui. En effet, la modélisation physique du voilier reste complexe, qu'elle soit expérimentale ou numérique, et elle demande encore un investissement important de la part des chercheurs du domaine. En plus d'être généralement coûteuse, cette modélisation de la physique du voilier est très souvent non linéaire et fait intervenir des problèmes d'interaction fluide-structure. L'optimisation de la performance est donc elle aussi très ardue de par la complexité de la physique mise en œuvre et du nombre important de paramètres pour représenter le fonctionnement du voilier; il s'agit pourtant d'un réel besoin pour le design d'un voilier de compétition. La résolution de ces problèmes d'optimisation nécessite alors des méthodes d'optimisation avancées afin de minimiser les coûts d'évaluations de la performance tout en étant capable de déterminer ses optimums globaux.

Dans ce contexte, l'IRENav s'est associé avec le « LIMSI-CNRS » (laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur) et l'« INRIA » (institut national de recherche en informatique et en automatique) dans le but de développer des méthodes d'optimisation avancées afin de réduire les coûts numériques de l'optimisation de la performance des voiliers. La rencontre de ces trois entités a permis d'aboutir au montage scientifique de cette thèse. Ainsi, ces travaux de thèse, réalisés physiquement à l'École-Navale, s'inscrivent dans la continuité des applications liées au domaine de la voile, mais ils sont une nouveauté pour

l'IRENav, dont les compétences scientifiques ne sont initialement pas du domaine des méthodes d'optimisation. Les objectifs de la thèse sont donc doubles :

- développer des méthodes et stratégies d'optimisation permettant d'obtenir un optimum pour un coût d'évaluation minimal ;
- valider ces développements dans le cadre d'applications réelles modélisant la performance des voiliers.

Les méthodes d'optimisation appliquées et développées dans cette thèse reposent sur l'utilisation de modèles mathématiques de substitution, appelés « méta-modèles », au problème d'optimisation direct d'une fonction objective (Büche et al., 2005; Laurenceau and Sagaut, 2008; Duvigneau et al., 2016). En effet, de par la complexité de la physique des voiliers, les évaluations de leur performance à optimiser, restent très coûteuses et rendent l'optimisation difficilement solvable en temps raisonnable. Le méta-modèle construit sur un ensemble fini d'observations de la fonction performance permet alors de réduire ce coût global d'optimisation. Cela est particulièrement vrai lorsque ses coûts de constructions sont bien inférieurs aux coûts d'évaluations de la performance. Les méta-modèles utilisés dans cette thèse ont la particularité d'avoir une vision statistique des évaluations, ce qui permet d'estimer les incertitudes de la prédiction de la performance donnée par le méta-modèle. L'optimisation par méta-modèles s'effectue alors itérativement, à l'aide d'un enrichissement séquentiel des observations de la performance, en prenant en compte la réponse prédictive de la performance ainsi que son incertitude. Cette approche séquentielle est associée à des mises à jour de la construction des méta-modèles, à chaque itération.

Le plan de ce mémoire de thèse est divisé en trois parties principales. La Partie I introduit le contexte général et spécifique du voilier. Le fonctionnement du voilier, ainsi que l'étude de sa performance sont détaillés. Les notions de la mécanique des fluides et d'interaction fluide-structure inhérentes à la physique du voilier sont également données dans cette Partie I. Enfin, quelques éléments concernant les essais expérimentaux et la modélisation numérique du voilier, permettant l'évaluation de sa performance, sont exposés.

Dans la Partie II, les méthodes d'optimisation employées dans cette thèse sont présentées. Spécifiquement, cela concerne l'utilisation de méta-modèles dans un processus d'optimisation itératif. La formulation de cette méthode aux cas d'optimisation sous contraintes, comprenant de possibles points non évaluables, est également traitée, de même que l'introduction de modèles de substitution, ayant différents niveaux de fidélité hiérarchisés, dans le processus d'optimisation itératif.

La Partie III, résume les contributions principales de cette thèse. Ces contributions sont présentées sous formes d'articles publiés ou soumis, au cours de la thèse, dans des revues scientifiques. Chaque contribution est exposée sous la forme d'un chapitre. Le Chapitre 7 présente une application originale de la méthode d'optimisation par méta-modèles publiée dans « *Journal of Fluids and Structures* » (Sacher et al., 2017c). Les réglages d'une grand-voile sont optimisés afin de maximiser une fonction performance. Le problème d'optimisation est résolu expérimentalement avec des mesures en soufflerie et numériquement avec des calculs d'interaction fluide-structure. Des comparaisons numériques-expérimentales sont réalisées avec une analyse de sensibilité. Cette étude de sensibilité permet d'analyser les variations de la fonction performance vis à vis des incertitudes

des essais expérimentaux. La robustesse du réglage optimum de la grand-voile est également traitée. Dans le Chapitre 8, l'utilisation d'une méthode permettant la prise en compte de points non évaluables dans l'optimisation par méta-modèles est exposée. Cette méthode est publiée dans « *Structural and Multidisciplinary Optimization* » (Sacher et al., 2018b). Des problèmes analytiques présentant des propriétés et niveaux de difficultés variés permettent de valider l'efficacité de cette méthode. Un cas d'optimisation réel, dans le contexte de la Coupe de l'America, concerne ensuite le Chapitre 9. Il porte sur une optimisation sous contraintes d'un hydrofoil flexible composé de paramètres géométriques et structurels. L'apport de la propriété flexible de l'hydrofoil est illustré à l'aide de comparaisons à un hydrofoil purement rigide et optimisé pour les mêmes conditions. L'application est publiée dans « *Ocean Engineering* » (Sacher et al., 2018a). Le Chapitre 10 contient des applications d'optimisation par méta-modèles ayant différents niveaux de fidélité hiérarchisés. Ces applications permettent d'illustrer l'utilisation de la méthode d'optimisation par méta-modèles multi-fidélité. L'optimisation des réglages d'un voilier est effectuée en faible et haute dimension avec des propriétés de fidélité variées. Des cas analytiques sont également utilisés afin de valider cette méthode, récemment soumise dans « *Computer Methods in Applied Mechanics and Engineering* ».

Finalement, la conclusion générale et quelques perspectives de ces travaux de thèse sont données à la fin du manuscrit.

PREMIÈRE PARTIE

CONTEXTE DU VOILIER

La vie est faite d'espérance.

ÉRIC TABARLY

CHAPITRE 1

Performance du voilier

Sommaire

1.1	Fonctionnement du voilier	10
1.1.1	Terminologie	10
1.1.2	Équilibre aéro-hydrodynamique	11
1.1.3	Vent apparent	12
1.2	Étude de la performance	15
1.2.1	Performance du voilier	15
1.2.2	Travaux scientifiques	17

L'étude de la performance du voilier est décrite dans ce Chapitre 1. Le voilier est un système physique complexe situé à l'interface de deux fluides, l'air et l'eau. Les voiles font office de « propulseur » captant l'énergie du vent (aérodynamique), qui est transmise à la carène (hydrodynamique) par l'intermédiaire du gréement, permettant ainsi le déplacement du voilier, lui-même soumis à la résistance de l'eau et de l'air : c'est donc un équilibre aéro-hydrodynamique et structurel, dont le fonctionnement est présenté dans la Section 1.1.

L'optimisation de la performance reste aujourd'hui très empirique et repose généralement sur les acquis et l'expérience des architectes/ingénieurs navals et maîtres voilier. Ceci est principalement dû à la complexité de l'étude des problèmes physiques du voilier, qui en plus d'être à l'interface de deux milieux fluides, fait intervenir des problèmes d'interactions fluide-structure dans ces deux milieux. En pratique, cela provient aussi du domaine spécifique de la voile où la confidentialité est généralement très présente. L'étude de la performance des voiliers est résumée dans la Section 1.2.

1.1 Fonctionnement du voilier

Dans cette Section, la terminologie générale du voilier est premièrement définie, les efforts aérodynamiques et hydrodynamiques auxquels est soumis le voilier sont ensuite présentés, et enfin, la propriété particulière du vent apparent vu par le voilier est illustrée.

1.1.1 Terminologie

Le domaine de la voile utilise un langage spécifique pour nommer chacun des éléments qui composent un voilier. La Figure 1.1 illustre la terminologie du voilier. Les voiles, le mât et l'ensemble des câbles qui composent le « gréement dormant » (câbles non ou peu réglés en navigation) et le « gréement courant » (câbles réglés en navigation, typiquement les câbles pour le réglages des voiles) sont représentés.

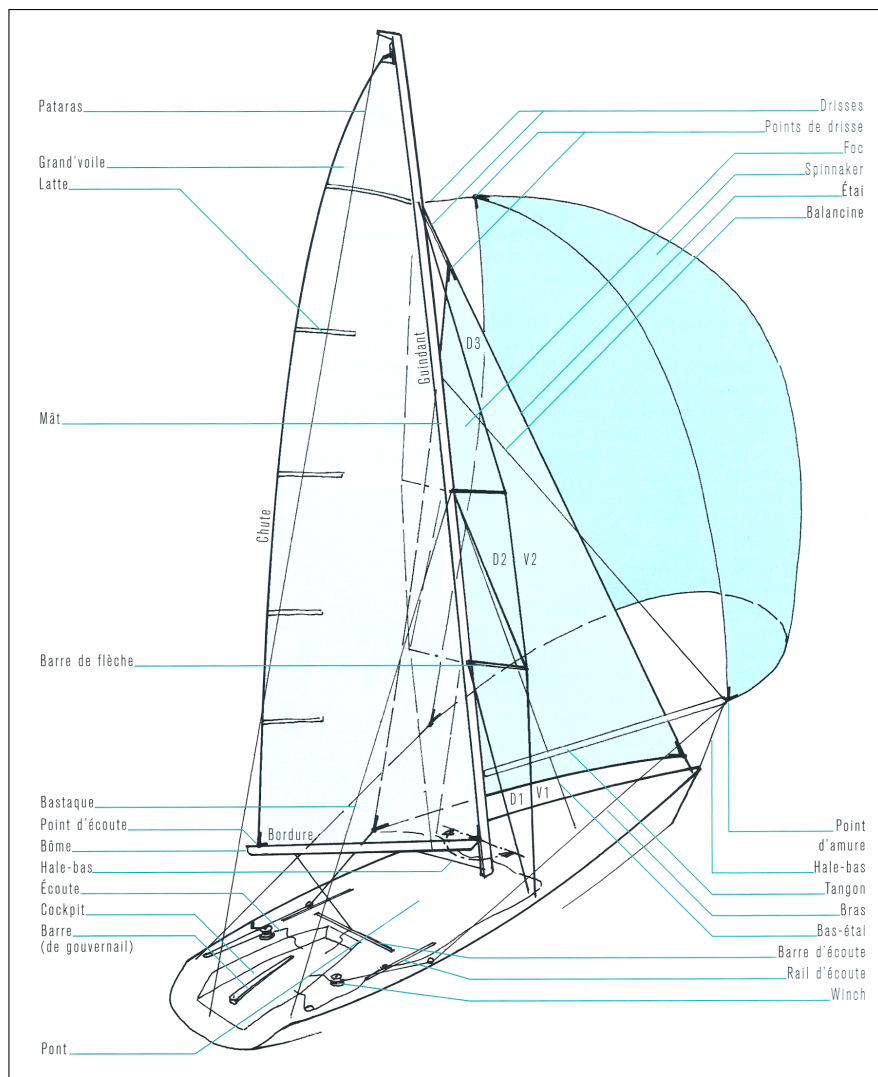


Figure 1.1 – La terminologie du voilier. Illustration provenant de (Paulet and Presles, 1998).

Le vocabulaire est également spécifique en ce qui concerne la position relative du voilier par rapport au vent. Cela s'appelle les « allures » et elles sont représentées sur la Figure 1.2. On distingue en particulier les deux modes de fonctionnements qui sont les cas « au vent », lorsque l'on navigue en direction du vent (angle du vent réel, ou « True Wind Angle (TWA) », $TWA < 90^\circ$) et « sous le vent », lorsque $TWA > 90^\circ$. On remarque sur cette Figure 1.2, que plus on remonte vers la direction du vent, plus l'angle entre le vent et la voile est fermé. Aux allures portantes ($TWA > 90^\circ$), la voile d'avant du voilier est généralement un gennaker ou un spinnaker, comme cela est illustré sur la Figure 1.2.

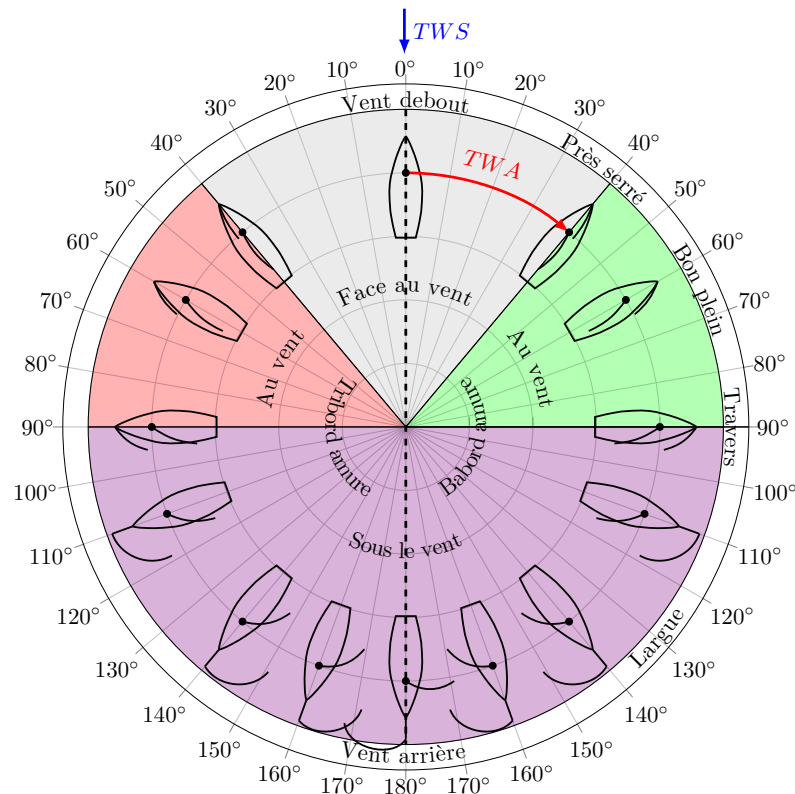


Figure 1.2 – Les différentes allures d'un voilier. Illustration originale dans (Aubin, 2017).

1.1.2 Équilibre aéro-hydrodynamique

Le voilier est soumis à des efforts aéro-dynamiques (le vent dans les voiles) et hydrodynamiques (l'eau sur la carène et les appendices). Lorsque sa vitesse et son attitude sont constantes, le voilier est à l'équilibre, et il y a alors une égalité des résultantes aéro-dynamique et hydrodynamique. L'équilibre aéro-hydrodynamique est décrit dans cette Section 1.1.2.

Le repère lié au voilier ainsi que les attitudes et les angles sont définis sur la Figure 1.3. L'origine du repère est généralement placé au centre du gravité, l'axe X est orienté suivant l'axe longitudinal du voilier (positif vers l'avant), l'axe Z est vertical (positif vers le haut) et l'axe Y complète le repère orthonormé. Les angles, vitesses de rotations et vitesses axiales sont aussi donnés pour les trois axes du repère sur la Figure 1.3.

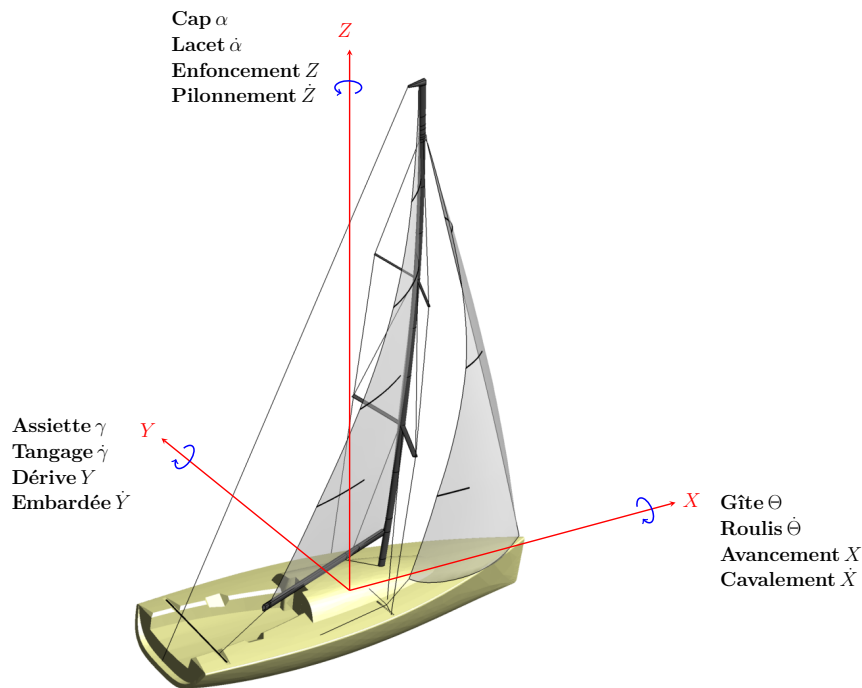


Figure 1.3 – Repère, attitudes et angles du voilier.

La Figure 1.4 présente l'équilibre aéro-hydrodynamique statique du voilier dans le plan (X,Y) , où il est supposé que les accélérations linéaires et angulaires sont nulles. La résultante aéro-dynamique provient du vent dans les voiles et elle peut se décomposer en un effort propulsif suivant X et un effort de dérive suivant Y . Le repère lié au vent (non illustré ici) est parfois utilisé pour décomposer la résultante aéro-dynamique en un effort de portance et de trainée. La résultante hydrodynamique est égale et opposée à la résultante aéro-dynamique. Les efforts hydrodynamiques proviennent de la carène et des appendices équipant le voilier (safran, dérive, hydrofoil, etc.). Le voilier progresse sur l'eau avec un angle de dérive, « leeway » en anglais, représenté sur la Figure 1.4 par λ . Cet angle est généralement inférieur à 5 deg et il correspond à l'angle d'attaque de l'eau vue par la carène et les appendices orientés suivant l'axe longitudinal du voilier. Une portance hydrodynamique est alors générée permettant un effort d'anti-dérive.

La Figure 1.5 illustre l'équilibre aéro-hydrodynamique statique du voilier dans le plan (Y,Z) , où les accélérations sont toujours nulles. Les points d'applications de la résultante aéro-dynamique (CE) et hydrodynamique (CLR) ont des altitudes différentes (voir Figure 1.5). Cela entraîne la création d'un couple de chavirement et donc la gîte du voilier. L'équilibre en gîte est assuré par la stabilité hydrostatique du voilier qui voit son centre de carène (CB) se déplacer sous le vent. Cela génère un couple de redressement qui équilibre l'angle de gîte du voilier.

1.1.3 Vent apparent

Le vent perçu par les voiles est appelé le « vent apparent », ou « Apparent Wind Speed (AWS) », et il se définit comme la somme vectorielle du « vent réel », ou « True Wind Speed (TWS) », et de l'opposée du vecteur

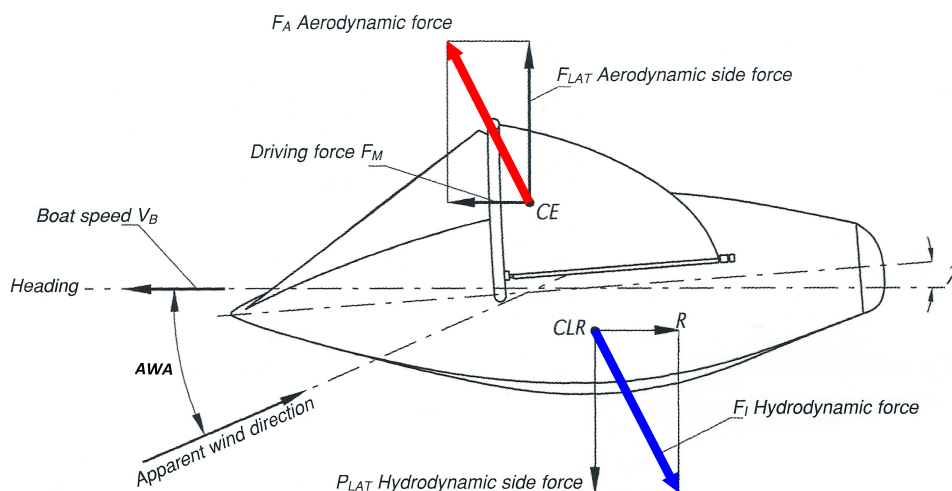


Figure 1.4 – Équilibre aéro-hydrodynamique du voilier dans le plan (X,Y). Illustration originale dans (Fossati, 2010).

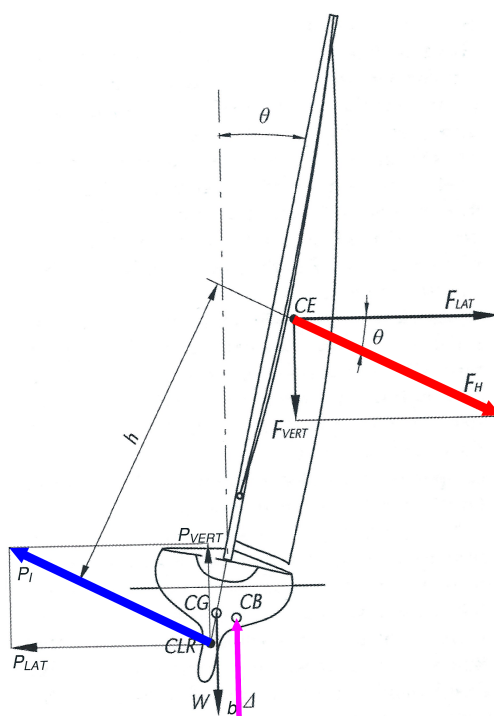


Figure 1.5 – Équilibre aéro-hydrodynamique du voilier dans le plan (Y,Z). Illustration originale dans (Fossati, 2010).

« vitesse bateau », ou « Boat Speed (BS) » : $\overrightarrow{AWS} = \overrightarrow{TWS} - \overrightarrow{BS}$. Cela s'appelle communément le « triangle des vitesses ». La Figure 1.6 illustre ce calcul vectoriel dans un cas de navigation au près. Les angles de vent réel et de « vent apparent », ou « Apparent Wind Angle (AWA) », sont également représentés.

La vitesse du vent réel vue par les voiles n'est pas constante en fonction de l'altitude. En effet, le voilier est situé dans la couche limite atmosphérique, comme montré sur la Figure 1.7(a), où la vitesse du vent réel augmente avec l'altitude. Ce gradient de vitesse peut être approché par une loi logarithmique (Hémon, 2006).

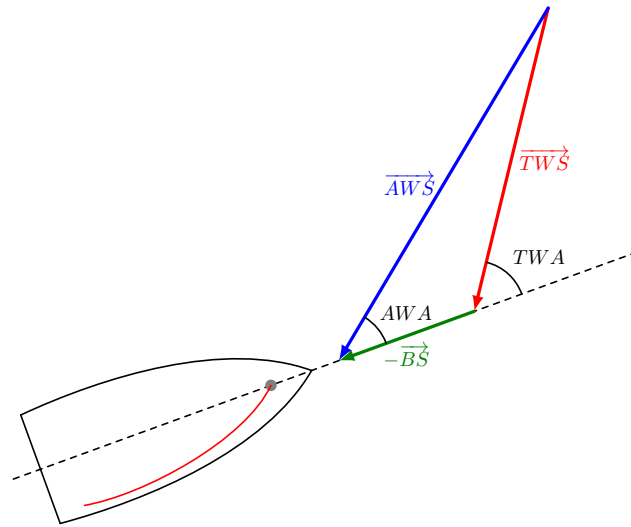


Figure 1.6 – Triangle des vitesses dans un cas de navigation au près ($\lambda = 0^\circ$).

La Figure 1.7(b) illustre l'effet du gradient de la vitesse du vent réel sur la distribution du vent apparent en fonction de l'altitude. En plus de varier en vitesse, le vent apparent est vrillé le long du mât. Cela implique que les voiles soient également vrillées pour s'adapter à l'évolution des angles du vent apparent. Sur un voilier en

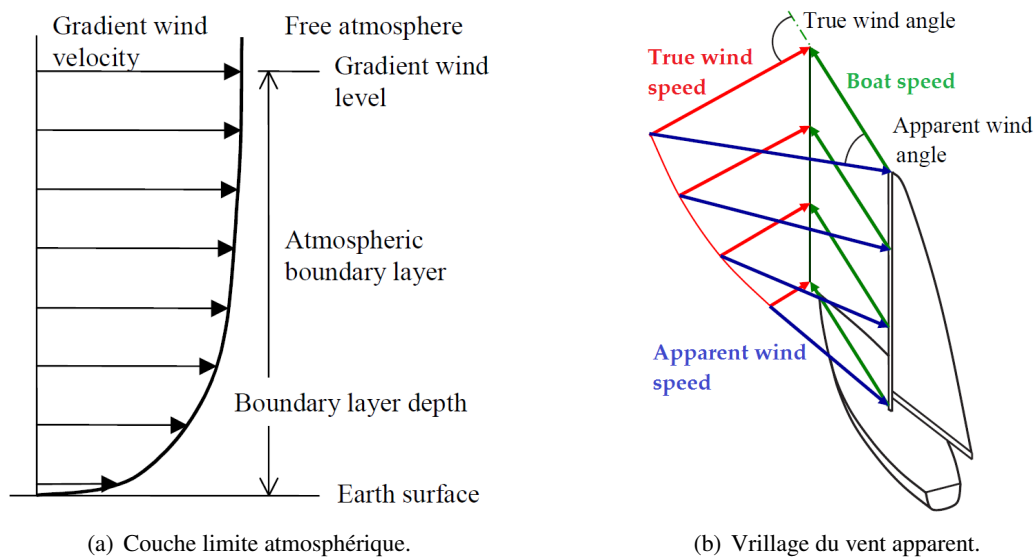


Figure 1.7 – Vrillage du vent apparent en fonction de l'altitude. Images originales dans (Hansen, 2006a).

navigation, le passage du vent apparent au vent réel est loin d'être aisé. En effet, le vent change continuellement en intensité et direction, et les attitudes du voilier sont très souvent instationnaires. Le calcul du vent réel doit alors prendre en considération les mouvements et la vitesse du voilier afin de les soustraire au vent apparent mesuré. La qualité de la reconstruction du vent réel est par conséquent fortement dépendante des capteurs embarqués et des algorithmes utilisés (Aubin, 2017; Kerhascoet, 2017).

1.2 Étude de la performance

Les principes de l'étude de la performance des voiliers sont présentés dans cette Section 1.2. On introduit en particulier les notions de « polaire » et de « programme de prévision de la vitesse », dans la Section 1.2.1, qui sont des outils d'analyse de la performance reposant sur la vitesse du voilier. Un résumé de travaux scientifiques portant sur l'analyse et l'optimisation de la performance des voiliers est également donné dans une seconde partie, à la Section 1.2.2.

1.2.1 Performance du voilier

L'étude de la performance d'un voilier passe généralement par l'analyse de sa vitesse à des points de fonctionnement préalablement choisis. Les « polaires de vitesses » sont des graphiques qui permettent de tracer la meilleure vitesse du voilier pour chaque angle et vitesse du vent réel. C'est une vision globale de la performance d'un voilier qui rend la comparaison entre plusieurs voiles, carènes, équipages, etc., possible. Les polaires de vitesses peuvent être prédites par des calculs ou bien construites directement sur l'eau. Dans une polaire, l'angle de vent réel est généralement compris dans l'intervalle $[30, 180]$ deg, ce qui implique d'adapter les réglages des voiles et du gréement, à chaque TWA et TWS, pour maximiser la vitesse du voilier. Un changement de voile est de plus généralement fait à $TWA \approx 90$ deg.

La Figure 1.8 présente un exemple de polaires réalisées pour une vitesse de vent réel et pour différents types de voiliers. La direction du vent est donnée par la verticale descendante du tracé et les allures correspondent aux traits obliques. Chaque courbe indique ici la vitesse d'un voilier (en kts) obtenue aux différentes allures et pour $TWS = 10$ kts. Les « VMG », de l'anglais « Velocity Made Good », sont également tracées pour le « Tornado » et le « Skiff » au près (en rouge) et au portant (en bleu) sur cette même Figure 1.8. La VMG peut se définir comme le meilleur gain au vent en terme de compromis cap/vitesse. Le Skiff est ainsi bien plus performant que le Tornado, en particulier au portant où sa VMG profite du départ au planning, comme le montre la singularité dans la courbe pour $TWA \approx 135$ deg.

Les polaires de vitesses peuvent aussi être calculées par des programmes informatiques, historiquement appelés « VPP » pour « Velocity Prediction Program ». Le schéma très simpliste de la vision d'un VPP est donné sur la Figure 1.9. Celui-ci comporte des entrées qui sont les paramètres du voilier (masse, dimensions, inerties, etc.) ainsi que le choix éventuel des modèles de prédiction des efforts aérodynamiques et hydrodynamiques et des degrés de liberté. Les conditions du vent réel (vitesse et angle) ainsi que l'état de mer sont aussi des entrées. L'objectif du VPP est alors de prédire la vitesse et les attitudes du voilier à l'équilibre aérohydrodynamique. Historiquement, les efforts aérodynamiques et hydrodynamiques étaient déterminés par des formulations analytiques ou sur la base d'essais expérimentaux (Van Oossanen, 1993). Les premiers VPP résolvaient l'équilibre du voilier en deux dimensions, pour la gîte et la vitesse, puis le nombre de degrés de liberté est monté jusqu'à six. Les récents développements des méthodes numériques en particulier dans le domaine de l'hydrodynamique (Queutey and Visonneau, 2007; Wackers et al., 2012) permettent aujourd'hui de construire des matrices d'efforts hydrodynamiques numériques préalablement calculées. Elles peuvent être très simples en ne dépendant que de la vitesse, ou au contraire très complètes lorsque les attitudes du voilier sont incluses ou

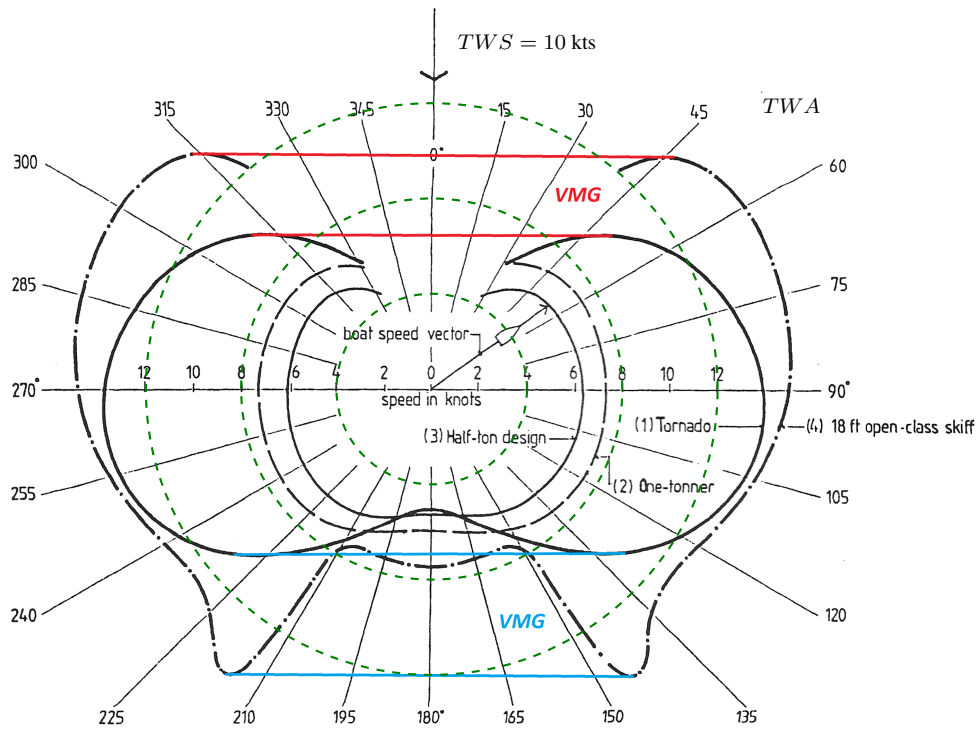


Figure 1.8 – Polaires de différents voiliers pour une vitesse du vent réel de 10 kts. Image originale dans (Garrett, 1996).

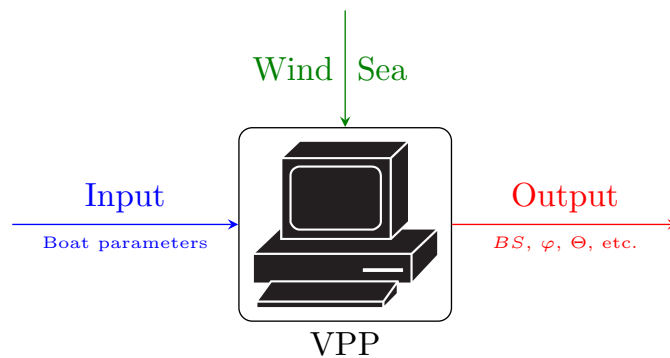


Figure 1.9 – Programme de prévision de la vitesse du voilier.

directement résolues. Les matrices aérodynamiques numériques sont aussi utilisées, principalement dans des cas simplifiés où les déformations des voiles ne sont pas prises en compte (Korpus, 2007). Dans ces derniers cas, le VPP est alors une simple résolution de l'équilibre aéro-hydrodynamique du voilier par une méthode de « Newton ». Récemment, Levin and Larsson (2017) ont aussi directement intégré des calculs hydrodynamiques dans un VPP. En pratique, il est préférable d'utiliser des matrices hydrodynamiques qui sont dans la plupart des cas facilement approchables (interpolations) par des polynômes ou des fonctions à base radiale. La réponse aérodynamique du voilier est quant à elle directement introduite dans le VPP car elle est plus complexe (non-linéaire) et dépendante d'un nombre important de réglages lorsque l'on modélise la structure complète

des voiles.

Finalement, lorsque l'on optimise la performance d'un voilier, le VPP est inclus dans une boucle d'optimisation itérative qui peut comporter plusieurs boucles imbriquées. Par exemple dans le cas où l'on cherche à optimiser les paramètres d'une voile (structure, forme, etc.), c'est bien une fonction objective basée sur la vitesse du voilier que l'on cherche à minimiser ou maximiser. Cependant, la réponse de la fonction objective, qui est dépendante des paramètres d'optimisation de la voile, est aussi dépendante des réglages du gréement, qui se doivent d'être optimaux pour chaque jeu de paramètres de la voile testé. Certaines optimisations peuvent également inclure des simulateurs de régates et/ou des notions tactiques (Tagliaferri et al., 2014).

1.2.2 Travaux scientifiques

On peut trouver de nombreux travaux scientifiques étudiant la physique et la performance des voiliers dans la bibliographie. Les études avancées où des problématiques d'interactions fluide-structure sont considérées et qui portent sur l'optimisation par méta-modèles de la performance des voiliers sont quant à elles bien plus rares. Un résumé non-exhaustif de publications se rapportant à la notion de la performance des voiliers est donné ci-après.

Le design structurel des voiles pour la Coupe de l'America de 1995 a été introduit dans (Fallow, 1996), où les modèles fluides sont basés sur des formulations empiriques et des essais en soufflerie. La formulation du problème d'optimisation n'est cependant pas détaillée, tout comme la méthode d'optimisation permettant la définition du design optimal des voiles.

Day (1996) propose une optimisation de la distribution de portance sur la hauteur du mât, afin de maximiser la vitesse d'un voilier. Le modèle aérodynamique repose sur une ligne portante et des données expérimentales nourrissent le modèle hydrodynamique. Les déformations de la structure des voiles ne sont ici pas considérées et un algorithme génétique est utilisé pour résoudre le problème d'optimisation.

Des simulations numériques sur des voiles prenant en compte des effets d'interactions fluide-structure sont présentées par Charvet et al. (1996). Le modèle fluide repose sur la théorie des écoulements potentiels et les auteurs présentent l'intérêt d'utiliser un sillage particulière en illustrant les effets d'interactions entre deux voiles. Les influences de la surface de la mer et la non-homogénéité du gradient de vent incident dans les voiles y sont également discutées. La notion d'optimisation ne fait pas partie du corps de cette publication.

L'utilité de la simulation numérique des écoulements dans le design des voiliers de la Coupe de l'America est également traitée dans (Rosen et al., 2000; DeBord et al., 2002). L'accent est particulièrement mis sur les applications hydrodynamiques (carène et appendices) qui nécessitent des essais expérimentaux coûteux en l'absence de calculs numériques. Les discussions restent cependant très générales dans ces deux papiers et la formulation mathématique du problème d'optimisation de la performance n'est pas exposée.

Les développements des méthodes d'optimisations appliquées à la performance des voiliers jusqu'aux années 2000 sont présentés dans (Philpott and Mason, 2002). La méthode du Simplex (Nelder and Mead, 1965) est proposée pour maximiser l'effort propulsif d'un voilier dans une soufflerie en optimisant les réglages des voiles. La thématique des incertitudes du vent (angles et vitesses) et donc de l'optimisation robuste des voiliers est aussi abordée. Le papier ne propose pas en soi des développements de méthodes d'optimisation avancées

pour améliorer la performance des voiliers. Il correspond plutôt à une vision des auteurs sur la réticence des designers à utiliser des méthodes mathématiques résolvant les problèmes d'optimisation des voiliers.

Doyle et al. (2002) et Rousselon (2008) s'intéressent à l'optimisation des formes de voiles soumises à des écoulements calculés numériquement. Le couplage du fluide aux structures des voiles n'est pas modélisé dans ces deux études. La performance à optimiser est bien définie et les auteurs utilisent des méthodes d'optimisation basées sur les calculs de gradients et stratégies d'évolution dans (Doyle et al., 2002), et sur un algorithme génétique dans (Rousselon, 2008). Le nombre d'évaluations des fonctions objectives sont par conséquent important dans ces deux études. Chapin et al. (2011) ont également réalisés des calculs d'interactions fluide-structure sur des voiles d'un 18 pieds. Un modèle fluide visqueux est couplé au code structure par un couplage faible et d'intéressantes études d'optimisation de formes de voiles sont effectuées. Dans (Lombardi et al., 2012), l'application de la simulation numérique des voiliers aux problèmes dynamiques, d'interactions fluide-structure et d'optimisations de formes est résumée. La description des méthodes d'optimisation reste « simpliste » car elles sont uniquement basée sur les méthodes à gradients et les algorithmes génétiques.

Les travaux de Ranzenbach et al. (2013) et Durand et al. (2014) font aujourd'hui parti des simulations d'interactions fluide-structure les plus avancées sur les voiles. Dans (Durand et al., 2014), les calculs sont réalisés sur deux spinnakers afin de comparer leurs performances respectives en terme de stabilité. Viola et al. (2014) ont aussi effectué des simulations numériques avancées (modèles fluide RANS et DES) sur un spinnaker rigide. Il n'est clairement pas question d'optimisation dans cette dernière publication où les calculs DES prennent environ 12 jours pour converger.

D'un point de vue expérimental, les études sur l'aérodynamique des voiles sont aussi très actives aujourd'hui. Motta et al. (2014) se sont par exemple intéressés aux efforts produits par un spinnaker lors d'essais en navigation au réel. En particulier, des mesures simultanées de pressions et de la forme du spinnaker ont été réalisées. Des essais en soufflerie ont aussi permis d'identifier les effets du « flapping » sur les efforts propulsifs de différents spinnakers dans (Aubin et al., 2017). Récemment, Persson et al. (2017) ont aussi comparé expérimentalement les performances hydrodynamiques de 7 safrans de « Finn ». Ces trois dernières études apportent des explications sur les phénomènes physiques des écoulements du spinnaker ou du safran, mais l'optimisation propre de la performance n'est pas traitée.

On présente finalement ci-après quelques récents travaux où la performance des voiliers est optimisée et/ou analysée par l'utilisation de méta-modèles. de Baar et al. (2015) proposent une étude de quantification d'incertitudes appliquée à des simulations hydrodynamiques d'une carène. Ils s'intéressent aux incertitudes provenant de la masse du voilier, de la position du centre de gravité et des effets de blocage lors d'essais en bassin de traction. Des méta-modèles Gaussiens à deux niveaux de fidélité sont utilisés pour estimer ces incertitudes. L'apport de la formulation en multi-fidélité pour réduire le coût calcul est clairement illustré dans ces travaux. Les auteurs mettent cependant l'accent dans la conclusion de l'article sur la question du choix pour l'utilisateur quant aux nombres et localisations des évaluations par niveau permettant de contrôler au mieux la précision des méta-modèles.

Dans (Lee et al., 2016), une optimisation par méta-modèles de trois ailes rigides dont les performances aérodynamiques sont estimées par des simulations numériques bidimensionnelles est menée. L'objectif dans cette

étude est de maximiser l'effort propulsif total produit par les trois ailes en optimisant 9 variables contrôlant les géométries des volets équipant les ailes ainsi que leurs angles d'attaque respectifs. Le design optimal des ailes permet d'augmenter l'effort propulsif d'un cas de référence non optimisé d'environ 14-22%. Des validations numériques sont aussi réalisées par des calculs tridimensionnels, qui conduisent à un gain de l'effort propulsif d'environ 10-17%.

Enfin, très récemment, **Bonfiglio et al. (2018)** ont mené une étude géométrique tridimensionnelle d'un hydrofoil super-cavitant par des simulations numériques incluant la modélisation de la cavitation dans un modèle fluide visqueux. La géométrie est contrôlée par 17 variables d'optimisation définissant une section unitaire (15 variables) et la forme tridimensionnelle de l'hydrofoil suivant sont envergure (2 variables). L'objectif est de maximiser le rapport portance/trainée de l'hydrofoil sous deux types de contraintes. La valeur du coefficient de portance est doublement bornée et trois épaisseurs de la poche de cavitation, situées à 25%, 50% et 75% de l'envergure, ont une même borne inférieure. Les travaux réalisés ici se définissent comme une optimisation multi-fidélité d'un hydrofoil super-cavitant. Il y a bien la notion de multi-fidélité car deux niveaux de modélisation numérique sont utilisés pour évaluer les performances de l'hydrofoil. Cependant, cette étude ne correspond pas à une réelle optimisation multi-fidélité par méta-modèles. En effet, les auteurs construisent des méta-modèles multi-fidélité à partir de deux plans d'expériences (un par niveaux) extrêmement denses : respectivement 2000 et 1000 observations. Ils recherchent ensuite le jeu de paramètres maximisant l'amélioration espérée et contrainte de la performance prédite par le méta-modèle de haute fidélité. La performance de cet hydrofoil est améliorée de 7% par rapport à une géométrie de référence. Il n'y a donc pas un réel processus itératif d'optimisation par méta-modèles dans ces travaux. De plus, la notion du choix du niveau de fidélité dans l'évaluation de la performance d'un nouvel hydrofoil n'est pas considérée.

Modéliser la physique du voilier

Sommaire

2.1	Physique du voilier	22
2.1.1	Mécanique des fluides	22
2.1.1.1	Milieu continu	22
2.1.1.2	Paramètres de similitude	22
2.1.1.3	Viscosité et notions de couche limite	23
2.1.1.4	Principes de conservation	24
2.1.1.5	Équations d'Euler et de Navier-Stokes	27
2.1.1.6	Écoulement autour d'un corps portant	27
2.1.2	Interaction fluide-structure	29
2.2	Outils numériques	31
2.2.1	Mécanique des fluides numérique	31
2.2.2	Simulations des voiles en IFS	35
2.3	Essais expérimentaux	37
2.3.1	Hydrodynamique en bassin	37
2.3.2	Aérodynamique en soufflerie	40
2.3.3	Mesure <i>in situ</i>	41

L'objectif du Chapitre 2 est de présenter les théories et outils permettant la modélisation de la physique du voilier. Le voilier évoluant à l'interface de deux milieux fluides, qui sont l'air et l'eau, sa modélisation fera alors intervenir des notions de mécanique des fluides et d'interactions fluide-structure (en particulier pour les voiles), qui sont décrites dans la Section 2.1. Dans la Section 2.2, sont introduits les outils numériques permettant de modéliser numériquement la physique du voilier. Enfin, dans la Section 2.3, sont présentés les moyens expérimentaux couramment utilisés pour étudier l'aérodynamique et l'hydrodynamique des voiliers.

2.1 Physique du voilier

On présente dans une première Section 2.1.1 la théorie de la mécanique des fluides dans un contexte général qui peut s'appliquer à l'aérodynamique ou l'hydrodynamique du voilier. Les notions données dans cette section sont une sous-partie résumée du vaste domaine de la mécanique des fluides, que l'on pense nécessaire à la compréhension des différents problèmes physiques appliqués au cours de cette thèse, dans le cadre de l'optimisation. Dans la Section 2.1.2 sont brièvement introduites les notions d'interactions fluide-structure qui sont utiles pour la compréhension des phénomènes physiques s'exerçant sur les voiles d'un voilier.

2.1.1 Mécanique des fluides

2.1.1.1 Milieu continu

On fait ici l'hypothèse que le milieu fluide est un milieu continu. Cela signifie que les propriétés du fluide sont des fonctions continues des variables d'espace (x, y, z) et du temps t . On peut donc appliquer des outils mathématiques reposant sur les fonctions continues et/ou dérivables. Pour rappel, un milieu fluide peut être considéré comme continu si le nombre de « Knudsen » est très inférieur à l'unité : $K_n = \frac{\lambda}{L} \leq 0.001$, où λ se définit comme le libre parcours moyen des particules et L est la longueur de référence de l'objet considéré. Ce régime continu permet en l'occurrence d'appliquer les équations d'« Euler » et de « Navier-Stokes » qui sont introduites par la suite.

2.1.1.2 Paramètres de similitude

L'objectif est de définir des paramètres de similitude pertinents dont vont dépendre les solutions. Soient L_r , V_r , ρ_r , p_r , T_r , μ_r , λ_r et t_r les grandeurs de référence, l'équation de la quantité de mouvement (détaillée par la suite) s'écrit alors

$$\left(\frac{\rho_r V_r}{t_r}\right) \bar{\rho} \frac{\partial \vec{V}}{\partial \bar{t}} + \left(\frac{\rho_r V_r^2}{L_r}\right) \bar{\rho} \vec{\nabla} \vec{V} \cdot \vec{V} = -\left(\frac{p_r}{L_r}\right) \vec{\nabla} \bar{p} + \left(\frac{\mu_r V_r}{L_r^2}\right) \vec{\nabla} \cdot \vec{\Sigma}_v + (\rho_r) \bar{\rho} \vec{F}, \quad (2.1.1)$$

avec, $\bar{x} = \frac{x}{L_r}$, $\bar{t} = \frac{t}{t_r}$, $\bar{V} = \frac{V}{V_r}$, $\bar{\rho} = \frac{\rho}{\rho_r}$, etc. En divisant l'expression (2.1.1) par la quantité $\frac{\rho_r V_r^2}{L_r}$, on obtient

$$\omega \bar{\rho} \frac{\partial \vec{V}}{\partial \bar{t}} + \bar{\rho} \vec{\nabla} \vec{V} \cdot \vec{V} = -\frac{1}{\gamma M_a^2} \vec{\nabla} \bar{p} + \frac{1}{Re} \vec{\nabla} \cdot \vec{\Sigma}_v + \frac{\bar{\rho}}{Fr^2} \frac{\vec{F}}{g}, \quad (2.1.2)$$

où l'on a introduit quatre paramètres sans dimension qui sont :

- La fréquence réduite $\omega = \frac{L_r}{V_r t_r}$ (nombre de « Strouhal » $S_t = \frac{f L_r}{V_r}$), qui représente les échelles de mouvements instationnaires périodiques ;
- Le nombre de « Froude » $Fr = \frac{V_r}{\sqrt{g L_r}}$, qui caractérise le rapport force d'inertie / force de gravité, et que l'on retrouve principalement dans l'hydrodynamique pour l'étude des carènes par exemple ;

- Le nombre de « Mach » $M_a = \frac{V_r}{a_r}$ ($a_r = \sqrt{\gamma \frac{p_r}{\rho_r}}$, célérité du son en gaz parfait), mesurant le rapport force liée au mouvement / compressibilité du fluide, qui est très utilisé dans l'aéronautique ;
- Le nombre de « Reynolds » $R_e = \frac{\rho_r V_r L_r}{\mu_r}$, représentant le rapport force d'inertie / force de viscosité, qui permet de caractériser la nature du régime d'un écoulement (laminaire, transitoire, turbulent), et que l'on retrouve dans la plupart des travaux de mécanique des fluides.

Ces paramètres de similitude sont particulièrement importants lors des essais expérimentaux. En effet, les résultats de mesures expérimentales ne sont transposables au réel que si les données définissant les problèmes posés satisfont les règles de similitudes géométrique, cinématique et dynamique.

En mécanique des fluides, la similitude de Reynolds et celle de Froude sont essentielles. Il est cependant bien connu que le respect parfait de ces deux similitudes est incompatible. En effet, le respect du nombre de Reynolds entraîne $R_{e1} = R_{e2} \Rightarrow \frac{V_{r2}}{V_{r1}} = \frac{L_{r1}}{L_{r2}}$ et celui du nombre de Froude $F_{r1} = F_{r2} \Rightarrow \frac{V_{r2}}{V_{r1}} = \sqrt{\frac{L_{r2}}{L_{r1}}}$, ce qui implique $L_{r1} = L_{r2}$ et rend impossible un essai expérimental à une échelle réduite. Dans la pratique, seule l'une de ces deux similitudes est respectée. Par exemple, lors d'un essai de résistance à l'avancement d'une carène dans un bassin de traction, on respecte généralement la similitude de Froude, car le comportement de la surface libre (force de gravité) est important, et on essaie de modéliser le régime d'écoulement au réel en jouant sur l'intensité turbulente du fluide amont ou bien en ajoutant des bandes de rugosité (écoulement turbulent) sur la carène proche de la proue.

Lorsque l'on s'intéresse aux problèmes d'interaction fluide-structure, il est également important de considérer le nombre de Cauchy. Ce nombre sans dimension s'écrit $C_a = \frac{\rho_r V_r^2}{E_r}$, où E_r est le module d'élasticité, et il caractérise le rapport entre les forces d'inertie et les forces élastiques. Pour des voiles, une faible valeur de C_a signifie par exemple, que l'écoulement fluide génère de faibles déformations des tissus.

2.1.1.3 Viscosité et notions de couche limite

Soit un fluide en mouvement. On considère la normale à la trajectoire $\vec{n}$, un parallélépipède élémentaire se déplaçant à la vitesse V , ainsi qu'un même élément adjacent à la vitesse $V + dV$. Ce dernier parallélépipède exerce sur le premier une action de contact qui comporte, dans le cas d'un fluide réel, une composante normale $df_n = p dS$ et une composante tangentielle df_t qui est due à la viscosité du fluide. Elle s'écrit comme $df_t = \mu \frac{\partial V}{\partial n} dS$, où μ est la viscosité dynamique qui dépend du fluide et de la température.

La couche limite est une zone d'interface mince entre un corps et un fluide, ou entre deux fluides, lorsqu'un mouvement relatif entre ces deux éléments est présent. De même que pour deux corps qui frottent l'un contre l'autre font apparaître une force de frottement, on observe le même phénomène, provenant des effets visqueux, dans la couche limite. La Figure 2.1(a) illustre l'évolution de la vitesse d'un fluide réel en mouvement le long d'une normale y à une paroi. Au niveau de la paroi il y a adhérence du fluide ($V = 0$), ensuite la variation de vitesse $\frac{\partial V}{\partial y}$ augmente et donne lieu à des frottements importants, finalement à partir d'une distance δ , appelée « épaisseur de couche limite », les gradients sont assez faibles pour que les frottements soient négligeables. Sur la Figure 2.1(b), on présente cette même évolution des vitesses dans le cas où le fluide est supposé « parfait »,

en d'autres termes, lorsque la viscosité est négligée. Dans ce cas, il y a une condition de glissement au niveau de la paroi.

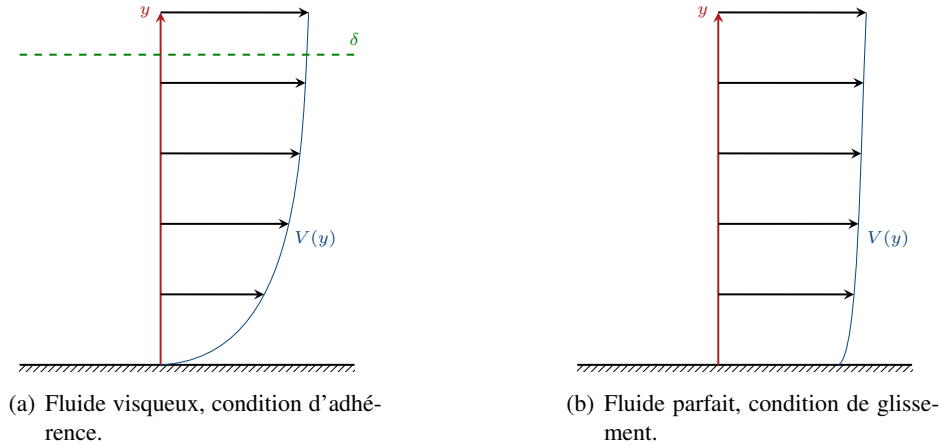


Figure 2.1 – Profils de vitesses au voisinage d'une paroi.

La description de la couche limite dans le domaine de la mécanique des fluides se fait généralement par des grandeurs caractéristiques non développées ici et nommées : « épaisseur de déplacement », « épaisseur de quantité de mouvement » et « épaisseur d'énergie cinétique ». Ces grandeurs caractérisent respectivement le déficit de débit, de quantité de mouvement et d'énergie cinétique dus à la viscosité dans la couche limite. Une couche limite est aussi caractérisée par la nature de l'écoulement : laminaire, transitoire ou turbulent. L'expérience de Reynolds permet d'illustrer ces différentes natures de l'écoulement. Elle consiste en une conduite circulaire dans laquelle un fluide visqueux est en mouvement à une vitesse V constante. A faible vitesse, les filets fluides observés sont parallèles et stables, on dit que l'écoulement est laminaire. Si on augmente la vitesse de l'écoulement, celui-ci reste laminaire jusqu'à une certaine vitesse critique V_c pour laquelle des instabilités apparaissent et rendent l'écoulement instationnaire. Le régime de l'écoulement est alors dit turbulent. Cette nature de l'écoulement (laminaire ou turbulent) implique des différences sur les profils de vitesse, sur le frottement à la paroi et sur l'acoustique de l'écoulement. En particulier, le régime turbulent augmente les quantités respectives de ces grandeurs.

2.1.1.4 Principes de conservation

En mécanique des fluides, il est commun de travailler en un point donné (description locale) où le mouvement est décrit par un système d'équations aux dérivées partielles, ou bien sur un volume de fluide, aussi appelé « volume de contrôle », où le mouvement est décrit par des équations intégrales (description globale). Une illustration de volume de contrôle V contenu dans un domaine fluide D est donnée par la Figure 2.2. Soit une quantité de matière donnée, de masse m , comprise dans un volume de contrôle V . On peut énoncer le principe de conservation d'une propriété extensive B (masse, quantité de mouvement, énergie) par le théorème

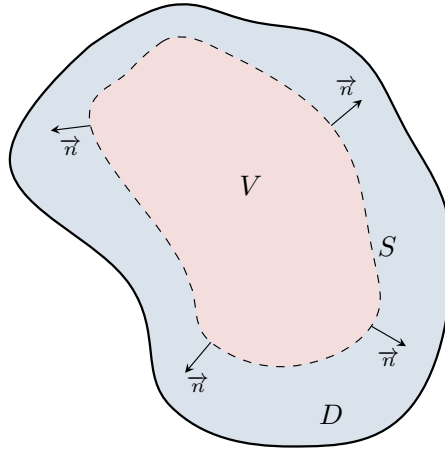


Figure 2.2 – Volume de contrôle V contenu dans un domaine fluide D .

de transport de Reynolds comme

$$\frac{dB}{dt} = \frac{d}{dt} \int_V \rho b dV + \int_S \rho b (\vec{V} \cdot \vec{n}) dS, \quad (2.1.3)$$

où $b = \frac{B}{m}$ est une propriété intensive, ρ est la masse volumique et $\vec{V}$ est la vitesse de déplacement du volume de contrôle. A partir du théorème de la divergence, où $\int_S \rho b (\vec{V} \cdot \vec{n}) dS = \int_V \nabla \cdot (\rho b \vec{V}) dV$, et pour un volume de contrôle de dimensions infinitésimales, le théorème (2.1.3) caractérise localement le principe de conservation et devient

$$\frac{d}{dt}(\rho b) = \frac{\partial}{\partial t}(\rho b) + \nabla \cdot (\rho b \vec{V}). \quad (2.1.4)$$

Il en découle alors le principe de conservation de la masse ($B = m$ et $b = 1$) qui s'écrit

$$\frac{dm}{dt} = 0, \quad (2.1.5a)$$

$$= \frac{d}{dt} \int_V \rho dV + \int_S \rho (\vec{V} \cdot \vec{n}) dS. \quad (2.1.5b)$$

De même, pour une particule infinitésimale, le principe de conservation de la masse devient

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{V}) = 0, \quad (2.1.6)$$

et se développe comme

$$\frac{\partial \rho}{\partial t} + \vec{V} \cdot \nabla \rho + \rho (\nabla \cdot \vec{V}) = \frac{D\rho}{Dt} + \rho (\nabla \cdot \vec{V}) = 0. \quad (2.1.7)$$

Pour un fluide incompressible, une particule fluide conserve sa masse volumique au cours de son mouvement. On a donc $\frac{D\rho}{Dt} = 0$ et le principe de conservation de la masse s'écrit simplement

$$\nabla \cdot \vec{V} = 0. \quad (2.1.8)$$

Le principe de conservation de la quantité de mouvement ($B = m\vec{V}$ et $b = \frac{B}{m} = \vec{V}$) s'écrit à partir de la seconde loi de Newton,

$$\frac{d}{dt}(m\vec{V}) = m\vec{a} = \sum \vec{F}_{\text{ext}}. \quad (2.1.9)$$

Ainsi,

$$\frac{d}{dt}(m\vec{V}) = \sum \vec{F}_{\text{volume}} + \sum \vec{F}_{\text{contact}}, \quad (2.1.10a)$$

$$= \int_V \frac{\partial}{\partial t}(\rho\vec{V})dV + \int_S \rho\vec{V}(\vec{V} \cdot \vec{n})dS, \quad (2.1.10b)$$

que l'on développe en

$$\frac{d}{dt}(m\vec{V}) = \int_V \rho \left(\frac{\partial \vec{V}}{\partial t} + \vec{V} \cdot \nabla \vec{V} \right) dV, \quad (2.1.11a)$$

$$= \int_V \rho \frac{D\vec{V}}{Dt} dV. \quad (2.1.11b)$$

La formulation locale des équations de la conservation de la quantité de mouvement s'écrit

$$\frac{\partial}{\partial t}(\rho\vec{V}) + \nabla \cdot (\rho\vec{V} \otimes \vec{V}) = \rho\vec{F} + \nabla \cdot \vec{\Sigma}, \quad (2.1.12)$$

où l'on a $\vec{F}$, la résultante des forces massiques, $\vec{\Sigma}$, le tenseur des forces de contact qui vaut $\vec{\Sigma}_p + \vec{\Sigma}_v$, avec le tenseur de pression $\vec{\Sigma}_p = -p\vec{I}$ et le tenseur des contraintes visqueuses (fluide newtonien) qui vaut $\vec{\Sigma}_v = \lambda \nabla \cdot (\vec{V}) \cdot \vec{I} + 2\mu \vec{d}$. Enfin, μ est la viscosité dynamique, $\vec{d} = \frac{1}{2}(\nabla \vec{V} + \nabla \vec{V}^T)$ et λ est la seconde viscosité (dimension de μ) qui n'est généralement pas considérée, excepté dans le domaine de l'acoustique.

Il existe également le principe de conservation de l'énergie totale, $E = e + \frac{V^2}{2}$ (interne + cinétique), que l'on présente ici directement sous sa formulation locale par

$$\frac{\partial}{\partial t}(\rho E) = \rho \vec{F} \cdot \vec{V} + \nabla \cdot (\vec{\Sigma} \cdot \vec{V}) + \rho \phi_v - \nabla \cdot \vec{\varphi}_c, \quad (2.1.13)$$

où $\rho \phi_v$ et $\vec{\varphi}_c$ sont les apports de chaleur massique et surfacique (conduction).

2.1.1.5 Équations d'Euler et de Navier-Stokes

Les équations d'Euler et de Navier-Stokes sont des équations aux dérivées partielles non linéaires qui permettent de décrire la physique des écoulements des fluides, qu'ils soient des liquides ou bien des gaz, dans le cadre de l'approximation des milieux continus.

Les équations de Navier-Stokes ont en fait déjà été introduites par les expressions (2.1.6), (2.1.12) et (2.1.13). Les équations d'Euler s'obtiennent en appliquant les hypothèses de fluide parfait (viscosité négligée) et en supposant qu'il n'y a pas d'échange de chaleur avec l'extérieur. Elle s'écrivent comme

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{V}) = 0, \quad (2.1.14)$$

$$\frac{\partial}{\partial t}(\rho \vec{V}) + \nabla \cdot (\rho \vec{V} \otimes \vec{V}) = \rho \vec{F} + \nabla \cdot (-p \vec{I}), \quad (2.1.15)$$

$$\frac{\partial}{\partial t}(\rho E) = \frac{\partial p}{\partial t} + \rho \vec{F} \cdot \vec{V}. \quad (2.1.16)$$

2.1.1.6 Écoulement autour d'un corps portant

Les forces exercées par un écoulement autour d'un corps portant sont principalement de deux natures :

- les forces de pression p , normales à la paroi ;
- les forces de frottement τ , tangentes à la paroi.

L'intégration de ces efforts le long des parois du corps portant donne lieu à la résultante R (voir Figure 2.3(a)), qui par définition se décompose généralement en une force de portance L et une force de traînée D , respectivement perpendiculaire et parallèle à la direction de l'écoulement incident. On rencontre parfois la convention locale au corps portant lorsque R est décomposée en F_y et F_x , comme le montre la Figure 2.3(a). Le calcul d'intégration des efforts élémentaires sur le corps portant se fait assez simplement en définissant, comme montré sur la Figure 2.3(b), une abscisse curviligne s_u pour l'extrados, s_l pour l'intrados, ainsi que l'angle d'inclinaison θ .

Ainsi, F_y et F_x sont déterminées pour une envergure unitaire par

$$F_y = - \int_{LE}^{TE} (p_u \cos \theta + \tau_u \sin \theta) ds_u + \int_{LE}^{TE} (p_l \cos \theta - \tau_l \sin \theta) ds_l, \quad (2.1.17)$$

$$F_x = \int_{LE}^{TE} (-p_u \sin \theta + \tau_u \cos \theta) ds_u + \int_{LE}^{TE} (p_l \sin \theta + \tau_l \cos \theta) ds_l. \quad (2.1.18)$$

Finalement, la portance et la traînée sont données par

$$L = F_y \cos \alpha - F_x \sin \alpha, \quad (2.1.19)$$

$$D = F_y \sin \alpha + F_x \cos \alpha, \quad (2.1.20)$$

où α est l'angle d'attaque du corps portant. En pratique, on utilise des coefficients de portance et de traînée

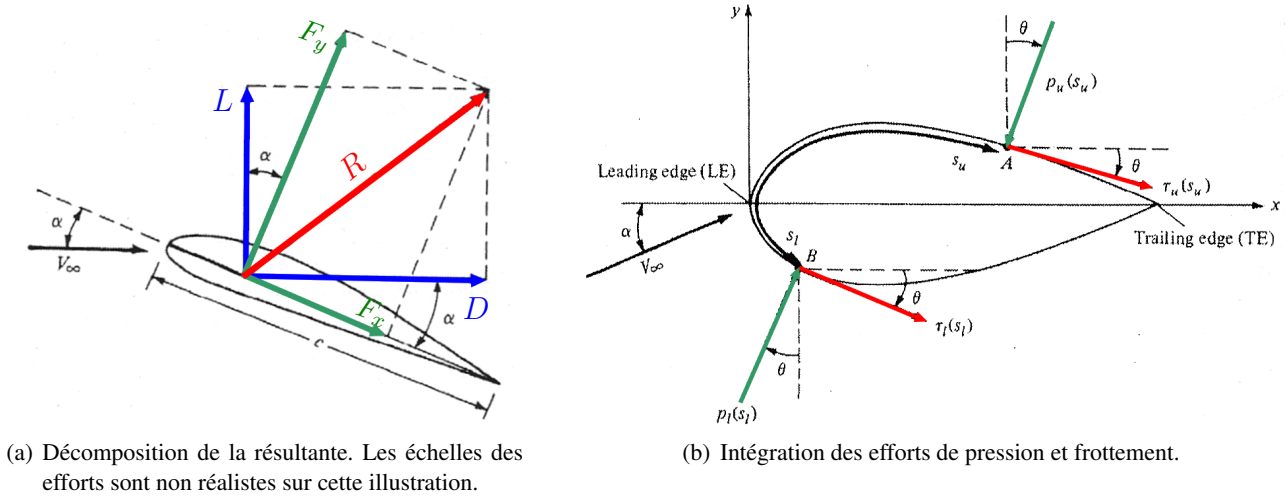


Figure 2.3 – Efforts sur un corps portant soumis à un écoulement.

adimensionnels qui s'écrivent

$$C_L = \frac{L}{\frac{1}{2}\rho_r V_r^2 S_r}, \quad (2.1.21)$$

$$C_D = \frac{D}{\frac{1}{2}\rho_r V_r^2 S_r}, \quad (2.1.22)$$

avec ρ_r , V_r et S_r des grandeurs de référence correspondant respectivement à la masse volumique du fluide, à la vitesse du fluide et à une surface du corps portant. La définition de la surface de référence est dépendante de l'application, mais dans notre cas elle vaut généralement $S_r = c_r b_r$ qui est le produit d'une corde de référence c_r par une envergure de référence b_r . Il est également commun d'écrire les efforts de pression et de cisaillement sous des formes adimensionnelles par

$$C_p = \frac{p - p_r}{\frac{1}{2}\rho_r V_r^2}, \quad (2.1.23)$$

$$c_f = \frac{\tau}{\frac{1}{2}\rho_r V_r^2}. \quad (2.1.24)$$

Très souvent, on compare les caractéristiques aérodynamiques ou hydrodynamiques des corps portants à partir de courbes $C_L = f(\alpha)$, $C_D = f(\alpha)$ ou $C_L = f(C_D)$ graduée en α . La dernière définition permet en l'occurrence de tracer le point de finesse maximale $\left(\frac{C_L}{C_D}\right)$ que l'on cherche généralement à maximiser. La Figure 2.4 illustre un exemple de l'évolution de $C_L = f(C_D)$ lorsque la cambrure maximale d'une voile varie. On peut ainsi remarquer que l'augmentation de la cambrure induira une augmentation de la portance, mais également de la trainée pour un angle d'attaque α fixé.

Les effets d'interactions entre une grand-voile et un géniois sur les différences de pressions entre l'extrados et l'intrados des deux voiles sont illustrés par les travaux de Garrett (1996) sur la Figure 2.5. On observe une augmentation de ces différences de pressions sur le géniois lorsque la grand-voile est présente. C'est l'inverse

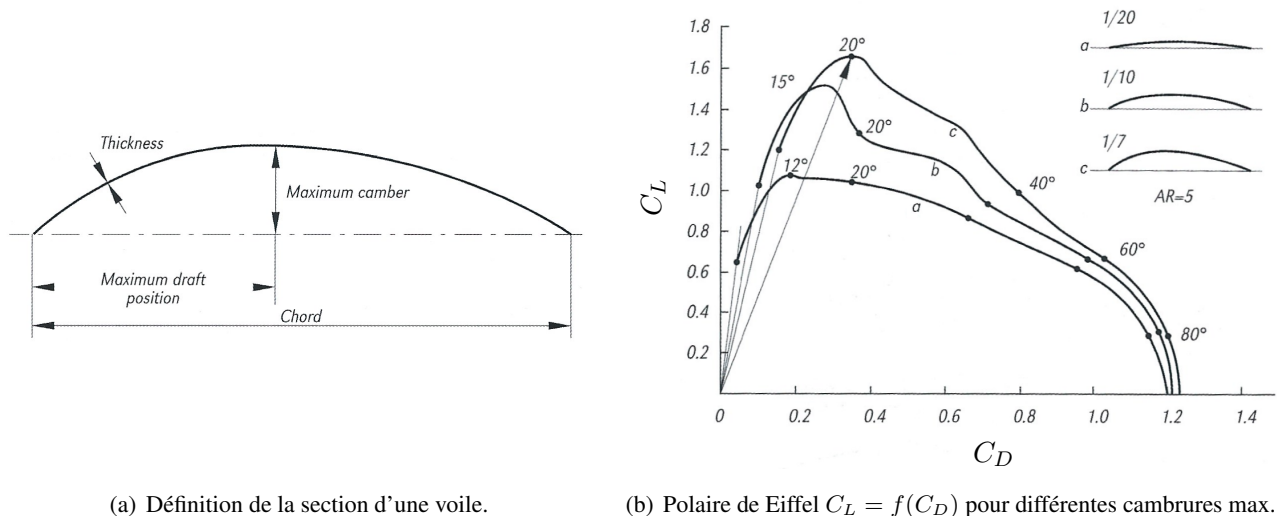


Figure 2.4 – Effets de la cambrure d'une voile. Images tirées de (Fossati, 2010)

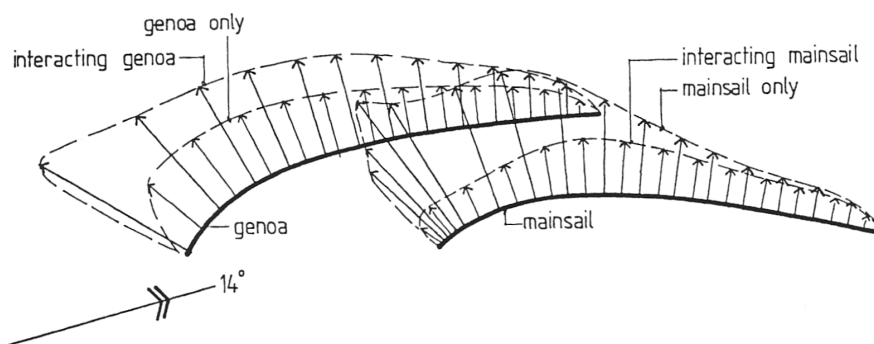


Figure 2.5 – Effets des interactions entre une grand-voile et un génois. Image issue de (Garrett, 1996)

sur la grand-voile lorsque le génois est présent. Cela montre bien l'importance de considérer les interactions entre les voiles lors d'une modélisation expérimentale ou numérique afin de correctement prédire les efforts aérodynamiques générés par les voiles.

2.1.2 Interaction fluide-structure

L'Interaction Fluide-Structure (IFS), ou « Fluid-Structure Interaction (FSI) » en anglais, concerne le phénomène de couplage entre une structure et un fluide composés d'une interface commune. Le fluide est liquide ou gazeux et la structure peut être fixe ou mobile ainsi que rigide ou déformable. Ce phénomène IFS est dit « couplé » car le couplage mécanique entre la structure et le milieu fluide s'effectue dans les deux sens au niveau de l'interface de contact. Plus concrètement, les déformations de la structure, sous l'effet des efforts générés par le fluide, modifient les caractéristiques de l'interface fluide-structure ; par conséquent, les conditions de l'écoulement sont modifiées et induisent une variation de l'effort fluide exercé sur la structure.

Les applications mettant en jeu un phénomène d'interaction fluide-structure sont très variées. Dans le do-

maine de la construction civile, la destruction tragique en novembre 1940 du premier pont de Tacoma Narrows, sollicité par une instabilité aéroélastique de torsion, est très certainement l'application IFS la plus populaire. Dans le naval, cela peut concerner le « tossage », ou « slamming » en anglais, qui se caractérise par des impacts hydrodynamiques sur l'étrave d'un navire naviguant à grande vitesse et/ou dans des conditions de mer difficiles. Dans l'aviation, le phénomène aéroélastique de « flutter » est considéré dans la phase de design des ailes ou des pales d'hélicoptère. Les pales des éoliennes sont également concernées par des problématiques IFS, en particulier lors des phases de maintenance ou des phénomène de « Vortex Induced Vibrations (VIV) » peuvent apparaître. L'interaction des voiles d'un voilier avec le vent est un type particulier d'IFS. En effet, la structure est dite « souple et légère » (Durand, 2012) de par la nature de ses grands déplacements, sa très faible rigidité de flexion ainsi que sa faible masse par rapport à la masse ajoutée du fluide. Ces caractéristiques rendent le couplage IFS « fort » car une faible variation de la structure ou du fluide peut entraîner de grands déplacements et donc des changements importants des efforts fluides sur les voiles. La modélisation numérique des phénomènes d'interaction fluide-structure rencontrés par les voiles en est par conséquent complexe.

La mise en équations d'un problème d'interaction fluide-structure nécessite d'écrire les équations décrivant le comportement de la structure, en plus des équations du fluide que l'on a précédemment présentées dans la Section 2.1.1.5. Suivant la nature de l'application et les propriétés physiques de la structure, des hypothèses plus ou moins simplificatrices peuvent être employées. Ces hypothèses modifient directement l'écriture mathématique du problème IFS. Dans le cadre des petites perturbations en élasticité linéaire, on peut écrire à chaque instant t , l'équation d'équilibre locale d'une partie s d'un système S comme

$$\nabla \cdot \vec{\Sigma} + \vec{F} = \rho \vec{a}, \quad (2.1.25)$$

où $\vec{\Sigma}$ est le tenseur symétrique des contraintes de « Cauchy » et $\vec{a}$ est le vecteur des accélérations. Des conditions de continuité au niveau de l'interface, notée ici Γ , définissant le couplage dynamique et cinématique doivent également être introduites. En effet, l'interface vue par la structure Γ_s et celle vue par le fluide Γ_f doivent satisfaire les conditions suivantes

$$\Gamma_s = \Gamma_f, \quad (2.1.26)$$

$$\vec{\Sigma}_s \cdot \vec{n} = \vec{\Sigma}_f \cdot \vec{n}, \quad (2.1.27)$$

$$\dot{\Gamma}_s = \dot{\Gamma}_f. \quad (2.1.28)$$

La première condition (2.1.26) assure que l'interface vue par la structure et le fluide est bien identique. L'expression (2.1.27) implique une continuité des efforts au niveau de l'interface. La dernière condition (2.1.28) permet d'assurer une continuité des vitesses au niveau de l'interface. Cette dernière expression devient $\dot{\Gamma}_s \cdot \vec{n} = \dot{\Gamma}_f \cdot \vec{n}$, dans le cas d'un fluide supposé parfait.

2.2 Outils numériques

Les outils et méthodes numériques permettant la modélisation de la physique du voilier sont décrits dans cette Section 2.2. Bien que le modèle physique puisse être considéré comme une boîte noire, sa formalisation et son implémentation restent souvent complexes et il est nécessaire d’avoir un minimum de notions sur ces méthodes afin de pouvoir mener à bien une optimisation. En effet, la connaissance *a priori* des hypothèses propres à chaque modèle numérique aidera l’utilisateur dans la compréhension et la validation des résultats d’une optimisation. Dans la suite, la discussion se veut non exhaustive de par la grande variété des méthodes et des modèles numériques disponibles aujourd’hui. La Section 2.2.1 résume un ensemble de méthodes appartenant au domaine de la mécanique des fluides numérique. La simulation numérique des voiles prenant en compte les phénomènes d’interaction fluide-structure concerne la Section 2.2.2.

2.2.1 Mécanique des fluides numérique

Les premières méthodes de mécanique des fluides numérique appliquées aux voiliers reposent sur la théorie des écoulements à potentiel de vitesse. Les effets de viscosité sont négligés, le fluide est considéré comme incompressible ($\rho = \text{cste}$) et irrotationnel ($\nabla \times \vec{V} = \vec{0}$), enfin, il n’y a pas de force de masse ($\vec{F} = \vec{0}$). Dans ces conditions, on montre que les vitesses dérivent d’un potentiel ϕ et que $\vec{V} = \nabla \phi$. A partir des hypothèses précédentes, on a $\nabla \cdot (\nabla \phi) = \nabla^2 \phi = 0$ et les équations de Navier-Stokes sont ramenées à la résolution de l’équation de « Laplace ». Dans le cas d’un écoulement autour d’un corps portant, le Laplacien de ϕ est nul partout sauf dans la couche limite et le sillage. Les méthodes de singularités permettent de formaliser cette singularité dans la couche limite, à l’aide d’écoulements potentiels élémentaires (doublets) répartis sur le corps portant. Des conditions d’imperméabilités (glissement) sur le corps portant sont également imposées afin de rendre le système d’équations solvable. La résolution seule de $\nabla^2 \phi = 0$ avec les conditions aux limites ne donne cependant pas accès à la portance du corps portant. Elle sera obtenue par le respect de la condition de « Kutta-Joukowski » qui fixe l’intensité de la circulation $\Gamma_C = \oint_C \vec{V} \cdot d\vec{s}$, de manière à obtenir une vitesse d’écoulement continue au bord de fuite : sa valeur en ce point est la même du côté intrados et extrados. Physiquement, cette condition s’impose naturellement par l’action de la viscosité dans la couche limite. Différentes théories et méthodes reposent sur la théorie des écoulements potentiels, comme les premières théories des profils minces et des lignes portantes (Prandtl, 1920; Munk, 1924; Abbott and Von Doenhoff, 1959) ou les méthodes à panneaux tourbillons et ses dérivées comme la méthode « VLM » (Falkner, 1947; Rehbach, 1977), pour « Vortex Lattice Method ». Ces différentes approches basées sur la théorie des écoulements potentiels ont l’avantage d’être relativement peu coûteuses en temps calcul : de l’ordre de la seconde ou de la minute. Il est cependant important de rappeler que le fluide est supposé non visqueux, ce qui implique une validité du modèle pour de faibles angles d’attaque : lorsque l’écoulement n’est pas ou peu détaché. Ces méthodes sont toujours très utilisées aujourd’hui dans l’industrie et les récents travaux scientifiques s’intéressent à de nouvelles approches où la méthode à écoulements potentiels est couplée avec d’autres modèles permettant la prise en compte des effets visqueux (Gallay, 2016; Sugar Gabor et al., 2016).

L’augmentation de performance et de mémoire des récents ordinateurs a permis le développement de nou-

veaux modèles et méthodes pour la mécanique des fluides numérique. Ces nouvelles approches ont la faculté de ne plus négliger les effets de viscosité du fluide. La modélisation de la viscosité implique directement une prise en compte de la turbulence. Cette turbulence se caractérise par la présence de tourbillons de différentes tailles et durées de vie dans l'écoulement. Les équations de Navier-Stokes décrivent complètement le comportement d'un écoulement de fluide visqueux turbulent. Malheureusement, la résolution directe de ces équations pour un cas concret d'écoulement autour d'un corps complexe et à haut Reynolds n'est pas encore possible aujourd'hui. Une approche statistique est alors employée afin de résoudre les équations de Navier-Stokes moyennées (« RANSE »). Les grandeurs (vitesse, pression, température, etc.) sont décomposées en une grandeur moyenne $\bar{X}$ et une grandeur fluctuante X' telle que

$$X = \bar{X} + X', \quad (2.2.1)$$

et $\bar{X}' = 0$. Les équations RANSE de la conservation de la masse et de la quantité de mouvement, s'écrivent alors comme

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot (\bar{\rho} \tilde{\vec{V}}) = 0, \quad (2.2.2)$$

$$\frac{\partial}{\partial t} (\bar{\rho} \tilde{\vec{V}}) + \nabla \cdot (\bar{\rho} \tilde{\vec{V}} \otimes \tilde{\vec{V}}) = \bar{\rho} \vec{F} + \nabla \cdot (-\bar{p} \vec{I}) + \nabla \cdot (\vec{\Sigma}_r + \vec{\Sigma}_v), \quad (2.2.3)$$

où l'on a noté l'opérateur de moyenne de « Favre » $\tilde{X} = \frac{\bar{\rho X}}{\bar{\rho}}$ et $\vec{\Sigma}_r$ qui correspond au tenseur symétrique des contraintes de Reynolds. L'équation de conservation de l'énergie totale en formulation RANSE n'est pas présentée ici pour éviter de surcharger la discussion. Le tenseur de Reynolds s'exprime par la relation $\vec{\Sigma}_r = -\bar{\rho \vec{V}'' \otimes \vec{V}''}$ et il introduit 6 inconnues supplémentaires qui posent un problème de fermeture. Cette fermeture sera réalisée par des lois phénoménologiques fournies par des modèles de turbulence. On distingue généralement deux classes de modèles de turbulence :

1. Les modèles du second ordre, du type « Reynolds Stress Model (RSM) », où les composantes de $\vec{\Sigma}_r$ sont directement calculées. Ces modèles donnent accès aux champs fluctuants moyennés, mais leur utilisation s'avère en pratique coûteuse et peu robuste.
2. Les modèles à viscosité turbulente μ_t , à 0 (longueur de mélange), 1 (k) ou en particulier 2 ($k - \epsilon$, $k - \omega$, etc.) équations. Ces modèles sont plus utilisés pour leur meilleure robustesse, comparée au RSM, ainsi que pour leur bon compromis entre le coût calcul et la précision obtenue.

On présente maintenant quelques modèles de turbulence couramment utilisés ainsi que les limites de leur application respective.

Modèle $k - \epsilon$. Ce modèle est constitué de deux équations qui vont s'ajouter aux équations RANSE. Elles reposent sur le transport de l'énergie cinétique turbulente k ainsi que du taux de dissipation de l'énergie cinétique

ϵ , et s'expriment comme

$$\frac{\partial}{\partial t}(\bar{\rho} \widetilde{V} k) + \nabla \cdot (\bar{\rho} \widetilde{V} k) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\alpha_k} \right) \nabla k \right] + \overrightarrow{\Sigma}_r : \nabla \widetilde{V} - \bar{\rho} \epsilon - K, \quad (2.2.4)$$

$$\frac{\partial}{\partial t}(\bar{\rho} \epsilon) + \nabla \cdot (\bar{\rho} \widetilde{V} \epsilon) = \nabla \cdot \left[\left(\mu + \frac{\mu_t}{\alpha_\epsilon} \right) \nabla \epsilon \right] + c_{\epsilon 1} \frac{\epsilon}{k} \overrightarrow{\Sigma}_r : \nabla \widetilde{V} - c_{\epsilon 2} f_t \frac{\epsilon^2}{k} + E, \quad (2.2.5)$$

où $f_t = \left(1 - 0.3 \exp \left(-\frac{\bar{\rho} k^2}{\mu \epsilon} \right) \right)$ et $\mu_t = c_\mu f_\mu \frac{k^2}{\epsilon}$. Les définitions des termes K , E , de la fonction de paroi f_μ ainsi que les valeurs des constantes α_k , α_ϵ , $c_{\epsilon 1}$, $c_{\epsilon 2}$ et c_μ sont dépendantes de la version du modèle $k - \epsilon$ choisi. Ce modèle $k - \epsilon$ qui a été développé dans les années 1970 par **Jones and Launder (1972)** reste le modèle le plus utilisé aujourd'hui dans l'industrie. Ses faiblesses connues concernent le manque de sensibilité aux gradients de pression adverses ainsi que sa tendance à prédire des intensités de viscosités turbulentes élevées.

Modèle $k - \omega$. Ce modèle ajoute également deux équations aux équations RANSE, l'une toujours basée sur le transport de k et la seconde basée sur une dissipation spécifique ω qui peut être considérée comme une échelle de temps turbulente $\omega = \frac{\epsilon}{k}$. Les deux équations du modèle $k - \omega$ développé par **Wilcox (1988)** s'écrivent

$$\frac{\partial}{\partial t}(\bar{\rho} \widetilde{V} k) + \nabla \cdot (\bar{\rho} \widetilde{V} k) = \nabla \cdot [(\mu + \sigma_* \mu_t) \nabla k] + \overrightarrow{\Sigma}_r : \nabla \widetilde{V} - \beta_* \bar{\rho} \omega k, \quad (2.2.6)$$

$$\frac{\partial}{\partial t}(\bar{\rho} \omega) + \nabla \cdot (\bar{\rho} \widetilde{V} \omega) = \nabla \cdot [(\mu + \sigma \mu_t) \nabla \omega] + \alpha \frac{\omega}{k} \overrightarrow{\Sigma}_r : \nabla \widetilde{V} - \beta \bar{\rho} \omega^2. \quad (2.2.7)$$

La viscosité turbulente est donnée par $\mu_t = \bar{\rho} \frac{k}{\omega}$ et les valeurs des constantes que propose **Wilcox (1988)** sont

$$\alpha = 5/9, \quad \beta = 3/40, \quad \beta_* = 9/100, \quad \sigma = 1/2, \quad \sigma_* = 1/2.$$

Ce modèle de turbulence $k - \omega$ fonctionne généralement mieux que le modèle $k - \epsilon$ dans les zones d'écoulement en proche paroi, mais c'est cependant l'inverse dans les zones lointaines des parois. C'est pourquoi en pratique il est préférable d'utiliser une combinaison des modèles $k - \epsilon$ et $k - \omega$ avec une fonction de mélange. C'est ce que proposent **Menter et al. (2003)** en ajoutant une fonction limitant les valeurs de la viscosité turbulente pour obtenir le modèle $k - \omega$ SST, aussi appelé « modèle de transport de contrainte de cisaillement » pour « Shear Stress Transport (SST) ». Ce modèle $k - \omega$ SST est très utilisé aujourd'hui pour sa précision de l'apparition d'une séparation de l'écoulement sous des gradients de pression négatifs. Il a également été validé dans des cas d'écoulements fortement décollés (**Guilmineau, 1995**).

De nouvelles approches ont également émergé avec l'augmentation de la puissance des calculs. Par exemple, la méthode « Large Eddy Simulation (LES) » (**Ghosal and Moin, 1995**) résout et modélise respectivement les grandes et petites échelles de la turbulence. Cette séparation des échelles qui repose sur l'hypothèse d'isotropie des petites structures permet d'accéder à un niveau de résolution très élevé en évitant le coût calcul bien supérieur de la résolution directe des équations de Navier-Stokes. La méthode LES reste encore coûteuse en particulier pour les cas d'écoulements comportant des parois, elle est donc plutôt appliquée dans le domaine

de la recherche ou alors pour des projets industriels bien spécifiques. L'approche « Detached Eddy Simulation (DES) » (Shur et al., 1999) est un bon compromis entre la résolution RANSE et LES, puisqu'elle correspond à une formulation hybride de ces deux méthodes. Les régions de l'écoulement où le fluide est attaché aux parois sont résolues par la méthode RANSE et celles où le fluide est détaché sont traitées par l'approche LES. La mise en place et l'exécution d'une simulation DES ne se font cependant pas sans difficultés. En effet, la solution présente des discontinuités en échelle et viscosité turbulente au niveau de l'interface entre les deux méthodes. La méthode DES rend cependant possible la résolution d'écoulements complexes, comme par exemple dans (Xu et al., 2017), pour un coût calcul bien inférieur à la méthode LES. Enfin, la méthode « Direct Numerical Simulation (DNS) » est la plus puissante puisqu'elle résout les équations de Navier-Stokes sans avoir de modèle de turbulence. Bien entendu, le coût calcul est ici beaucoup trop élevé pour que la méthode soit applicable à des cas industriels. Elle reste réservée à des cas d'écoulements très simples, par exemple dans un tuyau (Hattori and Nagano, 2004), ou pour de faibles valeurs de Reynolds (Cimarelli et al., 2018). En pratique, cette méthode est surtout utile pour les chercheurs qui travaillent sur la construction de modèles.

La résolution des équations de Navier-Stokes s'effectue sur un problème discret défini sur un domaine discrétisé. Une discrétisation temporelle doit être réalisée afin d'approcher les dérivées partielles. Cette discrétisation est du type différences finies et est issue des formules de « Taylor-Mac-Laurin ». Un schéma temporel explicite ou implicite d'ordre généralement égal à 2 est alors choisi ainsi qu'un pas de temps δ_t adapté pour respecter le nombre de « Courant-Friedrichs-Lewy (CFL) » (De Moura and Kubrusly, 2013) qui s'écrit $CFL = \delta_t \left(\frac{u}{\delta_x} + \frac{v}{\delta_y} + \frac{w}{\delta_z} \right)$, où (u, v, w) sont les composantes du vecteur vitesse et $(\delta_x, \delta_y, \delta_z)$ sont les intervalles dimensionnels. Le CFL doit en pratique être inférieur ou proche de 1 afin d'assurer la stabilité du schéma. Cela est souvent vérifié *a posteriori* à l'aide d'une étude de convergence en temps pour s'assurer que les phénomènes moyens essentiels sont correctement capturés. Une discrétisation spatiale est également effectuée. Le domaine continu du fluide est transformé en un maillage, structuré ou non structuré suivant le solveur utilisé, et les résultats de la simulation numérique sont donnés aux nœuds de ce maillage. La discrétisation spatiale est généralement réalisée par la méthode des volumes finis, qui est une extension de la méthode des différences finies, car elle est bien adaptée à la résolution des équations de Navier-Stokes. Les équations de Navier-Stokes sont écrites sous la forme intégrale pour toute cellule volumique composant le maillage discrétisé. La méthode des volumes finis consiste ensuite essentiellement à calculer les flux à travers les faces de ces cellules. La méthode de résolution la plus utilisée repose sur l'algorithme « SIMPLE » (Patankar and Spalding, 1983), où les équations de conservation de la masse et de la quantité de mouvement sont résolues itérativement. Une fois que les conditions des frontières du domaine ont été fixées, l'algorithme SIMPLE démarre de conditions initiales supposées des vitesses et pressions puis les utilisent dans les équations de conservations. Les résidus obtenus permettent de réévaluer les grandeurs de vitesse et pression avant de démarrer une nouvelle itération. Le suivi de la convergence est effectué par le calcul des résidus, des efforts et des bilans globaux des flux. Il est souvent nécessaire d'ajuster les paramètres de sous-relaxation du solveur avant d'atteindre la convergence du calcul.

2.2.2 Simulations des voiles en IFS

La simulation numérique de voiles soumises à un écoulement en interaction fluide-structure est loin d'être aisée. L'IRENav travaille activement sur ce sujet, et plus particulièrement sur la validation de modèles numériques par l'expérience (Augier et al., 2012), depuis maintenant quelques années. Un code numérique d'interaction fluide-structure a été développé pendant la thèse de Durand (2012), financée par la compagnie K-Epsilon. Cet outil est particulièrement bien adapté à la simulation numérique des écoulements sur les voiles en IFS (Durand et al., 2014), que ce soit en stationnaire (Sacher et al., 2017c), par une approche quasi-statique, ou instationnaire (Augier et al., 2013), par une approche dynamique. Dans cette thèse, l'outil de simulation IFS est uniquement utilisé pour des cas stationnaires.

Le code est composé du solveur structure « ARA », basé sur la méthode des éléments finis. Ce code structure utilise une grande variété d'éléments, permettant ainsi la modélisation structurelle complète d'un gréement par des poutres de Timoshenko (mat, bôme, lattes), des câbles (gréements dormant et courant) et des membranes du type « Constant Strain Triangles (CST) » (voiles). Un modèle de plis, développé dans (Durand, 2012), est également inclus dans ce solveur structure afin de capturer les déformations locales des voiles sur un maillage comportant un nombre raisonnable de membranes CST. La résolution du problème structure non linéaire est obtenue par une méthode itérative de Newton, associée à une relaxation adaptative de « Aitken », afin de modérer les déplacements de la structure entre deux itérations de couplage.

Le code ARA a été couplé avec deux modèles fluide. Le premier, qui repose sur la théorie des écoulements potentiels, est « AVANTI » (Roux et al., 2002). Ce modèle utilise une méthode particulière (Rehbach, 1978; Charvet et al., 1996) qui permet de simuler des écoulements tridimensionnels en instationnaire de fluides parfaits incompressibles autour de profils minces. Les voiles sont discrétisées par des maillages surfaciques composés de panneaux rectangulaires associés à des doublets, qui sont des solutions élémentaires de l'équation de Laplace. Les intensités des doublets sont ajustées de manière à respecter la condition de glissement ($\vec{V} \cdot \vec{n} = 0$) et un sillage particulière est émis au niveau d'un ou des bords de fuite des voiles permettant ainsi de respecter la condition de Kutta-Joukowski en jouant sur l'intensité des particules, et par la suite de calculer les efforts sur la structure. L'utilisation de ce modèle fluide parfait est donc uniquement valide pour des cas d'écoulements où la viscosité est négligeable, mais surtout pour des écoulements attachés, donc pour des angles de vent incident faibles ($\approx 15^\circ$) sur les voiles. Le second modèle fluide couplé avec ARA est « ISIS-CFD », qui résout les équations de Navier-Stokes en formulation RANSE pour des maillages structurés ou non structurés, à l'aide de la méthode des volumes finis. La formulation « Arbitrary Lagrangian Eulerian (ALE) » est utilisée dans ce code, ce qui lui permet d'accepter des déformations du maillage. ISIS-CFD est développé par l'équipe du « LHEEA » de l'École Centrale de Nantes et est distribué dans la suite logicielle commerciale « FINETM/Marine ». Ce code fluide est très utilisé dans l'industrie navale, en particulier pour les calculs hydrodynamiques avec surface libre (Queutey and Visonneau, 2007). Le couplage de ARA avec AVANTI et ISIS-CFD est réalisé par un algorithme original quasi-monolithique (Durand, 2012) qui correspond à un couplage implicite adapté aux codes partitionnés.

Les voiles simulées en IFS ont la particularité d'être sujettes à de grandes déformations. Cela est très souvent

le cas lors des premières itérations de couplage car la géométrie initiale des voiles peut être très éloignée de la géométrie soumise à un écoulement. Ces déformations de la structure des voiles doivent être prises en compte par le modèle fluide. Dans le cas du modèle fluide parfait, cette déformation est relativement simple puisque le maillage fluide est uniquement surfacique. Cela est beaucoup plus complexe et coûteux dans le cas du modèle fluide visqueux qui utilise la méthode des volumes finis. Une méthode efficace de déformation du maillage volumique par propagations en blocs a été proposée par [Durand \(2012\)](#); elle est utilisée dans les calculs IFS de cette thèse.

La Figure 2.6 illustre les différentes étapes nécessaires à la réalisation d'un calcul IFS visqueux sur des voiles. Un calcul purement structurel est tout d'abord effectué sur le gréement sans les voiles (Figure 2.6(a)). En effet, comme dans la réalité, les gréements des voiliers sont précontraints suivant les recommandations données par le « tuning guide » des constructeurs. Les maillages surfaciques (fluide parfait et structure) des voiles sont ensuite ajoutés au gréement précontraint (Figure 2.6(b)). Un premier calcul IFS peut alors être réalisé avec une pression uniforme imposée sur les voiles (Figure 2.6(c)). Cela permet de mettre l'ensemble du modèle numérique dans ses attitudes (gîte, cap et assiette) du calcul, mais également de mettre en forme les voiles par rapport aux réglages *a priori* et de les ajuster sur le gréement précontraint (ralinguage). Le calcul IFS avec le modèle fluide parfait peut alors démarrer (Figure 2.6(d)) en partant de la solution convergée en pression uniforme. Les étapes propres au modèle fluide visqueux viennent ensuite. Le domaine fluide visqueux est créé sous la forme d'un parallélépipède autour du modèle numérique convergé en IFS fluide parfait (Figure 2.6(e)). Le maillage volumique, non structuré dans cette thèse, est ensuite construit sur ce domaine (Figure 2.6(f)). Finalement, le calcul IFS fluide visqueux résolvant les équations RANSE est réalisé (Figure 2.6(g)). Cette procédure est bien entendu très résumée ici. En pratique, il existe des difficultés au niveau de la convergence et de la déformation du maillage. Il est alors nécessaire de jouer sur des paramètres de sous-relaxation du solveur fluide et du solveur structure. Dans certains cas extrêmes de déformation du maillage, des cellules ayant des volumes négatifs peuvent apparaître, ce qui implique un remaillage fluide d'une solution IFS non convergée, pour ensuite relancer le calcul IFS fluide visqueux.

Les simulations numériques ont l'avantage de pas souffrir des effets d'échelle, que l'on peut retrouver dans certains essais expérimentaux, tout en étant généralement moins coûteux. Ils permettent également d'avoir accès à des grandeurs physiques qui peuvent être difficilement mesurables, comme typiquement la pression. Néanmoins, les essais expérimentaux restent indispensables aujourd'hui afin de valider et calibrer ces modèles numériques, mais aussi pour étudier des phénomènes physiques complexes et non simulables. La prochaine Section 2.3 présente différents types d'essais expérimentaux, où la discussion est non exhaustive de par leur grande variété, qu'ils soient de nature aérodynamique et/ou hydrodynamique. L'objectif est de présenter un certain nombre de moyens de mesures, en liens avec la physique des navires, qui peuvent être utilisés comme des modèles d'une fonction performance dans un processus d'optimisation.

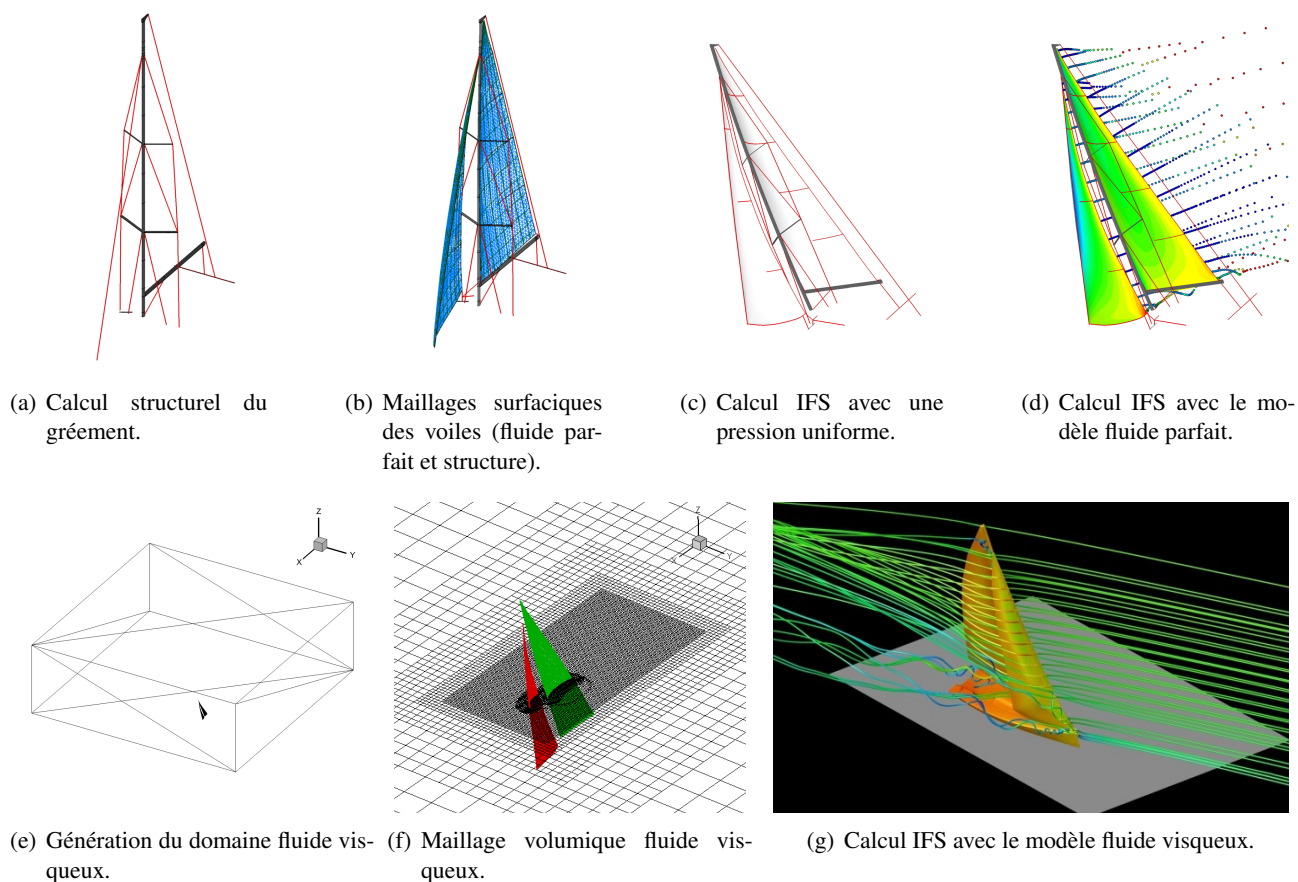


Figure 2.6 – Procédure d'un calcul IFS visqueux sur des voiles de Figaro.

2.3 Essais expérimentaux

Les différents dispositifs expérimentaux couramment utilisés aujourd'hui pour modéliser la physique des voiliers sont présentés dans cette Section 2.3. Les objectifs des essais sont variés. Ils permettent par exemple de mesurer des efforts aérodynamiques et/ou hydrodynamiques, d'étudier la cinématique et la dynamique des voiliers, ou encore de visualiser la dynamique des écoulements. Certaines mesures sont également très utiles pour valider et/ou nourrir des modèles numériques. Il existe également des phénomènes physiques qui sont encore trop complexes pour être correctement modélisés par la simulation numérique, les essais expérimentaux sont alors justifiés.

2.3.1 Hydrodynamique en bassin

Les essais hydrodynamiques sur les carènes et appendices sont très répandus et utilisés dans le développement des navires de commerce et militaire. Les moyens d'essais en hydrodynamique sont de différents types en fonction des informations recherchées.

La résistance à l'avancement d'une carène ainsi que son système propulsif sont souvent étudiés dans des

bassins de traction. La Figure 2.7 présente un exemple type de ce genre d'essai. Une maquette de carène est tractée par un chariot mobile qui impose un mouvement de translation rectiligne uniforme le long du bassin de traction. La maquette est reliée au chariot par l'intermédiaire d'un bras de pilonnement. Ce bras permet de



Figure 2.7 – Mesure de la résistance à l'avancement d'une carène. Image du « Wolfson Unit ».

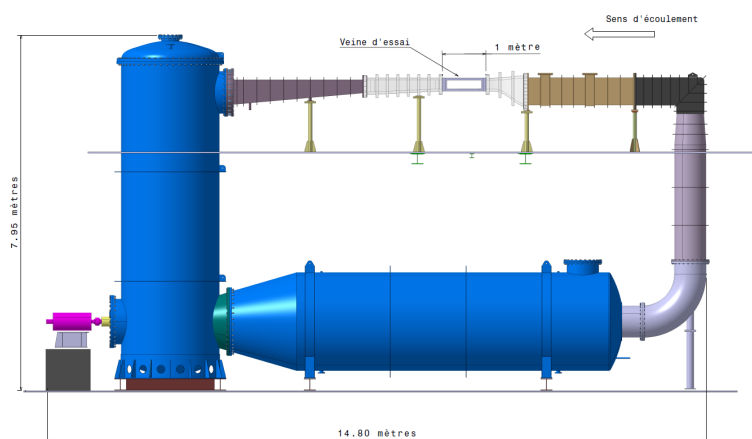
bloquer certains degrés de liberté, comme par exemple la gîte, tout en laissant certains autres libres : généralement l'enfoncement et l'assiette. Les efforts hydrodynamiques de la carène sont alors mesurés sur ce bras à l'aide d'une balance d'efforts. L'extrapolation des efforts au réel se fait en similitude de Froude et généralement par la méthode « ITTC 57 », adoptée par l'« International Towing Tank Conference » en 1957. Cette méthode consiste à déterminer le coefficient de résistance résiduaire C_{W_m} , à partir du coefficient de résistance total C_{T_m} , mesuré dans le bassin, et du coefficient de résistance de frottement C_{F_m} , estimé par la formule ITTC 57 avec $C_{F_{0m}} = \frac{0.075}{(\log R_{em} - 2)^2}$. Ce coefficient de résistance résiduaire est alors donné par $C_{W_m} = C_{T_m} - (1 + \phi)C_{F_{0m}}$, où $(1 + \phi)$ est le facteur de forme mesuré à très faible vitesse par, $(1 + \phi) = \lim_{V \rightarrow 0^+} \frac{C_{T_m}}{C_{F_{0m}}}$, et on a de par la similitude de Froude $C_{W_m} = C_{W_r}$, où C_{W_r} est le coefficient de résistance résiduaire au réel. Finalement le coefficient de résistance du navire au réel est calculé par $C_{T_r} = C_{W_r} + (1 + \phi)C_{F_{0r}}$, où $C_{F_{0r}}$ est estimé par la formule ITTC 57, avec $C_{F_{0r}} = \frac{0.3}{\log(14.7 \frac{L}{k})^{2.5}}$, et $\frac{L}{k}$ est la rugosité relative. Des modèles mathématiques permettant de prédire la résistance à l'avancement de différentes coques de voiliers ont été construits sur la base d'essais systématiques en bassins de traction. Les plus connus sont les « Delft Systematic Yacht Hull Series », que l'on doit aux travaux de « Keuning » à la « Delft University of Technology ». Initialement, **Keuning et al. (1996)** se sont concentrés sur la résistance à l'avancement de multiples géométries de carènes sans appendices. **Keuning (1999)** a ensuite pris en compte la contribution des appendices dans la résistance à l'avancement. Et très récemment, l'influence de la gîte sur la résistance à l'avancement a été étudiée dans (**Keuning and Katgert, 2010**).

Il existe aussi des essais dits « d'autopropulsion » pour lesquels, contrairement aux essais de traction où seule la maquette de carène est tractée (pas d'hélice), la maquette de carène est maintenant équipée d'un arbre d'hélice instrumenté permettant de mesurer les efforts de poussée et de couple de l'hélice à différents régimes de rotation. Ces essais permettent de déterminer les coefficients d'interaction entre l'hélice et la carène. Par exemple, le coefficient de sillage $(1 - \omega) = \frac{J_0}{J}$ est obtenu par le rapport des paramètres d'avance $J = \frac{V}{nD}$ en « eau libre » et en autopropulsion pour un même coefficient de poussée $K_T = \frac{T}{\rho n^2 D^4}$. Ces coefficients sont obtenus à partir d'essais de maquette d'hélice sur un arbre instrumenté en eau libre (hélice seule) réalisés dans un tunnel de cavitation et d'essais en autopropulsion (carène avec l'hélice) dans un bassin de traction. Pour

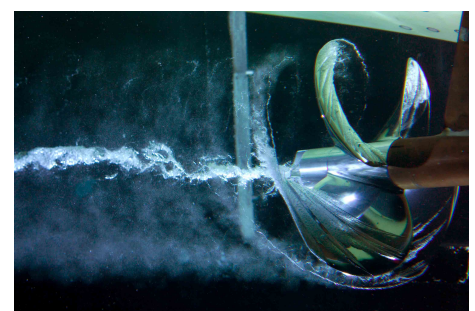
une vitesse fixée V (vitesse de l'écoulement dans le tunnel de cavitation ou vitesse de la maquette de carène dans le bassin de traction), le balayage en régime de rotation de l'hélice donne accès à ces coefficients. Pour le coefficient de succion $(1 - t) = \frac{R}{T}$, la carène est tout d'abord tractée sans l'hélice et à une vitesse fixée, afin de déterminer sa résistance à l'avancement R . Un essai d'autopropulsion est ensuite utilisé pour déterminer la vitesse de rotation n permettant d'atteindre la vitesse fixée. La poussée T de l'hélice est alors obtenue.

Certains bassins de traction sont capables de générer de la houle, généralement monodirectionnelle, permettant ainsi des mesures de résistance à l'avancement d'une carène en présence de houle. Cette houle est créée par un ensemble de volets reliés à un batteur et une plage d'amortissement est placée en fin du bassin afin d'amortir les effets de réflexion de la houle. Pour finir sur les bassins de traction, il en existe de multiples variétés, qu'ils soient académiques ou industriels. Le centre « DGA Techniques hydrodynamiques », anciennement « bassin d'essais des carènes (BEC) », dispose par exemple d'un très grand bassin de traction de 545 m de long, 15 m de large et 7 m de profondeur. Le « Hamburg Ship Model Basin (HSVA) » contient un bassin de traction de 30 m de long, 6 m de large et 1.2 m de profondeur, qui est contrôlable pour des températures allant de -15°C à $+15^{\circ}\text{C}$. Les conditions extrêmes de navires évoluant dans les mers australes peuvent être étudiées. Le « Maritime Research Institute Netherlands (MARIN) » est quant à lui équipé d'un bassin de traction de 240 m de long, 18 m de large et 8 m de profondeur, qui est pressurisable et capable de générer de la houle. Les phénomènes de cavitation et/ou de ventilation sur les hélices lors d'un essai de traction parmi des vagues sont observables.

Les études des hélices et des hydrofoils nécessitent généralement des moyens expérimentaux adaptés aux phénomènes de cavitation (voir Figure 2.8(b)). Ces mesures sont réalisées dans des tunnels de cavitation, contrôlables en pression et en vitesse, comme par exemple celui qui équipe l'IRENav, schématisé par la Figure 2.8(a). La veine d'essai du tunnel de l'IRENav mesure 1 m de long et est de section carrée de côté 192×192



(a) Tunnel hydrodynamique de l'IRENav.



(b) Mesure de cavitation d'une hélice. Image provenant de la « DGA ».

Figure 2.8 – Essais en tunnel de cavitation.

mm. On peut ainsi apprécier sur la Figure 2.8(a), les grandes dimensions de l'installation comparées à sa petite

veine d'essai. Le tunnel de cavitation de l'IRENav est capable de générer des vitesses d'écoulements de 1 à 17 m/s, tout en étant pressurisé à des pressions allant de 0.2 à 3 bar. Les moyens de mesures sont très variés, avec une balance d'efforts 5 composantes, un vibromètre laser à balayage pour les mesures de vibration, un télémètre laser pour les mesures de déplacement, des caméras rapides pour l'observation de la cavitation mais également pour faire des mesures de champ de vitesse instantanée par de la vélocimétrie par images de particules (PIV), un système de mesure de vitesse par vélocimétrie laser (LDV) et toute sorte de capteurs implantés sur les différents profils testés.

La tenue à la mer des navires, tout comme leurs sollicitations extrêmes lors d'impacts avec des vagues sont étudiées dans des cuves à houle. Dans ce type d'essai, la maquette d'un navire est à vitesse nulle (parfois à vitesse très faible) et maintenue au centre de la cuve par un ou plusieurs câbles. Une houle est ensuite générée à l'aide d'un batteur et la réponse dynamique du navire soumis à cette houle est mesurée (capteurs embarqués et/ou caméras externes). La maquette est parfois équipée de capteurs d'efforts ou de pression afin d'étudier les impacts des vagues sur la coque. Les caractéristiques des houles peuvent être très variées, comme obliques, déferlantes, gaufrees ou encore focalisées. Ces essais sont également très utiles pour l'étude des systèmes flottants ou ancrés en pleine mer. Cela concerne par exemple les éoliennes flottantes ou les plateformes pétrolières. A titre d'exemple, l'École Centrale de Nantes possède une cuve à houle de 50 m de long, 30 m de large et 5 m de profondeur. La hauteur maximale des houles régulières générées par cette installation est de 1 m et les périodes possibles sont de 0.5 à 5 s. Certains moyens d'essais permettent également de faire des mesures en présence de houle et de vent (bassin de l'Ifremer à Brest) ou bien en présence de houle et de courant (bassin de circulation de l'Ifremer à Boulogne-sur-mer).

2.3.2 Aérodynamique en soufflerie

Les essais en soufflerie sont très répandus dans l'industrie aéronautique. On y trouve des souffleries subsoniques, transsoniques, supersoniques et hypersoniques. La plus célèbre est très certainement la soufflerie transsonique « S1MA » du centre « Modane-Avrieux » de l'ONERA, aussi surnommée la « cathédrale » pour ses dimensions hors normes : 8 m de diamètre et 14 m de long pour la veine d'essais. Certaines d'entre elles sont pressurisées, cryogéniques, à rafales ou bien continues.

Les souffleries dédiées aux essais sur les voiles doivent être capables de générer un écoulement vrillé afin de correctement simuler le profil vertical du vent apparent dans la couche limite atmosphérique. Ces souffleries restent peu nombreuses aujourd'hui. Cela provient principalement du milieu de la voile qui reste petit en terme de taille et de moyen financier. On peut néanmoins citer la soufflerie de l'université d'Auckland (Flay, 1996), présentée par la Figure 2.9 et construite pour la Coupe de l'America de 1995 ou bien celle de l'université de Milan également construite pour la même course, mais lors des éditions de 2003 et 2007 (Fossati et al., 2006). Plus récemment, l'université de Kiel s'est également équipée d'une telle soufflerie afin de réaliser des mesures sur des spinnakers (Graf and Müller, 2009).

La qualité d'une mesure expérimentale, réalisée dans une soufflerie, est bien entendu dépendante des dimensions de la veine d'essai, ainsi que de son écoulement (turbulence, vitesse, etc.). Elle est également fortement dépendante des moyens de mesure associés à la soufflerie, et de la rigueur des expérimentateurs. Par exemple, la

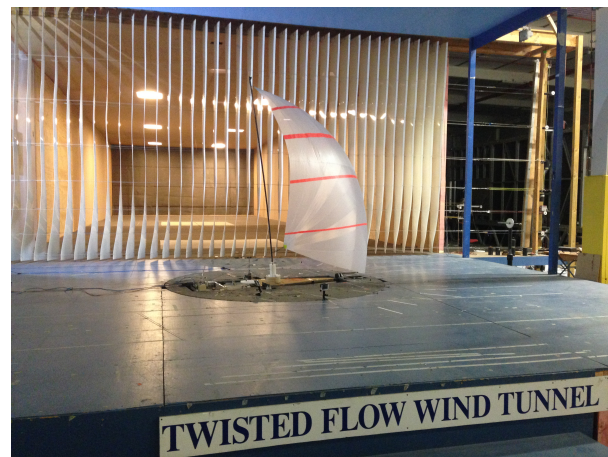


Figure 2.9 – Soufflerie de l’université d’Auckland à veine ouverte générant un écoulement vrillé. Les dimensions de la veine d’essai sont de 7 m de large par 3.5 m de haut ; la vitesse de l’écoulement peut varier entre 1 et 5 m/s. Image provenant de (Aubin, 2017).

précision d’une mesure des efforts aérodynamiques est dépendante de la précision de la balance, de l’amplitude de l’effort mesuré vis-à-vis de la gamme admissible par cette balance, et enfin de la qualité de la calibration de la balance. De même, la qualité de la mesure de l’écoulement incident, est dépendante du moyen de mesure utilisé (pression, laser, etc.) et de sa précision. La mesure de la forme des voiles, ou « flying shape » en anglais, nécessite aussi de la rigueur avant de pouvoir obtenir des mesures précises (Deparday, 2016) : positions des caméras, qualité de l’optique et du post-traitement, etc.

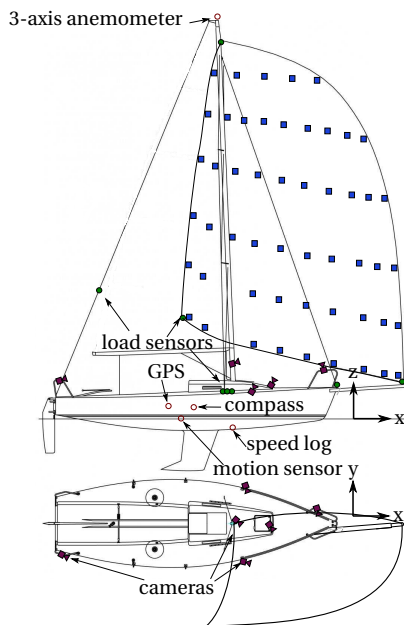
Les essais en soufflerie sur les voiles ont l’inconvénient d’être limités par l’impossibilité de respecter la similitude de Reynolds. En effet, cette condition demanderait une vitesse d’écoulement trop élevée pour être supportée par les matériaux de la maquette expérimentale.

2.3.3 Mesure *in situ*

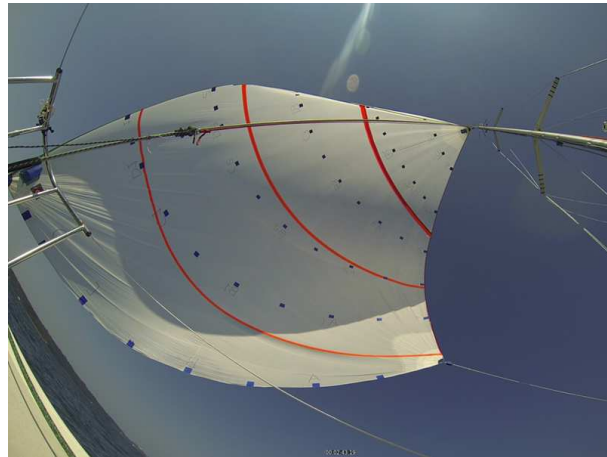
La mesure *in situ* est une expérience au réel, autrement dit en mer et directement réalisée sur un navire. Ces types d’essais ont l’avantage de ne pas être concernés par les conditions de similitude, si l’échelle est de 1/1 (« full scale »). Cependant, la mise en place, la réalisation, et enfin le post-traitement d’un essai *in situ* sur un voilier sont loin d’être aisés. En effet, les essais au réel sur l’eau sont complexes car les conditions de vent et de la mer ne sont pas contrôlables. La mer est aussi connue pour être un milieu hostile qui implique des précautions à prendre sur la mise en place des systèmes de mesure afin d’assurer la robustesse des essais. De plus, il est fréquent que ces mesures *in situ* utilisent un nombre important de capteurs variés dont les cadencements temporels sont généralement non identiques. Les incertitudes expérimentales (conditions météorologiques, capteurs, etc.) sont également non négligeables, en particulier pour la mesure du vent, dont la reconstruction complète demande un travail fastidieux (Kerhascoet, 2017). Toutes ces incertitudes peuvent être pénalisantes lorsque l’on souhaite faire des comparaisons entre différentes expériences ou de la confrontation à des modèles numériques.

Les essais *in situ* sur les voiliers sont étudiés depuis maintenant quelques années. Masuyama and Fukasawa (1997) ont par exemple effectué des mesures d’efforts sur des voiles en « full-scale » pour valider des

simulations numériques. Les mêmes auteurs ont réalisé dans (Masuyama et al., 2009), une base de données expérimentales comportant des formes de voiles mesurées en navigation. Les efforts aérodynamiques des voiles ainsi que les conditions de navigations ont aussi été mesurés en simultané. Des simulations numériques avec des modèles fluides parfait et visqueux font également parties de cette publication. Plus récemment, Le Pelley et al. (2012) se sont intéressés à des mesures de « flying shape » associées à des mesures de pression sur des voiles en « full scale ». Un voilier spécifiquement dédié aux mesures *in situ* a également été développé et construit par Fossati et al. (2013). Un système laser mesurant la forme des voiles ainsi que des capteurs de pression « MEMS » ont été développés et utilisés par cette même équipe (Fossati et al., 2015, 2016). Enfin, très récemment, des essais *in situ* ont été réalisés par l'IRENav (thèse de Deparday (2016), voir Figure 2.10) afin de mesurer la forme d'un spinnaker de J80 par une méthode de photogrammétrie (Deparday et al., 2016). Ces essais ont permis d'améliorer et d'adapter les systèmes de mesures du voilier instrumenté, initialement



(a) Set-up expérimental.



(b) Photographie du spinnaker en navigation.

Figure 2.10 – Mesures par photogrammétrie *in situ* d'un spinnaker. Images tirées de (Deparday, 2016).

développés pour l'étude des allures proches du vent (Augier, 2012), afin d'étudier les allures portantes, dans le but de mieux comprendre la dynamique du spinnaker, en particulier pour ce qui concerne le phénomène de « flapping ».

La description des mesures *in situ* clôt la Partie I de cette thèse où l'objectif est d'introduire le contexte spécifique du voilier, pour ce qui est de son fonctionnement ainsi que des moyens numériques et expérimentaux modélisant sa physique et sa performance. On va maintenant s'intéresser à l'optimisation de cette performance qui nécessite la mise en œuvre de méthodes d'optimisation efficaces, qui ont été développées dans le cadre de cette thèse et qui sont présentées dans la Partie II suivante.

DEUXIÈME PARTIE

MÉTA-MODÈLES POUR L'OPTIMISATION

All models are wrong, but some are useful.

GEORGE BOX

Généralités et contexte de l'optimisation

Sommaire

3.1	Motivations	47
3.2	Problème d'optimisation	48
3.3	Méthodes d'optimisation	49
3.3.1	Généralités	49
3.3.2	Algorithmes d'optimisation classiques	50
3.3.2.1	Descente de gradient	50
3.3.2.2	Méthode du « Simplex »	51
3.3.2.3	Stratégies d'évolution	51

Le Chapitre 3 introduit brièvement le contexte général de l'optimisation. Les raisons motivant son application, la définition d'un problème d'optimisation et des généralités sur les méthodes d'optimisation sont données dans les Sections 3.1 à 3.3 suivantes.

3.1 Motivations

L'utilité de l'optimisation est succinctement décrite dans cette Section. On rappelle en premier lieu la définition du verbe « Optimiser » donnée par le Larousse.

Définition 3.1.1 (Optimiser). *Donner à quelque chose, à une machine, à une entreprise, etc., le rendement optimal en créant les conditions les plus favorables ou en en tirant le meilleur parti possible.*

L'optimisation peut donc se voir comme un outil d'aide à la décision applicable dans de nombreux domaines. C'est par exemple le cas lorsque l'on souhaite inventer de nouveaux systèmes, où la conception fait

généralement intervenir des problèmes de dimensionnement (structure, coût, etc.) ou de contrôle optimal (commande, stabilité, etc.). En analyse numérique, les problèmes d'approximation ou de résolution de systèmes linéaires et non linéaires peuvent également se ramener à des problèmes d'optimisation. L'optimisation peut aussi être appliquée dans les domaines tels que la gestion de ressources, l'économie, ou bien encore dans le cadre de la surveillance opérationnelle (détection de dysfonctionnements, maintenance, etc.).

Dans toutes ces applications, on retrouve la notion de recherche d'optimalité dans la solution d'un problème d'optimisation. Le procédé qui permet la détermination de cette solution optimale est en soi une « méthode d'optimisation ». C'est un choix défini par l'utilisateur qui dépendra généralement de l'application. L'application est quant à elle bien évidemment connue par l'utilisateur. La difficulté pour l'utilisateur qui souhaite optimiser son application concerne la définition du problème d'optimisation. En effet, cette étape peut parfois être complexe et sa formulation impliquera qu'un problème d'optimisation sera dit « bien » ou « mal » posé (Hadamard, 1902). Concrètement, la définition d'un problème d'optimisation consiste à identifier les composantes (variables), les enjeux (objectifs) et les limites (contraintes) de l'application. Les variables sont liées aux données du problème d'optimisation et leur arrangement optimal correspondra à la solution optimale de l'application. Les objectifs sont les critères d'optimisation qui permettent la mesure de la performance de l'application pour un jeu de variables choisi. Enfin, les contraintes rendent possible la liaison entre les variables et les données du problème d'optimisation.

3.2 Problème d'optimisation

Le problème d'optimisation est vu comme un problème d'analyse fonctionnelle. On y cherche un extremum d'une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$, où $d < +\infty$ est la dimension finie du problème d'optimisation. La fonction f est appelée « fonction objective » et les variables $\mathbf{x} \mapsto f(\mathbf{x})$ sont recherchées sur Ω , un sous-ensemble non vide de $\mathbb{R}^d$, tel que $\mathbf{x} \in \Omega \subset \mathbb{R}^d$. L'ensemble Ω correspond au domaine des contraintes et $\forall \mathbf{x} \in \mathbb{R}^d$ vérifiant $\mathbf{x} \in \Omega$ est un point admissible du problème d'optimisation. La formulation générale d'un problème d'optimisation s'écrit alors simplement comme

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}). \quad (3.2.1)$$

La résolution de (3.2.1) revient à déterminer un minimum de f sur l'ensemble admissible Ω . On dira que $\mathbf{x}_*$ est un « minimum global » de f sur $\Omega \subset \mathbb{R}^d$ si $\mathbf{x}_*$ est la solution de (3.2.1). Il sera un « minimum local » de f sur $\Omega \subset \mathbb{R}^d$ si et seulement s'il existe un voisinage $\mathcal{V}_{\mathbf{x}_*}$ de $\mathbf{x}_*$ tel que $\forall \mathbf{x} \in \mathcal{V}_{\mathbf{x}_*} \cap \Omega, f(\mathbf{x}_*) \leq f(\mathbf{x})$. La recherche d'un maximum de f peut se ramener au problème (3.2.1) car $\min_{\mathbf{x} \in \Omega} f(\mathbf{x})$ et $\max_{\mathbf{x} \in \Omega} -f(\mathbf{x})$ sont équivalents. On pourra donc toujours se restreindre au cas de la minimisation.

3.3 Méthodes d'optimisation

Les généralités sur l'utilisation des méthodes d'optimisation sont brièvement données dans cette Section 3.3. Quelques exemples d'algorithmes d'optimisation courants sont également présentés.

3.3.1 Généralités

Il existe de nombreux types d'optimisations. Par exemple, l'optimisation est dite « discrète » lorsque Ω est dénombrable, on parle d'optimisation « stochastique » si les données sont aléatoires, ou encore « multi-critère » lorsque le problème d'optimisation est composé de plusieurs fonction objectives.

Suivant l'application, et donc de la nature de la fonction objective f , un choix de l'utilisateur devra être fait sur la méthode d'optimisation à utiliser pour résoudre le problème d'optimisation. En effet, comme le montre la Figure 3.1, les évolutions de la fonction f peuvent être extrêmement variées entre les cas convexe, multi-modale et bruitée. Une discussion plus complète sur les différentes natures de f peut être trouvée dans (Venderplaats, 1984; Fletcher, 1987). Les méthodes d'optimisation utilisant les calculs des dérivées (descentes de gradient,

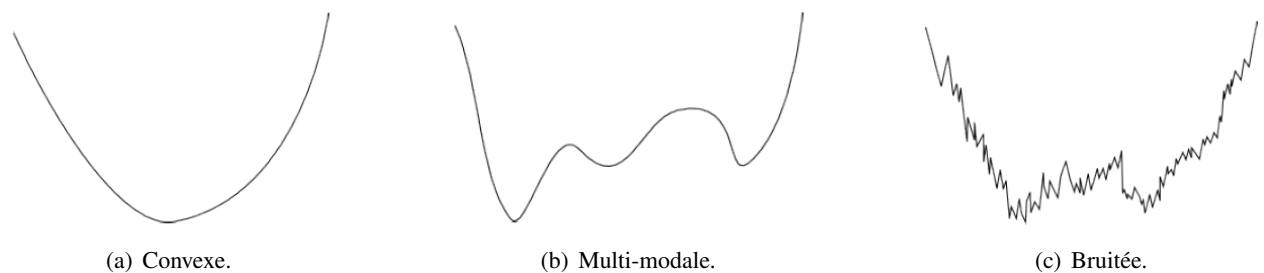


Figure 3.1 – Exemples d'évolutions de $f(x)$. Images issues de (Duvigneau, 2015).

lagrangiennes pour les cas sous contraintes, etc.) seront par exemple très bien adaptées au cas convexe de la Figure 3.1(a). On comprend aisément que ce ne sera pas forcément le cas avec les fonctions multi-modales et encore moins dans les cas bruités. Dans ces deux derniers cas, il sera préférable d'utiliser des méthodes d'optimisation globales qui ne reposent pas sur les calculs de dérivées.

Dans la plupart des applications réelles, la formulation analytique de la fonction objective est inconnue. Pour les cas d'algorithmes à descente de gradient, les calculs des dérivées doivent alors être approchés numériquement par différences finies. L'application est donc vue comme une « boîte noire », communément appelée « black-box » en anglais, et la méthode d'optimisation ne prend alors pas en considération le fonctionnement interne de l'application. La Figure 3.2 ci-dessous schématise cette vision. Les méthodes déterministes comme le « Simplex » (Nelder and Mead, 1965) ou le « Pattern Search » (Hooke and Jeeves, 1961), sont par exemple des méthodes qui ne reposent pas sur les calculs de dérivées et qui sont deux alternatives, très utilisées aujourd'hui, pour cette dernière propriété.

Les méthodes stochastiques comme les « Algorithmes Génétiques (GA) » (Goldberg and Holland, 1988), les « Stratégies d'Évolution (ES) » (Michalewicz and Schoenauer, 1996; Hansen and Ostermeier, 2001) ou encore les « Particle Swarm Optimization (PSO) » (Kennedy and Eberhart, 1995), sont de puissantes méthodes

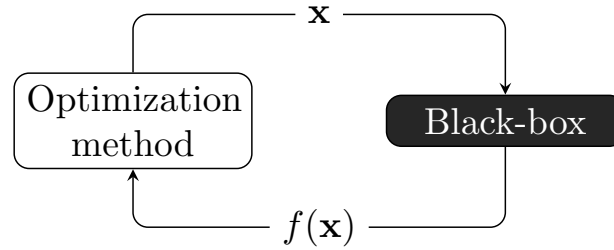


Figure 3.2 – Application à optimiser vue comme une boîte noire.

d'optimisation globale applicables aux cas multi-modaux et bruités (Figures 3.1(b) et 3.1(c)). Elles ont cependant le désavantage d'être généralement coûteuses en terme de nombre d'évaluations de la fonction objective.

Toutes ces méthodes d'optimisation sans gradient précédemment décrites sont applicables à des problèmes d'optimisations pour lesquels le coût d'évaluation de la fonction objective est relativement faible. Lorsque ce n'est pas le cas, il est judicieux d'utiliser des méthodes d'optimisation basées sur des « surfaces de réponse ». Dans ces approches, un modèle mathématique est construit et enrichi itérativement, sur la base des observations de la fonction objective, pour se substituer au problème d'optimisation directe. Le coût global de l'optimisation est alors fortement réduit. C'est l'une de ces dernières méthodes, considérant l'application à optimiser comme une boîte noire et basée sur des surfaces de réponse, qui est principalement appliquée et développée dans cette thèse. La méthode d'optimisation est expliquée et illustrée dans le Chapitre 4.

3.3.2 Algorithmes d'optimisation classiques

Trois types d'algorithmes qui sont couramment utilisés aujourd'hui sont présentés ci-après. Le premier repose sur les calculs de dérivées, contrairement aux deux autres, qui appartiennent à la famille des méthodes déterministes pour le second et aux algorithmes évolutionnaires pour le dernier.

3.3.2.1 Descente de gradient

Soit une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $\mathbf{x} \mapsto f(\mathbf{x})$ analytiquement connue et différentiable que l'on souhaite minimiser. Les algorithmes à descente de gradient sont des méthodes itératives qui procèdent par des améliorations successives de l'évaluation $f(\mathbf{x})$. Le principe est, à une itération donnée, de déterminer la direction du gradient $\nabla f(\mathbf{x})$, pour effectuer un déplacement dans la direction opposée et ainsi faire décroître la fonction à la prochaine itération. Il existe plusieurs variantes d'algorithmes à descente de gradient, comme la méthode de la plus grande descente, le gradient conjugué, la méthode de Newton ou quasi-Newton. La version la plus simple de l'algorithme à gradient est donnée par l'Algorithme 3.1. Dans cette version le pas de déplacement est déterminé à chaque itération ; cela peut se faire selon les critères de « Wolfe » (Wolfe, 1969).

Les algorithmes à descente de gradient ont la mauvaise faculté de pouvoir facilement converger vers des minima locaux si la fonction f est multimodale. Ils sont de plus dépendants des dérivées de f qui peuvent conduire à une divergence rapide de l'algorithme si les observations sont bruitées. Leur utilisation pour résoudre un problème d'optimisation d'une application réelle est donc difficilement envisageable en pratique.

Algorithme 3.1 : Algorithme de la plus grande descente.

```

Initialize :  $\mathbf{x}_1 \in \mathbb{R}^d, i \leftarrow 1$ 
1 while stopping criterion not reached do
2    $\alpha_i \leftarrow \arg \min_{\alpha \in \mathbb{R}^+} f(\mathbf{x}_i - \alpha \nabla f(\mathbf{x}_i))$            !Sélection du pas
3    $\mathbf{x}_{i+1} \leftarrow \mathbf{x}_i - \alpha_i \nabla f(\mathbf{x}_i)$                        !Déplacement
4    $i \leftarrow i + 1$ 

```

Lorsque le système étudié est régi par une équation aux dérivées partielles, la méthode la plus efficace pour calculer le gradient de la fonction objective est de résoudre une équation adjointe (Jameson, 1988; Keane and Nair, 2005). Cependant, pour des problèmes instationnaires ou en présence de couplages, l'approche adjointe est difficile à mettre en œuvre.

3.3.2.2 Méthode du « Simplex »

La méthode du « Simplex » est un algorithme d'optimisation développé par Nelder and Mead (1965). L'idée est de servir d'un simplexe de dimension d , qui est une enveloppe convexe de $d + 1$ sommets $(\mathbf{x}_1, \dots, \mathbf{x}_{d+1})$ de $\mathbb{R}^d$, avec les conditions $(\mathbf{x}_1 - \mathbf{x}_{d+1}, \dots, \mathbf{x}_d - \mathbf{x}_{d+1})$ qui soient linéairement indépendants. En partant d'un premier simplexe, l'algorithme consiste à itérativement le déformer et le déplacer dans $\mathbb{R}^d$ afin de faire converger ses $d + 1$ sommets vers la solution du problème d'optimisation. L'Algorithme 3.2 présente les étapes de cette méthode lorsque l'on souhaite minimiser la fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}, \mathbf{x} \mapsto f(\mathbf{x})$, analytiquement inconnue et possiblement non convexe. La procédure itérative consiste à déterminer le sommet du simplexe où la fonction f est maximale, pour ensuite le remplacer avec une étape de réflexion de ce sommet par rapport au centre de gravité $\mathbf{x}_0$ des d autres sommets. Si la réponse de f en ce nouveau point $\mathbf{x}_r$ est inférieure à l'ensemble des réponses aux différents sommets, alors le simplexe est étiré (étape d'expansion) dans cette direction. Sinon, le simplexe est contracté ou réduit par une homothétie centrée sur le sommet où la fonction est minimale.

La méthode du « Simplex » présente plusieurs avantages. En effet, sa mise en œuvre est assez simple et elle permet la minimisation de fonctions sans faire appel aux calculs de dérivées. Néanmoins, son efficacité est dépendante du choix initial du simplexe de départ, qui pourra faire converger le simplexe final vers des optimums locaux. Sa performance a de plus tendance à diminuer lorsque la dimension d du problème d'optimisation augmente. Enfin, il est possible que les simplexes dégénèrent au cours des itérations : par exemple dans $\mathbb{R}^2$, si trois sommets sont proches d'être colinéaires. La méthode de « Torczon » (Torczon, 1991) permet d'éviter cette dégénérescence en maintenant la géométrie du simplexe. Différentes variantes et améliorations de cette méthode du « Simplex » ont été développées depuis (Nelder and Mead, 1965), en particulier Kolda et al. (2003) donnent un état de l'art sur ces méthodes déterministes.

3.3.2.3 Stratégies d'évolution

Les « Stratégies d'Évolution (ES) » (Michalewicz and Schoenauer, 1996; Hansen and Ostermeier, 2001) sont des algorithmes évolutionnaires, qui tirent leur nom d'une analogie avec la théorie de l'évolution des es-

Algorithme 3.2 : Méthode du « Simplex ».

```

Initialize : Set  $(\mathbf{x}_1^{(1)}, \dots, \mathbf{x}_{d+1}^{(1)})$  by ranking  $f(\mathbf{x}_1^{(1)}) \leq \dots \leq f(\mathbf{x}_{d+1}^{(1)})$ ,  $i \leftarrow 1$ ,
                $\alpha \leftarrow 1, \gamma \leftarrow 2, \rho \leftarrow \frac{1}{2}, \sigma \leftarrow \frac{1}{2}$ 
1 while stopping criterion not reached do
2    $\mathbf{x}_0 \leftarrow \frac{1}{d} \sum_{j=1}^d \mathbf{x}_j^{(i)}$                                      !Centre de gravité sur  $d$  sommets
3    $\mathbf{x}_r \leftarrow \mathbf{x}_0 + \alpha(\mathbf{x}_0 - \mathbf{x}_{d+1}^{(i)})$ 
4   if  $f(\mathbf{x}_1^{(i)}) \leq f(\mathbf{x}_r) < f(\mathbf{x}_d^{(i)})$  then
5      $\mathbf{x}_{d+1}^{(i+1)} \leftarrow \mathbf{x}_r$                                      !Réflexion
6   else if  $f(\mathbf{x}_r) < f(\mathbf{x}_1^{(i)})$  then
7      $\mathbf{x}_e \leftarrow \mathbf{x}_r + \gamma(\mathbf{x}_r - \mathbf{x}_0)$ 
8     if  $f(\mathbf{x}_e) < f(\mathbf{x}_r)$  then
9        $\mathbf{x}_{d+1}^{(i+1)} \leftarrow \mathbf{x}_e$                                      !Expansion
10    else
11       $\mathbf{x}_{d+1}^{(i+1)} \leftarrow \mathbf{x}_r$                                      !Réflexion
12  else
13     $\mathbf{x}_c \leftarrow \mathbf{x}_0 + \rho(\mathbf{x}_{d+1}^{(i)} - \mathbf{x}_0)$                                      !On a  $f(\mathbf{x}_d^{(i)}) \leq f(\mathbf{x}_r)$ 
14    if  $f(\mathbf{x}_c) < f(\mathbf{x}_{d+1}^{(i)})$  then
15       $\mathbf{x}_{d+1}^{(i+1)} \leftarrow \mathbf{x}_c$                                      !Contraction
16    else
17      for  $k \leftarrow 2$  to  $d+1$  do
18         $\mathbf{x}_k^{(i+1)} \leftarrow \mathbf{x}_1 + \sigma(\mathbf{x}_k^{(i)} - \mathbf{x}_1)$                                      !Homothétie par  $\sigma$ 
19   $i \leftarrow i + 1$ 
20  Set  $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_{d+1}^{(i)})$  by ranking  $f(\mathbf{x}_1^{(i)}) \leq \dots \leq f(\mathbf{x}_{d+1}^{(i)})$                                      !Renumérotation

```

pièces de « Darwin » (Darwin, 1873). Ceux sont des algorithmes d'optimisation stochastiques très efficaces pour résoudre un problème multimodal. Leur principal inconvénient concerne le nombre important d'évaluations de la fonction à optimiser.

Soit une fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $\mathbf{x} \mapsto f(\mathbf{x})$ analytiquement inconnue et possiblement non convexe que l'on souhaite minimiser. Le principe des algorithmes ES est d'optimiser des individus $\mathbf{x}_i \in \mathbb{R}^d$, à l'aide d'opérateurs de « mutation » et de « sélection ». La version la plus simple des algorithmes ES est la $(1+1)$ -ES qui est présentée dans l'Algorithme 3.3 ci-après. L'étape de mutation est effectuée par l'ajout d'une variable aléatoire tirée selon une distribution normale et d'un paramètre de pas σ . L'adaptation de σ au cours des générations se fait ici selon la « règle de succès en 1/5 » (Rechenberg, 1978). La sélection est un choix déterministe des meilleurs individus au sens de la fonction objective f ; on dit que cette sélection est « élitiste ».

L'Algorithme $(1+1)$ -ES peut aussi se généraliser sous les formes canoniques $(\mu/\rho, \lambda)$ -ES et $(\mu/\rho + \lambda)$ -ES. Dans ces versions, μ est un nombre de « parents », $\rho \leq \mu$ est un nombre de « mixage » qui correspond au nombre de parents impliqués dans la création d'un nouvel individu et λ est un nombre « d'enfants » générés. La

Algorithme 3.3 : $(1 + 1)$ -ES.

```

Initialize :  $\mathbf{x} \in \mathbb{R}^d, \sigma \in \mathbb{R}^+$ 
1 while stopping criterion not reached do
2    $\mathbf{x}' \leftarrow \mathbf{x} + \sigma \mathcal{N}(\mathbf{0}_d, \mathbf{I}_d)$                                 !Mutation
3   if  $f(\mathbf{x}') \leq f(\mathbf{x})$  then
4      $\mathbf{x} \leftarrow \mathbf{x}'$                                                 !Sélection « élitiste »
5      $\sigma \leftarrow 1.5\sigma$                                           !Augmentation du pas
6   else
7      $\sigma \leftarrow 1.5^{-1/4}\sigma$                                     !Diminution du pas

```

dénomination « , » ou « + » définit le schéma de sélection utilisé. Dans l'algorithme $(\mu/\rho, \lambda)$ -ES, les nouveaux parents sont choisis parmi les enfants uniquement ($\mu \leq \lambda$), alors que la sélection des nouveaux parents dans l'algorithme $(\mu/\rho + \lambda)$ -ES se fait parmi les parents et les enfants. L'Algorithme 3.4 ci-après résume ces deux versions étendues de l'ES. L'adaptation du pas de mutation n'y est pas présentée afin de rester dans un cadre descriptif général. Davantage de détails sur les méthodes d'adaptations sont disponibles dans le résumé très complet de Hansen et al. (2015).

Algorithme 3.4 : $(\mu/\rho + \lambda)$ -ES.

```

Input :  $\mu, \rho, \lambda \in \mathbb{N}$ 
Initialize :  $\sigma \in \mathbb{R}^{d+}, \mathcal{P} = \{(\mathbf{x}_i, f(\mathbf{x}_i)) | 1 \leq i \leq \mu\}, \mathcal{P}_\lambda = \{\emptyset\}$ 
1 while stopping criterion not reached do
2   for  $i \leftarrow 1$  to  $\lambda$  do
3      $\mathbf{x}_i \leftarrow \text{recombine}(\rho, \mathcal{P})$                                 !Recombinaison de  $\rho \geq 2$  parents ( $\rho \leq \mu$ )
4      $\sigma \leftarrow \text{adapt\_s}(\sigma)$                                 !Adaptation du pas de mutation
5      $\mathbf{x}_i \leftarrow \mathbf{x}_i + \sigma \circ \mathcal{N}(\mathbf{0}_d, \mathbf{I}_d) \in \mathbb{R}^d$             !Mutation des individus
6      $\mathcal{P}_\lambda \leftarrow \mathcal{P}_\lambda \cup \{(\mathbf{x}_i, f(\mathbf{x}_i))\}$                     !Insertion dans la population des enfants
7   if  $(\mu/\rho, \lambda)$ -ES then
8      $\mathcal{P} \leftarrow \text{select\_best}(\mu, \mathcal{P}_\lambda)$                         !Sélection parmi les enfants
9   else
10     $\mathcal{P} \leftarrow \mathcal{P} \cup \mathcal{P}_\lambda$                                 !Insertion dans la population des parents
11     $\mathcal{P} \leftarrow \text{remove\_old}(\mathcal{P})$                                 !Supprime les individus de l'ancienne génération
12     $\mathcal{P} \leftarrow \text{select\_best}(\mu, \mathcal{P})$                         !Sélection parmi les enfants et parents
13   $\mathcal{P}_\lambda \leftarrow \{\emptyset\}$                                     !Supprime la population des enfants

```

L'algorithme CMA-ES, qui a été développé par Nikolaus Hansen (Hansen, 2006b), est une évolution de l'approche originale ES. La méthode CMA-ES utilise une distribution multi-normale dans l'étape de mutation des individus. Ceci est fait par l'intermédiaire d'une matrice de covariance $\mathbf{C}$, de dimension $[d \times d]$. La « grande force » de cette méthode repose sur l'adaptation, au cours des générations, de cette matrice de covariance. En particulier, cela permet d'adapter la génération des individus à l'anisotropie de la fonction objective. La Figure 3.3 illustre trois distributions de lois normales à deux dimensions pour lesquelles uniquement σ_{12}^2 varie. C'est

cette variation qui rend possible l'orientation des ellipses de la distribution lors de la mutation. Ainsi, le tirage

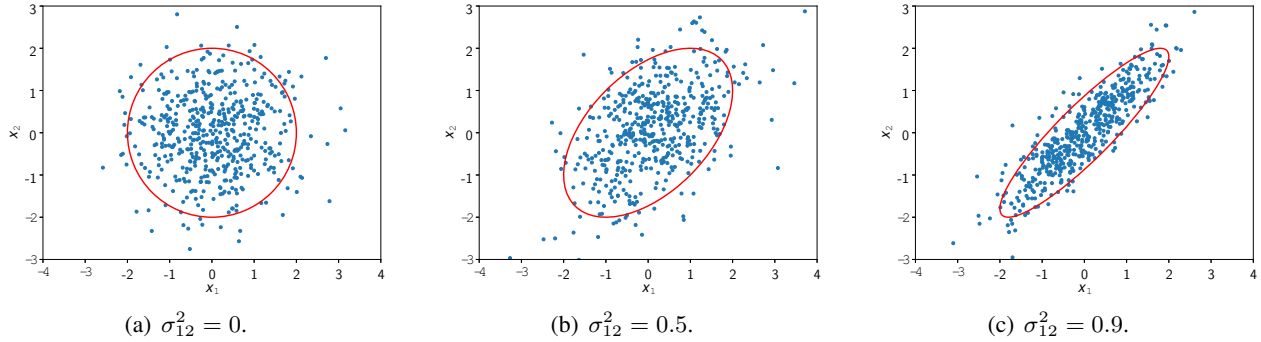


Figure 3.3 – Distributions de lois normales multidimensionnelles où $d = 2$, $\boldsymbol{\mu} = \mathbf{0}_2$ et $\sigma_{11}^2 = \sigma_{22}^2 = 1$, pour différentes valeurs de covariances σ_{12}^2 .

des enfants dans la méthode CMA-ES se fait selon $\mathbf{x}_i \leftarrow \mathbf{m} + \sigma \mathcal{N}(\mathbf{0}_d, \mathbf{C})$ pour $i = 1, \dots, \lambda$, où $\mathbf{m} \in \mathbb{R}^d$ et $\sigma \in \mathbb{R}^+$ sont respectivement la moyenne des meilleurs individus de la génération et le pas de mutation. Ces deux paramètres sont adaptés au cours des générations, tout comme la matrice de covariance $\mathbf{C} \in \mathbb{R}^{d \times d}$. Les différentes méthodes d'adaptation pour $\mathbf{m}$, σ et $\mathbf{C}$ sont expliquées dans (Hansen et al., 2003; Hansen, 2006b; Hansen et al., 2015).

Ce dernier algorithme CMA-ES est très utilisé dans les différents travaux de cette thèse. Il permet entre autres, de déterminer la calibration optimale des hyperparamètres, qui définissent les surfaces de réponse, utilisées ensuite dans les méthodes d'optimisation par modèles de substitution. Son utilité est également omniprésente au sein même de ces méthodes d'optimisation, afin de travailler directement sur les surfaces de réponse, comme par exemple, pour en déterminer les extrema. Dans la suite du manuscrit, la thématique porte sur l'utilisation de modèles de substitution pour l'application d'une méthode globale d'optimisation. Le prochain Chapitre 4 traite précisément de ce sujet, en détaillant la méthode de construction ainsi que différentes applications des « méta-modèles Gaussiens ».

Méta-modèles Gaussiens

Sommaire

4.1 Définitions générales	56
4.1.1 Variables et vecteurs aléatoires	56
4.1.2 Processus Gaussiens	57
4.2 Méta-modèles Gaussiens	60
4.2.1 Prédiction	61
4.2.2 Fonction de covariance	63
4.2.3 Détermination des hyperparamètres	64
4.2.4 Plan d'expériences	65
4.3 Applications des méta-modèles Gaussiens	67
4.3.1 Analyse de sensibilité	67
4.3.2 Classification	71
4.3.3 Optimisation globale	73
4.3.3.1 Algorithme de l'EGO	74
4.3.3.2 Fonctions de mérite	75
4.3.3.3 Application de l'EGO à la fonction Branin	77

Un phénomène physique est généralement modélisable par l'intermédiaire d'un modèle simplifié, qui peut être basé sur une expérience ou un code numérique. Dans la plupart des cas, on s'intéresse à l'influence de variables ou paramètres d'entrée sur une ou plusieurs sorties du modèle. Cette influence peut ensuite être utilisée dans un processus d'optimisation ou dans une étude de sensibilité par exemple. Cependant, l'exploitation du modèle est très souvent trop complexe ou coûteuse pour être directement appliquée. L'utilisation d'un modèle mathématique, que l'on nomme « méta-modèle » (« modèle de modèle »), pour se substituer au modèle physique est alors justifiée. Le méta-modèle, aussi appelé « surface de réponse », est une fonction mathématique

représentative du modèle physique, construit à partir de n observations ou simulations finies. On peut ainsi approcher les réponses du modèle physique à l'aide du méta-modèle dans l'ensemble du domaine de variabilité, et cela pour un temps d'évaluation négligeable. L'influence des variables d'entrée du phénomène physique initial peut alors être exploitée. Il existe différentes familles de méta-modèles comme le très simple modèle linéaire, les polynômes, les fonctions de base radiale (Rippa, 1999), les splines ou bien encore les réseaux de neurones (Chen, 1996). Dans ces travaux de thèse, on s'intéresse plus particulièrement aux méta-modèles Gaussiens (Rasmussen and Williams, 2006) et à leurs applications dans le cadre de l'optimisation. Les méta-modèles Gaussiens ont l'avantage d'approcher le modèle physique d'intérêt par des fonctions statistiques communément appelées « processus Gaussiens ». Ce contexte statistique donne accès à la variance des méta-modèles ce qui permet, d'une part une exploration efficace du domaine Ω , et d'autre part de travailler avec des observations possiblement bruitées.

Dans la suite du Chapitre 4, la construction des méta-modèles Gaussiens et quelques exemples de leurs applications, en plus de l'optimisation, sont détaillés.

4.1 Définitions générales

4.1.1 Variables et vecteurs aléatoires

Cette section contient des rappels élémentaires, sous formes de définitions, de la théorie des probabilités sur les variables et sur les vecteurs aléatoires.

Définition 4.1.1 (Espace de probabilité). *Un espace probabilisé est un espace mesurable $(\Omega, \mathcal{F})$ muni d'une application P appelée mesure de probabilité de $\mathcal{F}$ dans $[0,1]$ tel que $P(\Omega) = 1$ et $P(A \cup B) = P(A) + P(B)$ si $A \cap B = \emptyset$, où A et B sont des ensembles mesurables. Les ensembles mesurables $A \in \mathcal{F}$ sont appelés les événements de $\mathcal{F}$.*

Définition 4.1.2 (Variable aléatoire). *Une variable aléatoire est une application mesurable Z de l'espace probabilisé $(\Omega, \mathcal{F}, P)$ à valeurs dans $\mathbb{R}$, $Z : \Omega \rightarrow \mathbb{R}$.*

La loi de probabilité de Z est la mesure de probabilité P_Z définie sur $\mathbb{R}$ par $P_Z(\mathcal{A}) = P(Z \in \mathcal{A})$. Autrement dit, P_Z est l'image par l'application Z de la probabilité P définie sur $(\Omega, \mathcal{F})$.

On dit que Z est à densité s'il existe une fonction f définie sur $\mathbb{R}$, positive ou nulle, intégrable sur $\mathbb{R}$, continue *p.p.* et d'intégrale égale à 1, appelée densité de probabilité, telle que $P(Z \in [a,b]) = \int_a^b f(z)dz$.

L'espérance d'une variable aléatoire Z , notée $\mathbb{E}[Z]$, est définie par $\mathbb{E}[Z] = \int_{\Omega} Z dP$. Si Z est à densité f et d'intégrale bien définie, son espérance est alors donnée par $\mathbb{E}[Z] = \int_{-\infty}^{+\infty} z f(z) dz$.

La variance d'une variable aléatoire Z , notée $\text{Var}(Z)$, est définie par $\text{Var}(Z) = \mathbb{E}[(Z - \mathbb{E}[Z])^2]$. L'écart type $\sigma(Z)$ est la racine carrée de la variance $\sigma(Z) = \sqrt{\text{Var}(Z)}$.

La covariance de deux variables aléatoires Z et V ayant des variances finies, notée $\text{Cov}(Z, V)$, est définie par $\text{Cov}(Z, V) = \mathbb{E}[(Z - \mathbb{E}[Z])(V - \mathbb{E}[V])] = \mathbb{E}[ZV] - \mathbb{E}[Z]\mathbb{E}[V]$. Si Z et V sont indépendantes on a alors

$$\text{Cov}(Z, V) = 0.$$

Définition 4.1.3 (Vecteur aléatoire). *Un vecteur aléatoire $\mathbf{Z}_n$ de dimension n est une application mesurable d'un espace probabilisé $(\Omega, \mathcal{F}, P)$ dans $\mathbb{R}^n$. Il est composé d'un n -uplet $\mathbf{Z}_n = (Z_1, \dots, Z_n)^T$ où chaque Z_i est une variable aléatoire réelle (aussi appelée i -ème marginale).*

La loi du vecteur aléatoire $\mathbf{Z}_n$ est une mesure de probabilité sur $\mathbb{R}^n$ et se définit comme $P_{\mathbf{Z}_n}(\mathcal{A}_1 \times \dots \times \mathcal{A}_n) = P(Z_1 \in \mathcal{A}_1, \dots, Z_n \in \mathcal{A}_n)$.

Le vecteur $\mathbf{Z}_n$ de $\mathbb{R}^n$ admet une densité $f_{\mathbf{Z}_n}(z_1, \dots, z_n)$ si $P_{\mathbf{Z}_n}(\mathbf{A}_n) = \int_{\mathbf{A}_n} f(z_1, \dots, z_n) dz_1 \dots dz_n$, où $f_{\mathbf{Z}_n}$ est une fonction positive définie sur $\mathbb{R}^n$, continue $p.p.$ et d'intégrale égale à 1.

L'espérance d'un vecteur aléatoire $\mathbf{Z}_n$ est le vecteur constitué des espérances de chacune des coordonnées $\mathbb{E}[\mathbf{Z}_n] = (\mathbb{E}[Z_1], \dots, \mathbb{E}[Z_n])^T$.

La matrice de covariance du vecteur $\mathbf{Z}_n$ est la matrice carrée Σ_n de taille $[n \times n]$ dont les éléments sont donnés par $\Sigma_{ij} = \text{Cov}(Z_i, Z_j)$, $\forall i, j = 1, \dots, n$. Plus précisément, on a $\text{Cov}(Z_i, Z_j) = \mathbb{E}[(Z_i - \mathbb{E}[Z_i])(Z_j - \mathbb{E}[Z_j])] = \mathbb{E}[Z_i Z_j] - \mathbb{E}[Z_i]\mathbb{E}[Z_j]$. La matrice Σ_n est dite semi-définie positive si $\forall z, \sum_{i,j} \Sigma_{ij} z_i z_j \geq 0$ et définie positive si, de plus, $\det(\Sigma_n) \neq 0$. Si les variables Z_i sont toutes indépendantes, on a $\text{Cov}(Z_i, Z_j) = 0$, $\forall i \neq j$ et Σ_n est diagonale. La réciproque est cependant fausse.

4.1.2 Processus Gaussiens

Les définitions mathématiques des processus Gaussiens sont détaillées ci-après.

Définition 4.1.4 (Variable Gaussienne). *Z est dite variable Gaussienne et suit la loi normale $\mathcal{N}(\mu, \sigma^2)$ si elle admet pour densité $f(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right)$, $\forall z \in \mathbb{R}$.*

La Figure 4.1 présente deux exemples de densité $f(z)$ correspondant aux lois normales $\mathcal{N}(0, 0.5)$ et $\mathcal{N}(1, 0.75)$. Une variable aléatoire $Z \sim \mathcal{N}(\mu, \sigma^2)$ a pour espérance $\mathbb{E}[Z] = \mu$ et pour variance $\text{Var}(Z) = \sigma^2$. Z peut se

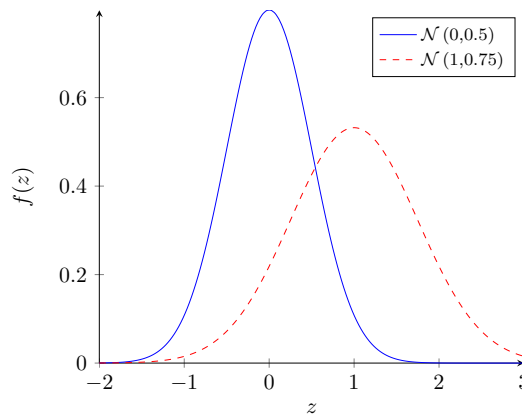


Figure 4.1 – Densité de lois normales.

définir comme la translatée et dilatée de la loi normale centrée réduite $Z_0 \sim \mathcal{N}(0,1)$ par la transformation $Z = \mu + \sigma Z_0$.

Définition 4.1.5 (Vecteur Gaussien). $\mathbf{Z}_n$ est un vecteur aléatoire Gaussien si toute combinaison linéaire à coefficients réels des Z_i suit une loi normale $\mathcal{N}(\mu, \sigma^2)$. Chaque marginale Z_i d'un vecteur Gaussien suit une loi normale. La réciproque est fausse, sauf dans le cas où les coordonnées sont indépendantes.

L'espérance de $\mathbf{Z}_n$ est le vecteur des espérances de ses marginales que l'on écrit comme $\boldsymbol{\mu}_n = \mathbb{E}[\mathbf{Z}_n] = (\mu_1, \dots, \mu_n)^T = (\mathbb{E}[Z_1], \dots, \mathbb{E}[Z_n])^T$. La matrice de covariance $\boldsymbol{\Sigma}_n$ d'un vecteur Gaussien est la matrice carrée symétrique, définie positive $\boldsymbol{\Sigma}_n = \text{Cov}(Z_i, Z_j)_{1 \leq i, j \leq n}$.

Comme pour les variables aléatoires Gaussiennes, un vecteur Gaussien $\mathbf{Z}_n \sim \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$, avec $\boldsymbol{\Sigma}_n$ inversible, peut se définir comme la translatée et dilatée du vecteur Gaussien standard $\mathbf{Z}_{0_n} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$, dont les composantes sont indépendantes, par $\mathbf{Z}_n = \sqrt{\boldsymbol{\Sigma}_n} \mathbf{Z}_{0_n} + \boldsymbol{\mu}_n$.

Les notions d'indépendance et de corrélation sont équivalentes pour les vecteurs Gaussiens. Cette importante affirmation se démontre en montrant la réciproque à l'aide de la fonction caractéristique. Ainsi, pour que les coordonnées Z_i du vecteur Gaussien $\mathbf{Z}_n$ soient indépendantes il faut et il suffit que la matrice de covariance $\boldsymbol{\Sigma}_n$ soit diagonale.

Considérons n variables aléatoire $Z_1, \dots, Z_n$ indépendantes et de loi $\mathcal{N}(\mu_1, \sigma_1^2), \dots, \mathcal{N}(\mu_n, \sigma_n^2)$. La variable Z_i admet donc pour densité $f_{Z_i}(z) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2}\left(\frac{z - \mu_i}{\sigma_i}\right)^2\right)$, $\forall z \in \mathbb{R}, \forall i = 1, \dots, n$.

En raison de l'indépendance des Z_i , la densité conjointe du vecteur $(Z_1, \dots, Z_n)^T$ s'écrit

$$f_{Z_1, \dots, Z_n}(z_1, \dots, z_n) = \frac{1}{(\sqrt{2\pi})^n \prod_{i=1}^n \sigma_i} \exp\left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{z_i - \mu_i}{\sigma_i}\right)^2\right).$$

Enfin, en notant $\mathbf{Z}_n = (Z_1, \dots, Z_n)^T \sim \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$, on peut réécrire la densité conjointe du vecteur Gaussien par

$$f_{\mathbf{Z}_n}(z_1, \dots, z_n) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\boldsymbol{\Sigma}_n)}} \exp\left(-\frac{1}{2}(\mathbf{z}_n - \boldsymbol{\mu}_n)^T \boldsymbol{\Sigma}_n^{-1} (\mathbf{z}_n - \boldsymbol{\mu}_n)\right), \quad \forall \mathbf{z}_n = (z_1, \dots, z_n)^T \in \mathbb{R}^n,$$

où l'on rappelle que $\boldsymbol{\Sigma}_n$ est symétrique et définie positive.

Soit un vecteur Gaussien $\mathbf{Z}_{n_1+n_2} = \begin{pmatrix} \mathbf{Z}_{n_1} \\ \mathbf{Z}_{n_2} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}\right)$ dont la matrice de covariance $\boldsymbol{\Sigma}_{11}$ de $\mathbf{Z}_{n_1}$ est définie positive. La loi conditionnelle de $\mathbf{Z}_{n_2}$ sachant $\mathbf{Z}_{n_1} = \mathbf{z}_1$, où $\mathbf{z}_1 = (z_1, \dots, z_{n_1})^T \in \mathbb{R}^{n_1}$, est alors une loi Gaussienne ayant comme espérance $\mathbb{E}[\mathbf{Z}_{n_2} | \mathbf{Z}_{n_1} = \mathbf{z}_1] = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{z}_1 - \boldsymbol{\mu}_1)$ et comme matrice de covariance $\boldsymbol{\Sigma}_{\mathbf{Z}_{n_2} | \mathbf{Z}_{n_1} = \mathbf{z}_1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}$ (Rasmussen and Williams, 2006).

Définition 4.1.6 (Processus stochastique). Un processus stochastique (ou aléatoire) $Z = (Z(\mathbf{x}))_{\mathbf{x} \in D \subset \mathbb{R}^d}$ est une famille de variables aléatoires indexées par un ensemble D . L'ensemble des indices D fait généralement référence au temps, mais il peut aussi correspondre à l'espace ou aux deux à la fois. Lorsque $D \subset \mathbb{R}^d$, où

$d \geq 2$, on parle aussi de champ stochastique. Si D est un ensemble fini, le processus est un vecteur aléatoire. Pour $\mathbf{x} \in D$ fixé, $\omega \in \Omega \mapsto Z(\omega)$ est une variable aléatoire sur l'espace de probabilité $(\Omega, \mathcal{F}, P)$ et pour $\omega \in \Omega$ fixé, $\mathbf{x} \in D \mapsto Z(\omega)$ est une fonction à valeurs réelles, appelée trajectoire du processus.

Définition 4.1.7 (Processus Gaussien). Soit Z un processus aléatoire. Z est dit Gaussien si et seulement si toute combinaison linéaire de ses marginales $a_1 Z_{\mathbf{x}_1} + \dots + a_n Z_{\mathbf{x}_n}$ suit une loi Gaussienne, $\forall n \in \mathbb{N}, \forall \mathbf{x}_1, \dots, \mathbf{x}_n \in D$ et $\forall a_1, \dots, a_n \in \mathbb{R}$.

Par conséquent, toutes les marginales ou toute combinaison linéaire de marginales d'un processus Gaussien sont Gaussiennes. Dans ce cas, comme on l'a montré précédemment pour un vecteur Gaussien, le processus Gaussien Z admet un vecteur espérance $\boldsymbol{\mu}_Z(\mathbf{x}) = \mathbb{E}[Z(\mathbf{x})]$ et une fonction de covariance $\Sigma_Z(\mathbf{x}, \mathbf{x}') = \text{Cov}(Z(\mathbf{x}), Z(\mathbf{x}'))$ qui rend possible la construction de la matrice de covariance.

On dit alors qu'un processus Gaussien est entièrement caractérisé par sa moyenne et sa fonction de covariance et l'on note $Z \sim \mathcal{PG}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$, où $\boldsymbol{\mu}_n = (\boldsymbol{\mu}_Z(\mathbf{x}_1), \dots, \boldsymbol{\mu}_Z(\mathbf{x}_n))^T$ et $\boldsymbol{\Sigma}_n = \Sigma_Z(\mathbf{x}_i, \mathbf{x}_j)_{1 \leq i, j \leq n}$.

Il est possible de simuler un processus Gaussien, ou plus formellement d'en générer une réalisation, en utilisant une méthode de décomposition et un vecteur de variables aléatoires indépendantes Gaussiennes centrées et réduites. La factorisation de Cholesky est une possibilité qui consiste, pour une matrice symétrique définie positive $\mathbf{C}$, à déterminer une matrice triangulaire inférieure $\mathbf{L}$ telle que $\mathbf{C} = \mathbf{L}\mathbf{L}^T$. Cette décomposition permet également de calculer la matrice inverse $\mathbf{C}^{-1}$ et le déterminant de $\mathbf{C}$.

Soit le processus centré $Z \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{\Sigma}_n)$, si l'on considère un vecteur $\mathbf{g} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$ et que $\boldsymbol{\Sigma}_n$ admet une décomposition de Cholesky $\mathbf{L}\mathbf{L}^T$, on peut alors construire une réalisation du processus Gaussien à partir du vecteur $\mathbf{g}$. En effet, si $Z = \mathbf{L}\mathbf{g}$, le vecteur espérance s'écrit $\mathbb{E}[Z] = \mathbf{L}\mathbb{E}[\mathbf{g}] = \mathbf{0}_n$ et la matrice de covariance est $\mathbb{E}[ZZ^T] = \mathbf{L}\mathbb{E}[\mathbf{g}\mathbf{g}^T]\mathbf{L}^T = \boldsymbol{\Sigma}_n$. Par conséquent, une réalisation non-conditionnelle du processus Gaussien est donnée par $\hat{Z} = \mathbf{L}\mathbf{g}$ où chaque $g_i \sim \mathcal{N}(0,1)$, $\forall i = 1, \dots, n$. Dans le cas où l'on a pu observer $\mathbf{z}_1 = (z_1, \dots, z_{n_1})^T \in \mathbb{R}^{n_1}$ et que l'on souhaite générer une réalisation conditionnelle du processus Gaussien, on utilise la loi Gaussienne conditionnelle. La réalisation d'écrit alors $\hat{Z} = \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{z}_1 + \mathbf{L}\mathbf{g}$, où $\mathbf{L}$ provient de la décomposition $\mathbf{L}\mathbf{L}^T = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ et $g_i \sim \mathcal{N}(0,1)$, $\forall i = 1, \dots, n_2$.

Un exemple de 50 réalisations d'un processus Gaussien est donné par la Figure 4.2, pour les deux cas non-conditionnel et conditionnel.

Dans certain cas, la décomposition de Cholesky peut devenir très coûteuse. Ceci est particulièrement vrai lorsque le nombre d'échantillons sur lesquels on construit la réalisation est grand : dimension du vecteur $\mathbf{g}$. On peut dans ce cas recourir à des méthodes de décomposition spectrale tronquées. La matrice $\mathbf{C}$ peut alors se factoriser comme $\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, où $\boldsymbol{\Lambda}$ est une matrice diagonale de dimension $[n \times n]$ composée des valeurs propres ordonnées ($\lambda_1 \geq \dots \geq \lambda_n \geq 0$) et $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ est la $[n \times n]$ matrice où les $\mathbf{u}_i$ sont les vecteurs propres de $\mathbf{C}$. Pour toute matrice de covariance $\boldsymbol{\Sigma}_n$, les valeurs propres sont non-négatives donc $\sqrt{\boldsymbol{\Lambda}}$ est bien définie. Par conséquent, $\boldsymbol{\Sigma}_n = (\mathbf{U}\sqrt{\boldsymbol{\Lambda}})(\mathbf{U}\sqrt{\boldsymbol{\Lambda}})^T$ et l'on peut procéder de la même manière que précédemment avec la décomposition de Cholesky, où ici $\mathbf{U}\sqrt{\boldsymbol{\Lambda}}$ joue le rôle de $\mathbf{L}$. Par exemple, on obtient une réalisation non-conditionnelle avec $\hat{Z} = \mathbf{U}\sqrt{\boldsymbol{\Lambda}}\mathbf{g} = \sum_{i=1}^n \sqrt{\lambda_i} \mathbf{u}_i g_i$, où $g_i \sim \mathcal{N}(0,1)$, $\forall i = 1, \dots, n$. Cette décomposition est aussi appelée décomposition de « Karhunen-Loève » (Dony et al., 2001). L'intérêt d'une telle factorisation est

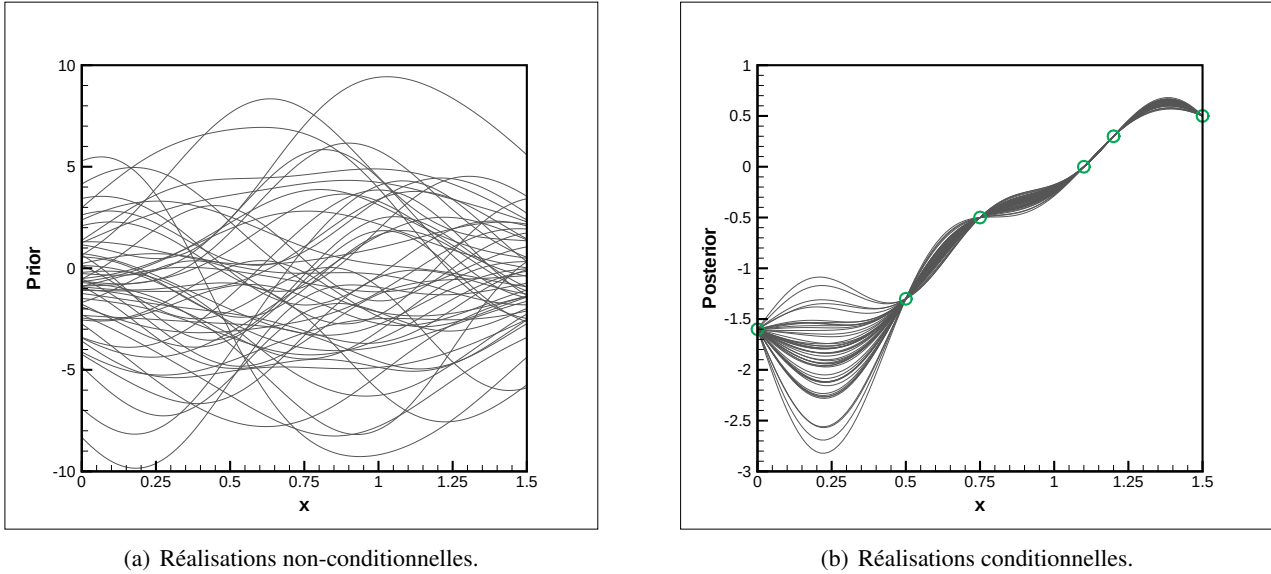


Figure 4.2 – Réalisations de processus Gaussiens dans un cas unidimensionnel.

justifié lorsque l'on utilise uniquement une partie m des valeurs et vecteurs propres ($m \ll n$). Dans ce cas, $\tilde{Z} \approx \hat{Z}$ et $\tilde{Z} = \sum_{i=1}^m \sqrt{\lambda_i} \mathbf{u}_i g_i$, où $g_i \sim \mathcal{N}(0,1)$, $\forall i = 1, \dots, m$. La qualité de l'approximation $\tilde{Z} \approx \hat{Z}$ est dépendante des valeurs propres tronquées $\lambda_{m+1}, \dots, \lambda_n$.

4.2 Méta-modèles Gaussiens

Les méta-modèles Gaussiens, ou régressions par processus Gaussiens, proviennent du domaine géostatistique qui applique des méthodes probabilistes à l'étude de phénomènes corrélés dans l'espace. La formalisation théorique a été élaborée par Georges Matheron (Matheron, 1963) dans les années 1960, qui a donné le nom « krigeage » (Cressie, 1990) pour honorer l'ingénieur sud-africain Danie Gerhardus Krige, connu pour ses travaux sur les gisements de minerai (Krige, 1981). L'utilisation des méta-modèles Gaussiens a depuis été étendue à de nombreux domaines (pétrole, océanographie, codes de calcul, etc.) où l'idée principale est d'utiliser la prédiction des méta-modèles, construits sur un nombre fini d'observations, pour réaliser une interpolation spatiale.

Deux principaux points de vue sont utilisés pour formaliser la construction des méta-modèles : « Weight space view » et « Function space view ». Dans le premier cas, l'idée est de construire un estimateur sans biais et de variance minimale qui soit une combinaison linéaire des observations. Dans le second cas, on travaille avec des vecteurs Gaussiens et des lois jointes et conditionnelles pour construire le méta-modèle processus Gaussien. Les deux approches sont différentes mais elles aboutissent au même méta-modèle Gaussien. Dans cette thèse on adoptera la seconde formulation, « Function space view ».

4.2.1 Prédiction

Soit un nombre fini d'observations $y_i \in \mathbb{R}$ telles que $y_i = f(\mathbf{x}_i) + \epsilon_i, \forall i = 1, \dots, n$, où f est la fonction d'intérêt et les ϵ_i sont des erreurs (bruits) aléatoires que l'on assume indépendantes et identiquement distribuées selon $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Chaque observation est associée à une composante $\mathbf{x}_i \in \Omega \subset \mathbb{R}^d$ que l'on regroupe dans la matrice (vecteur si $d = 1$) $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T$. On considère que la fonction d'intérêt $f(\mathbf{x})$ est une réalisation d'un processus Gaussien centré ($\boldsymbol{\mu}_n = \mathbf{0}_n$) ayant une fonction de covariance $C_f(\mathbf{x}, \mathbf{x}'; \boldsymbol{\Theta})$ connue. $\boldsymbol{\Theta}$ est un vecteur d'hyperparamètres à déterminer qui est inféré sur les observations (voir Section 4.2.3). On note également $\mathbf{C}(\boldsymbol{\Theta}) \in \mathbb{R}^{n \times n}$, la matrice de covariance composée des éléments $C_{i,j} = C_f(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\Theta}), 1 \leq i, j \leq n$ pour les points dans $\mathbf{X}_n$.

On souhaite prédire la réponse d'une observation $y(\mathbf{x})$ en un nouveau point non échantillonné $\mathbf{x}$ à l'aide du méta-modèle Gaussien. La loi jointe entre le vecteur des observations bruitées $\mathbf{Y}_n = (y_1, \dots, y_n)^T$ et la prédiction de l'observation $y(\mathbf{x})$ s'écrit comme

$$\begin{pmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{pmatrix} \sim \mathcal{N} \left(\mathbf{0}, \begin{bmatrix} \mathbf{C} + \sigma_\epsilon^2 \mathbf{I}_n & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^T(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\epsilon^2 \end{bmatrix} \right), \quad (4.2.1)$$

où $\kappa(\mathbf{x}) \doteq C_f(\mathbf{x}, \mathbf{x}; \boldsymbol{\Theta})$, $\mathbf{k}(\mathbf{x}) \doteq (C_f(\mathbf{x}, \mathbf{x}_1; \boldsymbol{\Theta}) \cdots C_f(\mathbf{x}, \mathbf{x}_n; \boldsymbol{\Theta}))^T$ et $\mathbf{I}_n$ est la matrice identité de $\mathbb{R}^n$. La loi de probabilité conditionnelle (Von Mises, 1964) permet d'écrire l'inférence de $y(\mathbf{x})$ sachant $\mathbf{Y}_n$, comme

$$P(y(\mathbf{x}) | \mathbf{Y}_n) = \frac{P(y(\mathbf{x}), \mathbf{Y}_n)}{P(\mathbf{Y}_n)}, \quad (4.2.2)$$

et en notant $\mathbf{C}_{n+1} = \begin{bmatrix} \mathbf{C} + \sigma_\epsilon^2 \mathbf{I}_n & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^T(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\epsilon^2 \end{bmatrix}$, on a alors la densité conditionnelle

$$P(y(\mathbf{x}) | \mathbf{Y}_n) \propto \exp \left(-\frac{1}{2} \begin{bmatrix} \mathbf{Y}_n & y(\mathbf{x}) \end{bmatrix} \mathbf{C}_{n+1}^{-1} \begin{bmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{bmatrix} \right). \quad (4.2.3)$$

La matrice inverse $\mathbf{C}_{n+1}^{-1}$ peut s'écrire de manière partitionnée (Barnett, 1979)

$$\mathbf{C}_{n+1}^{-1} = \begin{bmatrix} \mathbf{M}_n & \mathbf{m}_{n+1} \\ \mathbf{m}_{n+1}^T & m \end{bmatrix}, \quad (4.2.4)$$

où l'on a noté

$$\begin{aligned} m &= \left(\kappa(\mathbf{x}) + \sigma_\epsilon^2 - \mathbf{k}^T(\mathbf{x}) \left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}_n \right)^{-1} \mathbf{k}(\mathbf{x}) \right)^{-1}, \\ \mathbf{m}_{n+1} &= -m \left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}_n \right)^{-1} \mathbf{k}(\mathbf{x}), \\ \mathbf{M}_n &= \left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}_n \right)^{-1} + \frac{1}{m} \mathbf{m}_{n+1} \mathbf{m}_{n+1}^T. \end{aligned}$$

En substituant dans (4.2.3), on obtient la densité de probabilité en $\mathbf{x}$

$$P(y(\mathbf{x}) | \mathbf{Y}_n) \propto \exp \left(-\frac{1}{2} \left(\frac{y(\mathbf{x}) - \hat{y}(\mathbf{x})}{\hat{\sigma}_y(\mathbf{x})} \right)^2 \right), \quad (4.2.5)$$

où $y(\mathbf{x})$ est conditionnée comme $y(\mathbf{x}) | \mathbf{Y}_n, \mathbf{X}_n, \boldsymbol{\Theta}, \sigma_\epsilon^2 \sim \mathcal{N}(\hat{y}(\mathbf{x}), \hat{\sigma}_y^2(\mathbf{x}))$. On détermine ainsi la *meilleure* prédiction $\hat{y}(\mathbf{x})$ de $y(\mathbf{x})$ (moyenne de $y(\mathbf{x})$) et la prédiction de variance $\hat{\sigma}_y^2(\mathbf{x})$ avec

$$\hat{y}(\mathbf{x}) = \mathbf{k}^T(\mathbf{x}) \left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}_n \right)^{-1} \mathbf{Y}_n, \quad (4.2.6)$$

$$\hat{\sigma}_y^2(\mathbf{x}) = \kappa(\mathbf{x}) + \sigma_\epsilon^2 - \mathbf{k}^T(\mathbf{x}) \left(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}_n \right)^{-1} \mathbf{k}(\mathbf{x}). \quad (4.2.7)$$

La variance $\hat{\sigma}_y^2(\mathbf{x})$ peut s'interpréter comme une mesure d'incertitude de la prédiction $\hat{y}(\mathbf{x})$.

Une illustration de prédictions par processus Gaussien sur un cas à une dimension est donnée par la Figure 4.3. En particulier, les effets d'un bruit absent ($\sigma_\epsilon^2 = 0$) ou inféré sur les observations sont montrés. Dans le cas où $\sigma_\epsilon^2 = 0$, la prédiction du méta-modèle $\hat{y}(\mathbf{x})$ interpole les observations et conduit à des oscillations de $\hat{\sigma}_y^2(\mathbf{x})$ dans le domaine non-exploré (voir Figure 4.3(a)). Lorsque le bruit σ_ϵ^2 est considéré et inféré sur les observations, la réponse du méta-modèle n'est plus interpolante, et par conséquent elle est davantage lissée en $\hat{y}(\mathbf{x})$ et $\hat{\sigma}_y^2(\mathbf{x})$. On remarque également que l'incertitude du méta-modèle pour $x \approx 0.25$ se retrouve diminuée. Pour ces

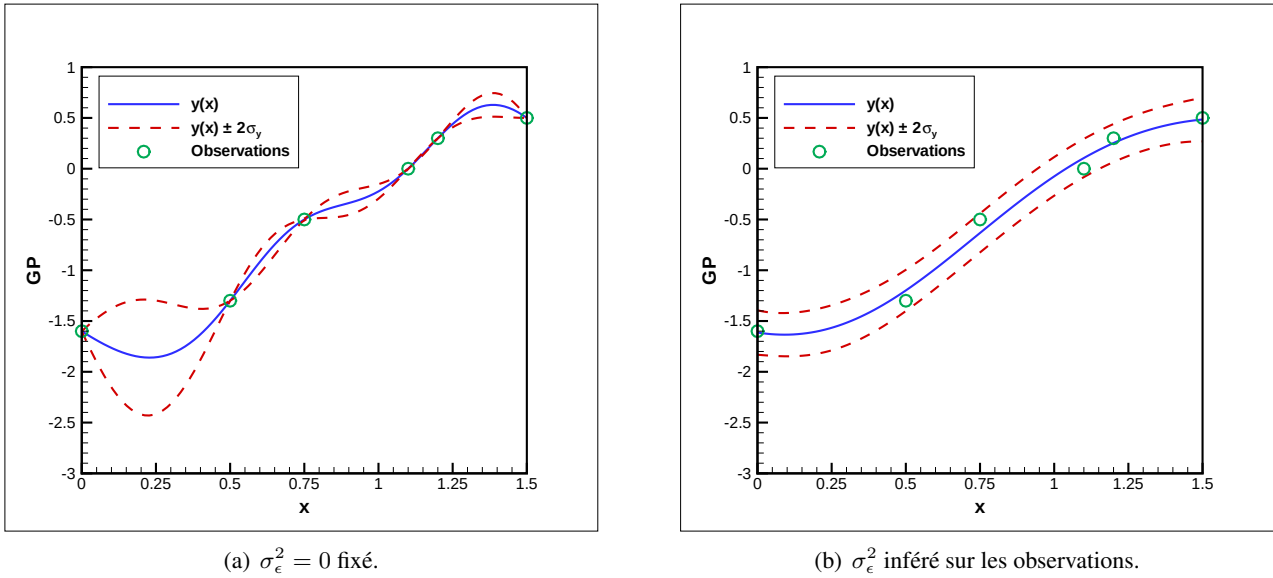


Figure 4.3 – Prédictions à l'aide d'un méta-modèle Gaussien dans un cas unidimensionnel.

raisons, mais également pour l'amélioration du conditionnement de la matrice de covariance, il est souvent bénéfique d'utiliser des méta-modèles Gaussiens qui comportent des bruits inférés, même si les observations ne sont pas bruitées.

4.2.2 Fonction de covariance

La fonction, ou noyau, de covariance permet de mettre en cohérence des observations en mesurant leurs « similarités ». C'est un élément essentiel du méta-modèle car il contient l'ensemble des hypothèses que l'on a *a priori* sur la fonction d'intérêt f . Bien qu'il n'existe pas de méthode optimale permettant de choisir la meilleure fonction de covariance, pour un problème donné, il est cependant possible, et conseillé, d'utiliser un noyau qui soit « le plus adapté », dès lors que l'on a un *a priori* sur la régularité de f . En effet, la régularité de la prédiction sera directement dépendante du choix de la fonction de covariance.

Dans le cadre des méta-modèles Gaussiens, on s'intéresse aux noyaux définis positifs. En d'autres termes, $C_f : \Omega \times \Omega \rightarrow \mathbb{R}$ est symétrique si $\forall (\mathbf{x}, \mathbf{x}') \in \Omega^2$, $C_f(\mathbf{x}, \mathbf{x}') = C_f(\mathbf{x}', \mathbf{x})$ et positif en satisfaisant $\forall n \in \mathbb{N}$, $\forall \mathbf{x}_1, \dots, \mathbf{x}_n \in \Omega$, $\forall a_1, \dots, a_n \in \mathbb{R}$, $\sum_{i=1}^n \sum_{j=1}^n a_i a_j C_f(\mathbf{x}_i, \mathbf{x}_j) \geq 0$. En pratique, il est difficile de vérifier la condition de positivité et on fait alors appel à des fonctions de covariances connues. La modélisation par processus Gaussiens fait généralement recours à des noyaux usuels, tels que exponentiels, Gaussiens ou de type Matern (Stein, 2012), comme présentés dans le Tableau 4.1 pour des cas univariés. Dans le cas multidimensionnel

Noyau	$C_f(x, x'; l), \forall x, x' \in \mathbb{R}, \forall l \in \mathbb{R}^+$
Exponentiel	$\exp\left(-\frac{ x - x' }{l}\right)$
Gaussien	$\exp\left(-\left(\frac{ x - x' }{l}\right)^2\right)$
Matern 3/2	$\left(1 + \frac{\sqrt{3} x - x' }{l}\right) \exp\left(-\frac{\sqrt{3} x - x' }{l}\right)$
Matern 5/2	$\left(1 + \frac{\sqrt{5} x - x' }{l} + \frac{5(x - x')^2}{3l^2}\right) \exp\left(-\frac{\sqrt{5} x - x' }{l}\right)$

Tableau 4.1 – Noyaux stationnaires univariés pour les méta-modèles Gaussiens.

($d > 1$), les fonctions de covariances sont construites comme le produit tensoriel de noyaux unidimensionnels. Dans ce cas, l devient l_i où $i = 1, \dots, d$ et rend possible l'utilisation d'un facteur d'échelle anisotrope associé à chaque dimension des variables d'entrée. La fonction de covariance se décompose alors comme

$$C_f(\mathbf{x}, \mathbf{x}'; \Theta) = \prod_{i=1}^d C_f(x_i, x'_i; l_i), \quad (4.2.8)$$

où ici $\Theta = \{l_1, \dots, l_d\}$ est le vecteur contenant les hyperparamètres que l'on infère sur les observations (voir Section 4.2.3). Dans cette thèse on utilisera principalement le noyau Gaussien sous la forme

$$C_f(\mathbf{x}, \mathbf{x}'; \Theta) = \theta_1 \prod_{i=1}^d \exp\left(-\left(\frac{|x_i - x'_i|}{l_i}\right)^2\right) + \theta_2, \quad (4.2.9)$$

avec $\Theta = \{\theta_1, \theta_2, l_1, \dots, l_d\}$, où θ_1 met à l'échelle la fonction de covariance et θ_2 sert d'ajustement depuis la

valeur 0. La Figure 4.4 présente des exemples d'évolutions de C_f dans les cas exponentiels et Gaussiens pour différentes valeurs de l .

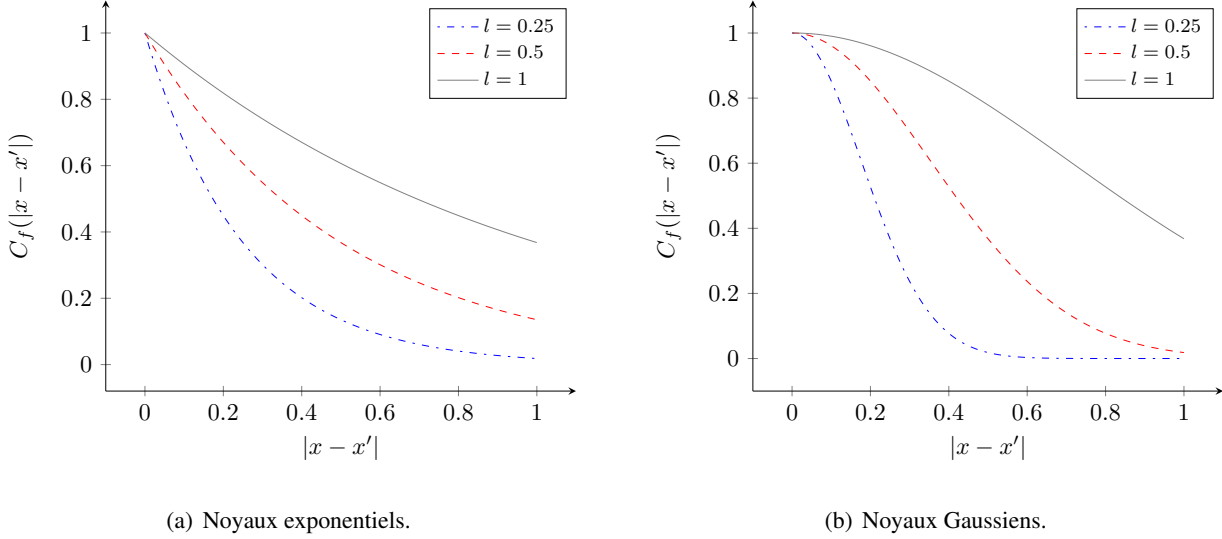


Figure 4.4 – Influences du paramètre l sur les noyaux exponentiels et Gaussiens.

4.2.3 Détermination des hyperparamètres

Comme on a pu le voir jusqu'ici, la construction d'un méta-modèle Gaussien fait intervenir un certain nombre d'hyperparamètres $\{\sigma_\epsilon, \Theta\}$ que l'on doit calibrer. Ces hyperparamètres peuvent être déterminés par une méthode de validation croisée (Dubrule, 1983), qui est généralement très coûteuse, ou par la méthode dite de « Maximum de Vraisemblance (MV) » décrite ci-après.

Idéalement, on aimerait que le méta-modèle Gaussien soit le plus représentatif possible des observations disponibles. Pour cela, la méthode MV (MacKay, 1992) consiste à déterminer les plus probables hyperparamètres $\{\sigma_\epsilon, \Theta\}$ qui maximisent la distribution $P(\{\sigma_\epsilon, \Theta\} | \mathbf{X}_n, \mathbf{Y}_n, \mathbf{C})$. Le théorème de Bayes permet de réécrire cette probabilité comme

$$P(\{\sigma_\epsilon, \Theta\} | \mathbf{X}_n, \mathbf{Y}_n, \mathbf{C}) = \frac{P(\mathbf{Y}_n | \mathbf{X}_n, \mathbf{C}, \{\sigma_\epsilon, \Theta\}) P(\{\sigma_\epsilon, \Theta\})}{P(\mathbf{Y}_n | \mathbf{X}_n, \mathbf{C})}. \quad (4.2.10)$$

La probabilité marginale du dénominateur dans (4.2.10) étant indépendante de $\{\sigma_\epsilon, \Theta\}$, cela revient à maximiser uniquement le produit $P(\mathbf{Y}_n | \mathbf{X}_n, \mathbf{C}, \{\sigma_\epsilon, \Theta\}) P(\{\sigma_\epsilon, \Theta\})$. De plus, $P(\{\sigma_\epsilon, \Theta\})$ est de distribution uniforme et peut être négligé, comme suggéré par Gibbs (1998). Finalement, il ne reste plus qu'à maximiser la densité de la probabilité conditionnelle, aussi appelée « likelihood »

$$f(\mathbf{Y}_n | \mathbf{X}_n, \mathbf{C}, \{\sigma_\epsilon, \Theta\}) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)}} \exp\left(-\frac{1}{2} \mathbf{Y}_n^T (\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)^{-1} \mathbf{Y}_n\right). \quad (4.2.11)$$

En pratique on utilise le logarithme du likelihood, aussi appelé « log-likelihood », qui évite de travailler avec de faibles valeurs de densité de probabilité :

$$\mathcal{L}(\{\sigma_\epsilon, \Theta\}) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(\det(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)) - \frac{1}{2} \mathbf{Y}_n^T (\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)^{-1} \mathbf{Y}_n. \quad (4.2.12)$$

La maximisation du log-likelihood (4.2.12) est un problème complet d'optimisation en soit. En effet, $\mathcal{L}(\{\sigma_\epsilon, \Theta\})$ est difficile à maximiser car elle est fréquemment multimodale et possiblement de haute dimension si $d \gg 1$. C'est pourquoi en pratique, il est conseillé d'utiliser un algorithme d'optimisation global pour résoudre $\max_{\{\sigma_\epsilon, \Theta\}}(\mathcal{L})$. Dans cette thèse, ce problème est résolu avec l'algorithme d'optimisation CMA-ES (Hansen, 2006b), introduit précédemment dans la Section 3.3.2.3, afin de déterminer les hyperparamètres du méta-modèle.

La matrice pleine $(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)$ est également responsable de l'important coût calcul dans la maximisation du log-likelihood. Ceci est particulièrement vrai pour les étapes de son inversion et la détermination de son déterminant. On peut cependant réduire ce coût numérique en implémentant intelligemment l'équation (4.2.12). La matrice pleine peut se décomposer à l'aide d'une méthode LU telle que $(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n) = \mathbf{L}\mathbf{U}$. On peut ensuite utiliser cette décomposition pour résoudre le système $(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n) \mathbf{W} = \mathbf{Y}_n$. Ceci détermine le produit

$$\mathbf{W} = (\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)^{-1} \mathbf{Y}_n. \quad (4.2.13)$$

Le logarithme du déterminant peut également se calculer efficacement à partir de la décomposition LU par

$$\log(\det(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)) = \sum_{i=1}^n (\log(\mathbf{L}_{ii}) + \log(\mathbf{U}_{ii})). \quad (4.2.14)$$

On comprend aisément que le coût calcul de la maximisation de (4.2.12) est aussi proportionnel au nombre d'observations n . Il peut même être une limite à la construction du méta-modèle Gaussien si $n \gg 1$. En pratique, l'utilisation d'un processus Gaussien devient impraticable lorsque le nombre d'observations vaut quelques milliers. Dans ce type de cas, il faut recourir à des méthodes dites « Sparse » où le principe est de travailler avec un sous-ensemble de $m \ll n$ observations. L'enjeu principal de ces méthodes concerne la sélection des m points. Cela peut se faire aléatoirement, ou d'une manière plus sophistiquée comme présentée par (Smola and Bartlett, 2001; Seeger et al., 2003). Cette complexe thématique de recherche est toujours d'actualité comme on a pu le voir récemment dans (Farmanesh and Pourhabib, 2015).

4.2.4 Plan d'expériences

Dans les Sections précédentes, le plan d'expériences $\mathbf{X}_n \in \mathbb{R}^{n \times d}$ à partir duquel le méta-modèle Gaussien est établi était supposé connu. On se pose à présent la question du choix de la construction de $\mathbf{X}_n$. Plus concrètement, on se demande comment choisir les valeurs des variables de $\mathbf{x}_i \in \Omega$, pour $i = 1, \dots, n$, afin de récolter un maximum d'informations sur le comportement de la fonction d'intérêt, et cela pour un minimum d'obser-

ventions n . En anglais cela s'appelle le « Space-Filling Design », qui correspond à une couverture optimale du domaine de variation Ω par $\mathbf{X}_n$. On supposera dans la suite que Ω est hypercubique et que chaque x_j , pour $j = 1, \dots, d$, est doublement bornée.

La méthode la plus simple pour construire un plan d'expériences est de répartir les observations sur une grille de points régulièrement espacés. C'est ce qu'on appelle alors un « plan factoriel complet », comme l'illustre la Figure 4.5 où le domaine comporte deux dimensions. Ce type de plan qui peut sembler évident

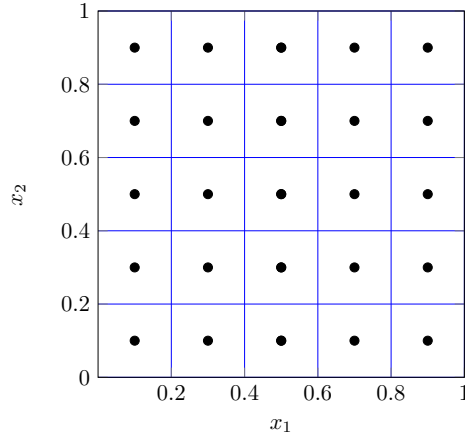


Figure 4.5 – Plan factoriel complet sur $\Omega = [0,1]^2$.

n'est cependant pas conseillé en pratique. En effet, si k est le nombre de points sur chaque axe factoriel, alors $n = k^d$ augmentera fortement avec la dimension du problème. De plus, si la fonction d'intérêt est essentiellement dépendante d'une sous-partie de d , alors le plan factoriel complet comportera beaucoup d'observations inutiles. En pratique, les plans d'expériences les plus utilisés pour la construction de méta-modèles Gaussiens sont les « hypercubes latins » (McKay et al., 1979); « Latin Hypercube Sampling (LHS) » en anglais. Soit $\Omega = [0,1]^d$, la construction d'un plan d'expériences $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T$ de type hypercubes latins se fait de la façon suivante :

$$x_j^{(i)} = \frac{\pi_j^{(i)} - U_j^{(i)}}{n}, \quad i = 1, \dots, n, \quad j = 1, \dots, d, \quad (4.2.15)$$

où $\mathbf{x}_i = (x_1^{(i)}, \dots, x_d^{(i)})^T$, $U_j^{(i)} \sim \mathcal{U}(0,1)$ est une valeur aléatoire, indépendante des $\pi_j^{(i)}$, de distribution uniforme sur $[0,1]$ et $\pi_j^{(i)}$ est une permutation aléatoire des entiers $\{1, \dots, n\}$. Cette construction revient à diviser chaque dimension de Ω en n segments réguliers de façon à obtenir un maillage du domaine, puis de sélectionner n cellules parmi les n^d disponibles afin d'avoir une observation, et une seule, dans chaque segment. Finalement, on tire une coordonnée au hasard dans chaque cellule sélectionnée. Ainsi, $(\pi_1^{(i)}, \dots, \pi_d^{(i)})^T$ donne la cellule dans laquelle se trouve l'observation $\mathbf{x}_i$ et $(U_1^{(i)}, \dots, U_d^{(i)})^T$ correspond aux coordonnées de $\mathbf{x}_i$ dans cette cellule. Il est aussi possible de centrer les observations dans les cellules en remplaçant les $U_j^{(i)}$ par la valeur 0.5.

Le plan d'expériences par hypercubes latins génère des observations uniformément distribuées sur les axes factoriels (voir Figure 4.6(a)). Cependant, cette propriété n'est pas suffisante pour obtenir un plan d'expériences exploratoire qui assure une uniformité des observations sur Ω (voir Figure 4.6(b)). Cependant, pour n fixé, il

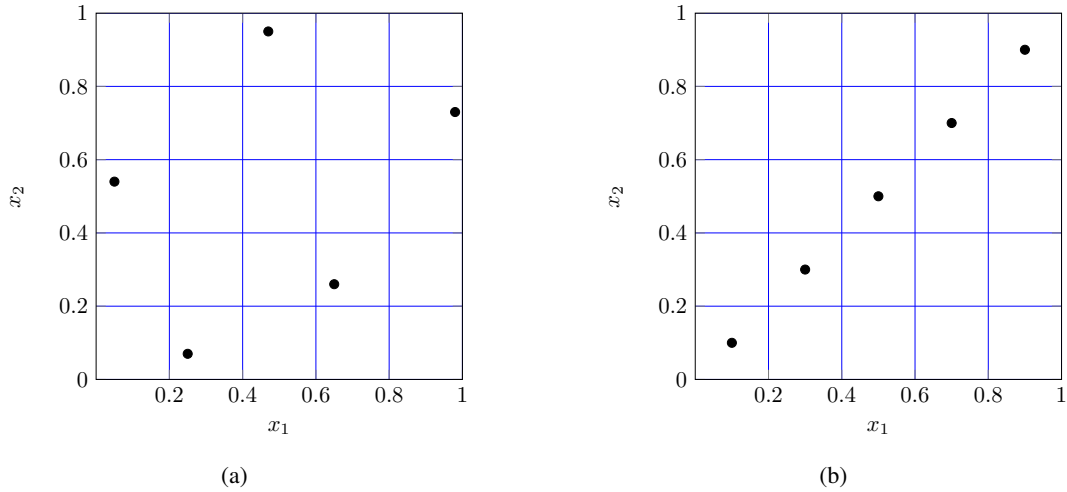


Figure 4.6 – Plans d'expériences par hypercubes latins sur $\Omega = [0,1]^2$.

existe $(n!)^{d-1}$ hypercubes latins possibles. On peut donc choisir le plan optimal à partir d'un critère statistique comme l'IMSE (Park, 1994), ou bien suivant un critère d'uniformité, comme le « maximin » où l'on cherche à maximiser la distance minimale entre deux observations (Johnson et al., 1990). En pratique, ces méthodes d'optimisations des plans par hypercubes latins sont coûteuses. Par conséquent, on préfère généralement générer un ensemble fini de plans candidats et sélectionner ensuite le plus performant, au sens d'un des critères énoncés ci-dessus.

4.3 Applications des méta-modèles Gaussiens

Dans cette Section, l'utilisation des méta-modèles Gaussiens pour l'analyse de sensibilité, la classification et enfin l'optimisation sont présentées. Le méta-modèle servira de prédicteur de la fonction d'intérêt en terme de réponse et/ou d'incertitude (variance) sur la prédiction.

4.3.1 Analyse de sensibilité

L'analyse de sensibilité d'une fonction d'intérêt $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $\mathbf{x} \mapsto f(\mathbf{x})$ est l'étude des incertitudes de $f(\mathbf{x})$ qui résultent de l'incertitude des variables de $\mathbf{x}$. L'objectif est de déterminer les principales sources d'incertitude parmi $x_1, \dots, x_d$, afin d'identifier les variables qui ont une forte influence sur $f(\mathbf{x})$. Ce type d'information est très utile lorsque l'on a aucun *a priori* sur les effets des variables d'un modèle. Dans le cadre de l'optimisation, cela permet par exemple d'identifier les variables d'intérêts qui doivent être optimisées en priorité. Si l'on souhaite faire de la simplification de modèle, alors l'analyse de sensibilité pourra aussi donner les variables à négliger.

On distingue en pratique deux types d'analyse de sensibilité : « locale » et « globale ». L'analyse locale est généralement basée sur le calcul de dérivés afin de déterminer les variations locales de la fonction d'intérêt « autour » de x_i . L'analyse globale étudie la variation de f lorsque x_i varie sur l'ensemble de son domaine d'incertitude. Elle consiste à définir des indices de sensibilité, puis de les calculer en faisant varier les variables incertaines $x_1, \dots, x_d$ sur leurs domaines respectifs dans Ω .

La méthode ANOVA (« Analysis Of Variance »), que l'on appelle aussi décomposition de « Sobol-Hoeffding » (Sobol, 2001) est basée sur une analyse fonctionnelle de la variance. Pour toute fonction $f(\mathbf{x})$ définie et intégrable sur $\Omega = \{\mathbf{x} \in \mathbb{R}^d, x_i \in [0,1], \forall i = 1, \dots, d\}$, la décomposition fonctionnelle ANOVA consiste à exprimer f comme une somme unique de fonctions de dimension croissante telle que

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^d f_i(x_i) + \sum_{1 \leq i < j \leq d} f_{i,j}(x_i, x_j) + \dots + f_{1,\dots,d}(x_1, \dots, x_d), \quad (4.3.1a)$$

$$= \sum_{u \in \mathbf{U}} f_u(\mathbf{x}_u), \quad (4.3.1b)$$

où $\mathbf{U}$ est l'ensemble de tous les sous-ensembles de $\{1, \dots, d\}$, f_0 est une constante, $f_i, i \in \{1, d\}$ sont les effets principaux, $f_{i,j}, i < j \in \{1, d\}$ correspondent aux effets d'ordre deux, etc. L'unicité de la décomposition est assurée par des contraintes d'orthogonalité sur les composantes comme

$$\int_0^1 f_u(\mathbf{x}_u) dx_i = 0, \quad \forall i \in u, \quad \forall u \in \mathbf{U}, \quad (4.3.2)$$

et les termes de la décomposition fonctionnelle sont alors deux à deux orthogonaux. Autrement dit,

$$\int_{[0,1]^d} f_u(\mathbf{x}_u) f_v(\mathbf{x}_v) d\mathbf{x} = 0, \quad \forall u \neq v \in \mathbf{U}. \quad (4.3.3)$$

Soit $Y = f(\mathbf{x})$, les termes de la décomposition fonctionnelle (4.3.1) peuvent s'écrire sous formes d'espérances conditionnelles où l'on a

$$\begin{aligned} f_0 &= \mathbb{E}[Y], \\ f_i(\mathbf{x}_i) &= \mathbb{E}[Y|\mathbf{x}_i] - f_0, \\ f_{i,j}(\mathbf{x}_i, \mathbf{x}_j) &= \mathbb{E}[Y|\mathbf{x}_i, \mathbf{x}_j] - f_i(\mathbf{x}_i) - f_j(\mathbf{x}_j) - f_0, \\ &\vdots \\ f_u(\mathbf{x}_u) &= \mathbb{E}[Y|\mathbf{x}_u] - \sum_{v \subset u} f_v(\mathbf{x}_v). \end{aligned}$$

La décomposition de Sobol (2001) peut alors s'appliquer à la variance globale $\text{Var}(Y)$ en la décomposant par une somme de variances partielles comme

$$\text{Var}(Y) = \sum_{i=1}^d \text{Var}(f_i(\mathbf{x}_i)) + \sum_{1 \leq i < j \leq d} \text{Var}(f_{i,j}(x_i, x_j)) + \dots + \text{Var}(f_{1,\dots,d}(x_1, \dots, x_d)), \quad (4.3.4a)$$

$$= \sum_{u \in \mathbf{U}} \text{Var}(f_u(\mathbf{x}_u)). \quad (4.3.4b)$$

Il en résulte alors la définition des indices de **Sobol (2001)**. S_u est l'indice de sensibilité de Sobol, d'ordre $|u|$, qui mesure la contribution de variabilité du groupe de variables $\mathbf{x}_u$ sur la fonction d'intérêt $Y = f(\mathbf{x})$, que l'on écrit comme

$$S_u = \frac{\text{Var}(f(\mathbf{x}_u))}{\text{Var}(Y)} = \frac{\text{Var}(\mathbb{E}[Y|\mathbf{x}_u]) - \sum_{v \subset u} \text{Var}(\mathbb{E}[Y|\mathbf{x}_v])}{\text{Var}(Y)}. \quad (4.3.5)$$

De plus, on a l'égalité suivante

$$1 = \sum_{i=1}^d S_i + \sum_{1 \leq i < j \leq d} S_{i,j} + \cdots + S_{1,\dots,d} = \sum_{u \in \mathbf{U}} S_u, \quad (4.3.6)$$

qui indique que si S_u est proche de 1, alors la part de variabilité de $\mathbf{x}_u$ sur Y est très importante et réciproquement, si $S_u \approx 0$, alors le groupe de variables $\mathbf{x}_u$ est très peu influent. Ainsi, le premier indice mesurant la contribution unique de x_i est donné par

$$S_i = \frac{\text{Var}(\mathbb{E}[Y|x_i])}{\text{Var}(Y)}, \quad (4.3.7)$$

et l'indice d'ordre deux $S_{i,j}$, $i \neq j$ qui exprime la sensibilité de $\text{Var}(Y)$ à l'interaction de x_i et x_j , sans prendre en compte les effets individuels des variables, s'écrit

$$S_{i,j} = \frac{\text{Var}(\mathbb{E}[Y|x_i, x_j]) - \text{Var}(\mathbb{E}[Y|x_i]) - \text{Var}(\mathbb{E}[Y|x_j])}{\text{Var}(Y)}. \quad (4.3.8)$$

On note également S_{T_i} , l'indice de sensibilité total de x_i qui mesure la contribution totale (individuelle + interactions) de la variable sur $\text{Var}(Y)$, et qui se calcule

$$S_{T_i} = 1 - \frac{\text{Var}(\mathbb{E}[Y|\mathbf{x}_{-i}])}{\text{Var}(Y)}, \quad (4.3.9)$$

où $\mathbf{x}_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d)^T$ est le vecteur de l'ensemble des variables, exceptée x_i . Contrairement aux indices S_u , on a pour les indices totaux $\sum_{i=1}^d S_{T_i} \geq 1$. Si $d = 3$, on a par exemple l'indice total de x_1 qui s'exprime comme $S_{T_1} = S_1 + S_{1,2} + S_{1,3} + S_{1,2,3}$.

L'estimation des indices de sensibilité se fait généralement par une méthode d'échantillonnage ou à partir d'une décomposition en polynômes de chaos (**Crestaux et al., 2009**). L'estimation par échantillonnage et selon la méthode « pick and freeze » de **Sobol (2001)** est présentée ci-après.

Soient un vecteur aléatoire $\mathbf{x} = (x_1, \dots, x_d)^T$ et un sous-ensemble $\mathbf{u}$ de $\{1, \dots, d\}$. On définit $\mathbf{x}^u = (x_1^u, \dots, x_d^u)^T$ tel que $x_i^u = x_i$ si $i \in \mathbf{u}$ et $x_i^u = x'_i$ si $i \notin \mathbf{u}$, où x'_i est une copie indépendante de x_i (x'_i et x_i ont la même distribution). L'objectif est ensuite d'exprimer les indices S_u en termes de covariance. Pour cela,

on utilise la Proposition suivante :

Proposition 4.3.1. *Pour tout $\mathbf{u} \subset \{1, \dots, d\}$,*

$$\text{Var}(\mathbb{E}[Y|x_i, i \in \mathbf{u}]) = \text{Cov}(Y, Y^{\mathbf{u}}), \quad (4.3.10)$$

où on a noté $Y^{\mathbf{u}} \doteq f(\mathbf{x}^{\mathbf{u}})$.

Les indices de Sobol peuvent alors être estimés en considérant deux m -échantillons aléatoires $\mathbf{X}_m = (\mathbf{x}_1, \dots, \mathbf{x}_m)^T$ et $\mathbf{X}_m^{\mathbf{u}} = (\mathbf{x}_1^{\mathbf{u}}, \dots, \mathbf{x}_m^{\mathbf{u}})^T$, où $m \in \mathbb{N}^+$, $\mathbf{x}_i, \mathbf{x}_i^{\mathbf{u}} \in \Omega \subset \mathbb{R}^d$ et $x_j^{\mathbf{u}} = x_j$ si $j \in \mathbf{u} \subset \{1, \dots, d\}$, $x_j^{\mathbf{u}} = x'_j$ sinon (x'_j est une copie indépendante de x_j). [Sobol \(2001\)](#) suit ce principe et propose l'estimation de (4.3.5) suivante

$$\hat{S}_u = \frac{\frac{1}{m} \sum_{i=1}^m f(\mathbf{x}_i) f(\mathbf{x}_i^{\mathbf{u}}) - \frac{1}{m} \sum_{i=1}^m f(\mathbf{x}_i) \frac{1}{m} \sum_{i=1}^m f(\mathbf{x}_i^{\mathbf{u}})}{\frac{1}{m} \sum_{i=1}^m f(\mathbf{x}_i)^2 - \left(\frac{1}{m} \sum_{i=1}^m f(\mathbf{x}_i) \right)^2}. \quad (4.3.11)$$

La formulation initiale (4.3.11) de Sobol a depuis été améliorée dans ([Janon et al., 2014](#)). L'échantillonnage aléatoire des deux $[m \times d]$ matrices $\mathbf{X}_m$ et $\mathbf{X}_m^{\mathbf{u}}$ peut se faire par une méthode de Monte-Carlo, quasi-Monte-Carlo ou bien par l'intermédiaire d'un plan d'expérience comme le type LHS.

L'utilité du méta-modèle Gaussien pour estimer les indices de sensibilité est assez explicite. En effet, on comprend que l'estimateur $\hat{S}_u$ sera proche de S_u lorsque $m \gg 1$. Cependant, dans les cas où le coût d'évaluation de la fonction d'intérêt est important, on ne peut se permettre d'évaluer directement $f(\mathbf{x})$. On peut alors avoir recours au méta-modèle Gaussien pour obtenir la prédiction de $f(\mathbf{x})$.

Une première approche est de directement remplacer $f(\mathbf{x})$ dans (4.3.11) par la prédiction du méta-modèle $\hat{y}(\mathbf{x})$, définie en (4.2.6). Cette méthode a l'avantage d'être peu coûteuse, mais elle n'inclut pas l'erreur de prédiction (variance) du méta-modèle. [Oakley and O'Hagan \(2004\)](#) et [Marrel et al. \(2009\)](#) proposent une seconde approche où $f(\mathbf{x})$ est substituée par le processus Gaussien $Z \sim \mathcal{PG}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$ construit à partir des n observations $\mathbf{X}_n$ de $f(\mathbf{x})$. L'estimateur des indices de sensibilité $\hat{S}_u$ est alors considéré comme une variable aléatoire $\tilde{S}_u$. La prédiction (moyenne) de $\tilde{S}_u$ est considérée pour déterminer l'estimateur de S_u et sa prédiction d'incertitude (variance) donne accès à un intervalle de confiance de $\tilde{S}_u$. Cette méthode prend en compte l'erreur du méta-modèle, cependant, elle nécessite des calculs d'intégrales multiples qui restent coûteux. De plus, comme l'ont indiqué [Le Gratiet et al. \(2014\)](#), les définitions des estimateurs (moyenne et variance) de S_u ne correspondent formellement pas aux calculs des indices de Sobol.

[Le Gratiet et al. \(2014\)](#) ont développé une troisième approche pour prendre en compte l'erreur du méta-modèle Gaussien dans le calcul des indices de sensibilité. Ils proposent, comme dans ([Oakley and O'Hagan, 2004](#); [Marrel et al., 2009](#)), de substituer $f(\mathbf{x})$ par le processus Gaussien $Z(\mathbf{x})$, puis d'estimer $\text{Var}(\mathbb{E}[Z(\mathbf{x})|x_i, i \in \mathbf{u}])$ et $\text{Var}(Z(\mathbf{x}))$ par une méthode de Monte-Carlo, comme présentée initialement dans ([Sobol, 1993](#)). L'estimation de S_u se déroule de la façon suivante :

1. Construction du méta-modèle Gaussien $Z(\mathbf{x})$ à partir de n observations $\mathbf{Y}_n = f(\mathbf{X}_n)$;
2. Génération des $[m \times d]$ matrices $\mathbf{X}_m$ et $\mathbf{X}_m^u$;
3. Pour $j = 1, \dots, k$;
 - (a) Génération d'une réalisation conditionnelle $\hat{Z}_j(\mathbf{x})$;
 - (b) Calcul de $\hat{S}_{u,j} = \frac{\frac{1}{m} \sum_{i=1}^m \hat{Z}_j(\mathbf{x}_i) \hat{Z}_j(\mathbf{x}_i^u) - \frac{1}{m} \sum_{i=1}^m \hat{Z}_j(\mathbf{x}_i) \frac{1}{m} \sum_{i=1}^m \hat{Z}_j(\mathbf{x}_i^u)}{\frac{1}{m} \sum_{i=1}^m \hat{Z}_j(\mathbf{x}_i)^2 - \left(\frac{1}{m} \sum_{i=1}^m \hat{Z}_j(\mathbf{x}_i) \right)^2}$;
4. Calcul de l'estimateur de S_u par $\hat{S}_u = \frac{1}{k} \sum_{j=1}^k \hat{S}_{u,j}$.

Une technique de bootstrap (tirages aléatoires avec remise de l'échantillon) est également employée dans (Le Gratiet et al., 2014) pour prendre en compte l'incertitude due aux intégrations par Monte-Carlo. Pour plus de détails sur cette dernière méthode, le lecteur est invité à se référer aux contributions (Le Gratiet et al., 2014; Gratiet et al., 2016). En particulier, les calculs des incertitudes du méta-modèle et de l'intégration par Monte-Carlo y sont détaillés, ainsi qu'une procédure qui détermine le nombre minimal d'échantillons m pour ne pas avoir une erreur d'intégration dominante.

4.3.2 Classification

On considère un ensemble d'observations dont on connaît certaines propriétés permettant de définir un nombre fini de classes C^i . Chaque observation est alors attribuée d'une de ces classes. L'objectif de la classification est d'identifier la classe d'une nouvelle observation inconnue à partir d'une fonction mathématique, que l'on appelle « classifieur », construite sur le groupe des observations initialement classées. Il existe de nombreuses méthodes pour faire la classification, comme par exemple les k-voisins (Kowalski and Bender, 1972; Dudani, 1976), les réseaux de neurones (Hepner, 1990; MacKay, 2003), la régression logistique (Zhu and Hastie, 2012) ou encore les « Support Vector Machine (SVM) » (Suykens and Vandewalle, 1999). Dans cette Section, on s'intéresse plus particulièrement à l'application des méta-modèles Gaussiens dans le cadre de la classification. Pour plus de clarté dans l'explication, on se restreindra ici au cas de classification binaire.

Soit un ensemble d'observations $\mathcal{D} = \{(\mathbf{x}_i, z_i) | \forall i = 1, \dots, n\}$ étant binairement classifié $C = (C^+, C^-)$ par les étiquettes $z_i = \pm 1$. On souhaite prédire la classe C d'une nouvelle observation $\mathbf{x}$, à partir d'un classifieur processus Gaussien construit sur l'ensemble $\mathcal{D}$. Le classifieur définit une frontière séparatrice des classes C^+ et C^- et la classe de $\mathbf{x}$ est donnée par son évaluation. La Figure 4.7 ci-dessous présente un exemple de classification ($d = 2$) où les losanges blancs appartiennent à la classe C^+ et les triangles noirs à la classe C^- . La fonction séparatrice est ici représentée par les contours colorés $z \in [-1, 1]$ et les iso-lignes $z = 0$ correspondent aux frontières séparatrices des deux classes. Contrairement au cas classique de la régression par processus Gaussiens (Section 4.2.1), les observations z_i sont maintenant discrètes. Il serait possible d'appliquer la prédiction d'un méta-modèle Gaussien, construit directement sur $\mathcal{D}$, pour faire de la classification. Cette approche est cependant inefficace en pratique car les processus Gaussiens se comportent généralement mal lorsque les observations présentent des discontinuités. On utilise alors une projection des observations par une fonction $f(\mathbf{x})$ (« latent function » en anglais) que l'on transforme ensuite par une sigmoïde $\pi : \mathbb{R} \rightarrow [0, 1]$

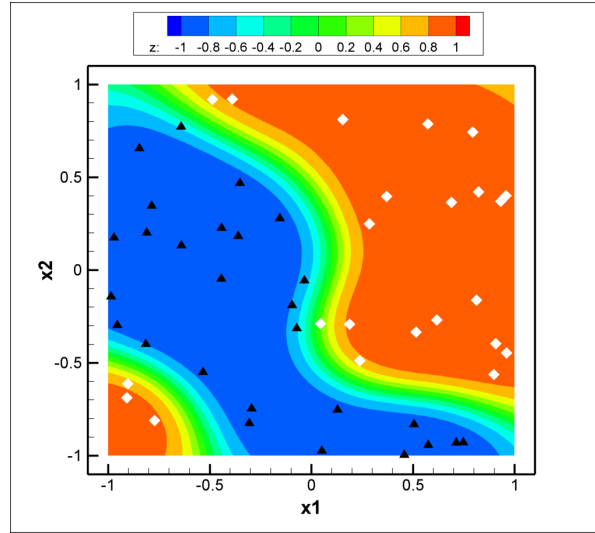


Figure 4.7 – Classification binaire non linéaire et non séparable.

pour se ramener dans l'intervalle unitaire. Cette dernière transformation peut s'interpréter comme une mesure de probabilité d'appartenance à la classe C^+ , où $P(z = 1|\mathbf{x}) = P(C^+|\mathbf{x}) = \pi(f(\mathbf{x}))$.

La fonction sigmoïde π correspond généralement à la fonction de répartition de la loi normale $\Phi(f)$. Soit la fonction $f(\mathbf{x})$, on suppose que les étiquettes z_i sont des réalisations indépendantes de Bernoulli, on peut alors écrire la probabilité jointe conditionnelle

$$P(\mathbf{z}_n|\mathbf{f}_n) = \prod_{i=1}^n P(z_i|f_i) = \prod_{i=1}^n \Phi(z_i f_i), \quad (4.3.12)$$

où l'on a noté $f_i = f(\mathbf{x}_i)$ et $\mathbf{f}_n = (f_1, \dots, f_n)^T$. Un processus Gaussien est considéré pour modéliser la fonction f telle que $\mathbf{f}_n \sim \mathcal{N}(\mathbf{0}_n, \mathbf{K})$. Le théorème de Bayes permet ensuite d'écrire la probabilité conditionnelle sur $\mathbf{f}_n$, qui n'est ici pas de distribution Gaussienne :

$$P(\mathbf{f}_n|\mathcal{D}, \Theta) = \frac{P(\mathbf{z}_n|\mathbf{f}_n)P(\mathbf{f}_n|\mathbf{X}_n, \Theta)}{P(\mathcal{D}|\Theta)} = \frac{\mathcal{N}(\mathbf{0}_n, \mathbf{K})}{P(\mathcal{D}|\Theta)} \prod_{i=1}^n \Phi(z_i f_i). \quad (4.3.13)$$

La prédiction de f est calculée par marginalisation sur les observations,

$$P(f|\mathcal{D}, \Theta, \mathbf{x}) = \int P(f, \mathbf{f}_n|\mathcal{D}, \Theta, \mathbf{x}) d\mathbf{f}_n = \int P(f|\mathbf{f}_n, \mathbf{X}_n, \Theta, \mathbf{x}) P(\mathbf{f}_n|\mathcal{D}, \Theta) d\mathbf{f}_n, \quad (4.3.14)$$

et finalement, la prédiction de probabilité d'appartenance à la classe C^+ est donnée par

$$P(z|\mathcal{D}, \Theta, \mathbf{x}) = \int P(z|f) P(f|\mathcal{D}, \Theta, \mathbf{x}) df = \int \Phi(zf) P(f|\mathcal{D}, \Theta, \mathbf{x}) df. \quad (4.3.15)$$

Malheureusement, cette dernière intégrale n'est pas calculable analytiquement à cause de la probabilité conditionnelle non Gaussienne (4.3.13). Il est alors nécessaire d'utiliser une méthode d'approximation comme les

méthodes de Monte-Carlo par chaînes de Markov (MCMC), ou bien d'approcher $P(\mathbf{f}_n|\mathcal{D}, \Theta)$ par un processus Gaussien avec $q(\mathbf{f}_n|\mathcal{D}, \Theta) = \mathcal{N}(\mathbf{f}_n|\boldsymbol{\mu}_n, \mathbf{A})$. En supposant que $\mathcal{N}(\mathbf{f}_n|\boldsymbol{\mu}_n, \mathbf{A})$ est connue, on peut écrire l'approximation de la distribution (4.3.14) comme $q(f|\mathcal{D}, \Theta, \mathbf{x}) = \mathcal{N}(f|\mu, \sigma_\mu^2)$, dont la moyenne et la variance sont données par

$$\mu(\mathbf{x}) = \mathbf{k}^T(\mathbf{x})\mathbf{K}^{-1}\boldsymbol{\mu}_n, \quad (4.3.16)$$

$$\sigma_\mu^2(\mathbf{x}) = \kappa(\mathbf{x}) - \mathbf{k}^T(\mathbf{x})(\mathbf{K}^{-1} - \mathbf{K}^{-1}\mathbf{A}\mathbf{K}^{-1})\mathbf{k}(\mathbf{x}), \quad (4.3.17)$$

où $\mathbf{K}$ est la matrice de covariance sur $\mathbf{X}_n$, $\kappa(\mathbf{x}) \doteq C_f(\mathbf{x}, \mathbf{x}; \Theta)$ et $\mathbf{k}(\mathbf{x}) \doteq (C_f(\mathbf{x}, \mathbf{x}_1; \Theta) \cdots C_f(\mathbf{x}, \mathbf{x}_n; \Theta))^T$. Ces paramètres sont dépendants des hyperparamètres Θ que l'on détermine par la méthode MV. Finalement, l'approximation de la probabilité d'appartenance à C^+ (4.3.15) se calcule analytiquement avec la fonction $\Phi(f)$ (Rasmussen and Williams, 2006) :

$$q(z = 1|\mathcal{D}, \Theta, \mathbf{x}) = \int \Phi(zf)\mathcal{N}(f|\mu, \sigma_\mu^2)df = \Phi\left(\frac{\mu(\mathbf{x})}{\sqrt{1 + \sigma_\mu^2(\mathbf{x})}}\right). \quad (4.3.18)$$

Les paramètres $\boldsymbol{\mu}_n$ et $\mathbf{A}$ sont donnés par la méthode de « Laplace » (Williams and Barber, 1998) ou par la méthode « Expectation Propagation » développée pendant la thèse de Minka (2001). Ces deux méthodes fournissent également une approximation de la vraisemblance

$$P(\mathcal{D}, \Theta) = \int P(\mathbf{z}_n|\mathbf{f}_n)P(\mathbf{f}_n|\mathbf{X}_n, \Theta)d\mathbf{f}_n, \quad (4.3.19)$$

qu'il faut maximiser afin de déterminer Θ .

L'« Expectation Propagation » est généralement plus précise (Rasmussen and Williams, 2006) pour un coût de calcul comparable. Ces deux approches restent cependant coûteuses et il est nécessaire d'utiliser une formulation « Sparse » si le nombre d'observations dépasse quelques centaines. La méthode « Fully Independent Training Conditional (FITC) » (Quiñonero-Candela and Rasmussen, 2005) est particulièrement efficace pour ce genre d'application. Davantage de détails sur la classification par méta-modèles Gaussiens et son implémentation sont disponibles dans (Kuss and Rasmussen, 2005; Rasmussen and Williams, 2006).

4.3.3 Optimisation globale

L'optimisation par méta-modèles, aussi appelée « Efficient Global Optimization (EGO) », est une méthode d'optimisation itérative où les méta-modèles sont utilisés pour se substituer à l'évaluation d'une fonction objective f , généralement coûteuse. Cette approche est potentiellement très efficace si le coût de construction des méta-modèles est faible par rapport au coût d'évaluation de f .

Le principe de l'EGO est d'enrichir de manière itérative un ensemble d'observations $\mathbf{X}_n$ avec de nouveaux points $\mathbf{x}$, qui, à chaque itération, donnent lieu à la reconstruction du méta-modèle. Cela s'appelle aussi « adaptive sampling ». Les $\mathbf{x}$ sont sélectionnés par un critère d'enrichissement bien défini. Intuitivement, ce critère peut correspondre à la prédiction de la fonction objective donnée par le méta-modèle où l'on détermine son

minimum, dans le cas où le problème d'optimisation est la minimisation de f . Les méta-modèles Gaussiens ont l'avantage de fournir une prédiction de la fonction objective mais également une mesure d'incertitude (variance) de cette prédiction. Il est alors possible de définir des critères d'enrichissement statistiques, que l'on appelle « fonctions de mérite », qui prennent ces deux composantes en compte. On pourra alors bénéficier de la prédiction de f pour rechercher les zones prometteuses, par exemple en terme de minimum, mais également de l'incertitude du méta-modèle pour permettre la découverte du domaine Ω , jusqu'ici potentiellement non-exploré.

Dans la suite de cette Section 4.3.3, la méthode d'optimisation par méta-modèles Gaussiens est résumée sous la forme d'un algorithme. On présentera ensuite différentes définitions de fonctions de mérite en détaillant leurs avantages et inconvénients respectifs. Finalement un exemple d'application de l'EGO sera mené sur la fonction analytique « Branin » où la dimension est $d = 2$.

4.3.3.1 Algorithme de l'EGO

L'Algorithme 4.1 ci-après résume les principales étapes de l'optimisation par méta-modèles Gaussiens. Les entrées de l'algorithme sont la dimension du problème d'optimisation d , le nombre d'observations initiales n_{LHS} pour créer le plan d'expériences, des valeurs d'initialisations des hyperparamètres $\{\sigma_\epsilon, \Theta\}$, le choix de la fonction de covariance C_f et le choix de la fonction de mérite $\hat{f}_{\text{merit}}$. En sortie, la prédiction de la solution du problème d'optimisation $\hat{\mathbf{x}}_{\text{opt}}$ est donnée.

La première étape (ligne 1) consiste à définir le plan d'expérience, ici par hypercubes latins, pour obtenir $\mathbf{X}_n$. On évalue ensuite la fonction objective f en chacun des points du plan (lignes 3 à 5).

Les lignes 6 à 13 correspondent à la principale boucle d'optimisation itérative. Le critère d'arrêt de ces itérations n'est ici pas détaillé, mais il peut par exemple être défini comme un seuil à atteindre sur la valeur de la fonction de mérite, ou bien un écart type sur la convergence d'un paramètre (minimum du méta-modèle, distances consécutives entre les nouveaux points $(\mathbf{x}_n, \mathbf{x}_{n+1})$, incertitude du méta-modèle, etc.), ou encore à une limite du nombre d'itérations préalablement fixé.

Dans la boucle d'optimisation itérative, la première étape est de construire le méta-modèle Gaussien à partir des observations $\mathbf{Y}_n$ évaluées en $\mathbf{X}_n$. Pour cela, un problème d'optimisation est résolu (ligne 7) par l'algorithme CMA-ES (Hansen, 2006b) où l'objectif est de déterminer les hyperparamètres du méta-modèle $\{\sigma_\epsilon, \Theta\}$ en maximisant le log-likelihood défini précédemment en (4.2.12). On utilise ici une factorisation LU de la matrice de covariance afin de réduire le temps de calcul de $\mathcal{L}(\{\sigma_\epsilon, \Theta\})$, comme présenté par (4.2.13) et (4.2.14).

Une fois que les $\{\sigma_\epsilon, \Theta\}$ sont déterminés, l'inverse de la matrice de covariance est calculée à la ligne 8. La matrice étant pleine, c'est cette inversion, en plus des évaluations de $\mathcal{L}(\{\sigma_\epsilon, \Theta\})$, qui demandera le plus important coût calcul de l'EGO. La matrice inverse est sauvegardée pour être par la suite réutilisée dans les calculs des prédictions $\hat{y}(\mathbf{x})$ et $\hat{\sigma}_y^2(\mathbf{x})$.

Un second problème d'optimisation est résolu à la ligne 9 pour déterminer les coordonnées de la nouvelle observation $\mathbf{x}_{n+1}$. On maximise (parfois minimise) ici la fonction de mérite $\hat{f}_{\text{merit}}$ en utilisant l'algorithme CMA-ES.

La fonction objective est évaluée ensuite en $\mathbf{x}_{n+1}$ à la ligne 10 et enfin, on complète $\mathbf{X}_n$ et $\mathbf{Y}_n$ avec la nouvelle observation (lignes 11 à 13).

Finalement, la prédiction de la solution du problème global d'optimisation $\hat{\mathbf{x}}_{\text{opt}}$ est déterminée par un dernier problème d'optimisation résolu à l'aide de l'algorithme CMA-ES. On suppose alors que $\hat{\mathbf{x}}_{\text{opt}} \approx \mathbf{x}_{\text{opt}} = \arg \min_{\mathbf{x} \in \Omega} f(\mathbf{x})$ si l'erreur d'approximation du méta-modèle est suffisamment faible.

Algorithme 4.1 : EGO par méta-modèles Gaussiens.

Input : $d, n_{\text{LHS}}, \{\sigma_\epsilon, \Theta\}, C_f, \hat{f}_{\text{merit}}$
Output : $\hat{\mathbf{x}}_{\text{opt}}$

```

!Plan d'expériences!
1  $\mathbf{X}_n \leftarrow \text{sampling\_LHS}(n_{\text{LHS}}, d)$                                 !Génération des  $n_{\text{LHS}}$  points initiaux
2  $\mathbf{Y}_n \leftarrow \{\emptyset\}$ 
3 for  $i \leftarrow 1$  to  $n_{\text{LHS}}$  do
4    $y_i \leftarrow f(\mathbf{x}_i)$                                 !Évaluations de la fonction objective
5    $\mathbf{Y}_n \leftarrow \mathbf{Y}_n \cup y_i$ 

!Boucle d'optimisation itérative!
6 while stopping criterion not reached do
   !Construction du méta-modèle!
7   solve  $\{\sigma_\epsilon, \Theta\} \leftarrow \arg \max \mathcal{L}$                                 ! $\mathcal{L}$  défini par (4.2.12), avec une décomposition LU
8   compute  $(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I}_n)^{-1}$ 
   !Optimisation de la fonction de mérite!
9   solve  $\mathbf{x}_{n+1} \leftarrow \arg \max \hat{f}_{\text{merit}}(\mathbf{x}, \hat{y}(\mathbf{x}), \hat{\sigma}_y^2(\mathbf{x}))$                                 !Prédictions données par (4.2.6) et (4.2.7)
   !Évaluation de la fonction objective!
10   $y_{n+1} \leftarrow f(\mathbf{x}_{n+1})$ 
   !Ajout du nouveau point!
11   $\mathbf{X}_{n+1} \leftarrow \mathbf{X}_n \cup \mathbf{x}_{n+1}$ 
12   $\mathbf{Y}_{n+1} \leftarrow \mathbf{Y}_n \cup y_{n+1}$ 
13   $n \leftarrow n + 1$ 

!Fin de l'optimisation!
14  $\hat{\mathbf{x}}_{\text{opt}} \leftarrow \arg \min_{\mathbf{x} \in \Omega} \hat{y}(\mathbf{x})$ 

```

D'un point de vue pratique, cet algorithme EGO se couple très facilement aux possibles modèles de f grâce à sa forme itérative. En effet, comme l'évaluation de f est généralement coûteuse, il n'est pas nécessaire d'insérer l'EGO dans un code de calcul (ou inversement) pour mener à bien l'optimisation. La boucle itérative est alors interrompue à la ligne 10 pendant l'évaluation de $f(\mathbf{x}_{n+1})$, puis une reprise de l'EGO est faite en ajoutant le nouveau point « à la main » dans $\mathbf{X}_n$ et $\mathbf{Y}_n$.

4.3.3.2 Fonctions de mérite

On présente quatre définitions de fonctions de mérite ci-après. Davantage de formulations sont disponibles et comparées dans (Picheny et al., 2013).

Borne inférieure de confiance. Cette première définition de fonction de mérite a été introduite par [Cox and John \(1992\)](#) sous le nom de « Lower Confidence Bound (LCB) ». Elle consiste à minimiser la borne inférieure de confiance définie comme

$$LCB(\mathbf{x}) = \hat{y}(\mathbf{x}) - \rho \hat{\sigma}_y(\mathbf{x}), \quad (4.3.20)$$

où $\rho \in \mathbb{R}$ permet la pondération des choix d'exploration et de minimisation du méta-modèle. Ainsi, si $\rho \approx 0$, le critère LCB revient à directement minimiser la prédiction $\hat{y}(\mathbf{x})$, et inversement si $\rho \gg 1$, alors l'exploration du méta-modèle est favorisée. La difficulté d'utilisation de cette fonction de mérite réside dans le choix de ρ .

Probabilité d'amélioration. La probabilité d'amélioration ([Jones, 2001](#)), ou « Probability of Improvement (PI) », est la probabilité d'obtenir une valeur de la fonction f qui soit plus faible que le minimum courant $y_{\min} = \min(\mathbf{Y}_n)$. Elle se calcule comme

$$PI(\mathbf{x}) = \Phi\left(\frac{y_{\min} - \hat{y}(\mathbf{x})}{\hat{\sigma}_y(\mathbf{x})}\right), \quad (4.3.21)$$

avec $\Phi(\cdot)$ qui est la fonction de répartition de la loi normale centrée réduite. $y_{\min}$ peut aussi être remplacée par une valeur cible $T \in \mathbb{R}$ donnée. Cette fonction de mérite a le principal inconvénient de générer des nouveaux points qui restent dans le voisinage du minimum courant. [Jones \(2001\)](#) propose d'améliorer l'exploration en utilisant un enrichissement où plusieurs points sont générés pour différentes valeurs de T , à chaque itération de l'EGO.

Amélioration espérée. Cette fonction de mérite est très certainement celle qui a été la plus utilisée dans l'application de l'EGO au cours de ces dernières années. Elle a été introduite sous le nom d'« Expected Improvement (EI) » dans la thèse de [Schonlau \(1997\)](#), puis plus largement développée dans ([Jones et al., 1998](#)). L'idée est de prendre en compte la probabilité et l'amplitude de progrès qu'une observation pourrait apporter. On définit alors l'amélioration $I(\mathbf{x})$ par

$$I(\mathbf{x}) \doteq \max(y_{\min} - f(\mathbf{x}), 0), \quad (4.3.22)$$

où $y_{\min} = \min(\mathbf{Y}_n)$ et en supposant que $f(\mathbf{x}) \sim \mathcal{N}(\hat{y}(\mathbf{x}), \hat{\sigma}_y^2(\mathbf{x}))$, on peut écrire la probabilité de réaliser cette amélioration $I(\mathbf{x})$ comme

$$P(f(\mathbf{x}) = y_{\min} - I(\mathbf{x})) = \frac{1}{\sqrt{2\pi}\hat{\sigma}_y(\mathbf{x})} \exp\left(-\frac{1}{2}\left(\frac{y_{\min} - I(\mathbf{x}) - \hat{y}(\mathbf{x})}{\hat{\sigma}_y(\mathbf{x})}\right)^2\right), \quad (4.3.23)$$

qui donne lieu, après une intégration par parties, au critère d'amélioration espérée

$$EI(\mathbf{x}) = \hat{\sigma}_y(\mathbf{x}) [u(\mathbf{x})\Phi(u(\mathbf{x})) + \phi(u(\mathbf{x}))], \quad u(\mathbf{x}) = \frac{y_{\min} - \hat{y}(\mathbf{x})}{\hat{\sigma}_y(\mathbf{x})}. \quad (4.3.24)$$

Dans (4.3.24), $\Phi(\cdot)$ et $\phi(\cdot)$ représentent respectivement la fonction de répartition et la densité de probabilité de la loi normale centrée réduite. Contrairement aux précédentes fonctions de mérite, on remarque que l'EI n'est pas dépendant d'un paramètre fixant le niveau d'amélioration ou l'arbitrage entre l'exploration et la minimisation. Le compromis entre les zones prometteuses et incertaines est cependant bien pris compte car l'EI est croissant avec la variance et décroissant avec la prédiction moyenne. De plus, dans le cas d'observations non bruitées, l'EI sera nul au niveau des points explorés. Il existe également une version pondérée de l'amélioration espérée (Sóbestor et al., 2005) permettant de privilégier l'exploration ou la minimisation du méta-modèle.

Amélioration espérée augmentée. L'amélioration espérée augmentée, ou l'« Augmented Expected Improvement (AEI) » est une évolution de l'EI développée par Huang et al. (2006) pour les cas d'observations bruitées. En effet, l'EI est dépendant de $y_{\min} = \min(\mathbf{Y}_n)$ qui n'est cependant pas connue avec certitude en présence de bruit. Huang et al. (2006) proposent alors d'adapter l'EI définie en (4.3.24) en introduisant deux modifications où interviennent,

1. L'« effective best solution » $\mathbf{x}_{*,n} = \arg \min_{\mathbf{x} \in \mathbf{X}_n} (\hat{y}(\mathbf{x}) + \hat{\sigma}_y(\mathbf{x}))$;
2. Le bruit inféré σ_ϵ .

L'AEI est alors définie comme

$$AEI(\mathbf{x}) = EI(\mathbf{x}) \left(1 - \frac{\sigma_\epsilon}{\sqrt{\hat{\sigma}_y^2(\mathbf{x}) + \sigma_\epsilon^2}} \right), \quad (4.3.25)$$

où l'EI dans (4.3.25) est donnée par

$$EI(\mathbf{x}) = \hat{\sigma}_y(\mathbf{x}) [u(\mathbf{x})\Phi(u(\mathbf{x})) + \phi(u(\mathbf{x}))], \quad u(\mathbf{x}) = \frac{\hat{y}(\mathbf{x}_{*,n}) - \hat{y}(\mathbf{x})}{\hat{\sigma}_y(\mathbf{x})}. \quad (4.3.26)$$

Huang et al. (2006) proposent également un critère d'arrêt relatif de l'EGO exprimé comme

$$\frac{\max AEI(\mathbf{x})}{\max(\mathbf{Y}_n) - \min(\mathbf{Y}_n)} < \Delta_r, \quad (4.3.27)$$

où $\Delta_r \in \mathbb{R}^+$ est une tolérance sans dimension. Cette dernière fonction de mérite AEI est très efficace comme cela a été montré dans le benchmark (Picheny et al., 2013).

4.3.3.3 Application de l'EGO à la fonction Branin

L'optimisation par méta-modèles Gaussiens appliquée à la fonction analytique « Branin-Hoo » (Jones et al., 1998) de dimension $d = 2$ est illustrée ci-après. La fonction Branin est définie pour $(x_1, x_2) \in [-5, 10] \times [0, 15]$ comme

$$f(x_1, x_2) = \left(x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5}{\pi} x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos(x_1) + 10. \quad (4.3.28)$$

Cette fonction possède trois minimum globaux $\mathbf{x}_{\text{opt}_1} = (-\pi, 12.275)$, $\mathbf{x}_{\text{opt}_2} = (\pi, 2.275)$ et enfin $\mathbf{x}_{\text{opt}_3} = (9.425, 2.475)$ ayant la même valeur $f(\mathbf{x}_{\text{opt}_{1,2,3}}) \approx 0.398$. Une représentation surfacique de la fonction Branin est donnée par la Figure 4.8(a).

La première étape de l'EGO consiste à créer un plan d'expériences. On utilise ici un plan LHS composé de 10 observations, comme présentées sur la Figure 4.8(b). La Figure 4.9 illustre les 3 premières itérations globales

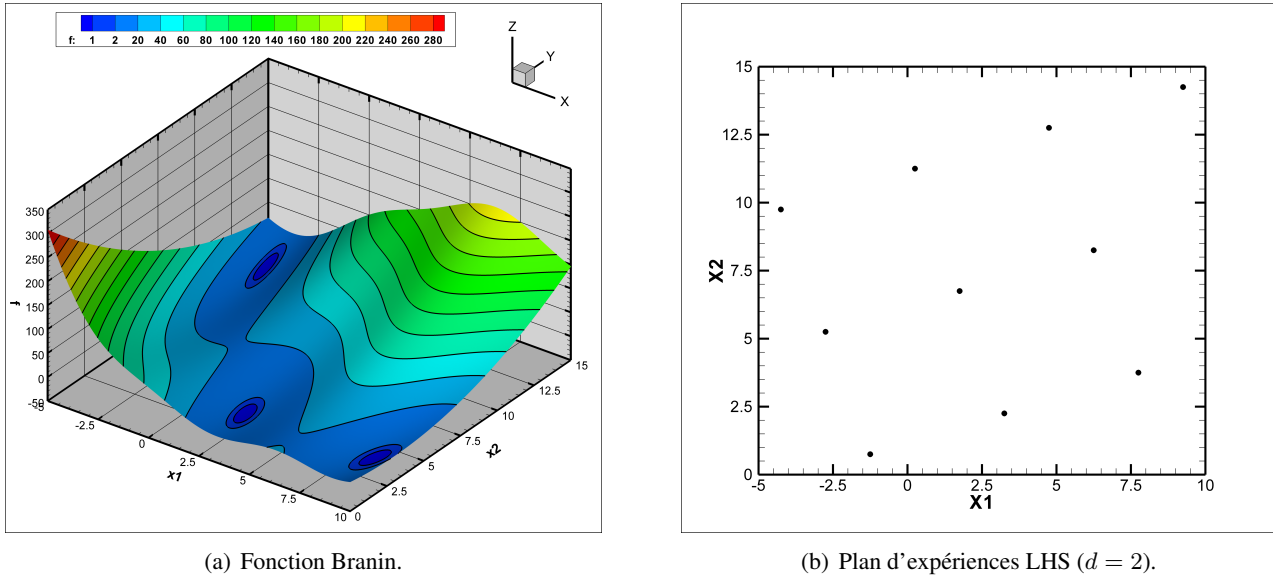


Figure 4.8 – Fonction Branin et plan d'expériences LHS.

de l'EGO. Les itérations 15, 20 et 50 sont également reportées sur la Figure 4.10. Pour chaque itération, la prédiction moyenne $\hat{y}(\mathbf{x})$ est donnée sur la Figure de gauche, la variance $\hat{\sigma}_y^2(\mathbf{x})$ concerne la Figure du centre et enfin la fonction de mérite $AEI(\mathbf{x})$ est tracée sur la Figure de droite.

On remarque qu'à l'itération 1, la prédiction du méta-modèle (Figure 4.9(a)) contient un seul minimum, la variance (Figure 4.9(b)) est faible dans les zones proches des observations et la fonction de mérite est composée d'un maximum global très marqué au niveau du minimum de $\hat{y}(\mathbf{x})$. A l'itération 2, un nouveau point a été ajouté au niveau du maximum de l'AEI construit lors de l'itération 1. Le second méta-modèle créé avec cette nouvelle observation produit une prédiction de $f(x_1, x_2)$ qui contient toujours un seul minimum global, mais celui-ci s'est légèrement déplacé et affiné. La variance a quant à elle augmenté sur l'ensemble de Ω . Cela est dû à l'augmentation du bruit inféré σ_e^2 . Le maximum de l'AEI s'est également déplacé et affiné autour du minimum de la prédiction. On observe un étalement du minimum $\hat{y}(\mathbf{x})$ à l'itération 3. La fonction de mérite reste prometteuse au niveau de ce minimum, mais on peut remarquer qu'elle a diminué en amplitude. Cette diminution de l'AEI est ici plutôt due à la meilleure connaissance de la prédiction $\hat{y}(\mathbf{x})$ (augmentation de l'amplitude du minimum) qu'à la réduction de la variance locale.

Après 15 itérations de l'EGO (Figure 4.10), $\hat{y}(\mathbf{x})$ comporte toujours un seul minimum global, mais un second minimum commence à se dessiner. La variance sur la Figure 4.10(b) est maintenant très faible au niveau des points explorés. Cela est induit par la forte concentration d'observations au niveau du minimum

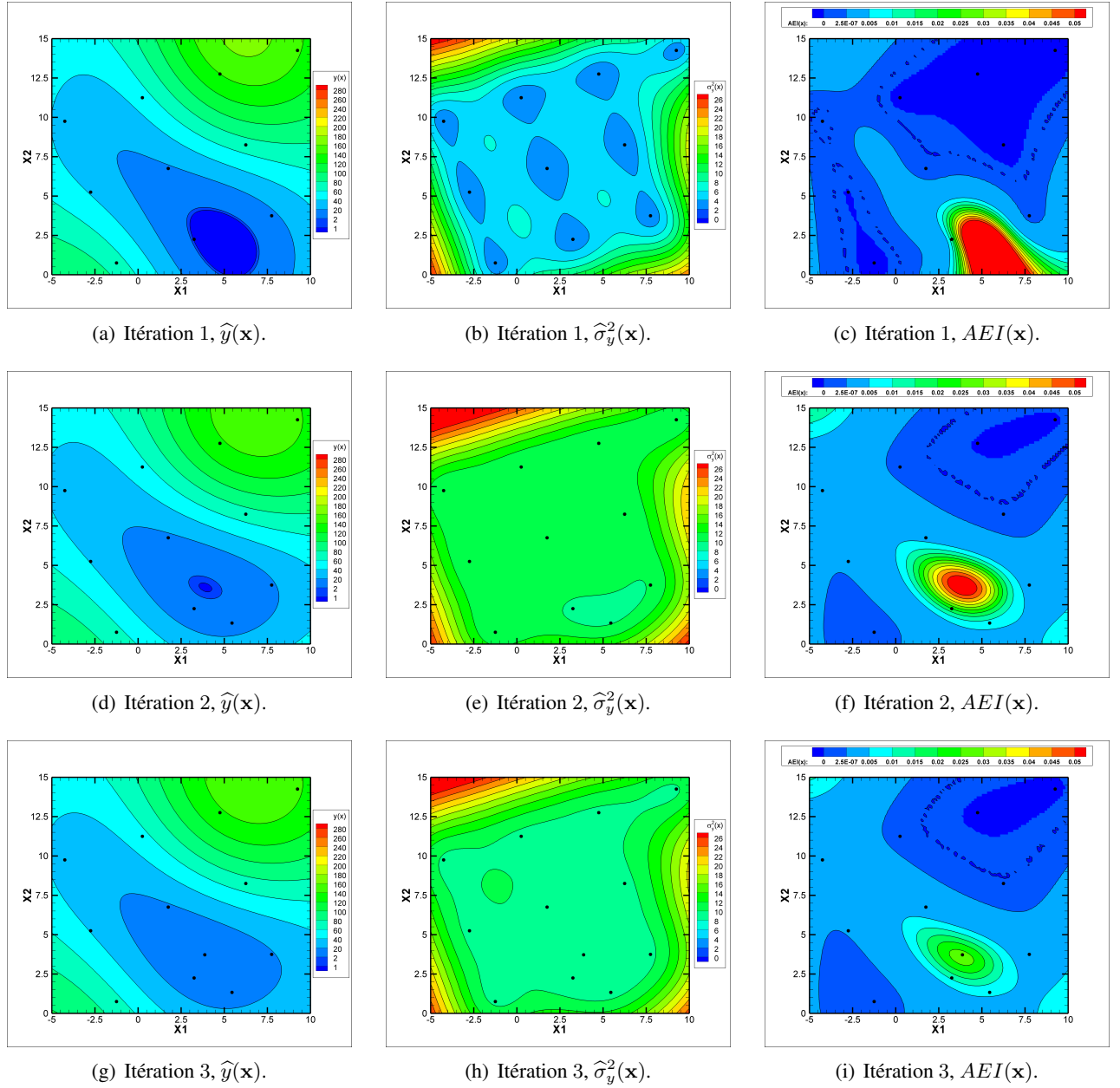


Figure 4.9 – Optimisation par méta-modèles Gaussiens de la fonction Branin, itérations 1, 2 et 3.

global du méta-modèle. On remarque par le tracé de la fonction de mérite que les deux zones non-explorées, qui contiennent les deux autres minimum globaux de la fonction Branin, deviennent des zones prometteuse. Et en effet, ces deux minimum globaux sont explorés entre les itérations 15 et 20 (voir les Figures 4.10(d)-4.10(f)). Finalement, à l'itération 50 les trois minimum globaux ont bien été explorés par l'EGO, la variance du méta-modèle est faible dans les zones d'intérêt du domaine Ω et l'AEI maximale est négligeable ($\approx 10^{-4}$). On peut alors considérer que l'algorithme EGO a convergé.

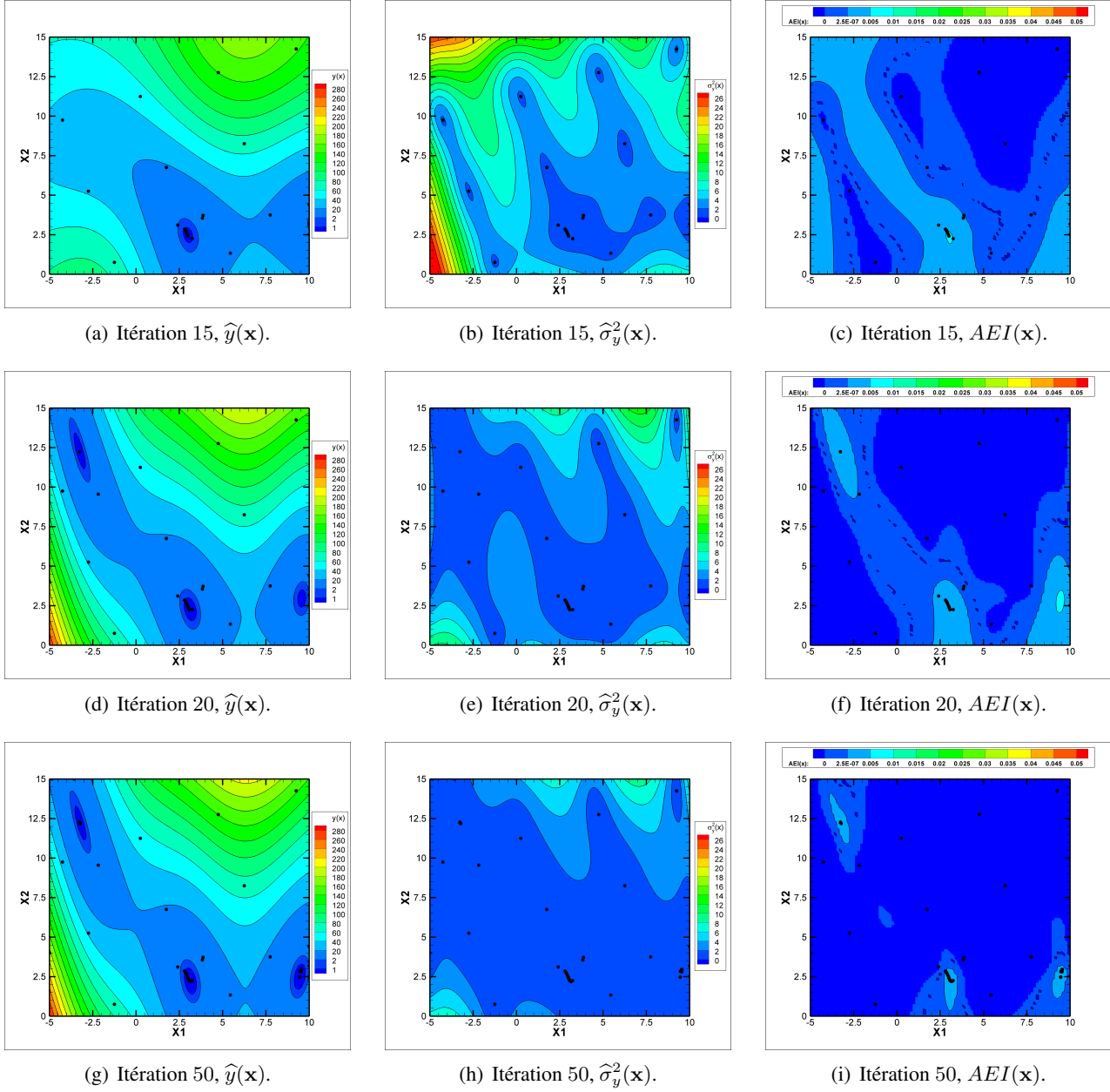


Figure 4.10 – Optimisation par méta-modèles Gaussiens de la fonction Branin, itérations 15, 20 et 50.

Optimisation sous contraintes

Sommaire

5.1	Points non évaluable	82
5.2	Modélisation des contraintes	83
5.3	Classification des contraintes	85
5.3.1	Support Vector Machines, méthodes SVM et LS-SVM	85
5.3.1.1	Méthode SVM	86
5.3.1.2	Méthode LS-SVM	92
5.3.1.3	Détermination des hyperparamètres	94
5.3.1.4	Sur-apprentissage	95
5.3.2	Méthode EGO-LS-SVM	97
5.4	EGO sous contraintes évaluable et non évaluable	99

Le Chapitre 5 concerne les problèmes d’optimisations par méta-modèles sous contraintes. On s’intéresse en particulier à des méthodes numériques permettant la prise en compte de contraintes d’inégalité. Les cas particuliers de calculs « non évaluable » sont également traités.

La formulation de l’EGO avec contraintes a été initialement évoquée dans la thèse de [Schonlau \(1997\)](#) et récemment approfondie par [Gardner et al. \(2014\)](#). L’idée est de construire de nouveaux méta-modèles qui modélisent la réponse des contraintes. Un calcul de probabilité d’admissibilité est ensuite introduit par l’intermédiaire de ces nouveaux méta-modèles afin de pénaliser la fonction de mérite dans l’EGO.

Le problème d’optimisation par méta-modèles sous contraintes a aussi été abordé à l’aide d’une méthode de classification ([Basudhar et al., 2012](#)). Un modèle de classification binaire est construit pour séparer les zones admissibles et non-admissibles de la fonction objective. L’admissibilité de la fonction objective est alors définie à l’aide d’un unique modèle qui se retrouve indépendant du nombre et du type de contraintes (inégalité, égalité,

etc.). Une fois établie, la classification permet de pénaliser la fonction de mérite à l'aide d'une probabilité d'admissibilité comme dans le cas classique où l'on modélise les contraintes.

Dans la suite du Chapitre 5, la problématique des points non évaluables est tout d'abord définie dans la Section 5.1. La définition de l'EGO sous contraintes est ensuite présentée dans les cas de « modélisation » (Section 5.2) et de « classification » (Section 5.3) des contraintes. Finalement, la Section 5.4 contient un exemple analytique d'optimisation sous contraintes comprenant des points non évaluables. Les deux méthodes de modélisation et de classification sont utilisées simultanément pour résoudre le problème d'optimisation par l'EGO.

Problème d'optimisation sous contraintes. Soit Ω un sous-ensemble non vide de $\mathbb{R}^d$, on considère le problème d'optimisation sous contraintes d'inégalité suivant

$$\begin{aligned} \min_{\mathbf{x} \in \Omega} \quad & f(\mathbf{x}), \\ \text{s.t.} \quad & \mathbf{c}(\mathbf{x}) \leq \mathbf{\Lambda}. \end{aligned} \tag{5.0.1}$$

$f : \Omega \rightarrow \mathbb{R}$ est la fonction objective à minimiser et l'ensemble des m contraintes d'inégalité est défini par $\mathbf{c} : \Omega \rightarrow \mathbb{R}^m$, $\mathbf{c}(\mathbf{x}) = (c_1(\mathbf{x}), \dots, c_m(\mathbf{x}))^T$ et $\mathbf{\Lambda} = (\lambda_1, \dots, \lambda_m)^T$.

5.1 Points non évaluables

On définit $\mathbf{x}$ comme « non évaluable » si $f(\mathbf{x}) = \emptyset$, autrement dit lorsque la fonction objective f ne peut être évaluée en $\mathbf{x}$. Ceci est une réelle difficulté pour l'application de l'EGO car si le point $\mathbf{x}$, optimum de la fonction de mérite à l'itération n , est non évaluable, alors le processus d'optimisation itératif se retrouve interrompu. En effet, comme il n'est pas possible d'évaluer $f(\mathbf{x})$, les ensembles $\mathbf{X}_n$ et $\mathbf{Y}_n$ restent inchangés et le méta-modèle n'est pas reconstruit. Ces points non évaluables ajoutent alors une contrainte dans le problème d'optimisation et ils le rendent particulièrement plus complexe car on ne connaît pas *a priori* leur domaine de variabilité, lorsque l'on entame le processus d'optimisation.

Dans le cadre particulier de la simulation numérique sur les voiles, il n'est pas inhabituel d'obtenir des points non évaluables. Il peut par exemple y avoir des combinaisons de réglages qui aboutissent à une interaction ou un contact entre deux voiles, grand-voile et génois par exemple comme l'illustre la Figure 5.1(a). Il est également possible d'obtenir une non convergence du solveur numérique (fluide ou structure), ce qui peut se traduire par des oscillations de la position des voiles lors d'un calcul quasi-statique entre deux itérations de couplage (voir Figure 5.1(b)). Les cas de non convergence sont très pénalisants dans le processus d'optimisation car ils coûtent cher en temps de calcul avant de pouvoir considérer que $f(\mathbf{x}) = \emptyset$.

Les points non évaluables sont généralement présents lorsque la dimension du problème d'optimisation est élevé. Par exemple lorsque l'on optimise la géométrie d'une section d'un hydrofoil avec un nombre de paramètres important, on peut se retrouver dans le cas de la Figure 5.1(c) où le calcul numérique d'un écoulement est difficile à résoudre.

Numériquement ces cas particuliers de point non évaluables provoquent généralement une divergence ou

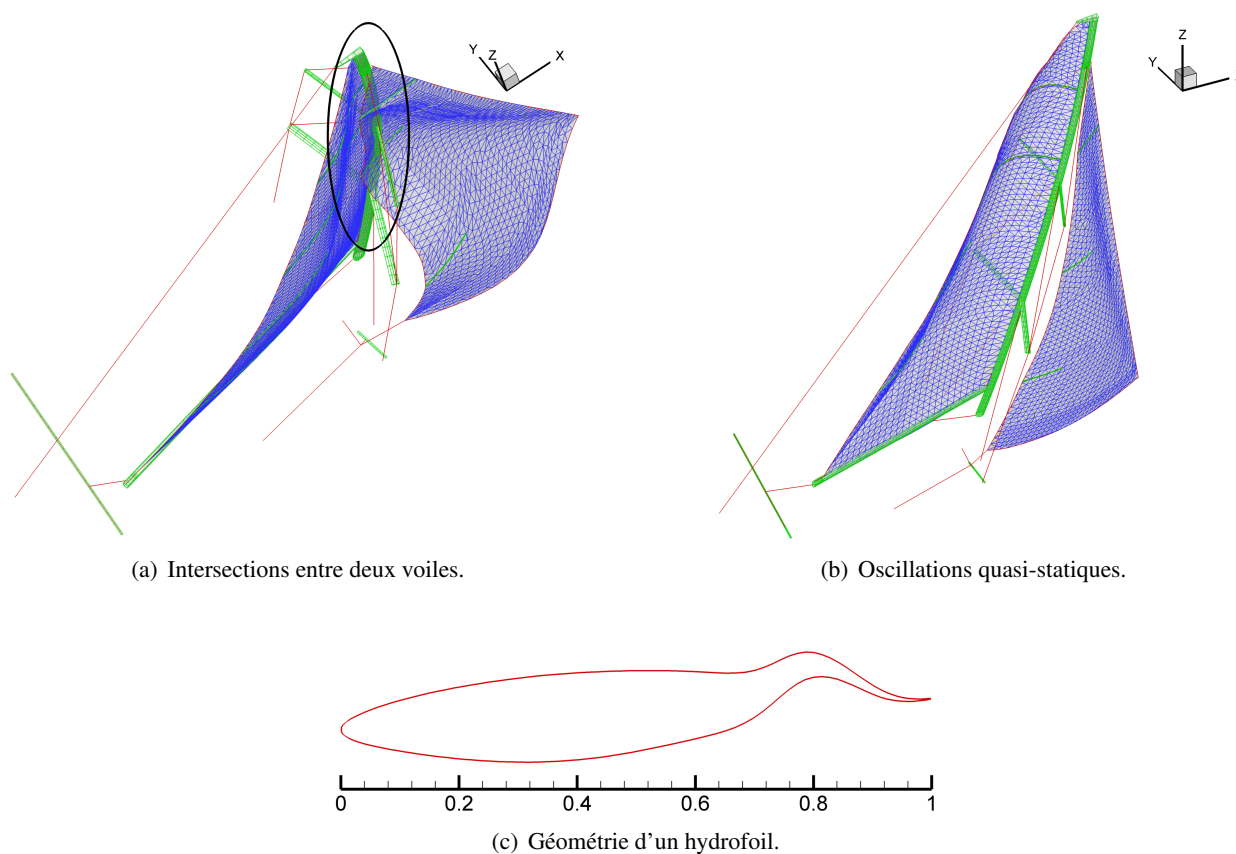


Figure 5.1 – Exemples de cas numériquement non évaluables.

une non convergence du calcul. Il n'est donc pas possible d'évaluer numériquement la fonction objective et le processus d'optimisation itératif est stoppé si aucune méthode n'est utilisée pour prendre en compte ces cas particuliers.

Dans le cadre expérimental, la présence de points non évaluables est également possible. En effet, les réglages des voiles étant souvent corrélés, une modification de la position d'un chariot de grand-voile sur un rail droit, pourra augmenter fortement la tension dans l'écoute et provoquer une rupture du tissu si la longueur d'écoute reste inchangée.

5.2 Modélisation des contraintes

On souhaite résoudre le problème (5.0.1) en pénalisant la fonction de mérite dans l'EGO lorsque $\mathbf{x}$ est non admissible : non respect des contraintes d'inégalité. La méthode de modélisation des contraintes nécessite la construction de m méta-modèles supplémentaires, associés à chaque contrainte d'inégalité, afin de calculer les probabilités de respecter chacune des m contraintes.

L'amélioration I est tout d'abord modifiée (Schonlau, 1997), dans le but d'y introduire la notion des

contraintes, par l'amélioration sous contraintes I_C

$$I_C(\mathbf{x}) \doteq \begin{cases} y_{\min} - f(\mathbf{x}) & \text{si } f(\mathbf{x}) < y_{\min}, \quad c_i(\mathbf{x}) \leq \lambda_i, \quad \forall i = 1, \dots, m \\ 0 & \text{sinon} \end{cases}, \quad (5.2.1)$$

où $y_{\min}$ est le minimum admissible à l'itération n de $\mathbf{Y}_n$. L'amélioration sous contraintes peut également s'écrire comme le propose [Gardner et al. \(2014\)](#)

$$I_C(\mathbf{x}) \doteq \max(y_{\min} - f(\mathbf{x}), 0) \Delta(\mathbf{x}), \quad (5.2.2a)$$

$$\doteq I(\mathbf{x})\Delta(\mathbf{x}), \quad (5.2.2b)$$

avec $\Delta(\mathbf{x})$ défini comme indicateur d'admissibilité qui vaut 1 si $\mathbf{c}(\mathbf{x}) \leq \mathbf{\Lambda}$ et 0 sinon. En supposant que les m contraintes sont statistiquement indépendantes, l'amélioration espérée sous contraintes s'exprime alors

$$EI_C(\mathbf{x}) = \mathbb{E}[I_C(\mathbf{x})], \quad (5.2.3a)$$

$$= \mathbb{E}[I(\mathbf{x})\Delta(\mathbf{x})], \quad (5.2.3b)$$

$$= \mathbb{E}[I(\mathbf{x})|\mathbf{x}] \mathbb{E}[\Delta(\mathbf{x})|\mathbf{x}], \quad (5.2.3c)$$

$$= EI(\mathbf{x})P_C(\mathbf{x}), \quad (5.2.3d)$$

où $P_C \in [0,1]$ est défini par

$$P_C(\mathbf{x}) = \prod_{i=1}^m P_i(c_i(\mathbf{x}) \leq \lambda_i). \quad (5.2.4)$$

P_C est le produit des probabilités de respecter les m contraintes d'inégalité. Les P_i se calculent à l'aide de la fonction de répartition

$$P_i(c_i(\mathbf{x}) \leq \lambda_i) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\lambda_i - \hat{y}_{c_i}}{\hat{\sigma}_{c_i} \sqrt{2}} \right) \right], \quad (5.2.5)$$

avec $\hat{y}_{c_i}$ et $\hat{\sigma}_{c_i}^2$ qui sont les prédictions de la moyenne et de la variance données par le modèle Gaussien de la contrainte c_i . On obtient ainsi une fonction de mérite $EI_C(\mathbf{x})$ qui correspond à l'amélioration espérée pondérée par la probabilité que le point $\mathbf{x}$ est admissible.

La méthode a l'avantage d'être « auto-adaptative » dans le sens où la variabilité des méta-modèles de chaque contrainte aura un impact sur les P_i et donc sur les choix de $\mathbf{x}_{n+1}$. Ainsi, l'enrichissement de $\mathbf{X}_n$ permet également d'enrichir les modèles associés aux c_i , dans les zones où l'amélioration espérée n'est pas nulle et/ou lorsque $P_C(\mathbf{x}) \approx 0.5$.

Des méthodes d'optimisations qui reposent sur la même idée de construire des méta-modèles pour les contraintes, ont aussi été développées par ([Audet et al., 2000](#); [Sasena et al., 2002a,b](#)).

La modélisation des contraintes peut néanmoins devenir très coûteuse lorsque le nombre de contraintes m

est important, elle peut également être difficilement applicable lorsque les contraintes sont corrélées (Forrester et al., 2008) et elle ne résout pas les cas spécifiques des points non évaluables.

5.3 Classification des contraintes

Comme introduit dans la Section 4.3.2, l'objectif de la classification binaire est de construire une fonction mathématique appelée « classifieur » qui, à partir d'un ensemble d'observations $\mathcal{D} = \{(\mathbf{x}_i, z_i) | \forall i = 1, \dots, n\}$, binairement classifié $C^{\pm 1}$ par les étiquettes $z_i = \pm 1$, prédit la classe d'une nouvelle observation $\mathbf{x}$.

Dans le cadre de l'EGO sous contraintes, on peut, comme le propose Basudhar et al. (2012) avec des « Support Vector Machines (SVM) », utiliser le classifieur pour séparer les zones admissibles et non admissibles du domaine Ω . En introduisant ensuite une vision statistique du classifieur, on pourra, comme avec la modélisation des contraintes (Section 5.2), adapter l'EGO non contraint à la prise en compte des contraintes, qui seront ici classifiées par $C^{\pm 1}$. Dans le cas des contraintes, la classification est généralement non linéaire et non séparable, comme dans l'exemple présenté dans la Section 4.3.2 avec la Figure 4.7.

La classification par processus Gaussiens (Section 4.3.2) est un choix possible de classifieur pour l'EGO sous contraintes. En pratique, ce type de classification n'est pas forcément bien adapté à l'EGO sous contraintes car son coût de construction devient trop important lorsque le nombre d'observations dépasse quelques centaines. Une autre méthode de classification, qui a un coût de construction relativement faible, est utilisée pour l'EGO sous contraintes dans cette thèse : la méthode dite « Least Squares Support Vector Machine (LS-SVM) », développée par Suykens and Vandewalle (1999), qui est une variante de la célèbre méthode SVM (Vapnik, 1995).

Dans la suite de cette Section 5.3, le classifieur SVM et sa variante LS-SVM sont tout d'abord détaillés dans la Section 5.3.1. La méthode « EGO-LS-SVM », développée au cours de cette thèse, est ensuite présentée par la Section 5.3.2.

5.3.1 Support Vector Machines, méthodes SVM et LS-SVM

Les méthodes SVM sont inspirées du perceptron, inventé dans les années 60 par Rosenblatt (1958). Le perceptron est un classifieur linéaire à d observations (nombres réels ou entiers) et une sortie définie par $d + 1$ poids, aussi appelés coefficients synaptiques, $w_0, \dots, w_d$. Un schéma de perceptron est visible sur la Figure 5.2, où chaque entrée x_i de $\mathbf{x}$ (nouvelle observation à classifier) est pondérée par w_i et le résultat z qui vaut $+1$ ou -1 , suivant le signe de la somme pondérée, définit la classe de la nouvelle observation. Le perceptron divise donc l'espace des $\mathbf{x}$ en deux régions correspondants aux classes C^+ et C^- . Une étape d'apprentissage est nécessaire pour que la classification soit la plus parfaite possible. Pour cela, une fonction objective qui traduit la performance de la classification est définie et les poids w_i sont optimisés dans le but de minimiser les erreurs de classification.

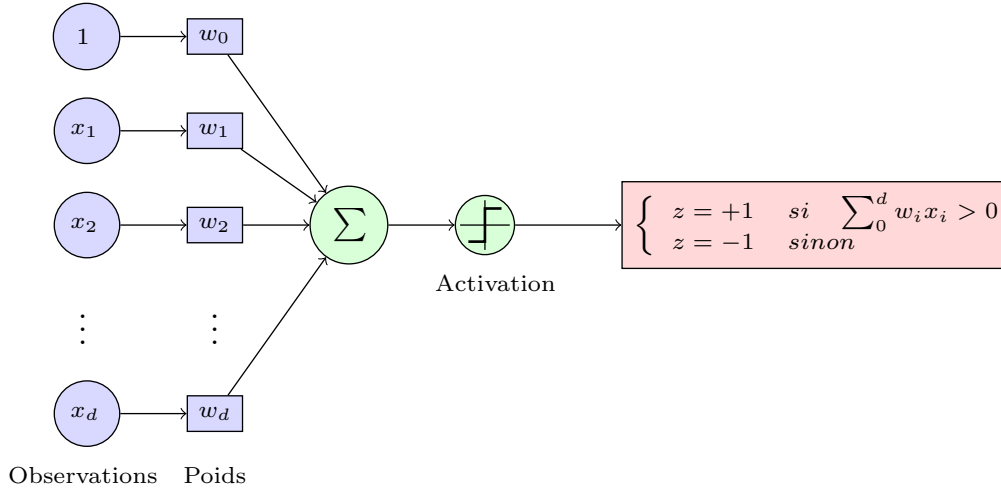


Figure 5.2 – Schéma d'un perceptron.

5.3.1.1 Méthode SVM

Les SVM (Vapnik, 1995, 1998b,a) appartiennent aux algorithmes d'apprentissage et ils sont principalement utilisés dans des problèmes de partitionnement ou de classification de données, mais ils peuvent également être appliqués à la régression. On présentera ces méthodes SVM et variantes dans des cas de classifications binaires. L'objectif est de rechercher la frontière de décision qui sépare les classes C^+ et C^- . À partir de chaque observation $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T$ étiquetée par $\mathbf{Z}_n = (z_1, \dots, z_n)^T$, avec $z_i = +1$ si $\mathbf{x}_i$ est de classe C^+ et $z_i = -1$ sinon (de classe C^-), on souhaite construire la fonction séparatrice $h : \mathbf{x} \in \Omega \rightarrow (-1, 1)$. En pratique, on utilise une fonction $h : \mathbf{x} \in \Omega \mapsto \mathbb{R}$ et la classe z d'une nouvelle observation $\mathbf{x}$ est donnée par le signe de la fonction séparatrice : $z = \text{sign}(h(\mathbf{x}))$. Dans le cas d'une classification binaire, la fonction séparatrice linéaire peut être définie comme

$$h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \quad (5.3.1)$$

où $\mathbf{w} \in \mathbb{R}^d$ est un vecteur de poids et $b \in \mathbb{R}$ un biais. Le séparateur (5.3.1) est, dans un cas à deux dimensions, une droite séparant « au mieux » les classes C^+ et C^- (voir Figure 5.3). On appelle cette fonction un « hyperplan séparateur », qui existe si les observations sont linéairement séparables. L'hyperplan optimal est celui pour lequel la distance minimale aux données d'apprentissage est maximale. Cette distance à maximiser est appelée la « marge » ; elle est représentée sur la Figure 5.3. La distance d'un point à l'hyperplan est donnée par

$$d(\mathbf{x}) = \frac{|\mathbf{w}^T \mathbf{x} + b|}{\|\mathbf{w}\|}. \quad (5.3.2)$$

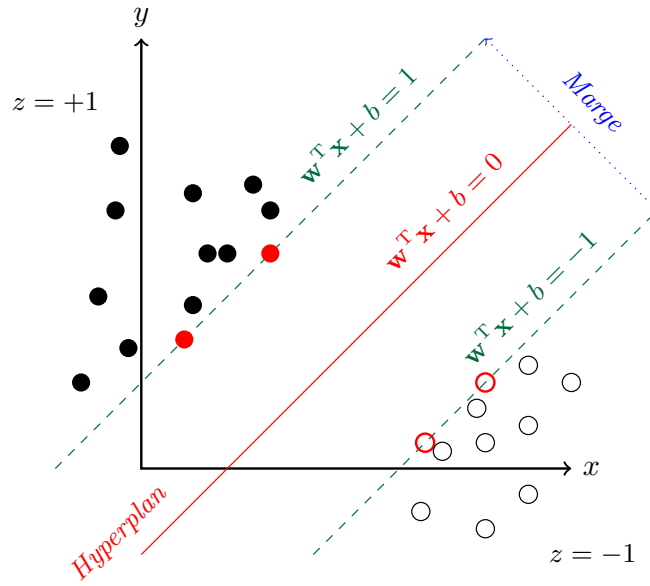


Figure 5.3 – Classification binaire de données séparables avec un SVM linéaire.

Maximiser la marge est donc équivalent à minimiser $\|\mathbf{w}\|$ sous les contraintes

$$\begin{cases} \mathbf{w}^T \mathbf{x}_i + b \geq 1 & \text{si } z_i = +1 \\ \mathbf{w}^T \mathbf{x}_i + b \leq -1 & \text{si } z_i = -1 \end{cases}, \quad i = 1, \dots, n, \quad (5.3.3)$$

que l'on peut réécrire par

$$z_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, \dots, n. \quad (5.3.4)$$

La formulation classique des SVM correspond donc à la résolution d'un problème d'optimisation sous contraintes et elle s'écrit comme

$$\begin{aligned} \min_{\mathbf{w}, b} \quad & \frac{1}{2} \|\mathbf{w}\|^2, \\ \text{s.t.} \quad & z_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \quad i = 1, \dots, n. \end{aligned} \quad (5.3.5)$$

Les observations qui vérifient l'égalité de la contrainte sont les vecteurs supports, ils suffisent à eux seuls à la détermination de l'hyperplan séparateur et ils sont représentés par les points rouges sur la Figure 5.3. L'écriture (5.3.5), appelée « formulation primale », est un problème de programmation quadratique qui requiert la résolution de $d + 1$ variables ($\mathbf{w}$ et b). Cette formulation est difficile à résoudre lorsque la dimension du problème d est importante. Il est donc nécessaire de passer à la « formulation duale » en introduisant les multiplicateurs de Lagrange $\alpha_i \geq 0$ associés à chaque contrainte comme

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (z_i (\mathbf{w}^T \mathbf{x}_i + b) - 1), \quad (5.3.6)$$

où $\alpha \in \mathbb{R}^n$ est le vecteur des multiplicateurs de Lagrange. Le produit des contraintes aux α_i et les dérivées par rapport à $\mathbf{w}$ et b s'annulent à l'optimum de $\mathcal{L}$, et on obtient ainsi

$$\left\{ \begin{array}{ll} \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = \mathbf{0} & \longrightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i z_i \mathbf{x}_i \\ \frac{\partial \mathcal{L}}{\partial b} = 0 & \longrightarrow \sum_{i=1}^n \alpha_i z_i = 0 \\ \alpha_i (z_i (\mathbf{w}^T \mathbf{x}_i + b) - 1) = 0 & \longrightarrow \mathcal{I} = \{i \in [1, n] \mid z_i (\mathbf{w}^T \mathbf{x}_i + b) = 1\} \\ \alpha_i \geq 0, \quad i = 1, \dots, n \end{array} \right. . \quad (5.3.7)$$

$\mathcal{I}$ contient les vecteurs supports qui définissent l'hyperplan pour lesquels $|\mathbf{w}^T \mathbf{x}_i + b| = 1$, $i \in \mathcal{I}$. Lorsque $i \notin \mathcal{I}$, l'observation $\mathbf{x}_i$ est inutile dans la définition de la séparatrice et $\alpha_i = 0$. A partir des équations (5.3.6) et (5.3.7), on a

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^n \alpha_i z_i \mathbf{x}_i^T \sum_{j=1}^n \alpha_j z_j \mathbf{x}_j - \sum_{i=1}^n \alpha_i \left(z_i \left(\sum_{j=1}^n \alpha_j z_j \mathbf{x}_j^T \mathbf{x}_i + b \right) - 1 \right), \quad (5.3.8a)$$

$$= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j z_i z_j \mathbf{x}_i^T \mathbf{x}_j + \sum_{i=1}^n \alpha_i - \sum_{i=1}^n \alpha_i z_i \sum_{j=1}^n \alpha_j z_j \mathbf{x}_j^T \mathbf{x}_i - b \sum_{i=1}^n \alpha_i z_i, \quad (5.3.8b)$$

$$= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j z_i z_j \mathbf{x}_i^T \mathbf{x}_j. \quad (5.3.8c)$$

La formulation duale à maximiser s'écrit finalement comme le problème (5.3.9) ci-dessous

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j z_i z_j \mathbf{x}_i^T \mathbf{x}_j, \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i z_i = 0, \\ & \alpha_i \geq 0, \quad i = 1, \dots, n. \end{aligned} \quad (5.3.9)$$

Le problème de programmation quadratique est bien de dimension n et la fonction séparatrice peut alors s'écrire comme

$$h(\mathbf{x}) = \sum_{i \in \mathcal{I}} \alpha_i z_i \mathbf{x}_i^T \mathbf{x} + b, \quad (5.3.10)$$

où les α_i sont solutions du problème (5.3.9) et b est calculé à partir d'un vecteur support par l'expression

$$b = z_i - \mathbf{w}^T \mathbf{x}_i. \quad (5.3.11)$$

Toute nouvelle observation $\mathbf{x}$ peut ainsi être classifiée à l'aide de l'équation (5.3.10) par $z = \text{sign}(h(\mathbf{x}))$.

Dans le cas de données non linéairement séparables, les contraintes (5.3.4) ne peuvent être respectées. Il est néanmoins possible de suivre la même méthode que dans le cas linéairement séparable en introduisant des variables de relaxation ξ_i sur les contraintes. Ces nouvelles variables permettent de modéliser des erreurs associées à chaque observation $\mathbf{x}_i$ en acceptant une « mal-classification » de certaines données. Elles sont définies comme

$$\xi_i = \max \left(1 - z_i \left(\mathbf{w}^T \mathbf{x}_i + b \right), 0 \right), \quad i = 1, \dots, n. \quad (5.3.12)$$

L'écart $\xi_i = 0$ si la contrainte $z_i \left(\mathbf{w}^T \mathbf{x}_i + b \right) \geq 1$ est vérifiée et les contraintes du cas de données non séparables deviennent

$$z_i \left(\mathbf{w}^T \mathbf{x}_i + b \right) \geq 1 - \xi_i, \quad i = 1, \dots, n. \quad (5.3.13)$$

L'objectif est maintenant de rechercher l'hyperplan séparateur qui maximise la marge tout en minimisant la somme des erreurs $\sum_{i=1}^n \xi_i$. Le problème (5.3.5) est modifié par

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i, \\ \text{s.t.} \quad & z_i \left(\mathbf{w}^T \mathbf{x}_i + b \right) \geq 1 - \xi_i, \quad i = 1, \dots, n, \\ & \xi_i \geq 0, \quad i = 1, \dots, n. \end{aligned} \quad (5.3.14)$$

$\xi \in \mathbb{R}^n$ est le vecteur des variables de relaxation et $C \in \mathbb{R}$ est un paramètre fixé qui permet le contrôle de la pénalisation par le dépassement de la contrainte. Comme précédemment, les multiplicateurs de Lagrange $\alpha_i \geq 0$ et $\beta_i \geq 0$ sont introduits afin d'obtenir le problème dual dans le cas non séparable

$$\mathcal{L}(\mathbf{w}, b, \alpha, \xi, \beta) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i \left(z_i \left(\mathbf{w}^T \mathbf{x}_i + b \right) - 1 + \xi_i \right) - \sum_{i=1}^n \beta_i \xi_i, \quad (5.3.15)$$

où $\beta \in \mathbb{R}^n$ est, comme α , un vecteur des multiplicateurs de Lagrange. Le calcul des gradients est identique au cas séparable par rapport à $\mathbf{w}$ et b et la dérivée par rapport aux variables de relaxation ξ donne l'équation

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \xi} = \mathbf{0} \\ \alpha_i \geq 0; \beta_i \geq 0; \xi_i \geq 0, \quad i = 1, \dots, n \end{cases} \longrightarrow C - \alpha_i - \beta_i = 0, \quad (5.3.16)$$

qui permet d'écrire la condition $\alpha_i \leq C$. On obtient finalement le problème dual dans le cas d'observations non séparables (équation (5.3.17)) ayant comme unique différence avec le cas séparable la borne supérieure sur les

α_i

$$\begin{aligned}
& \max_{\alpha} \quad \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j z_i z_j \mathbf{x}_i^T \mathbf{x}_j, \\
& \text{s.t.} \quad \sum_{i=1}^n \alpha_i z_i = 0, \\
& \quad \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n.
\end{aligned} \tag{5.3.17}$$

La résolution du problème (5.3.17) consiste finalement à déterminer les α_i . Dans le cas où le nombre d'observations reste faible, des méthodes classiques comme quasi-Newton sont applicables. Dans la plupart des autres cas, le problème est résolu à l'aide de la méthode « Sequential Minimal Optimization (SMO) » développée par Platt (1999a).

Les outils de classifications binaire avec les SVM dans les cas de séparateurs linéaires ont été détaillés jusqu'ici. Or, comme on a pu le voir sur l'exemple donné avec la Figure 4.7, il existe des jeux de données qui sont non linéairement classifiables. La fonction séparatrice doit dans ce cas être non linéaire et « coller au mieux » les observations comme sur l'exemple de la Figure 5.4 où la courbe rouge représente le séparateur non linéaire ($z = 0$). Il est cependant très difficile de construire une telle fonction dans son espace de description. L'idée

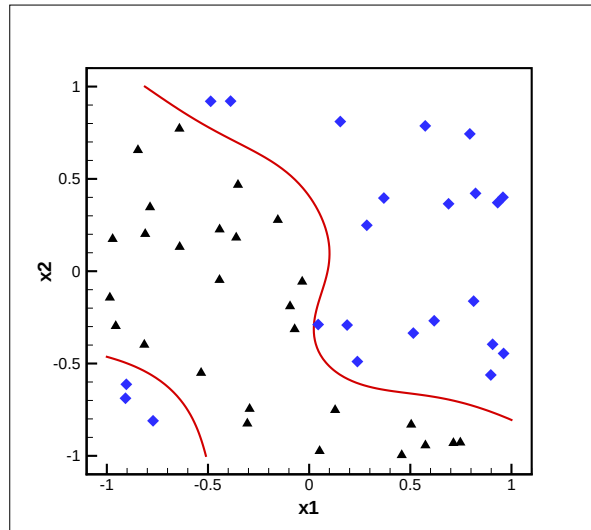


Figure 5.4 – Fonction séparatrice non linéaire.

des SVM (Vapnik, 1998b) a alors été de transformer l'espace des observations en un espace de plus grande dimension où la fonction séparatrice est linéaire. En effet, la probabilité de construire une fonction séparatrice linéaire sera plus importante dans un espace de grande dimension. On utilise la transformation non linéaire Φ , appelée « fonction de mapping », de l'espace des observations vers l'espace de caractéristiques (« Feature space » en anglais). Le vecteur image $\phi(\mathbf{x}) = (\phi(x_1), \dots, \phi(x_d), \dots)$ est ainsi de dimension supérieure à d en

pratique, et l'équation (5.3.17) devient dans l'espace de redescription

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j z_i z_j \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle, \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i z_i = 0, \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n, \end{aligned} \quad (5.3.18)$$

avec la notation $\langle \cdot, \cdot \rangle$ qui correspond au produit scalaire. L'expression du séparateur est alors

$$h(\mathbf{x}) = \sum_{i \in \mathcal{S}} \alpha_i z_i \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle + b, \quad (5.3.19)$$

où $\mathcal{S}$ est l'ensemble des vecteurs supports. Malheureusement, le calcul du produit scalaire $\langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle$ peut devenir très coûteux lorsque la dimension de ϕ augmente et l'équation du séparateur devient difficilement évaluable. L'astuce consiste ensuite à introduire l'utilisation de « fonctions noyaux » symétriques et semi-définies positives dont on peut montrer l'équivalence à un produit scalaire dans un espace de grande dimension (méthode « kernel trick » en anglais). Plusieurs théorèmes permettent de caractériser les fonctions noyaux, mais le plus connu est le théorème de Mercer (Mercer, 1909) qui donne les conditions pour qu'une fonction symétrique $k : \Omega \times \Omega \rightarrow \mathbb{R}$ soit une fonction noyau. Le noyau $k(\mathbf{x}, \mathbf{x}')$ associé à un couple $(\mathbf{x}, \mathbf{x}')$ mesure la similarité réciproque entre les observations $\mathbf{x}$ et $\mathbf{x}'$. Les fonctions noyaux sont aussi appelées des « noyaux de covariance » et dans la pratique ils composent la matrice $\mathbf{K}$ (matrice de Gram, semi-définie positive) par les éléments $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. Le Tableau 5.1 propose quelques exemples de noyaux positifs courants pour l'application des SVM. Les variables c , p et σ sont des hyperparamètres qui sont fixés par l'utilisateur, ou bien lorsque

Noyau	$k(\mathbf{x}, \mathbf{x}')$
Linéaire	$\mathbf{x}^T \mathbf{x}'$
Polynomial	$(\mathbf{x}^T \mathbf{x}' + c)^p$
Gaussien	$\exp\left(-\frac{\ \mathbf{x} - \mathbf{x}'\ ^2}{2\sigma^2}\right)$

Tableau 5.1 – Exemples de noyaux pour la classification par SVM.

cela est possible, ils sont optimisés lors d'une phase d'apprentissage, que l'on détaillera par la suite. Lorsque les fonctions noyaux sont utilisées, le problème (5.3.18) est équivalent au problème (5.3.20) suivant

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j z_i z_j k(\mathbf{x}_i, \mathbf{x}_j), \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i z_i = 0, \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n, \end{aligned} \quad (5.3.20)$$

dont l'équation de la séparatrice est alors donnée par

$$h(\mathbf{x}) = \sum_{i \in \mathcal{S}} \alpha_i z_i k(\mathbf{x}_i, \mathbf{x}) + b, \quad (5.3.21)$$

où $\mathcal{S}$ est l'ensemble des vecteurs supports et les coefficients α_i et b sont obtenus comme précédemment à partir de la solution du problème (5.3.20). L'attribution d'une classe à une nouvelle observation $\mathbf{x}$ est calculée par $z = \text{sign}(h(\mathbf{x}))$.

5.3.1.2 Méthode LS-SVM

Les LS-SVM (Suykens and Vandewalle, 1999) sont une variante des SVM de Vapnik (1995). Le principe de maximisation de la marge pour construire un hyperplan séparateur optimal est gardé, mais les modifications apportées à la formulation des SVM permettent de résoudre un système linéaire plutôt qu'un problème de programmation quadratique. La formulation primale des LS-SVM est la suivante

$$\begin{aligned} \min_{\mathbf{w}, b, \mathbf{e}} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2, \\ \text{s.t.} \quad & z_i = \mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, n, \end{aligned} \quad (5.3.22)$$

avec $\gamma \in \mathbb{R}^+$ un paramètre fixé et les erreurs notées e_i . On remarque que les contraintes d'inégalité sont remplacées par des égalités et que la pénalisation devient quadratique. Le Lagrangien de (5.3.22) se définit comme

$$\mathcal{L}(\mathbf{w}, b, \alpha, \mathbf{e}) = \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2 - \sum_{i=1}^n \alpha_i \left(\mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i - z_i \right), \quad (5.3.23)$$

où les e_i composent le vecteur $\mathbf{e} \in \mathbb{R}^n$ et les α_i peuvent être positifs ou négatifs car les contraintes sont des égalités. Les conditions d'optimalité du problème produisent les équations

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = \mathbf{0} \longrightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i \phi(\mathbf{x}_i) \\ \frac{\partial \mathcal{L}}{\partial b} = 0 \longrightarrow \sum_{i=1}^n \alpha_i = 0 \\ \frac{\partial \mathcal{L}}{\partial \mathbf{e}} = \mathbf{0} \longrightarrow \alpha_i = \gamma e_i, \quad i = 1, \dots, n \\ \frac{\partial \mathcal{L}}{\partial \alpha} = \mathbf{0} \longrightarrow \mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i - z_i = 0, \quad i = 1, \dots, n \end{array} \right., \quad (5.3.24)$$

qui peuvent se réécrire sous la forme matricielle

$$\begin{bmatrix} \mathbf{I}_n & 0 & 0 & -\phi_n^T \\ 0 & 0 & 0 & -\mathbf{1}^T \\ 0 & 0 & \gamma \mathbf{I}_n & -\mathbf{I}_n \\ \phi_n & \mathbf{1} & \mathbf{I}_n & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ b \\ \mathbf{e} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 0 \\ \mathbf{0} \\ \mathbf{z} \end{bmatrix}, \quad (5.3.25)$$

en ayant posé les notations

$$\begin{aligned} \phi_n &= (\phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_n))^T, \\ \mathbf{z} &= (z_1, \dots, z_n)^T, \\ \mathbf{e} &= (e_1, \dots, e_n)^T, \\ \boldsymbol{\alpha} &= (\alpha_1, \dots, \alpha_n)^T, \\ \mathbf{1} &= (1, \dots, 1)^T, \\ \mathbf{0} &= (0, \dots, 0)^T, \end{aligned}$$

et $\mathbf{I}_n$ est la $[n \times n]$ matrice identité. En utilisant les fonctions noyaux (condition de Mercer, $k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}), \phi(\mathbf{x}') \rangle$), la solution est finalement obtenue en résolvant le système linéaire suivant

$$\begin{bmatrix} \mathbf{K} + \gamma^{-1} \mathbf{I}_n & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ 0 \end{bmatrix}. \quad (5.3.26)$$

L'équation de l'hyperplan séparateur avec le méthode LS-SVM est finalement donnée par

$$h(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b, \quad (5.3.27)$$

et comme dans le cas des SVM, la classification d'une nouvelle observation $\mathbf{x}$ est donnée par $z = \text{sign}(h(\mathbf{x}))$.

La méthode LS-SVM a donc l'avantage d'éviter la résolution du problème quadratique présent dans le cas des SVM. Cependant, les contraintes d'égalité imposent que la grande majorité des α_i soient non nuls, car $\alpha_i = \gamma e_i$. Seuls les points situés au niveau de la marge ne sont pas des vecteurs supports ($e_i = 0$). De ce fait, la méthode LS-SVM perd le côté « sparse » que l'on avait avec les SVM et sa phase de test est plus complexe. Plusieurs études ont été produites dans le but de comparer les SVM aux LS-SVM comme par exemple le travail très complet de [van Gestel et al. \(2004\)](#) qui fait des « benchmarks » sur une vingtaine de jeux de données publics, dans des cas de classifications binaires et multi-classes, avec différents noyaux et en utilisant des méthodes de « sparsification » sur certains tests en LS-SVM. Les résultats ont montré une très bonne concordance entre les deux méthodes et l'utilisation des LS-SVM « sparses » semble bien fonctionner dans le sens où les résultats sont très proches du même cas sans la « sparsification », mais avec une réduction non négligeable de la difficulté

rencontrée lors de l'étape d'apprentissage. Il existe également des améliorations que l'on peut ajouter à la formulation de base des LS-SVM. [Suykens et al. \(2002\)](#) proposent une formulation plus robuste pour améliorer l'estimation des vecteurs supports dans le cas où les erreurs sur les variables ne suivent pas une loi Normale. [Cawley \(2006\)](#) donne une implémentation efficace de la résolution du système linéaire (5.3.26), dans sa version de base et robuste, par une factorisation de type Cholesky.

Les travaux sur la « sparsification » des LS-SVM sont assez denses dans la bibliographie. Cela est particulièrement vrai pour les cas d'applications où le nombre d'observations est élevé. Les premières méthodes pour dégrader le modèle LS-SVM consistent à supprimer les données pour lesquelles $|\alpha_i|$ est faible ([Suykens et al., 2000](#)). La dégradation de la performance du séparateur peut cependant être très soudaine et il est souvent nécessaire de l'adapter au « cas par cas ». On peut aussi citer le travail de [López Lázaro et al. \(2011\)](#) qui ont adapté les développements de [Huang et al. \(2010\)](#) réalisés initialement dans le cadre des SVM pour imposer $\alpha_i = 0$ aux observations inutiles, à l'aide d'une procédure itérative du type régularisation en norme L_0 . La méthode est efficace dans la réduction du nombre de vecteurs supports, mais elle est néanmoins très couteuse en calculs : entre 15 et 20 itérations de complexité $\mathcal{O}(n^3)$. Ces deux méthodes ont le désavantage d'utiliser la totalité des observations au début du processus de « sparsification ». Par conséquent, l'étape de dégradation du LS-SVM est couteuse lorsque le nombre de données est important. D'autres auteurs ([Baudat and Anouar, 2003](#); [Jiao et al., 2007](#)) proposent alors de construire $\mathbf{X}_n$ en ajoutant itérativement des observations selon un critère de sélection. La méthode « Feature Vector Selection » de [Baudat and Anouar \(2003\)](#) qui sélectionne les observations dans l'espace de caractéristiques est notamment bien éprouvée dans le domaine ([Cawley and Talbot, 2004](#)).

5.3.1.3 Détermination des hyperparamètres

Comme on a pu le voir précédemment, les méthodes SVM font parfois intervenir des noyaux non linéaires définis par un ou plusieurs hyperparamètres (Tableau 5.1). Dans le cas par exemple du noyau Gaussien avec la méthode LS-SVM, les hyperparamètres σ (du noyau) et γ (du classifieur) doivent être fixés avant de résoudre le système (5.3.26), pour ensuite évaluer l'équation (5.3.27). Les hyperparamètres sont parfois directement fixés par l'utilisateur, dans le cas contraire il est nécessaire de passer par une étape d'apprentissage (appelée « Model Selection » en anglais) pour déterminer leurs valeurs optimales.

Les méthodes de « cross-validation » ([Stone, 1974](#)) sont très souvent employées lors de la phase d'apprentissage. Elles déterminent les hyperparamètres en évaluant la performance du modèle de classification. La méthode « k -Fold cross-validation » consiste à partitionner un ensemble de données de manière aléatoire en k groupes de tailles identiques. Puis chaque groupe est utilisé pour valider le modèle construit à partir des $k - 1$ autres groupes. La performance de la classification évaluée sur chacun des k groupes fournit une performance globale moyenne du classificateur construit sur l'ensemble des données. Le cas extrême est lorsque chaque groupe est constitué d'une seule observation, cela s'appelle la « Leave-One-Out (LOO) cross-validation ». Un critère doit aussi être choisi pour définir la performance d'une classification : le critère de sélection des hyperparamètres. On peut citer le critère PRESS (pour « Predicted Residual Sum-of-Squares ») de [Allen \(1974\)](#) à

minimiser et qui s'écrit

$$\text{PRESS} = \sum_{i=1}^n \left(z_i - \hat{z}_i^{(-i)} \right)^2, \quad (5.3.28)$$

où $\hat{z}_i^{(-i)}$ est la prédiction du modèle LS-SVM (équation (5.3.27)) pour la $i^{\text{ième}}$ observation du $i^{\text{ième}}$ groupe dans la procédure LOO. Autrement dit, c'est l'évaluation $h(\mathbf{x}_i)$ à partir du classifieur LS-SVM construit sur l'ensemble des observations sauf $\mathbf{x}_i$. En posant

$$\mathbf{C} = \begin{bmatrix} \mathbf{K} + \gamma^{-1} \mathbf{I} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix}, \quad (5.3.29)$$

dans l'équation (5.3.26), [Cawley and Talbot \(2003\)](#) ont montré que l'erreur résiduelle sur l'observation $\mathbf{x}_i$ pouvait s'écrire comme

$$r_i^{(-i)} = z_i - \hat{z}_i^{(-i)} = \frac{\alpha_i}{C_{ii}^{-1}}, \quad i = 1, \dots, n, \quad (5.3.30)$$

où C_{ii}^{-1} est le $i^{\text{ième}}$ composant de la diagonale de l'inverse de la matrice $\mathbf{C}$. Les auteurs proposent dans le même travail une implémentation efficace par une factorisation de Cholesky pour déterminer les coefficients de la diagonale de $\mathbf{C}^{-1}$. Cela permet de réduire fortement le coût de calcul de la procédure LOO même si le nombre d'observations qui composent $\mathbf{X}_n$ est important. On peut trouver différents critères de sélection dans la bibliographie, en particulier le travail de [Cawley \(2006\)](#) en compare cinq avec plusieurs types de noyaux, qu'il teste sur cinq « benchmarks » publics. Comme pour les critères de sélection, il existe plusieurs modèles de sélection des hyperparamètres comme par exemple la « Generalized Cross-Validation (GCV) » appliquée aux LS-SVM dans ([Alvarez Meza et al., 2012](#)) qui cherche à minimiser un ratio entre l'erreur de classification et la « complexité » du modèle. Il y a également des méthodes qui utilisent des notions de probabilité de « mal-classification » ([Ayat et al., 2005](#); [Adankon and Cheriet, 2009](#)) ou bien encore plusieurs niveaux d'inférence bayésienne ([Van Gestel et al., 2002](#); [Adankon et al., 2011](#)).

5.3.1.4 Sur-apprentissage

Appelé « overfitting » en anglais, le sur-apprentissage est une difficulté que l'on peut rencontrer lors de la phase de détermination des hyperparamètres. Considérant deux sous-ensembles, de validation et d'apprentissage, d'un jeu de données, le risque de sur-apprentissage peut apparaître si l'erreur sur l'ensemble d'apprentissage diminue alors que celle sur l'ensemble de validation augmente, lorsque la complexité du modèle augmente. C'est ce que représente la Figure 5.5, avec la courbe continue bleu qui correspond à l'erreur de l'ensemble d'apprentissage et la courbe rouge discontinue qui est l'erreur de l'ensemble de validation. La complexité optimale du modèle se situe dans la zone précédant « l'overfitting », où les deux erreurs sont minimales.

Dans le cadre de l'optimisation par méta-modèle sous contraintes avec une méthode de classification des contraintes, le sur-apprentissage peut apparaître lorsque l'optimum commence à être « sur-défini » par le méta-

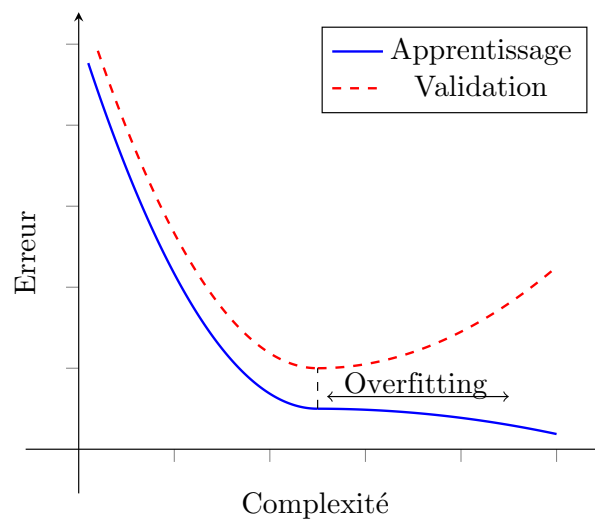


Figure 5.5 – Risque de surapprentissage.

modèle : en cours ou en fin de convergence du processus d'optimisation. En effet, l'optimisation peut rendre la répartition des observations très inégale dans le domaine de variabilité et comme l'optimum est généralement situé sur la limite d'une contrainte d'inégalité, il peut y avoir une forte concentration de données au niveau de cette limite. Cela est pénalisant pour le modèle de sélection car la performance globale du séparateur sur l'erreur de classification est par conséquent très dépendante des points locaux situés autour de l'optimum. La longueur de corrélation σ du noyaux Gaussien devient très faible pour que le classifieur puisse capter le « saut de classe » au niveau de l'optimum, et même la méthode k -Fold cross-validation ne peut dans certains cas éviter « l'overfitting ». Le problème du sur-apprentissage n'est cependant pas toujours bien détaillé dans les publications, en partie car son apparition est souvent dépendante des observations à classifier plutôt que du type de séparateur ou du modèle de sélection des hyperparamètres employé. Certains auteurs y consacrent tout de même de l'intérêt, comme [Cawley and Talbot \(2007\)](#) qui proposent l'introduction d'une régularisation bayésienne des hyperparamètres dans le modèle de sélection, ou bien toujours avec les mêmes auteurs [Cawley and Talbot \(2010\)](#) qui donnent quelques bonnes pratiques de base pour éviter « l'overfitting », tout en montrant que cette difficulté peut exister quel que soit le modèle ou critère de sélection employé. Ils incitent également les chercheurs à poursuivre ou entamer des développements sur cette problématique dans cette dernière publication. [Yan et al. \(2013\)](#) proposent quant à eux une étude de la distribution des hyperparamètres d'un SVM avec un noyau Gaussien. Ils définissent ainsi un range optimal des hyperparamètres qui permet d'éviter les minima locaux lors de la phase d'optimisation du critère de sélection. La publication de [Guyon et al. \(2010\)](#) est particulièrement intéressante car les auteurs y ont réalisé une synthèse des dernières méthodes concernant les modèles et critères de sélection, en abordant aussi la problématique de l'évitement de « l'overfitting ».

Pour résumer, il n'existe pas vraiment de méthode de sélection de modèle optimale pour les SVM et ses variantes. Certaines méthodes seront plus performantes avec un premier jeu de données mais cela ne sera pas forcément vrai avec un deuxième. Une bonne pratique consiste à utiliser la LOO de ([Cawley and Talbot, 2003](#)) avec le critère PRESS et le noyau Gaussien composé d'une unique longueur de corrélation σ . Le fait de prendre

en considération des bornes de recherche adaptées des hyperparamètres (γ, σ) est aussi souvent très utile pour éviter les problèmes « d'overfitting ».

5.3.2 Méthode EGO-LS-SVM

Le principe de gestion des contraintes par la classification LS-SVM dans l'EGO est assez proche du cas de modélisation des contraintes. La différence concerne la « vision des contraintes » qui sont ici, peu importe leur nombre, toutes rassemblées en deux classes : admissible (C^+) ou non admissible (C^-). Si une contrainte d'inégalité n'est pas respectée, alors l'observation est classée comme non admissible. Par conséquent, la notion de distance aux limites de chaque contrainte est perdue. Cela permet cependant la gestion d'un seul type de contraintes (vision binaire), indépendante de m , sans avoir à connaître les corrélations possibles entre les contraintes. De plus, tous les types de contraintes peuvent être utilisées, telles que binaires ou discontinues. Ainsi le problème des points non évaluable est résoluble avec cette méthode. Le point $\mathbf{x}$ non évaluable ne pourra pas renseigner le méta-modèle de la fonction objective, mais il renseignera la classification comme étant non admissible et la taille des matrices $\mathbf{X}_n$ (fonction objective et classification) sera alors différente. La méthode de classification aura néanmoins plus de difficultés que la méthode de modélisation des contraintes à déterminer un point qui respecte les contraintes, lorsque le domaine admissible n'aura pas encore été exploré lors de la première phase d'initialisation (étape du plan d'expériences).

L'objectif est de calculer la probabilité d'appartenance à la classe admissible qui puisse pénaliser la fonction de mérite de l'EGO. Van Calster et al. (2007) proposent une comparaison de méthodes pour modéliser les probabilités d'appartenance à une classe dans un cas de décision médical avec des LS-SVM. On souhaite pénaliser la fonction de mérite par une probabilité définie comme

$$P(z = +1|h) = \begin{cases} 1 & \text{si } h(\mathbf{x}) \geq +1 \\ 0 & \text{si } h(\mathbf{x}) \leq -1 \\ P \in]0,1[& \text{sinon} \end{cases} \quad (5.3.31)$$

On applique pour cela les développements de Platt (1999b) qui estime P par la fonction sigmoïde suivante

$$P(z = +1|h) \approx P_{A,B}(h) = \frac{1}{1 + \exp(Ah(\mathbf{x}) + B)}. \quad (5.3.32)$$

$A < 0$ et B sont des paramètres à déterminer et $h(\mathbf{x})$ est donnée par le classifieur LS-SVM. Lin et al. (2007) ont proposé une implémentation plus robuste que celle initiale de Platt (1999b) pour l'apprentissage des paramètres (A, B) . Ils sont déterminés sur l'ensemble $\mathcal{A} = \{(\mathbf{x}_1, t_1), \dots, (\mathbf{x}_n, t_n)\}$, où t_i est la probabilité que $z_i(\mathbf{x}_i) = +1$ comme

$$t_i = \begin{cases} \frac{N_+ + 1}{N_- + 2} & \text{si } z_i = +1 \\ \frac{1}{N_- + 2} & \text{si } z_i = -1 \end{cases} \quad i = 1, \dots, n, \quad (5.3.33)$$

avec N_+ et N_- le nombre d'observations z_i positives et négatives. (A, B) sont estimés en maximisant la probabilité de bon classement sur $\mathcal{A}$, ce qui revient en introduisant le logarithme, à résoudre le problème de maximisation log-likelihood suivant

$$\max_{A, B} - \sum_{i=1}^n (t_i \log(P_i) + (1 - t_i) \log(1 - P_i)), \quad (5.3.34)$$

avec $P_i = P_{A, B}(h(\mathbf{x}_i))$. Une fois A et B déterminés, la fonction de mérite avec prise en compte des contraintes par la classification LS-SVM s'écrit comme

$$EI_C(\mathbf{x}) = EI(\mathbf{x})P_{A, B}(h(\mathbf{x})). \quad (5.3.35)$$

Ainsi, comme dans le cas de modélisation des contraintes (Section 5.2), l'EI est pénalisé avec une probabilité, qui est ici modélisée par $P_{A, B} \in [0, 1]$, et une nouvelle observation $\mathbf{x}$ admissible ou non, renseignera le GP de la fonction objective et le modèle de classification. Les sous-domaines non explorés par l'optimisation à une itération n auront une probabilité $P_{A, B} \approx 0.5$, ce qui est une bonne propriété de la méthode pour ne pas empêcher le processus d'y placer ultérieurement des observations.

La méthode de classification SVM qui n'utilise qu'une faible partie des observations (les vecteurs supports) peut, dans certains cas, donner une mauvaise prédiction de la probabilité d'admissibilité, par conséquent, elle peut avoir besoin d'un raffinement local de la classification, comme l'indique [Basudhar et al. \(2012\)](#) dans leurs travaux. Pour cela, les auteurs ajoutent n_p observations auxiliaires ($n_p = d$, la dimension du problème) à chaque itération de l'EGO autour du maximum de la fonction de mérite courant, suivant un critère de distance de probabilité de « mal-classification » à celui-ci. Ils modifient également l'équation (5.3.32), à partir d'un critère de distances pour corriger ces possibles erreurs, malheureusement cette modification entraîne des discontinuités dans la fonction de mérite. Dans ces travaux de thèse, la méthode EGO-LS-SVM est appliquée sans avoir à modifier l'équation (5.3.32), mais avec également une possibilité de raffinement local du classifieur LS-SVM. L'ajout de n_p observations n'est pas applicable car chaque évaluation de la fonction objective est considérée comme très coûteuse. Par conséquent, on adopte une stratégie séquentielle en déterminant $\mathbf{x}_{n+1}$ par les fonctions de mérite suivantes

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \Omega} EI_C(\mathbf{x}), \quad (5.3.36)$$

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \Omega} EI(\mathbf{x}) \quad \text{s.t.} \quad P_{A, B} \geq \rho, \quad (5.3.37)$$

$$\mathbf{x}_{n+1} = \arg \min_{\mathbf{x} \in \Omega} \hat{y}(\mathbf{x}) \quad \text{s.t.} \quad P_{A, B} \geq \rho, \quad (5.3.38)$$

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \Omega} [EI(\mathbf{x})P_{A, B}(1 - P_{A, B})]. \quad (5.3.39)$$

La première définition (5.3.36) est l'extension classique de l'EI qui favorise les points ayant une forte probabilité d'admissibilité. Dans (5.3.37), on maximise directement l'EI en respectant une probabilité minimale ρ qui vaut généralement 0.5. L'expression (5.3.38) minimise la fonction objective en prenant en compte une proba-

bilité minimale d'admissibilité ρ . Dans (5.3.39), on cherche à favoriser les zones incertaines de classification ($P_{A,B} \approx 0.5$) pour lesquelles l' AEI est maximale.

5.4 EGO sous contraintes évaluable et non évaluable

Les méthodes de modélisation et de classification pour la gestion des contraintes dans l'optimisation par méta-modèles Gaussiens ont été décrites dans les Sections précédentes. Cette Section 5.4 a pour objectif d'illustrer l'EGO sur un exemple simple d'optimisation composé de plusieurs contraintes d'inégalité et d'un domaine non évaluable défini lui aussi sous la forme d'une contrainte. Dans cet exemple, on utilisera simultanément la méthode de modélisation des contraintes pour les contraintes d'inégalité évaluable et la méthode EGO-LS-SVM pour le domaine non évaluable. Dans le cas où les deux méthodes de gestion des contraintes sont appliquées, on peut réécrire la fonction de mérite AEI sous sa forme contrainte par

$$AEI_C(\mathbf{x}) = AEI(\mathbf{x})P_C(\mathbf{x})P_{A,B}(h(\mathbf{x})), \quad (5.4.1)$$

où $P_C \in [0,1]$ est la probabilité d'admissibilité calculée avec les méta-modèles des m contraintes (voir (5.2.4) et (5.2.5)) et $P_{A,B} \in [0,1]$ est la probabilité d'évaluabilité estimée par la classification LS-SVM (voir (5.3.32)).

Problème d'optimisation. Afin d'illustrer le processus d'optimisation sous contraintes évaluable et non évaluable, on définit le problème d'optimisation à deux dimensions dans $\Omega = [-1,1]^2$ et contraint par $\mathbf{c}(\mathbf{x}) = (c_1(\mathbf{x}), \dots, c_4(\mathbf{x}))^T$ suivant

$$\begin{aligned} \min \quad & f(\mathbf{x}) = x_1^2 + x_2^2, \\ \text{s.t.} \quad & \begin{cases} c_1(\mathbf{x}) = 0.015 - (x_1 - 0.05)^2 - (x_2 - 0.05)^2 \leq 0 \\ c_2(\mathbf{x}) = 0.015 - (x_1 - 0.075)^2 - (x_2 + 0.05)^2 \leq 0 \\ c_3(\mathbf{x}) = 0.020 - (x_1 + 0.075)^2 - (x_2 - 0.05)^2 \leq 0 \\ c_4(\mathbf{x}) = 0.010 - x_1^2 - (x_2 + 0.05)^2 > 0 \implies f(\mathbf{x}) = \emptyset \end{cases} \end{aligned} \quad (5.4.2)$$

Le problème d'optimisation (5.4.2) est représenté sur la Figure 5.6 ci-dessous où la fonction $f(\mathbf{x})$ est tracée à l'aide des contours colorés, les zones blanches entourées par les courbes en couleurs correspondent aux domaines non admissibles $c_{1,2,3}(\mathbf{x}) > 0$ et la zone grisée entourée par la courbe violette est le domaine non évaluable $c_4(\mathbf{x}) > 0$. Comme l'indique la quatrième contrainte, qui définit ici le domaine non évaluable, on a l'implication $c_4(\mathbf{x}) > 0 \implies f(\mathbf{x}) = \emptyset$. Cela signifie qu'une observation $\mathbf{x}$ pour laquelle $c_4(\mathbf{x}) > 0$ est considérée comme non évaluable. Par conséquent aucune valeur de la fonction objective ou des contraintes $c_{1,2,3}$ n'est retournée à l'algorithme, si ce n'est que l'observation n'est pas évaluable. La donnée $\mathbf{x}$ est alors uniquement présente dans l'ensemble des observations utilisé pour la classification LS-SVM. Pour ce problème d'optimisation, on utilise le critère d'enrichissement auxiliaire séquentiel défini par (5.3.36) – (5.3.39). Un plan d'expérience par hypercubes latins composé de 15 observations initialise le calcul.

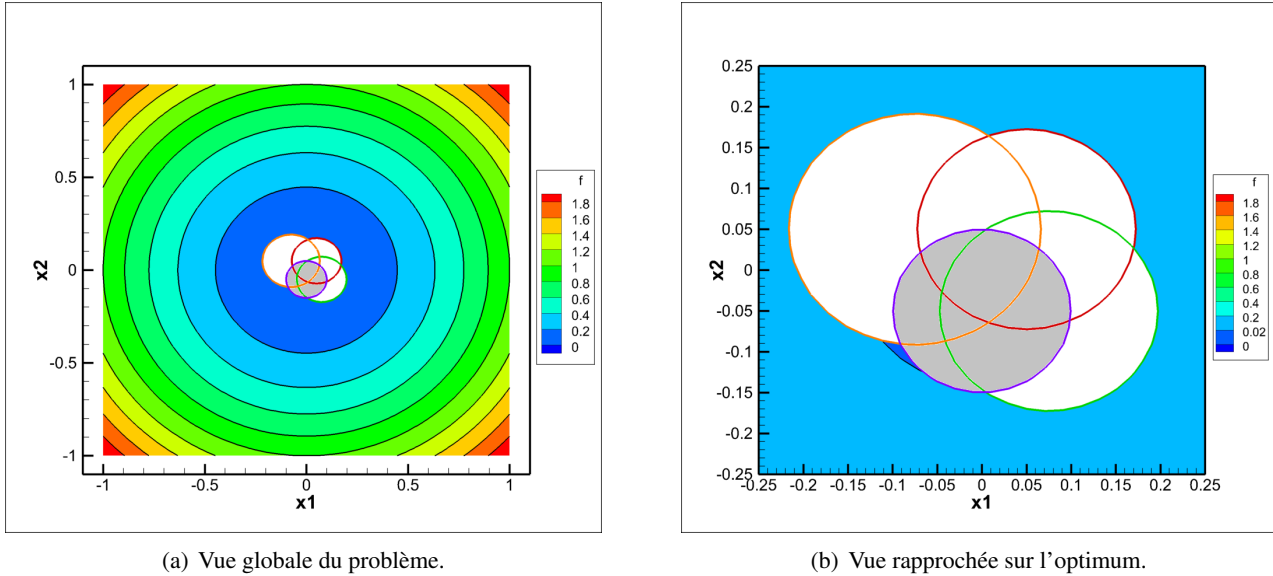


Figure 5.6 – Problème d'optimisation ($d = 2$) sous contraintes évaluable et non évaluable.

Les itérations 5, 25, 50 et 100 de la fonction de mérite AEI_C et des probabilités P_C et $P_{A,B}$ sont tracées en vue rapprochée autour de l'optimum sur la Figure 5.7. La courbe rouge continue correspond à l'iso-valeur $P_C(\mathbf{x}) = 0.5$ et celle en rouge pointillée à l'iso-valeur $P_{A,B}(\mathbf{x}) = 0.5$. Les courbes blanches continues sont les iso-valeurs $c_{1,2,3}(\mathbf{x}) = 0$ et la courbe blanche pointillée est l'iso-valeur $c_4(\mathbf{x}) = 0$. Les carrés blancs (classe évaluable) et rouges (classe non évaluable) sont les points explorés au cours du processus itératif.

A l'itération 5, on remarque que la fonction de mérite AEI_C est déjà pénalisée par les probabilités P_C et $P_{A,B}$. Cela provient principalement de la modélisation des contraintes qui sont ici « assez simples » à approcher et induisent $P_C \approx 0$, dès l'itération 5, dans la zone non admissible de Ω . La zone non évaluable est encore incertaine à cette itération car la classe C^- n'est composée que de deux observations.

On peut remarquer une diminution plus importante de $P_{A,B}$ à l'itération 25 dans la zone non évaluable, ce qui a un effet local direct sur l' AEI_C . La probabilité P_C a quant à elle peu évolué car elle n'est pas dépendante de l'ajout des nouveaux points non évaluable. De plus, les bruits inférés des méta-modèles des contraintes sont encore trop élevés (voir Figure 5.8(b) par la suite).

Après 50 itérations de l'EGO, on peut observer un changement de la « trajectoire » d'exploration des points évaluable (carrés blancs). Cela est dû à $P_{A,B}$ qui pénalise la fonction de mérite dans la zone non évaluable. On remarque que de nouveaux points ont été explorés dans cette zone afin d'enrichir le modèle de classification. Les méta-modèles des contraintes commencent maintenant à gagner en précision comme le montre l'évolution de l'iso-valeur $P_C(\mathbf{x}) = 0.5$, sur la Figure 5.7(g), qui se rapproche de la limite $c_{1,2,3}(\mathbf{x}) = 0$. On note aussi que la fonction de mérite est maintenant très faible excepté très localement autour de la solution.

A l'itération 100, les observations sont semblables à l'itération 50, mais celles-ci sont encore plus accentuées. En effet, les méta-modèles qui construisent P_C sont suffisamment raffinées (les bruits inférés sont faibles, voir Figure 5.8(b)) pour capter le « saut de classe » au niveau de la limite d'admissibilité. Le modèle de classi-

fication conduit également à une transition très brève de $P_{A,B}$ au niveau de la limite d'évaluabilité. Cela est dû à la répartition très locale des observations qui construisent le classifieur. Le maximum de la fonction de mérite AEI_C est négligeable et le processus d'optimisation itératif est convergé.

La Figure 5.8 contient des résultats sur les convergences de l'optimisation au cours des itérations de l'EGO. La Figure 5.8(a) montre les évolutions des distances à la solution du problème (5.4.2) au cours des itérations. Les carrés rouges correspondent à la distance $\|\mathbf{x}_{\min} - \mathbf{x}_{\min}^*\|_2$, où $\mathbf{x}_{\min}^*$ est le minimum du méta-modèle de f qui satisfait les contraintes $\mathbf{c}(\mathbf{x}) \leq \mathbf{0}$ par l'intermédiaire des probabilités P_C et $P_{A,B}$. Le minimum est déterminé à l'aide de la méthode d'optimisation sous contraintes (1+1)-CMA-ES de Arnold and Hansen (2012). Les triangles bleus donnent $|f_{\min} - f_{\min}^*|$, où $f_{\min}^* = f(\mathbf{x}_{\min}^*)$. On peut observer une diminution nette de la distance à la solution du problème (5.4.2) au cours des itérations de l'EGO.

Les bruits inférés du méta-modèle de f et des trois contraintes d'inégalité sont tracés en fonction des itérations de l'EGO sur la Figure 5.8(b). On observe une phase de forte diminution des bruits peu avant l'itération 50. C'est cette transition qui permet d'augmenter la précision locale des méta-modèles et ainsi de correctement capter les frontières d'admissibilités des contraintes.

Enfin, comme la classification est utilisée pour ce problème d'optimisation qui contient une zone non évaluable, on trace les évolutions des hyperparamètres du classifieur LS-SVM sur la Figure 5.8(c). Les valeurs des hyperparamètres oscillent durant les 40 premières itérations. La classification est en cours de construction durant cette phase et le modèle est sensible à l'ajout de nouveaux points non évaluables. Après approximativement 80 itérations, les hyperparamètres convergent et le modèle de classification est stabilisé.

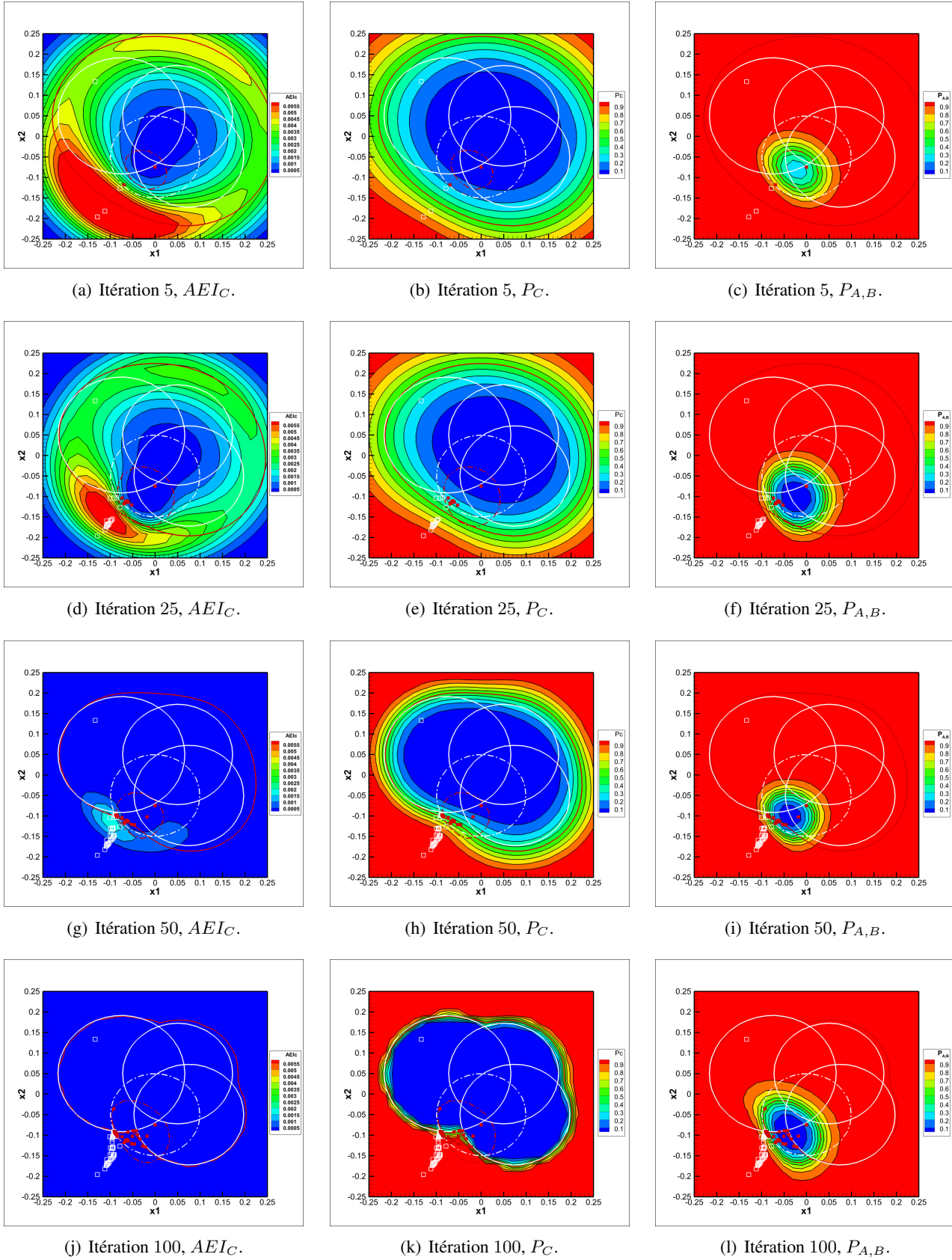
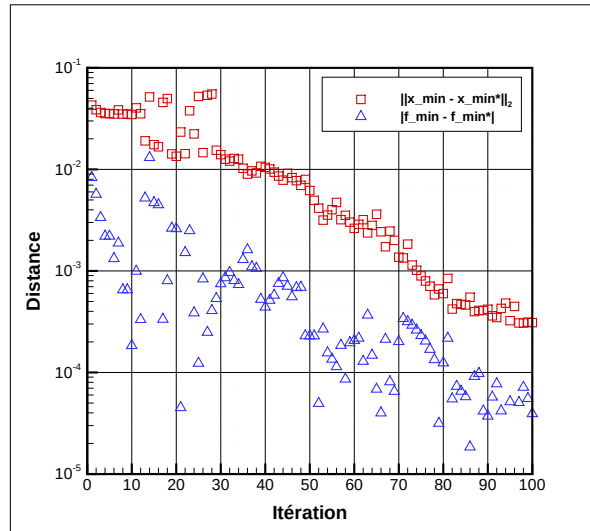
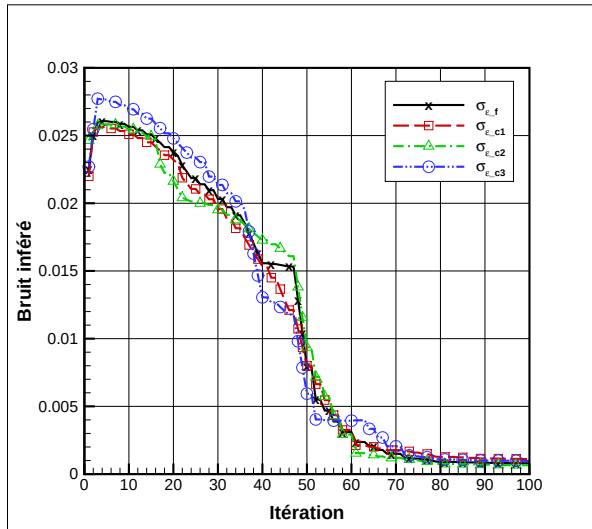


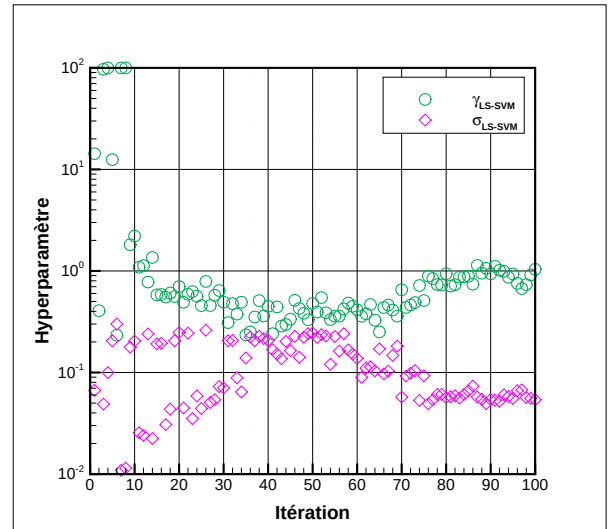
Figure 5.7 – Itérations 5, 25, 50 et 100 de la fonction de mérite AEI_C et des probabilités P_C et $P_{A,B}$ du problème d'optimisation (5.4.2).



(a) Distances $\|\mathbf{x}_{\min} - \mathbf{x}_{\min}^*\|_2$ et $|f_{\min} - f_{\min}^*|$ en fonction des itérations.



(b) Bruits inférés en fonction des itérations.



(c) Hyperparamètres du classifieur LS-SVM en fonction des itérations.

Figure 5.8 – Convergences du problème d'optimisation (5.4.2).

Optimisation multi-fidélité

Sommaire

6.1	Construction des méta-modèles	106
6.1.1	Méthode auto-régressive	106
6.1.2	Méthode récursive	108
6.2	EGO en multi-fidélité	110
6.2.1	Amélioration espérée augmentée en multi-fidélité	110
6.2.2	Points non évaluables	113

Un même phénomène physique est généralement modélisable par des modèles ayant des niveaux de précision ou de modélisation différents. Si l'on prend par exemple la mécanique des fluides numérique, la précision du calcul sera dépendante des choix de modèles (type, ordre, conditions aux limites, etc.), de la réalité géométrique (taille du domaine de calcul, simplification de l'objet étudié, etc.) ou bien encore de la discrétisation du maillage (nombre et taille de mailles, zones de raffinements, etc.). Ces différents choix de modélisation feront directement varier le coût global de la simulation. La plupart du temps, le coût de modélisation est connu car il est dépendant des choix de l'utilisateur qui met en place le calcul numérique. La précision de la modélisation est quant à elle généralement inconnue. En pratique, il est tout même possible de hiérarchiser qualitativement des niveaux de précision de manière *a priori*. Par exemple en CFD, on classe généralement les modèles « potentiel », « RANSE », « LES » et « DNS » dans un ordre de précision ascendant, sans pour autant avoir de valeur précise sur les différences de précisions entre ces modèles.

Un méta-modèle « multi-fidélité (MF) » se construit à partir d'observations évaluées par différents modèles ayant des niveaux de fidélité respectifs. Le coût de modélisation n'intervient pas dans la définition du modèle MF, mais le classement hiérarchique des niveaux doit lui être défini *a priori*. La construction de méta-modèles Gaussiens en multi-fidélité est assez récente puisque la première approche est la méthode « auto-régressive » de Kennedy and O'Hagan (2000). Le Gratiet (2013) en a ensuite proposé une formulation originale « récursive »

pendant sa thèse. L'application de l'EGO en multi-fidélité (Forrester et al., 2007) a quant à elle très peu été explorée jusqu'à aujourd'hui.

Dans la suite du Chapitre 6, la méthode de construction d'un méta-modèle Gaussien ayant des observations évaluées par différents niveaux de fidélité est tout d'abord détaillée dans la Section 6.1. L'optimisation par méta-modèles Gaussiens en multi-fidélité est ensuite présentée dans la Section 6.2. Dans l'EGO, la notion de fidélité doit intervenir dans la construction des modèles, mais également dans la fonction de mérite. Une méthode d'enrichissement des observations qui prend en compte les niveaux de fidélité est proposée dans cette thèse ; elle est décrite et illustrée avec un exemple analytique dans cette même Section 6.2.

6.1 Construction des méta-modèles

Comme dans le cas classique d'un méta-modèle Gaussien construit sur un ensemble d'observations évaluées par un même niveau de fidélité, on fait l'hypothèse *a priori* que les évaluations provenant de différents niveaux de fidélité sont aussi des réalisations de processus Gaussiens. L'objectif est de construire un méta-modèle Gaussien sur l'ensemble de toutes ces observations. Cette construction est présentée dans la Section 6.1.1 par la méthode « auto-régressive » de Kennedy and O'Hagan (2000) et par la formulation « récursive » de Le Gratiet and Garnier (2014) dans la Section 6.1.2.

6.1.1 Méthode auto-régressive

Considérons s niveaux de fidélité et des observations $y_s(\mathbf{x})$ que l'on souhaite modéliser avec des approximations de niveaux inférieurs $y_l(\mathbf{x})_{l=1,\dots,s-1}$. Les observations sont modélisées par des processus Gaussiens $Z^l(\mathbf{x})_{l=1,\dots,s}$, $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ et elles sont hiérarchisées par l'ordre de fidélité suivant : 1 (moins précise) à s (plus précise). Le modèle récursif s'écrit comme

$$\begin{cases} Z^l(\mathbf{x}) = \rho_{l-1} Z^{l-1}(\mathbf{x}) + \delta^l(\mathbf{x}) \\ Z^{l-1}(\mathbf{x}) \perp \delta^l(\mathbf{x}) \end{cases}, \quad \forall l = 2, \dots, s. \quad (6.1.1)$$

Dans (6.1.1), ρ_{l-1} est un paramètre de régression et $\perp$ signifie que les processus Gaussiens $\delta^l(\mathbf{x})$ et $Z^{l-1,\dots,1}(\mathbf{x})$ sont indépendants. On définit également $\mathbf{Y}_s = \mathbf{Y}_{n_1,\dots,n_s}^{1,\dots,s}$ et $\mathbf{Y}_{n_l}^l = y_l(\mathbf{X}_{n_l}^l)$, avec $\mathbf{X}_{n_l}^l$ l'ensemble des n_l observations données par le niveau de fidélité l . L'expression (6.1.1) qui provient de (Kennedy and O'Hagan, 2000) est issue de l'hypothèse de Markov suivante

$$\text{cov} \left(Z^l(\mathbf{x}), Z^{l-1}(\mathbf{x}') | Z^{l-1}(\mathbf{x}) \right) = 0, \quad \forall \mathbf{x} \neq \mathbf{x}'. \quad (6.1.2)$$

Autrement dit, on ne peut rien apprendre de plus sur $Z^l(\mathbf{x})$ à partir des observations $Z^{l-1}(\mathbf{x}')$ si l'on connaît $Z^{l-1}(\mathbf{x})$, pour $\mathbf{x} \neq \mathbf{x}'$.

Soit $\mathcal{Z}$, le vecteur Gaussien contenant les valeurs des processus $Z^l(\mathbf{x})_{l=1,\dots,s}$ aux points d'observations

$\mathbf{X}_{n_1, \dots, n_s}^{1, \dots, s}$. Le méta-modèle multi-fidélité est alors donné par

$$[Z^s(\mathbf{x}) | \mathcal{Z}, \mathbf{Y}_s] \sim \mathcal{N}(\hat{y}_s(\mathbf{x}), \hat{\sigma}_{y_s}^2(\mathbf{x})), \quad (6.1.3)$$

où la moyenne du méta-modèle de plus haute fidélité s'écrit

$$\hat{y}_s(\mathbf{x}) = \mathbf{k}_s^T(\mathbf{x}) \left(\mathbf{C}_s(\underline{\Theta}_s) + \sigma_{\epsilon_s}^2 \right)^{-1} \mathbf{Y}_s, \quad (6.1.4)$$

avec comme estimation d'erreur de modèle

$$\hat{\sigma}_{y_s}^2(\mathbf{x}) = \kappa_s(\mathbf{x}) + \sigma_{\epsilon_s}^2 - \mathbf{k}_s^T(\mathbf{x}) \left(\mathbf{C}_s(\underline{\Theta}_s) + \sigma_{\epsilon_s}^2 \right)^{-1} \mathbf{k}_s(\mathbf{x}). \quad (6.1.5)$$

La matrice $\mathbf{C}_s \in \mathbb{R}^{\sum_1^s n_l \times \sum_1^s n_l}$ correspond à la matrice de covariance du vecteur $\mathcal{Z}$ et le vecteur de covariance entre $Z^s(\mathbf{x})$ et $\mathcal{Z}$ est donné par $\mathbf{k}_s(\mathbf{x})$. La matrice diagonale $\sigma_{\epsilon_s}^2$ contient l'ensemble des bruits inférés $\sigma_{\epsilon_{1, \dots, s}}^2$ et $\kappa_s(\mathbf{x})$ est la variance de $Z^s(\mathbf{x})$. L'ensemble des hyperparamètres est contenu dans la matrice $\underline{\Theta}_s$ de taille $n_{hyp} \times s$, avec n_{hyp} le nombre d'hyperparamètres par niveau de modèle. La covariance entre $Z^l(\mathbf{x})$ et $Z^{l'}(\mathbf{x}')$ se calcule par

$$\text{cov}(Z^l(\mathbf{x}), Z^{l'}(\mathbf{x}')) = \left(\prod_{i=l'}^{l-1} \rho_i \right) \text{cov}(Z^{l'}(\mathbf{x}), Z^{l'}(\mathbf{x}')), \quad \forall l > l', \quad (6.1.6)$$

avec

$$\text{cov}(Z^{l'}(\mathbf{x}), Z^{l'}(\mathbf{x}')) = \sum_{j=1}^{l'} \left(\prod_{i=j}^{l'-1} \rho_i^2 \right) C_{f_j}(\mathbf{x}, \mathbf{x}'; \underline{\Theta}_j). \quad (6.1.7)$$

C_{f_j} est une fonction de covariance choisie dont le vecteur des hyperparamètres $\underline{\Theta}_j$ est à déterminer. Les paramètres de régression ρ_i peuvent aussi être dépendant de $\mathbf{x}$ comme proposé dans les travaux de (Qian and Wu, 2008; Le Gratiet and Garnier, 2014).

Les hyperparamètres $(\rho_{l-1}, \underline{\Theta}_l, \sigma_{\epsilon_l}^2)$ peuvent être estimés de manière indépendante de $(\underline{\Theta}_{l-1}, \sigma_{\epsilon_{l-1}}^2)$ grâce à l'hypothèse de Markov (Kennedy and O'Hagan, 2000). On maximise le log-marginal likelihood (Rasmussen and Williams, 2006) de chaque niveau l (voir (6.1.8)) en commençant par le plus faible niveau 1 jusqu'au niveau final s . Lorsque les paramètres d'un niveau de fidélité l sont déterminés, ils sont alors assumés comme fixés.

$$\begin{aligned} \mathcal{L}_l(\rho_{l-1}, \underline{\Theta}_l, \sigma_{\epsilon_l}^2) &= -\frac{n_l}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{C}_l(\underline{\Theta}_l) + \sigma_{\epsilon_l}^2 \mathbf{I}| - \frac{1}{2} \mathbf{U}_l^T \left(\mathbf{C}_l(\underline{\Theta}_l) + \sigma_{\epsilon_l}^2 \mathbf{I} \right)^{-1} \mathbf{U}_l, \\ \mathbf{U}_l &= \begin{cases} \mathbf{U}_1 = \mathbf{Y}_{n_1}^1 \\ \mathbf{U}_{l>1} = \mathbf{Y}_{n_l}^l - \rho_{l-1} y_{l-1}(\mathbf{X}_{n_l}^l) \end{cases}, \quad l = 1, \dots, s. \end{aligned} \quad (6.1.8)$$

Le vecteur $y_{l-1}(\mathbf{X}_{n_l}^l)$ représente les évaluations du modèle $l-1$ aux points de $\mathbf{X}_{n_l}^l$ du modèle l . En effet, dans la plupart des travaux sur la multi-fidélité (Kennedy and O'Hagan, 2000; Forrester et al., 2007; Le Gratiet and

Cannamela, 2012; Le Gratiet and Garnier, 2014), les auteurs considèrent des observations imbriquées que l'on écrit comme $\mathbf{X}_{n_l}^l \subseteq \mathbf{X}_{n_{l-1}}^{l-1}$, $\forall l = 2, \dots, s$.

6.1.2 Méthode réursive

La méthode auto-régressive présentée dans la Section 6.1.1 a l'inconvénient de pouvoir devenir très coûteuse lorsque le nombre total d'observations est important. En effet, le problème de conditionnement de la matrice de covariance $\mathbf{C}_s$ est d'autant plus présent lorsque $\sum_1^s n_l$ est grand. De plus, les équations (6.1.4) et (6.1.5) fournissent uniquement les estimations du modèle de plus haute fidélité s . Ceci est pénalisant lorsque l'on souhaite faire de l'EGO en multi-fidélité car on ne peut mesurer l'impact d'une évaluation faite par un modèle intermédiaire l sur le méta-modèle multi-fidélité complet de niveau s .

La formulation réursive proposée par Le Gratiet and Garnier (2014) qui est présentée dans cette Section décompose la construction du méta-modèle multi-fidélité à s niveaux en s constructions de méta-modèles indépendants. L'inversion unique de $\mathbf{C}_s$ se transforme alors en s inversions de sous-matrices prises à chaque niveau de fidélité. Cette approche a également l'avantage de rendre les estimations des moyennes et variances possibles à chaque niveau l . Ceci se fait à partir des $l - 1$ niveaux, d'où l'expression « réursive » dans l'appellation de la méthode. Le Gratiet and Garnier (2014) considère des ensembles d'observations imbriquées et la modélisation pour $l = 2, \dots, s$ est donnée par

$$\begin{cases} Z^l(\mathbf{x}) = \rho_{l-1} \hat{Z}^{l-1}(\mathbf{x}) + \delta^l(\mathbf{x}), \\ \hat{Z}^{l-1}(\mathbf{x}) \perp \delta^l(\mathbf{x}), \\ \hat{Z}^{l-1}(\mathbf{x}) \sim [Z^{l-1}(\mathbf{x}) | \mathbf{Y}_{n_{l-1}}^{l-1}]. \end{cases} \quad (6.1.9)$$

On note ici la différence avec la formulation (6.1.1) au niveau de l'expression de $Z^l(\mathbf{x})$. Le processus Gaussien de fidélité l est ici exprimé en fonction du processus $Z^{l-1}(\mathbf{x})$ conditionné par les observations $\mathbf{Y}_{n_{l-1}}^{l-1}$. On a maintenant

$$[Z^l(\mathbf{x}) | \mathbf{Y}_{n_l}^l] \sim \mathcal{N}(\hat{y}_l(\mathbf{x}), \hat{\sigma}_{y_l}^2(\mathbf{x})), \quad (6.1.10)$$

où les expressions des deux estimateurs du processus Gaussien sont

$$\hat{y}_l(\mathbf{x}) = \rho_{l-1} \hat{y}_{l-1}(\mathbf{x}) + \mathbf{k}_l^T(\mathbf{x}) \left(\mathbf{C}_l(\boldsymbol{\Theta}_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l \right)^{-1} \left(\mathbf{Y}_{n_l}^l - \rho_{l-1} y_{l-1}(\mathbf{X}_{n_l}^l) \right), \quad (6.1.11)$$

$$\hat{\sigma}_{y_l}^2(\mathbf{x}) = \rho_{l-1}^2 \hat{\sigma}_{y_{l-1}}^2(\mathbf{x}) + \kappa_l(\mathbf{x}) + \sigma_{\epsilon_l}^2 - \mathbf{k}_l^T(\mathbf{x}) \left(\mathbf{C}_l(\boldsymbol{\Theta}_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l \right)^{-1} \mathbf{k}_l(\mathbf{x}). \quad (6.1.12)$$

Dans cette thèse, on se place dans le cadre général d'observations non imbriquées, on substitue alors $y_{l-1}(\mathbf{x})$ par l'estimation de la moyenne du méta-modèle $\hat{y}_{l-1}(\mathbf{x})$ dans la phase de détermination des hyperparamètres (6.1.8) et dans la prédiction (6.1.11). On verra dans la suite du manuscrit, au Chapitre 10, que cette substitution est légitime.

Comme pour la méthode auto-régressive, les hyperparamètres $\boldsymbol{\Theta}_l$, pour $l = 1, \dots, s$, doivent être déterminés. On applique pour cela la méthode de maximisation du log-likelihood (6.1.8), avec cependant des matrices

et vecteurs de dimensions bien inférieures. Par exemple, la matrice de covariance $\mathbf{C}_l(\boldsymbol{\Theta}_l)$ est ici de taille $[n_l \times n_l]$ et est composée des entrées $C_{l,i,j} = C_{F_l}(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\Theta}_l)_{\mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}_{n_l}^l}$, où l'on a noté

$$C_{F_l}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\Theta}_l) = \sum_{j=1}^l \left(\prod_{i=j}^{l-1} \rho_i^2 \right) C_{f_l}(\mathbf{x}, \mathbf{x}'; \boldsymbol{\Theta}_l). \quad (6.1.13)$$

Les matrices $(\mathbf{C}_l(\boldsymbol{\Theta}_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1}$ sont sauvées au cours du processus de construction du méta-modèle car elles servent ensuite dans les calculs des estimateurs (6.1.11) et (6.1.12).

Un exemple de méta-modèle unidimensionnel, $x \in [0, 1]$, ayant deux niveaux de fidélité est construit à partir de la fonction de [Forrester et al. \(2007\)](#) définie par

$$f_e(x) = (6x - 2)^2 \sin(12x - 4), \quad (6.1.14a)$$

$$f_c(x) = \frac{1}{2}f_e(x) + 10\left(x - \frac{1}{2}\right) + 5, \quad (6.1.14b)$$

est présenté sur la Figure 6.1. Dans (6.1.14), $f_e(x)$ est considérée comme la fonction coûteuse et de haute fidélité que l'on cherche à modéliser par un méta-modèle multi-fidélité. $f_c(x)$ est une estimation de $f_e(x)$ ayant un coût d'évaluation faible. Dans la Figure 6.1, $f_e(x)$ et $f_c(x)$ sont respectivement tracées par les courbes rouge et bleue. Les observations de ces fonctions sont représentées par des carrés rouges et cercles bleus. La courbe discontinue noire est la prédiction $\hat{y}_{l=e}(x)$ du méta-modèle multi-fidélité et celle en vert discontinue-pointillée est la prédiction $\hat{y}_e(x)$ d'un méta-modèle construit uniquement sur les observations de $f_e(x)$ (méta-modèle simple fidélité). On peut ainsi apprécier l'apport de la modélisation multi-fidélité qui fournit une prédiction $\hat{y}_{l=e}(x)$ très proche de $f_e(x)$ et cela pour les deux minima local et global. Le méta-modèle simple fidélité est quant à lui très loin de $f_e(x)$ dans les zones non explorées.

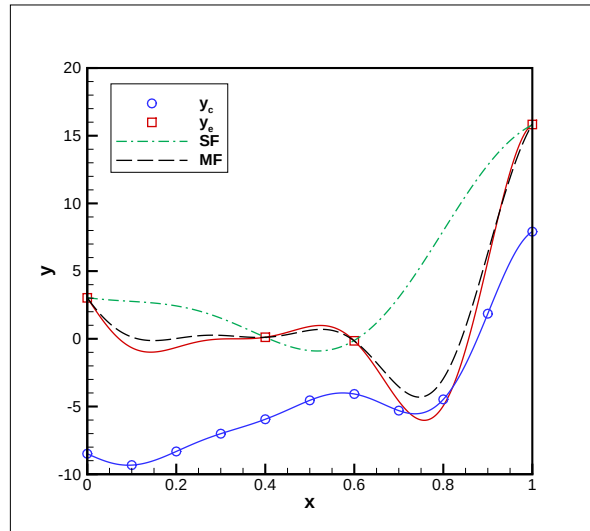


Figure 6.1 – Fonction unidimensionnelle ([Forrester et al., 2007](#)) illustrant un méta-modèle multi-fidélité.

6.2 EGO en multi-fidélité

La formulation de l'EGO en multi-fidélité a été initialement introduite par [Forrester et al. \(2007\)](#). Dans ces travaux, les auteurs considèrent des méta-modèles ayant deux niveaux de fidélité, qu'ils appellent « co-kriging », et ils appliquent la méthode de construction originale de [Kennedy and O'Hagan \(2000\)](#). Ils adaptent alors la fonction de mérite EI (4.3.24) au cas multi-fidélité en remplaçant les prédictions $\hat{y}(\mathbf{x})$ et $\hat{\sigma}_y(\mathbf{x})$ par les prédictions données par le méta-modèle multi-fidélité, (6.1.4) et (6.1.5). Plus récemment, [Elsayed \(2015\)](#) ou [Bonfiglio et al. \(2018\)](#) ont réalisé des simulations numériques incluant de l'optimisation en multi-fidélité. Cependant, comme dans ([Forrester et al., 2007](#)), la question du choix du niveau de fidélité l , pour lequel $f_l(\mathbf{x}_{n_l+1})$ doit être évaluée, n'est pas considérée dans la définition de la fonction de mérite. Dans les travaux de [Le Gratiet and Cannamela \(2012\)](#), le niveau l est inclus dans une fonction de mérite qui repose sur l'incertitude (variance) du méta-modèle de plus haute fidélité (niveau s). Les auteurs utilisent des ensembles d'observations imbriquées et ils définissent une approximation de l'« integrated variance predictor (IMSE) » pour déterminer la potentielle réduction de variance dans le modèle de niveau s qui serait due à l'évaluation des l premiers niveaux. Ils équilibrent le poids de cette réduction de variance par un ratio de coût d'évaluation entre les niveaux l et $l - 1$. Bien que le niveau de fidélité l soit considéré dans ([Le Gratiet and Cannamela, 2012](#)), cette dernière approche comporte des désavantages. En effet, elle requiert l'utilisation d'observations imbriquées qui n'est pas une formulation efficace dans les cas où les coûts d'évaluations des niveaux intermédiaires sont non négligeables. De plus, de par sa formulation imbriquée, cette méthode favorise les évaluations des niveaux de faible fidélité même si ceux-ci apportent une plus faible réduction de variance que l'évaluation unique d'un niveau supérieur à l dans le méta-modèle complet de niveau s . Enfin, comme présenté dans ([Picheny et al., 2013](#)), il est connu que l'utilisation d'une fonction de mérite basée uniquement sur l'incertitude du méta-modèle est moins efficace que les critères basés sur l'amélioration espérée.

Une nouvelle définition de fonction de mérite pour l'application de l'EGO en multi-fidélité est développée dans cette thèse. En particulier, on considère des ensembles d'observations non imbriquées et une fonction de mérite étendue basée sur le critère d'amélioration espérée dans le niveau de plus haute fidélité s . Cette nouvelle fonction de mérite prend en compte le niveau de fidélité l pour lequel $f_l(\mathbf{x}_{n_l+1})$ doit être évaluée à partir d'un compromis entre le coût d'évaluation de ce niveau et la potentielle réduction d'incertitude dans le méta-modèle de niveau s . En plus de la réduction du coût numérique de l'EGO, la formulation non imbriquée permet d'éviter des évaluations de potentiels niveaux qui n'apporteraient pas de réduction d'incertitude dans le modèle s . La méthode est décrite dans la Section 6.2.1 ci-après.

6.2.1 Amélioration espérée augmentée en multi-fidélité

L'objectif de l'EGO formulé en multi-fidélité est d'optimiser une fonction $f_s(\mathbf{x})$ en utilisant des évaluations de modèles $f_{l < s}(\mathbf{x})$, approchant $f_s(\mathbf{x})$, ayant des coûts d'évaluation inférieurs au modèle s , et cela, pour un coût total d'optimisation qui soit le plus faible possible. On souhaite utiliser une fonction de mérite qui dépend de la localisation $\mathbf{x} \in \Omega$ pour mener à bien l'EGO comme dans le cas simple fidélité, mais également du niveau $l \in \{1, \dots, s\}$ pour lequel $f_l(\mathbf{x}_{n_l+1})$ est évaluée. On considère des ensembles d'observations non imbriquées

pour ne pas être dépendants des $l - 1$ premiers niveaux lorsque l'on évalue $f_l(\mathbf{x}_{n_l+1})$. On définit ensuite le coût d'évaluation β_l associé à chaque niveau $l = 1, \dots, s$ et le ratio $\frac{\beta_s}{\beta_l}$. Ce ratio de coûts permet de favoriser les évaluations des niveaux de faible fidélité dans la fonction de mérite. On a ensuite besoin de contrebalancer le poids $\frac{\beta_s}{\beta_l}$ pour favoriser les évaluations des « meilleurs » niveaux qui apportent une réduction de l'incertitude (variance) dans les zones d'améliorations espérées maximales du modèle s . Pour cela, on définit la prédiction de variance du modèle de plus haute fidélité s en $\mathbf{x}$, en considérant que $\mathbf{x}$ a été évalué par le niveau q comme

$$\begin{aligned} \hat{\sigma}_{y_s}^2|_q(\mathbf{x}) &= \rho_{s-1}^2 \hat{\sigma}_{y_{s-1}}^2(\mathbf{x}) + \kappa_s(\mathbf{x}) + \sigma_{\epsilon_s}^2 - \mathbf{k}_s^T(\mathbf{x}) (\mathbf{C}_s(\Theta_s) + \sigma_{\epsilon_s}^2 \mathbf{I}_s)^{-1} \mathbf{k}_s(\mathbf{x}), \\ \hat{\sigma}_{y_l}^2(\mathbf{x}) &= \begin{cases} \rho_{l-1}^2 \hat{\sigma}_{y_{l-1}}^2(\mathbf{x}) + \kappa_l(\mathbf{x}) + \sigma_{\epsilon_l}^2 - \mathbf{k}_l^T(\mathbf{x}) (\mathbf{C}_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1} \mathbf{k}_l(\mathbf{x}), & \forall l \neq q, \\ \rho_{q-1}^2 \hat{\sigma}_{y_{q-1}}^2(\mathbf{x}) + \kappa_q(\mathbf{x}) + \sigma_{\epsilon_q}^2 \\ \quad - [\mathbf{k}_q^T(\mathbf{x}) \quad \kappa_q(\mathbf{x})] \begin{bmatrix} \mathbf{C}_q(\Theta_q) + \sigma_{\epsilon_q}^2 \mathbf{I}_q & \mathbf{k}_q(\mathbf{x}) \\ \mathbf{k}_q^T(\mathbf{x}) & C_{F_q}(\mathbf{x}, \mathbf{x}; \Theta_q) + \sigma_{\epsilon_q}^2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{k}_q(\mathbf{x}) \\ \kappa_q(\mathbf{x}) \end{bmatrix}, & l = q. \end{cases} \end{aligned} \quad (6.2.1)$$

Plus précisément, la variance $\hat{\sigma}_{y_s}^2|_q(\mathbf{x})$ du méta-modèle s se calcule comme dans le cas général $\forall l = 1, \dots, s$, si $l \neq q$ avec la formulation réursive (6.1.12), et si $l = q$, la dernière équation dans (6.2.1) est appliquée. Bien entendu, on considère dans (6.2.1) que les hyperparamètres des méta-modèles restent inchangés après l'ajout de l'observation $\mathbf{x}$ de niveau q . Le calcul de (6.2.1) requiert une inversion de la matrice étendue par l'observation $\mathbf{x}$. Ceci est réalisé en utilisant une méthode de matrice partitionnée pour éviter d'augmenter le coût numérique de l'EGO par le Lemme suivant

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}. \quad (6.2.2)$$

Dans (6.2.2), on a $A^{-1} = (\mathbf{C}_q(\Theta_q) + \sigma_{\epsilon_q}^2 \mathbf{I}_q)^{-1}$ qui a déjà été calculée $\forall q = 1, \dots, s$ au cours de la phase de construction des méta-modèles. Finalement, la formulation originale de l'AEI (Huang et al., 2006) devient dans le cas étendu à l'EGO multi-fidélité

$$AEI_l(\mathbf{x}) = EI_s(\mathbf{x}) \left(1 - \frac{\sigma_{\epsilon_l}}{\sqrt{\hat{\sigma}_{y_s}^2(\mathbf{x}) + \sigma_{\epsilon_l}^2}} \right) \frac{\beta_s}{\beta_l} \max \left(\left(1 - \frac{\hat{\sigma}_{y_s}^2|_l(\mathbf{x})}{\hat{\sigma}_{y_s}^2(\mathbf{x})} \right), 0 \right), \quad (6.2.3)$$

où l'amélioration espérée en multi-fidélité est définie par

$$EI_s(\mathbf{x}) = \hat{\sigma}_{y_s}(\mathbf{x}) [u_s(\mathbf{x})\Phi(u_s(\mathbf{x})) + \phi(u_s(\mathbf{x}))], \quad u_s(\mathbf{x}) = \frac{\hat{y}_s(\mathbf{x}_{*,n_1,\dots,n_s}) - \hat{y}_s(\mathbf{x})}{\hat{\sigma}_{y_s}(\mathbf{x})}, \quad (6.2.4)$$

et $\mathbf{x}_{*,n_1,\dots,n_s} \in \mathbf{X}_{n_1,\dots,n_s}^{1,\dots,s}$ est l'« effective best solution » de toutes les observations.

Contrairement au cas de simple fidélité, on a maintenant une fonction de mérite qui est dépendante du niveau de fidélité l . Cela implique qu'une optimisation de (6.2.3) doit être réalisée à chaque niveau $l = 1, \dots, s$ pour déterminer la solution $\mathbf{x}_{n_l+1} = \arg \max_{l \in \{1,\dots,s\}, \mathbf{x} \in \Omega} AEI_l(\mathbf{x})$. Finalement, $f_l(\mathbf{x}_{n_l+1})$ est évaluée et $\mathbf{x}_{n_l+1}$ est ajouté dans $\mathbf{X}_{n_l}^l$. En appliquant $n_l \leftarrow n_l + 1$, une nouvelle itération de l'EGO en multi-fidélité peut être

effectuée pour déterminer la prochaine observation de niveau l à évaluer.

Une illustration de l'EGO en multi-fidélité que l'on applique aux fonctions de (6.1.14) est proposée sur la Figure 6.2. Les coûts utilisés pour cet exemple sont $\beta_1 = 1$ et $\beta_2 = 10$. Sur chaque Figure, qui correspond à une itération de l'EGO, on trace les observations $f_e(x)$ (carrés rouges), $f_e(x)$ (cercles bleus) et la prédiction $\hat{y}_{l=2}(x)$ (courbe noire) du méta-modèle multi-fidélité donnée par (6.1.11), qui se réfèrent à l'ordonnée de gauche. Les deux fonctions de mérite $AEI_2(x)$ (courbe rouge discontinue) et $AEI_1(x)$ (courbe bleue discontinue-pontillée) sont également tracées et elles se réfèrent à l'ordonnée de droite. On remarque que durant les deux premières itérations de l'EGO, la formulation multi-fidélité favorise des observations du premier niveau. Cela permet d'améliorer la prédiction du méta-modèle dans la zone de l'optimum global de $f_e(x)$ et ainsi d'augmenter l'amélioration espérée du niveau de haute fidélité. A l'itération 3 et 4, c'est le modèle de haute fidélité qui est choisi pour affiner le méta-modèle dans la zone de l'optimum global. A l'itération 10 (Figure 6.2(f)), la fonction $f_e(x)$ est également reportée. On peut ainsi observer que la prédiction $\hat{y}_{l=2}(x)$ est proche de $f_e(x)$, particulièrement au niveau de l'optimum global, mais également sur l'ensemble du domaine de variabilité $\Omega = [0,1]$.

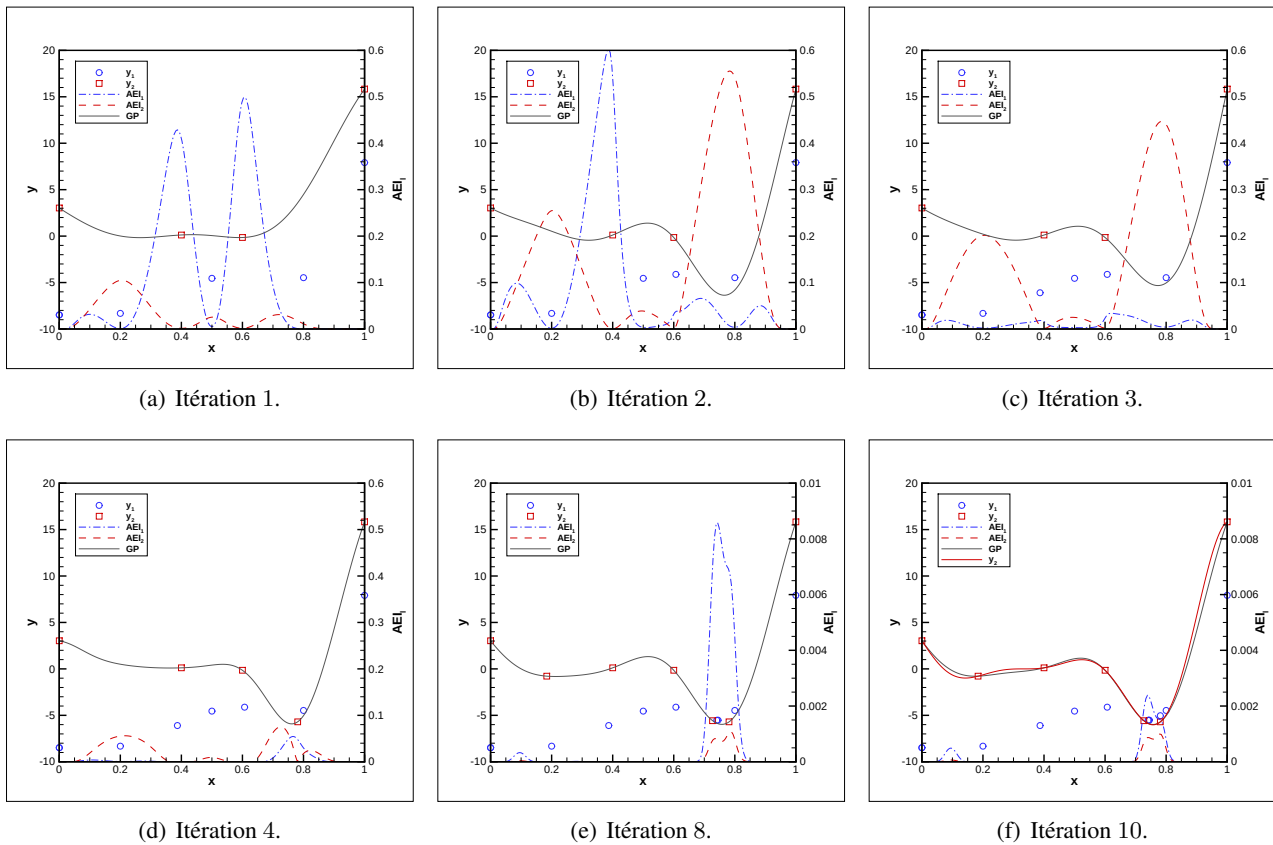


Figure 6.2 – Illustration de l'EGO en multi-fidélité sur le cas unidimensionnel défini par (6.1.14).

6.2.2 Points non évaluables

La présence de points non évaluables, que l'on a décrit à la Section 5.1, est également possible lorsque l'on réalise une optimisation par méta-modèles en multi-fidélité. Il est alors nécessaire de construire un modèle de classification binaire formulé en multi-fidélité afin d'éviter les possibles zones non évaluables du domaine Ω . Pour cela, on propose ici d'utiliser un modèle de classification construit par des processus Gaussiens (voir Section 4.3.2) et formulé en multi-fidélité. Ce classifieur permet de calculer la probabilité qu'une observation $\mathbf{x}$ appartienne à la classe évaluable et cela pour chaque niveau de fidélité l . On peut alors étendre la fonction de mérite multi-fidélité (6.2.3) aux applications contenant des observations non évaluables avec

$$AEI_{l_C}(\mathbf{x}) = AEI_l(\mathbf{x})\Phi_l(\mathbf{x}), \quad (6.2.5)$$

où $AEI_l(\mathbf{x})$ est donnée par (6.2.3) et $\Phi_l(\mathbf{x})$ est la probabilité multi-fidélité d'appartenance à la classe évaluable de niveau l , que l'on note C_l^+ . La question subsidiaire concerne la détermination de $\Phi_l(\mathbf{x})$. On considère tout d'abord la construction de modèles de classification par processus Gaussiens indépendants à chaque niveau l , $q(f_l|\mathcal{D}_l, \Theta_l, \mathbf{x}) = \mathcal{N}(f_l|\mu_l, \sigma_{\mu_l}^2)$. À partir des prédictions « locales », $\mu_l(\mathbf{x})$ et $\sigma_{\mu_l}^2(\mathbf{x})$, on définit $\Gamma_{n_l}^l = \gamma_l(\mathbf{X}_{n_l}^l)$, $\forall l = 1, \dots, s$ où γ_l s'écrit

$$\gamma_l(\mathbf{x}) = \frac{\mu_l(\mathbf{x})}{\sqrt{1 + \sigma_{\mu_l}^2(\mathbf{x})}}. \quad (6.2.6)$$

Les prédictions $\Gamma_{n_l}^l$ calculées par les classifieurs par processus Gaussiens sont ensuite modélisées par des processus Gaussiens $I^l(\mathbf{x})_{l=1,\dots,s}$ et le méta-modèle multi-fidélité pour $l = 2, \dots, s$ est donné par

$$[I^l(\mathbf{x})|\Gamma_{n_l}^l] \sim \mathcal{N}(\hat{\gamma}_l(\mathbf{x}), \hat{\sigma}_{\gamma_l}^2(\mathbf{x})), \quad (6.2.7)$$

où les prédictions $\hat{\gamma}_l(\mathbf{x})$ et $\hat{\sigma}_{\gamma_l}^2(\mathbf{x})$ sont calculées respectivement par (6.1.11) et (6.1.12). C'est cette dernière équation (6.2.7) qui permet le passage de s modèles de classifications indépendants à $s - 1$ modèles de classification formulés en multi-fidélité. Finalement, la prédiction de la probabilité multi-fidélité d'appartenance à la classe évaluable C^+ (et non pas C_l^+) se calcule comme

$$\Phi_l(\mathbf{x}) = \Phi_l(\hat{\gamma}_l(\mathbf{x})), \quad \forall l = 2, \dots, s. \quad (6.2.8)$$

Afin d'illustrer cette prise en compte de points non évaluables dans l'EGO en multi-fidélité, on reprend l'application de la fonction Branin (4.3.28) en considérant trois niveaux de fidélité et des zones non évaluables. La méthode « Expectation Propagation » (Minka, 2001) est utilisée pour déterminer les paramètres μ_n et $\mathbf{A}$ dans la prédiction de la classification à chaque niveau l . Les coûts des trois niveaux sont $\beta_1 = 1$, $\beta_2 = 5$ et $\beta_3 = 10$. Le problème d'optimisation est écrit ci-après en (6.2.9), où l'on peut voir qu'un offset de 10 ($s - l$) a été ajouté à la réponse de la fonction $f_l(\mathbf{x})$. Les zones non évaluables sont exprimées en terme de contraintes d'inégalité, mais leur application correspond bien à un critère d'évaluabilité dans le sens où aucune valeur n'est

retournée à l'EGO si l'une des contraintes n'est pas vérifiée. Dans (6.2.9), la première contrainte rend la zone $\|\mathbf{x} - \mathbf{x}_{\text{opt}_1}\|_2 \leq 2$ non évaluable par les trois niveaux de fidélité. La seconde contrainte concerne uniquement le niveau de plus haute fidélité, pour lequel $\|\mathbf{x} - \mathbf{x}_{\text{opt}_2}\|_2 \leq 2$ n'est pas évaluable. Enfin, la dernière contrainte dit que $\|\mathbf{x} - \mathbf{x}_{\text{opt}_3}\|_2 \leq 2$ n'est pas évaluable par les deux premiers niveaux uniquement. Pour cet exemple illustratif, on utilise uniquement la fonction de mérite (6.2.5) et non pas une approche séquentielle comme cela a été fait dans la résolution du problème (5.4.2).

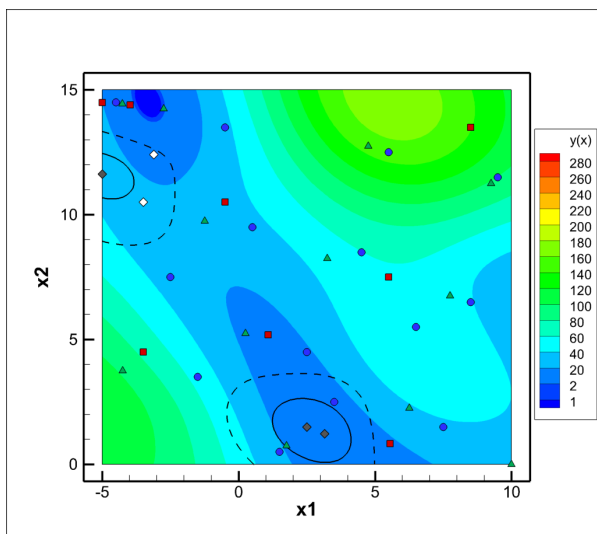
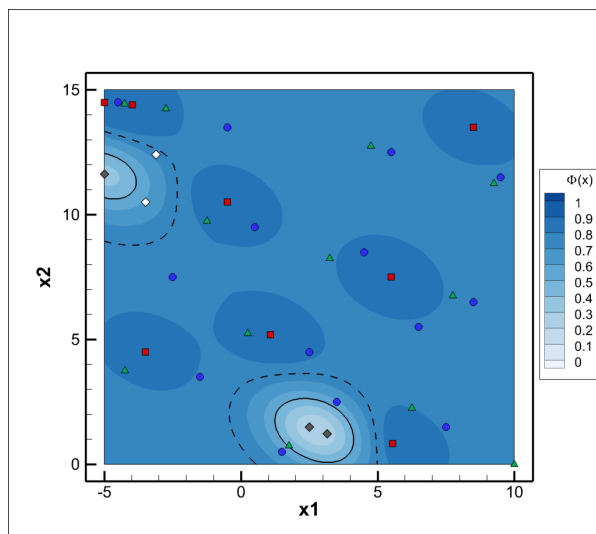
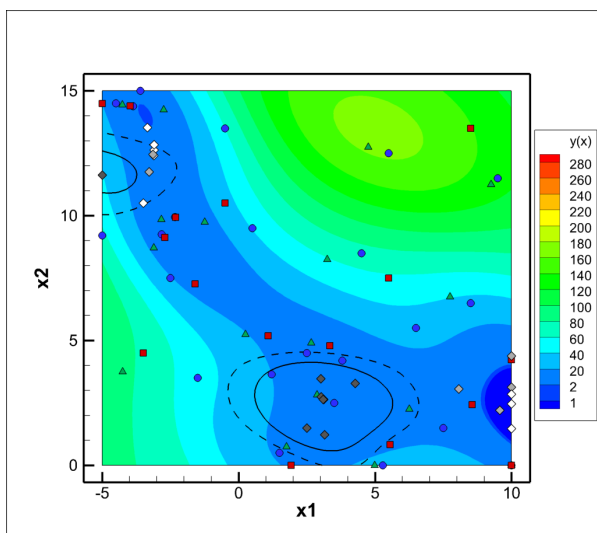
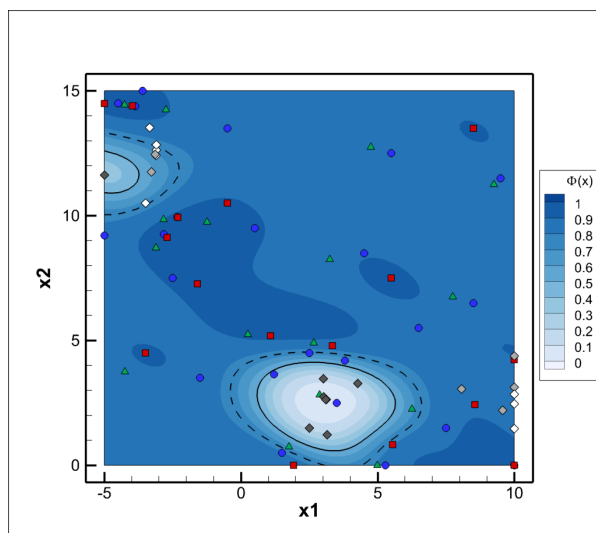
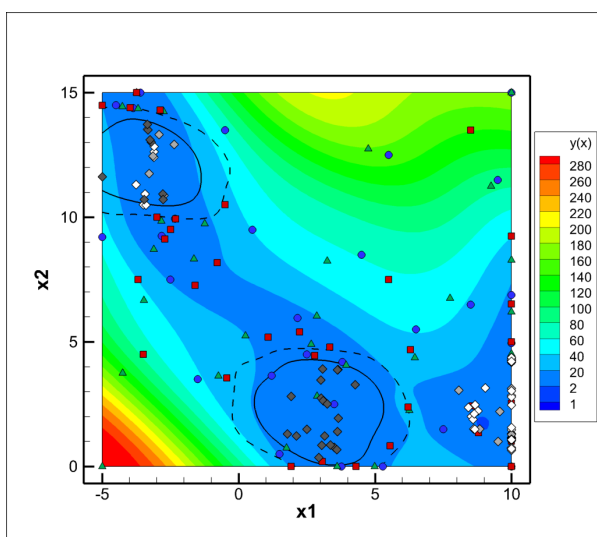
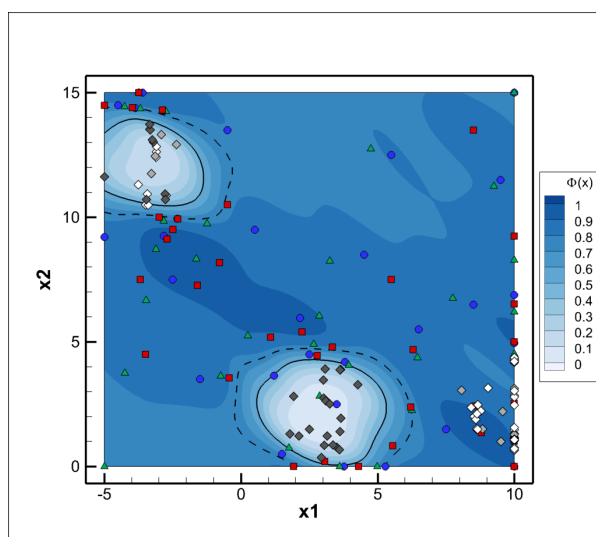
$$\begin{aligned} \min_{\mathbf{x} \in \Omega} \quad & f_l(\mathbf{x}) = f(\mathbf{x}) + 10(s - l), \\ \text{s.t.} \quad & \begin{cases} \forall l, f_l(\mathbf{x}) = \emptyset & \text{si } \|\mathbf{x} - \mathbf{x}_{\text{opt}_1}\|_2 \leq 2 \\ f_3(\mathbf{x}) = \emptyset & \text{si } \|\mathbf{x} - \mathbf{x}_{\text{opt}_2}\|_2 \leq 2 \\ f_{1,2}(\mathbf{x}) = \emptyset & \text{si } \|\mathbf{x} - \mathbf{x}_{\text{opt}_3}\|_2 \leq 2 \end{cases} \end{aligned} \quad (6.2.9)$$

La Figure 6.3 présente les tracés de $\hat{y}_3(\mathbf{x})$ et $\Phi_3(\mathbf{x})$ pour les itérations 10, 50 et 150 de l'EGO. Les isovaleurs de $\Phi_3(\mathbf{x}) = 0.5$ et $\Phi_3(\mathbf{x}) = 0.75$ sont respectivement tracées avec les courbes noires continues et noires discontinues. Les évaluations de niveaux 1, 2 et 3 sont respectivement représentées par les cercles bleus, les triangles verts et les carrés rouges. Les observations non évaluables sont également reportées avec respectivement pour les niveaux 1, 2 et 3 des losanges blancs, gris et noirs.

A l'itération 10, on note la présence de trois points non évaluables de niveau 3 et deux de niveau 1. La probabilité $\Phi_3(\mathbf{x})$ est déjà réduite (≈ 0.25) dans les zones non évaluables par le niveau 3. La prédiction $\hat{y}_3(\mathbf{x})$ est quant à elle pas encore assez raffinée pour distinguer les courbures caractéristiques de la fonction Branin.

Après 50 itérations de l'EGO multi-fidélité, on peut maintenant commencer à reconnaître la fonction Branin dans $\hat{y}_3(\mathbf{x})$. De nouveaux points non évaluables par le niveau 3 et évaluables par les deux autres niveaux ont été ajoutés dans la zone proche de l'optimum $\mathbf{x}_{\text{opt}_2}$. On voit que la formulation multi-fidélité de la classification apporte ici une correction de la prédiction de la classification donnée par les deux premiers niveaux car $\Phi_3(\mathbf{x}) \approx 0.05$ dans cette zone. De même, dans la zone proche de $\mathbf{x}_{\text{opt}_3}$, on remarque $\Phi_3(\mathbf{x}) \approx 0.9$ bien que des points non évaluables des niveaux 1 et 2 sont présents.

Finalement après 150 itérations, les deux zones non évaluables par le niveau 3 ont bien été explorées et la probabilité $\Phi_3(\mathbf{x})$ associée est proche de 0 dans ces zones. La prédiction moyenne $\hat{y}_3(\mathbf{x})$ a également été nettement améliorée.

(a) Itération 10, $\hat{y}_3(x)$.(b) Itération 10, $\Phi_3(x)$.(c) Itération 50, $\hat{y}_3(x)$.(d) Itération 50, $\Phi_3(x)$.(e) Itération 150, $\hat{y}_3(x)$.(f) Itération 150, $\Phi_3(x)$.**Figure 6.3** – Illustration de la classification multi-fidélité dans un processus EGO. Cas du problème (6.2.9).

TROISIÈME PARTIE

CONTRIBUTIONS DE LA THÈSE

Après-demain, demain sera hier.

PEUPLE TZIGANE

Réglages optimaux d'une grand-voile, applications expérimentales et numériques

Sommaire

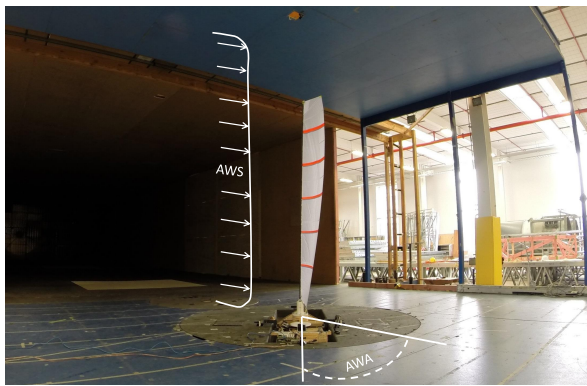
7.1	Introduction	119
7.2	Publication dans « <i>Journal of Fluids and Structures</i> »	122
7.3	Conclusion	145

7.1 Introduction

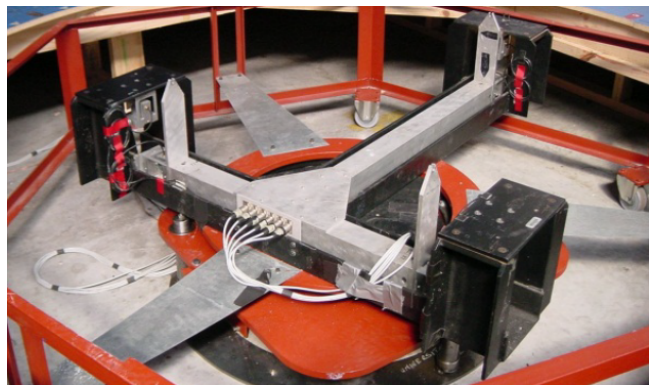
Le Chapitre 7 présente la première contribution de cette thèse, à savoir l'application de l'optimisation par méta-modèles Gaussiens (EGO) aux réglages d'une grand-voile dans le but d'en maximiser sa performance. Ces travaux sont réalisés pendant la première année de thèse afin de mettre en place et de valider les développements des outils d'optimisation. La validation présentée ici se fait sur la méthode EGO dans sa version originale à fidélité unique et non contrainte, que l'on a introduite dans la Section 4.3.3. La performance (fonction objective) de la grand-voile est optimisée par la détermination des deux réglages principaux, appelés « longueur d'écoute » et « position du chariot ». La faible dimension $d = 2$ de ce problème d'optimisation simplifie la compréhension et l'analyse des résultats de l'EGO. En plus de l'originalité de l'application, la contribution apportée par ce Chapitre 7 concerne les modèles utilisés pour mesurer la performance de la grand-voile. En effet, la réponse de la performance en fonction des deux réglages, est ici modélisée expérimentalement et numériquement. Une comparaison des réponses de ces deux modèles est également effectuée.

Les mesures expérimentales sont réalisées dans la soufflerie du « Yacht Research Unit (YRU) » de l'« University of Auckland », et financées par le programme européen « Sailing Fluid (FP7 2007-2013) ». Cette validation expérimentale a bénéficié des travaux expérimentaux de la thèse de Nicolas Aubin (Aubin, 2017). Une

grand-voile d'Imoca à l'échelle 1/13, fabriquée par la voilerie « Incidence-Sails », et son support qui comporte les deux réglages (écoute et chariot), construit en partie à l'IRENav, sont envoyés en Nouvelle-Zélande afin d'y réaliser des mesures expérimentales pour cette validation expérimentale de l'EGO. Ces mesures sont détaillées dans la thèse de [Aubin \(2017\)](#). Une illustration des moyens expérimentaux du YRU est donnée par la Figure 7.1. En particulier, la Figure 7.1(a) est une photographie de la soufflerie prise au niveau de la veine d'essai qui a ici la particularité d'être une veine ouverte. Les grandeurs AWS (vitesse du vent apparent) et AWA (angle du vent apparent) y sont également illustrées. Les dimensions de la section d'essai sont 3.5 m de hauteur et 7 m de largeur. La Figure 7.1(b) montre la balance qui mesure les 6 composantes des efforts aérodynamiques générés par la grand-voile et utilisés pour définir la fonction performance. Celle-ci est située sous le plancher de la maquette expérimentale. La robustesse de la méthode EGO est ainsi évaluée sur un problème physique



(a) Soufflerie à veine ouverte.



(b) Balance aérodynamique à 6 composantes.

Figure 7.1 – Illustration de la soufflerie du YRU.

réel qui comporte des incertitudes. Les évaluations de la performance sont donc des observations bruitées qui proviennent des imperfections du modèle expérimental et du bruit inhérent aux mesures. Dans ces travaux expérimentaux, l'intérêt d'utiliser des modèles Gaussiens qui prennent en compte un bruit $\sigma_\epsilon \neq 0$ est montré. Cela est illustré par une première application de l'EGO, où $\sigma_\epsilon = 0$, qui présente des difficultés à converger vers l'optimum expérimental. Une seconde optimisation où σ_ϵ est fixé sur la base du bruit expérimental montre ensuite l'utilité du bruit dans la construction des méta-modèles de l'EGO.

Un modèle numérique est construit sur la base de ces expériences effectuées à Auckland. L'objectif est de réaliser une optimisation des deux réglages de la grand-voile avec une fonction performance simulée par un code IFS. La géométrie et les caractéristiques mécaniques du modèle expérimental (voile, mat, lattes, etc.) sont modélisées afin de reproduire le plus fidèlement possible l'expérience. Le profil de vitesse du vent apparent en amont de la grand-voile, que l'on impose dans les simulations numériques, provient lui aussi des mesures expérimentales. Une illustration du modèle numérique (maillages fluide et structure) et d'un résultat montrant un champ de pression sur la grand-voile est donnée sur la Figure 7.2. Dans cette expérience numérique, le bruit σ_ϵ est également considéré en étant directement inféré sur les observations. Le nombre d'évaluations de la fonction objective pour déterminer l'optimum est comparable au cas expérimental et la réponse numérique est satisfaisante vis-à-vis des mesures expérimentales compte tenu des incertitudes de ces expériences. En effet, en

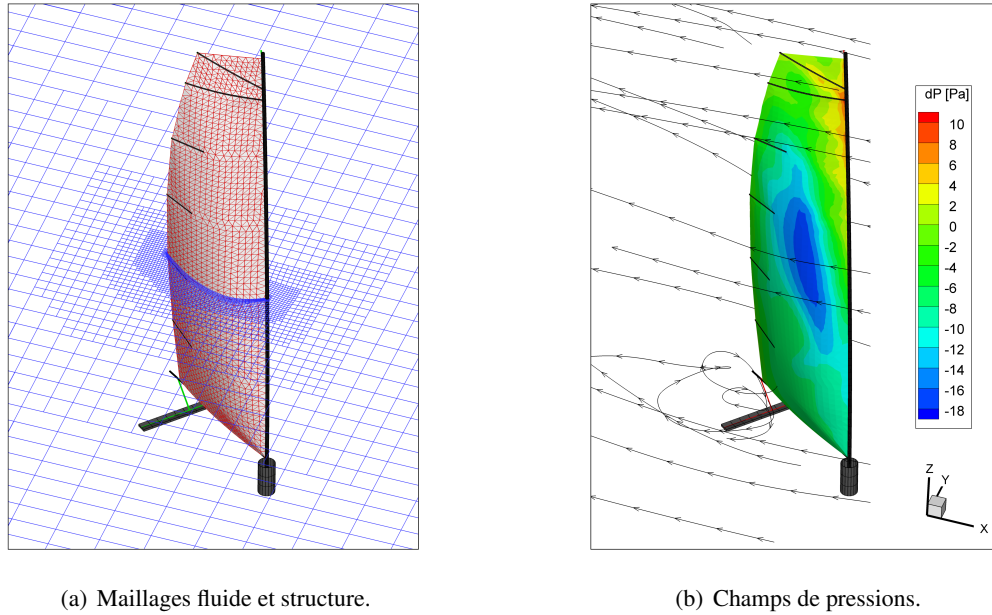


Figure 7.2 – Illustration d’un calcul IFS réalisé pendant la validation numérique de la méthode EGO.

plus de valider la méthode EGO sur un cas purement numérique, cette application donne accès à des comparaisons numériques et expérimentales. On peut ainsi valider l’outil numérique IFS sur l’ensemble du domaine de variabilité des réglages Ω , et faire de la propagation d’incertitudes expérimentales dans l’expérience numérique. Pour cela, une analyse de sensibilité est menée afin de déterminer les principales sources d’incertitude sur la fonction performance. La robustesse des réglages optimaux de la grand-voile par rapport à ces incertitudes est également caractérisée dans ces travaux.

Les travaux préliminaires de cette validation expérimentale et numérique de la méthode EGO sont présentés à la conférence « Chesapeake Sailing Yacht Symposium » (Sacher et al., 2016). Après les avoir approfondies, ces validations sont publiées dans (Sacher et al., 2017c), dont fait objet la Section 7.2 à suivre.

7.2 Publication dans « *Journal of Fluids and Structures* »

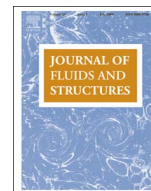
Journal of Fluids and Structures 69 (2017) 209–231



Contents lists available at ScienceDirect

Journal of Fluids and Structures

journal homepage: www.elsevier.com/locate/jfs



Efficient optimization procedure in non-linear fluid-structure interaction problem: Application to mainsail trimming in upwind conditions



Matthieu Sacher^{a,*}, Frédéric Hauville^a, Régis Duvigneau^b, Olivier Le Maître^c,
Nicolas Aubin^a, Mathieu Durand^d

^a Naval Academy Research Institute, IRENAV CC600, 29240 Brest Cedex 9, France

^b INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902 Valbonne, France

^c LIMSI - CNRS, Rue John Von Neumann, 91400 Orsay, France

^d K-EPSILON, 1300 Route des Crêtes, 06560 Valbonne, France

ARTICLE INFO

Keywords:

Sail trimming optimization
Fluid-structure interaction
Gaussian process model
Experimental
Numerical
Uncertainty quantification

ABSTRACT

This paper investigates the use of Gaussian processes to solve sail trimming optimization problems. The Gaussian process, used to model the dependence of the performance with the trimming parameters, is constructed from a limited number of performance estimations at carefully selected trimming points, potentially enabling the optimization of complex sail systems with multiple trimming parameters. The proposed approach is tested on a two-parameter trimming for a scaled IMOCA mainsail in upwind sailing conditions. We focus on the robustness of the proposed approach and study especially the sensitivity of the results to noise and model error in the point estimations of the performance. In particular, we contrast the optimization performed on a real physical model set in a wind tunnel with a fully non-linear numerical fluid-structure interaction model of the same experiments. For this problem with a limited number of trimming parameters, the numerical optimization was affordable and found to require a comparable amount of performance estimation as for the experimental case. The results reveal a satisfactory agreement for the numerical and experimental optimal trimming parameters, considering the inherent sources of errors and uncertainties in both numerical and experimental approaches. Sensitivity analyses have been eventually performed in the numerical optimization problem to determine the dominant source of uncertainties and characterize the robustness of the optima.

1. Introduction

Research on sailing yachts has fostered the advancement of methods to predict and improve the performances of racing yachts. Yacht performance is usually assessed using so-called Velocity Prediction Programs (VPPs) (Oossanen, 1993), which by equilibrating loads on hull, appendages and sails, determine several performance indicators, such as Boat Speed (BS) and Velocity Made Good (VMG). The loads estimation in VPPs can be based on empirical formulas, experimental data and/or numerical models of various complexity level (Hansen et al., 2003; Korpus, 2007). Due to the complexity and multi-physics character of yacht dynamics, performance studies are often separated in hydrodynamic (Huetz and Guillem, 2014) and aerodynamic (Augier et al., 2012;

* Corresponding author.

E-mail address: matthieu.sacher@ecole-navale.fr (M. Sacher).

<http://dx.doi.org/10.1016/j.jfluidstructs.2016.12.006>

Received 15 June 2016; Received in revised form 5 December 2016; Accepted 16 December 2016

0889-9746/ © 2016 Elsevier Ltd. All rights reserved.

Trimarchi, 2012; Menotti et al., 2013) aspects. In the present work, we focus on the aerodynamics optimization for the performance of a sail system; however, the procedure developed below can be applied to hydrodynamic optimization and even fully coupled (hydro-aero) yacht performance optimization.

The physics of sail systems involves very complex phenomena, such as nonlinear Fluid-Structure Interaction (FSI) effects and aero-elastic instabilities. Moreover, the modeling of real sailing conditions is still an open challenge because of the large uncertainties in the prediction of wind and sea states. To our knowledge, the sails optimization has thus been limited so far to idealized situations. For instance, sail shape optimizations (without accounting for the full FSI problem) were reported in Rousselon (2008), while the numerical trimming of two-dimensional sails was considered in Chapin et al. (2008). Regarding three-dimensional FSI problems, the authors of Ranzenbach et al. (2013) mention an optimization of the trimming of sails, but within an inviscid flow approximation and few details are provided on the actual optimization procedure used.

The present work aims toward the development of efficient numerical optimization procedures, capable of dealing with complex sail systems with realistic physical models (e.g. nonlinear FSI and turbulent flows) and large number of optimization variables (i.e. trimming parameters). Abstractly, the optimization problem can be expressed as

$$\mathbf{x}_{\text{opt}} = \underset{\mathbf{x} \in \Omega}{\operatorname{argmin}} - \mathcal{P}(\mathbf{x}), \quad (1)$$

where $\mathbf{x}$ are the optimization variables, Ω the domain of variation of the optimization variables, and $\mathcal{P}: \Omega \mapsto \mathbb{R}$ is the measurement of performance to be maximized. The main difficulty in solving problem (1), in the context of sail-trimming, is related to the cost of numerically estimating the performance $\mathcal{P}$ at tentative values $\mathbf{x}$ of the parameters.

Indeed, the estimation of $\mathcal{P}(\mathbf{x})$ involves the resolution of a complex nonlinear FSI problem, with typically several convergence iterations between the nonlinear elastic and flow solvers. Further, adjoint-based techniques are hardly amenable to non-linear FSI problems, particularly when the resolution is based on the coupling of distinct solvers. This fact precludes the use of efficient gradient-based optimization methods, and favors the use of derivative free optimization algorithms such as the simplex based (Nelder and Mead, 1965) and evolutionary (Bäck and Schwefel, 1993; Hansen, 2006) methods. These approaches classically require many performance evaluations of $\mathcal{P}(\mathbf{x})$, making applications to sail systems very costly as a single evaluation of $\mathcal{P}$ can routinely require several hours of CPU-time even on modern parallel computers.

From these observations, we advocate the use of meta-models to mitigate the large computational cost of optimizing the trimming parameters of sail systems. Meta-models-based optimization methods have been experiencing a growing interest for the last years, and are currently used in several other disciplines, such as aerodynamic drag reduction (Laurenceau and Sagaut, 2008; Jeong et al., 2005), vibration minimization for rotating aircrafts (Glaz et al., 2009), civil engineering for the design of water distribution network (di Piero et al., 2009) and geological carbon sequestration (Espinete and Shoemaker, 2013), or FSI problems (Aghajari and Schäfer, 2015; Degroote et al., 2012). Specifically, we rely in this work on Gaussian Process (GP) to approximate the mapping $\mathcal{P}: \Omega \mapsto \mathbb{R}$. This statistical approach uses a coarse set of performance evaluations at some selected parameters values $\mathbf{x} \in \Omega$ to infer a GP $\mathcal{G}(\mathbf{x}) \approx \mathcal{P}(\mathbf{x})$. Given the GP approximation one can apply his favorite optimization procedure substituting $\mathcal{G}$ to $\mathcal{P}$ in (1), and obtain an approximation of $\mathbf{x}_{\text{opt}}$. This surrogate-based optimization procedure is embedded in an iterative scheme, where new evaluations of the performance at carefully selected new points $\mathbf{x}$ are introduced in order to refine the GP approximation in regions of Ω of interest, that is susceptible to include the optimum. The GP approach is then expected to improve the direct optimization of $\mathcal{P}$ by a) requiring an overall lower number of performance evaluations, compared to direct gradient-free approaches, and b) enabling the use of efficient global optimization tools.

Another interest of GP-based optimization is that it naturally accommodates for errors and noise in the performance evaluation. This feature is especially attractive in the case of optimizations relying on complex numerical simulations, where both modeling and numerical errors are expected to be significant and hardly reducible. To illustrate the interest of the robustness of the GP-based optimization to inherent error, this work focuses on the optimization of a scaled IMOCA mainsail in upwind conditions. The objective is to find the optimal trimming of the sail, for a performance criterion combining the drive and side aerodynamic force coefficients. The GP-based optimization is performed first on a physical model using measurements of $\mathcal{P}(\mathbf{x})$ performed in the wind tunnel of the Yacht Research Unit (the University of Auckland), for the sequence of trimming points $\mathbf{x}$ requested by the iterative optimization procedure. In this case, the error in the observed values of the performance is due to the imperfections in the experimental apparatus and the inherent noise in the measurements. This experiment shows that the significant error in the measurements of $\mathcal{P}(\mathbf{x})$ compromises the convergence of descent methods (Saul'ev and Samoilova, 1975) without its implicit account in the GP reconstruction.

The experimental optimization is subsequently used to assess the relevance of numerical optimizations, where a non-linear FSI solver is used to compute the performance. To this end, a high fidelity numerical model was created from measurements of the experimental model (e.g. dimensions, and sail geometry, mechanical characteristics of mast and boom, wind tunnel inflow conditions, ...). A state of the art FSI solver is then used to solve the numerical model at the sequence of trimming points requested by the GP-based optimizer. The FSI solver involves a nonlinear structural solver coupled to a finite volume flow solver with an Unsteady Reynolds-Average Navier-Stokes Equations (URANS) model. The comparison of the experimental and numerical optimal trimming parameters shows a significant discrepancy, as expected from both numerical modeling and experimental errors. However, it is found that the agreement between the numerical and experimental optimal performances is consistent with the error and noise levels estimated by the GP constructions. This observation has motivated further uncertainty quantification studies to determine the dominant source of model error affecting the optimal trimming parameters and the performance. Because of the highly non-linear and coupled nature of the FSI solver, the uncertainty analyses rely again on local surrogates of the performance, using a Polynomial

Chaos (PC) expansion (Le Maître and Knio, 2010) constructed on sparse grids. These surrogates allow to quantify the variability in the achieved performance and in the optimal trimming parameters. They are also used to propose a robust optimum.

The paper is organized as follows. Section 2 reviews the construction of the Gaussian process approximation and the iterative optimization procedure detailing the selection of the new parameters. We then detail the experimental set-up and report the results of the experimental optimization in Section 3, contrasting the cases of GP with and without noise assumption. The numerical modeling of the experiment and the FSI solver are discussed in Section 4, which also presents the corresponding optimization results. Section 5 discusses the various possible source of disagreement between the experimental and numerical optima, and presents several uncertainty quantification analyses to assess their respective importance and impact. Finally, conclusions of this work and direction for future developments are provided in Section 6.

2. GP-based optimization

The use of surrogate models is a classical approach to reduce the computational burden related to the optimization of complex systems (Simpson et al., 2001). Polynomial models were firstly used as surrogates, thanks to their ease of construction. However, more sophisticated meta-models such as Gaussian processes (GP) (Kleijnen, 2009) or Support Vector Machines (SVM) (Vapnik, 1995) have emerged so far. GP models have been found especially appealing for optimization purpose, because their statistical nature allows to provide both a prediction of the performance function, in terms of model mean, and an error estimate, in terms of model variance. Optimization procedures need also to be adapted to the statistical nature of the approximation. Here, we shall rely on the rigorous optimization framework based on the maximization of the Expected Improvement (EI) criterion, referred as the Efficient Global Optimization (EGO) (Jones et al., 1998) method. An additional key feature of the GP construction exploited in this work, is that it can naturally deal with observation noise. We also mention the recent extensions to multi-point optimization proposed (Ginsbourger et al., 2010; Liu et al., 2016).

In this Section we start by introducing the definitions of a Gaussian process and briefly summarizing the construction of a Gaussian process to model a function from noisy observations. More details on GP models can be found, e.g., in Gibbs (1997), Rasmussen and Williams (2006). We then describe the GP-based optimization procedure (Duvigneau and Chandrashekar, 2012), detailing the selection of successive optimal candidates.

2.1. Gaussian process

Gaussian Process modeling is a statistical method to approximate functions from a finite set of observations, possibly noisy, at arbitrary points. The method takes advantage of the full definition of a Gaussian process from its second-order characteristics, namely its mean and covariance functions. Using the observations, it explicitly updates the prior mean and covariance functions into their posterior counterparts, minimizing the mean squared-error of the estimator. For Gaussian observation noise, the estimator corresponds to the optimal Bayesian posterior. Compared to alternative regression-type approaches, the Gaussian process modeling also avoids the need to prescribe explicitly a basis for the approximation space, which can be infinite. Instead, one needs to specify the prior covariance structure, which is usually selected using objective criteria (see below) in a family of parametrized kernels.

For a formal definition of the Gaussian processes, we are considering a probability space Ω , a measurable space E and a random variable defined by the measurable application $X: \Omega \rightarrow E$. The space $E \subset \mathbb{R}^d$ and X is called a real-valued random variable when $d=1$. Gaussian variables are a particular class of the random ones and the real-valued random variable X is a Gaussian variable if $\exists (\mu, \Sigma) \in \mathbb{R} \times \mathbb{R}^+, Z \sim \mathcal{N}(0, 1)$ and $X = \mu + \Sigma Z$. We write $X \sim \mathcal{N}(\mu, \Sigma^2)$.

Definition 1. A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution (Rasmussen and Williams, 2006). The mean function $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$, that fully parameterize a Gaussian process of a real process $f(\mathbf{x})$, are defined by

$$m(\mathbf{x}) = \mathbb{E}\{f(\mathbf{x})\}, \quad (2)$$

$$k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\{f(\mathbf{x}) - m(\mathbf{x}), f(\mathbf{x}') - m(\mathbf{x}')\}, \quad (3)$$

and a Gaussian process is written as

$$f(\mathbf{x}) \sim \mathcal{G}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')). \quad (4)$$

2.2. Gaussian process model

We first focus on the construction of a GP to approximate a mapping $f: \mathbf{x} \in \Omega \mapsto f(\mathbf{x}) \in \mathbb{R}$, from a set $\mathbf{X}_n = (\mathbf{x}_1, \dots, \mathbf{x}_n)^T$ of n training inputs vectors, or observation points, $\mathbf{x}_i \in \Omega$. Each component $\mathbf{x}_i \in \mathbf{X}_n$ is associated to an observation (or measurement) $y_i \in \mathbb{R}$ which is assumed to be dependent on the latent function $f(\mathbf{x})$ through

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (5)$$

where ε_i is a random measurement error (i.e. the measurement noise). In this work, the ε_i are assumed *independent* and identically

M. Sacher et al.

Journal of Fluids and Structures 69 (2017) 209–231

distributed, with (centered) Gaussian distribution:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2). \quad (6)$$

In the following, σ_ε^2 is referred to as the noise variance.

In the GP model, the latent function is considered as a realization of a zero-mean multivariate Gaussian process F , with unknown covariance function C_F ; that is $F \sim \mathcal{N}(0, C_F)$, with

$$C_F(\mathbf{x}, \mathbf{x}') \doteq \mathbb{E}\{F(\mathbf{x}), F(\mathbf{x}')\}, \quad (7)$$

where $\mathbb{E}\{\cdot\}$ denotes the expectation operator.

The covariance function of F must be specified. We shall consider the Matérn class (Stein, 2012) of stationary covariance functions, given by one-dimensional generator with form

$$M_\nu(r, l) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu} r}{l} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu} r}{l} \right). \quad (8)$$

Here $r \doteq \|\mathbf{x} - \mathbf{x}'\|$, ν and l are two positive parameters, and K_ν is the modified Bessel function of the second kind. We shall further restrict ourselves to covariances with $\nu \rightarrow \infty$, leading to the classical squared exponential covariance family with generator

$$M_\infty(r, l) = \exp\left(-\frac{r^2}{2l^2}\right). \quad (9)$$

Other kernel functions could be used as well. However, as shown in the benchmark from Picheny et al. (2013), the influence of the kernel choice is low in general, and only noticeable for multi-modal objective functions, which is not the case in the present study. Various other definitions of kernel are given in Rasmussen and Williams (2006).

The multidimensional counterpart is defined as the product of the one-dimensional generator. The final expression of the covariance function for the GP approximation is

$$C_F(\mathbf{x}, \mathbf{x}'; \boldsymbol{\Theta}) \doteq \theta_1 \prod_{i=1}^d \exp\left(-\frac{(x_i - x'_i)^2}{2l_i^2}\right) + \theta_2. \quad (10)$$

In the expression (10) of C_F , $\boldsymbol{\Theta} = \{\theta_1, \theta_2, l_1, l_2, \dots, l_d\}$ is a vector of hyper-parameters that will have to be inferred from the observations. The first hyper-parameter, θ_1 , scales the distance-dependent correlation, while θ_2 is an offset from zero. The remaining parameters l_i are the anisotropic correlation lengths associated to the d directions of Ω . We denote $\mathbf{C}(\boldsymbol{\Theta}) \in \mathbb{R}^{n \times n}$ the covariance matrix for the observation points in $\mathbf{X}_n$, having entries

$$C_{i,j}(\boldsymbol{\Theta}) \doteq C_F(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\Theta}), \quad 1 \leq i, j, \leq n. \quad (11)$$

The vector $\mathbf{Y}_n = (y_1, \dots, y_n)^T$ of noisy observations and the predicted observation $y(\mathbf{x})$ at a point $\mathbf{x} \in \Omega$ have for joint Gaussian distribution

$$\begin{pmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{pmatrix} \Big| \mathbf{X}_n, \boldsymbol{\Theta}, \sigma_\varepsilon^2 \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{C} + \sigma_\varepsilon^2 \mathbf{I} & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^T(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\varepsilon^2 \end{bmatrix}\right). \quad (12)$$

The dependence of $\mathbf{C}$ on the hyper-parameters has been removed in (12) to simplify the notation; in addition, we denoted

$$\kappa(\mathbf{x}) \doteq C_F(\mathbf{x}, \mathbf{x}; \boldsymbol{\Theta}), \quad \mathbf{k} \doteq (C_F(\mathbf{x}, \mathbf{x}_1; \boldsymbol{\Theta}), \dots, C_F(\mathbf{x}, \mathbf{x}_n; \boldsymbol{\Theta}))^T,$$

while $\mathbf{I}$ is the identity of $\mathbb{R}^n$. The rule of conditional probabilities (Von Mises, 1964; Rasmussen and Williams, 2006) allows to write the inference of $y(\mathbf{x})$ given the noisy observations $\mathbf{Y}_n$ as

$$p(y(\mathbf{x})|\mathbf{Y}_n) = \frac{p(y(\mathbf{x}), \mathbf{Y}_n)}{p(\mathbf{Y}_n)}, \quad (13)$$

and noting $\mathbf{C}_{n+1} = \begin{bmatrix} \mathbf{C} + \sigma_\varepsilon^2 \mathbf{I} & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^T(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\varepsilon^2 \end{bmatrix}$, that predictive posterior distribution is

$$p(y(\mathbf{x})|\mathbf{Y}_n) \propto \exp\left(-\frac{1}{2}[\mathbf{Y}_n \ y(\mathbf{x})]\mathbf{C}_{n+1}^{-1} \begin{bmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{bmatrix}\right). \quad (14)$$

Using the partitioned inverse equation (Barnett and Barnett, 1979), $\mathbf{C}_{n+1}^{-1}$ is written

$$\mathbf{C}_{n+1}^{-1} = \begin{bmatrix} \mathbf{M}_n & \mathbf{m}_{n+1} \\ \mathbf{m}_{n+1}^T & \mu \end{bmatrix}, \quad (15)$$

where

$$\begin{aligned}\mu &= (\kappa(\mathbf{x}) + \sigma_\epsilon^2 - \mathbf{k}^T(\mathbf{x})(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}))^{-1}, \\ \mathbf{m}_{n+1} &= -\mu(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}), \\ \mathbf{M}_n &= (\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} + \frac{1}{\mu} \mathbf{m}_{n+1} \mathbf{m}_{n+1}^T.\end{aligned}$$

Then, we substitute in Eq. (14) and we find

$$p(y(\mathbf{x})|\mathbf{Y}_n) \propto \exp\left(-\frac{(y(\mathbf{x}) - \hat{y}(\mathbf{x}))^2}{2\hat{\sigma}_y^2(\mathbf{x})}\right). \quad (16)$$

To the end, we have $y(\mathbf{x})|\mathbf{Y}_n, \mathbf{X}_n, \boldsymbol{\Theta}, \sigma_\epsilon^2 \sim \mathcal{N}(\hat{y}(\mathbf{x}), \hat{\sigma}_y^2(\mathbf{x}))$, with *best* prediction (mean of $y(\mathbf{x})$) $\hat{y}(\mathbf{x})$ and prediction variance $\hat{\sigma}_y^2(\mathbf{x})$ given by

$$\hat{y}(\mathbf{x}) = \mathbf{k}^T(\mathbf{x})(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{Y}_n, \quad (17)$$

$$\hat{\sigma}_y^2(\mathbf{x}) = \kappa(\mathbf{x}) + \sigma_\epsilon^2 - \mathbf{k}^T(\mathbf{x})(\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}). \quad (18)$$

The hyper-parameters $\boldsymbol{\Theta}$ and noise variance σ_ϵ^2 are unknown *a priori* and must to be learned from the observations. It is done by maximization of the evidence, which involves to maximize the probability density of the observations:

$$f(\mathbf{Y}_n|\mathbf{X}_n, \boldsymbol{\Theta}, \sigma_\epsilon^2) = (2\pi)^{\frac{n}{2}} |\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mathbf{Y}_n^T (\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{Y}_n\right). \quad (19)$$

The likelihood function is given by $L(\boldsymbol{\Theta}, \sigma_\epsilon^2) \doteq L(\boldsymbol{\Theta}, \sigma_\epsilon^2; \mathbf{Y}_n) \doteq f(\mathbf{Y}_n|\mathbf{X}_n, \boldsymbol{\Theta}, \sigma_\epsilon^2) \in \mathbb{R}^+$. In the present work, $(\boldsymbol{\Theta}, \sigma_\epsilon^2)$ are determined by maximizing the log-marginal likelihood (Rasmussen and Williams, 2006) defined by

$$\mathcal{L}(\boldsymbol{\Theta}, \sigma_\epsilon^2) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I}| - \frac{1}{2} \mathbf{Y}_n^T (\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\epsilon^2 \mathbf{I})^{-1} \mathbf{Y}_n. \quad (20)$$

$\mathcal{L}(\boldsymbol{\Theta}, \sigma_\epsilon^2) \doteq \log[L(\boldsymbol{\Theta}, \sigma_\epsilon^2)] \in \mathbb{R}$ is the natural logarithm of the likelihood and it avoids to work with very small probability density values. An evolution strategy algorithm (Hansen, 2006) is used to maximize $\mathcal{L}(\boldsymbol{\Theta}, \sigma_\epsilon)$ with respect to its arguments. The GP model can be used to predict values at new points $\mathbf{x}$ using (17) and (18). The most computationally demanding part of the GP construction is the inversion (and determinant calculation) of the (full) matrix $(\mathbf{C} + \sigma_\epsilon^2 \mathbf{I})$, required in the definition of the log-marginal likelihood and for new predictions. For the problems treated in this work, this is efficiently performed by LU decomposition.

The effect of the observation noise σ_ϵ on the constructed GP model is illustrated in Fig. 1 for a one-dimensional function. The construction is based on 6 observations points, depicted with circles in the plots, and the covariance hyper-parameters $\boldsymbol{\Theta}$ are determined by maximizing the log-marginal likelihood. However, in Fig. 1(a) a value $\sigma_\epsilon = 0$ is imposed, while in Fig. 1(b) the noise level is also optimized. The two plots report the mean $\hat{y}$ of the GP models with the standard $\pm 3\hat{\sigma}$ uncertainty range ($\approx 95\%$ confidence interval). In the noise-free ($\sigma_\epsilon = 0$) case, the resulting GP model interpolates the observations with a variance of the prediction equal

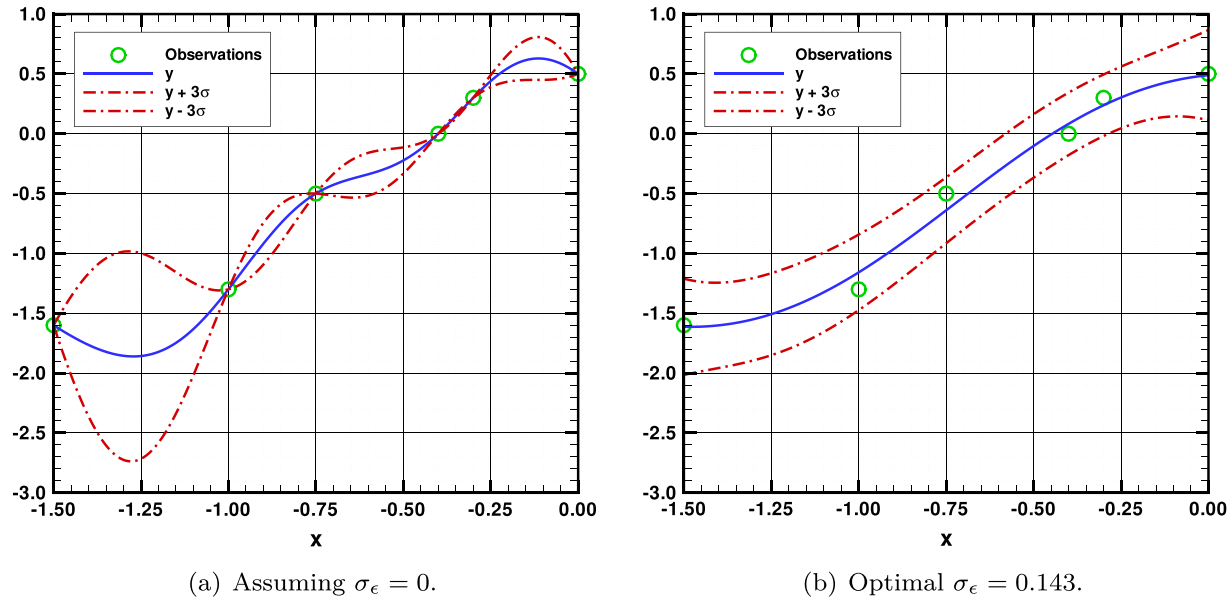


Fig. 1. Effect of σ_ϵ on the GP model.

to zero at the observation points. However, the mean of the GP model exhibits significant oscillations such that over-fitting can be suspected. In contrast, optimizing the noise level σ_ϵ in addition to Θ results in a mean process mostly free of oscillations but no more interpolating, as it can be appreciated from Fig. 1(b). The optimal value $\sigma_\epsilon = 0.143$ is characteristic of the averaged distance between the best prediction and the observations.

2.3. Optimization using Gaussian process models

GP models constitute an attractive alternative to approximate performance functions in complex optimization problems where $\mathcal{P}(\mathbf{x})$ is costly to evaluate. More precisely, from a set of (possibly noisy) estimates of $\mathcal{P}(\mathbf{x}_i)$, at n observation points $\mathbf{x}_i \in \mathbf{X}_n$, we construct a GP model. This GP model can then be used in the optimization procedure, to estimate at a low computational cost the performance $\mathcal{P}(\mathbf{x})$ at new tentative points $\mathbf{x}$. In the following, we shall denote $\hat{\mathcal{P}}_n$ the predictive mean of the performance, and $\hat{\mathbf{x}}_n$ the corresponding optimum

$$\hat{\mathbf{x}}_n = \underset{\mathbf{x} \in \Omega}{\operatorname{argmin}} - \hat{\mathcal{P}}_n(\mathbf{x}). \quad (21)$$

It is expected that $\hat{\mathbf{x}}_n \approx \mathbf{x}_{\text{opt}}$ if the approximation error $\hat{\mathcal{P}}_n - \mathcal{P}$ is small enough. This claim can be made more formal assuming sufficient regularity on $\mathcal{P}(\mathbf{x})$ and $\hat{\mathcal{P}}_n$. The advantage is that $\hat{\mathcal{P}}_n(\mathbf{x})$ is inexpensive to evaluate (compared to the evaluation of $\mathcal{P}(\mathbf{x})$) and can be easily differentiated (see (17)), enabling efficient gradient-based optimization methods. The error between $\hat{\mathbf{x}}_n$ and $\mathbf{x}_{\text{opt}}$ is a concern; in the approach below, the error on the optimum is iteratively reduced by completing the set of observation points $\mathbf{X}_n$.

It is clear that the error between $\mathcal{P}$ and $\hat{\mathcal{P}}_n$ only matters in the neighborhood of the optimum, so the completion of $\mathbf{X}_n$ with a new point $\mathbf{x}_{n+1}$ should aim at reducing the predictive error close to $\mathbf{x}_{\text{opt}}$ while ensuring that $\hat{\mathcal{P}}_n(\mathbf{x})$ has no other spurious optimum. A deterministic optimization procedure would classically choose $\mathbf{x}_{n+1}$ as the optimal point given our current approximation, that is $\mathbf{x}_{n+1} = \hat{\mathbf{x}}_n$ with $\hat{\mathbf{x}}_n$ given in (21). However, the GP model of $\mathcal{P}(\mathbf{x})$ provides not just a predictive mean, but a whole distribution of values at any $\mathbf{x}$. This probabilistic information can be exploited to construct more robust approaches for the selection of $\mathbf{x}_{n+1}$, particularly during the initial stages of the optimization when $\hat{\mathcal{P}}_n$ may have significant errors. Such improved strategies are based on merit functions which usually are composite functions involving the predictive mean ($\hat{\mathcal{P}}_n$) and the variance ($\hat{\sigma}_p^2$) of the GP model; the new point $\mathbf{x}_{n+1}$ is selected as the optimum of such composite merit function. Specifically, the merit functions should balance an optimization of $\hat{\mathcal{P}}_n$ (optimality) with the selection of $\mathbf{x}_{n+1}$ in areas of large variance $\hat{\sigma}_p^2$ to reduce the GP model error. A review of various merit functions proposed in the literature is provided in Jones (2001). In this work, we use the Augmented Expected Improvement (AEI) merit function (Huang et al., 2006), which is an extension of the popular Expected Improvement (EI) of Jones et al. (1998) to the case of noisy estimations. The AEI merit function $AEI(\mathbf{x})$ estimates the expected increase in the performance, taking into account the noise in the observed values and penalizing areas where the variance $\hat{\sigma}_p^2$ is small. It writes as

$$AEI(\mathbf{x}) = EI(\mathbf{x}) \left(1 - \frac{\sigma_\epsilon}{\sqrt{\hat{\sigma}_p^2(\mathbf{x}) + \sigma_\epsilon^2}} \right), \quad (22)$$

where the Expected Improvement is defined by

$$EI(\mathbf{x}) = \hat{\sigma}_p^2(\mathbf{x}) [u(\mathbf{x}) \Phi(u(\mathbf{x})) + \phi(u(\mathbf{x}))], \quad u(\mathbf{x}) = \frac{\hat{\mathcal{P}}_n(\mathbf{x}_{*,n}) - \hat{\mathcal{P}}_n(\mathbf{x})}{\hat{\sigma}_p(\mathbf{x})}, \quad (23)$$

with Φ and $\phi = \Phi'$ the cumulative distributions (Erf-function) and density of the standard Gaussian distribution, and $\mathbf{x}_{*,n}$ is the effective best solution over the set of current observation points $\mathbf{X}_n$:

$$\mathbf{x}_{*,n} \doteq \underset{\mathbf{x}_i \in \mathbf{X}_n}{\operatorname{argmin}} [\hat{\mathcal{P}}_n(\mathbf{x}_i) + \hat{\sigma}_p(\mathbf{x}_i)]. \quad (24)$$

Then, we set the new points $\mathbf{x}_{n+1}$ as the optimum point of the AEI, and add the new point to $\mathbf{X}_n$ setting $n \leftarrow n + 1$. The iteration is completed by evaluating the performance $\mathcal{P}(\mathbf{x}_n)$ of the new point, solving the full model. The next iteration can then proceed, starting by updating the GP model using the new observation point and its performance. Overall, each iteration requires one numerical model resolution to compute $\mathcal{P}(\mathbf{x}_n)$, and the resolution of two optimization problems, the first one to determine the covariance parameters of the GP model, and the second one to find AEI optimum. For these optimization problems, the results reported in this work were obtained using the nonlinear non-convex black-box optimization library based on the Covariance Matrix Adaptation Evolution Strategy (Hansen and Ostermeier, 2001; Hansen, 2006). The iterations of the GP-based optimization problem are continued until a stopping criterion is satisfied or the resources allocated to the optimization procedure have been exhausted (e.g., fixing a maximum number of performance evaluations). Classically, the stopping criterion compares the distance between successive iterates. The distance can be measured from the successive AEI optimal points, $\mathbf{x}_n$, successive predictive optimal points, $\hat{\mathbf{x}}_n$ given by (21), or based on the differences in successive AEI values, $AEI(\mathbf{x}_{n+1})$, or optimal predictive performance $\hat{\mathcal{P}}_n(\hat{\mathbf{x}}_n)$.

The optimization procedure is schematically presented in Fig. 2, in the case of the numerical optimization considered in Section 4. In the top block, labeled sampling loop, an initial set of observation points $\mathbf{x}_i \in \Omega$ are generated. In the present work we rely on a Latin Hypercube Sample (LHS) scheme (McKay et al., 2000) to randomly generate the initial sample set. The performance $\mathcal{P}(\mathbf{x}_i)$ is also evaluated for all the initial sample points. The lower block, label iterative loop in Fig. 2, is the core of the optimization procedure consisting in the repetition of the sequence of GP model construction, AEI optimization and performance evaluation at the new point

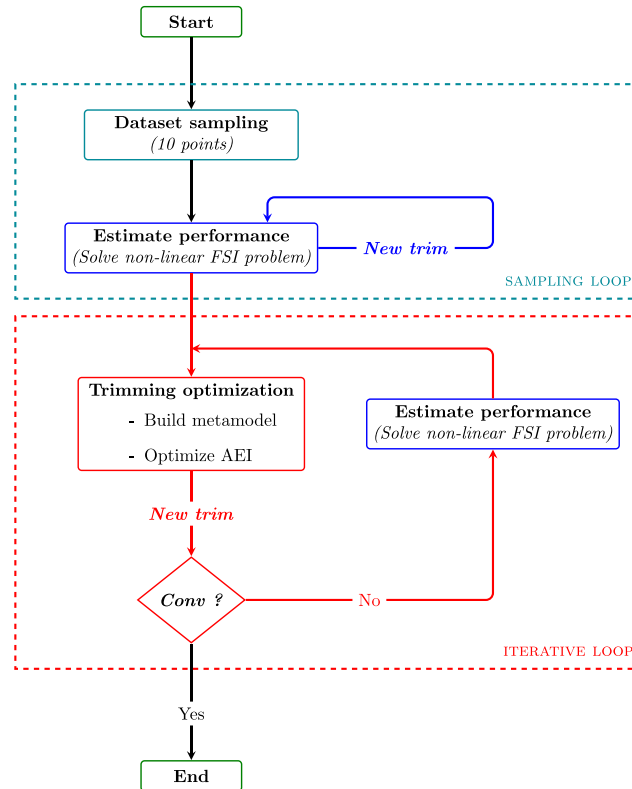


Fig. 2. Schematic of the GP-model based optimization procedure.

$\mathbf{x}_n$ selected, as well as the convergence check.

As explained above, the performance evaluations, which represent the bulk of the computational cost, are divided between the initial sampling of observation points and the additional entries governed by the AEI criterion. As with most derivative-free approaches, the number of evaluations required to determine a satisfactory approximation of the optimum can grow rapidly with the number of parameters. Obviously, this cost depends strongly on the characteristics of the performance function, such as its smoothness, anisotropy, separability, making the prescription of guidelines tedious. As discussed in references provided, GP models are usually restricted to few dozens of parameters. The reader interested in a more precise quantification of the computational cost and the best strategy to allocate resources, in particular for problems in larger dimension, can refer to Chandrashekar and Duvigneau (2009).

3. Experimental optimization

This Section concerns the sail trimming optimization performed in the wind tunnel of the Yacht Research Unit (YRU) (Flay, 1996).

3.1. Experimental setup

The experimental sail is inspired by an IMOCA 60-foot design mainsail at 1:13 scale, designed and manufactured by INCIDENCE SAILS company. The surface area is 1 m^2 , for a height of $h=2 \text{ m}$; it is supported by a rig made of a flexible circular section carbon mast (constant diameter 14 mm). The mast is simply clamped at its base and has no spreader, backstay or forestay. The sail and rig are set in the open jet test section of the YRU wind tunnel, see Fig. 3(a). The test section is 7.2 m wide and 3.5 m high.

Three stepper motors and a control card are used to remotely control the main sheet length (L_{sheet}) and main car position (L_{car}), which are the two trimming parameters to be optimized. The remote control system, shown in Fig. 3(b), allows us to enforce new trimmings with a precision of $\pm 2 \text{ mm}$, without switching off the wind tunnel flow and recalibrating the measurement devices. These devices consist in a six-component force balance, located under the floor of the wind tunnel, to measure the aerodynamic forces, a load sensor ($\pm 0.02 \text{ N}$ precision, range 5 daN) to measure the load in the sheet and a V-SPARS acquisition system (Le Pelley and Modral, Dec, 2–4. 2008) to track the position of five fluo-red stripes across the sail (see Fig. 3(a)) and record the sail's flying shape. The force balance measures the thrust force in the X-direction corresponding to the model longitudinal forward direction, the side force in the Y-direction defined as the positive port-side. The Z-direction of the balance is pointing upward. After careful calibration

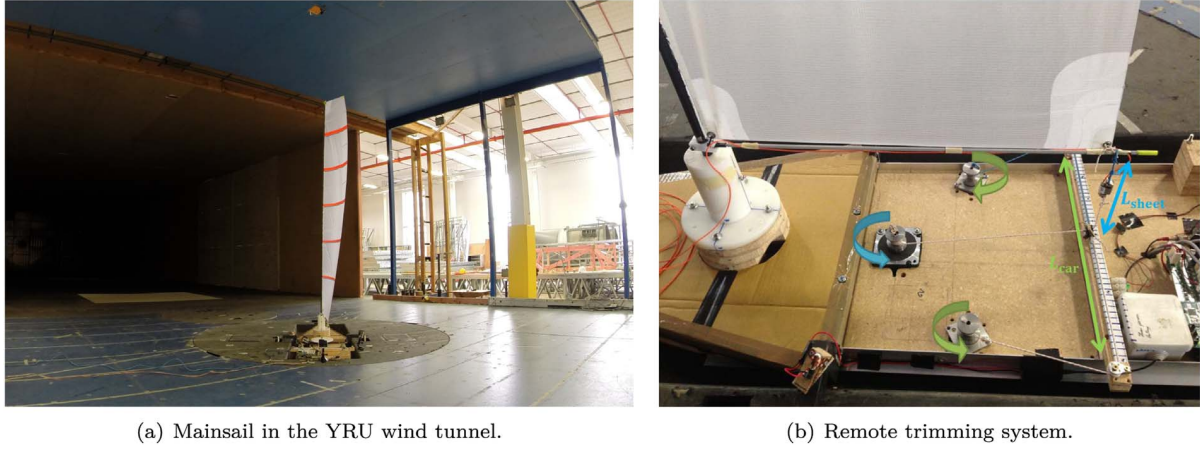


Fig. 3. Experimental setup.

checking, the balance precisions in the X , Y and Z directions were estimated to be ± 0.09 N, ± 0.11 N and ± 0.27 N respectively.

The wind tunnel inflow velocity was measured and found to have an apparent wind speed (AWS) of 3.5 ± 0.15 m/s for an apparent wind angle (AWA) of 40 ± 2 deg. The experimental Reynolds number, based on the sail chord length $c = S/h = 0.5$ m, is $Re = 1.2 \cdot 10^5$. A multi-hole pressure probe (Cobra Probe) was used to measure profiles of the flow velocity at several locations inside the wind tunnel, in particular in the inflow section. These measurements were repeated with and without the sail model to observe its effects on the flow field.

The optimization procedure is used to maximize $\mathcal{P}(\mathbf{x})$, where $\mathbf{x} = (L_{\text{sheet}}, L_{\text{car}})$. The performance (objective) function is a linear combination of the thrust and side force aerodynamic coefficients, denoted respectively C_X and C_Y :

$$\mathcal{P}(\mathbf{x}) = C_X(\mathbf{x}) + 0.1C_Y(\mathbf{x}). \quad (25)$$

The aerodynamic coefficients are the aerodynamic forces normalized by $S \times p_\infty$, where p_∞ is the measured dynamic pressure (precision ± 1 Pa) of the inflow. In addition to the thrust coefficient to be maximized, the objective function in (25) penalizes the (negative) side force coefficient C_Y , with a weight coefficient 1/10, to account for the hydrodynamic drag and leeway that are detrimental to yacht performance. The weight value is selected on the basis of previous numerical hydrodynamic studies on a Figaro sailing yacht. Changing the weight value affects the objective function and optimal trimmings such that a sensitivity analysis would be necessary if it is not well known. Alternatively, a more complex Velocity Prediction Program can be used to estimate the performance in lieu of (25). Nevertheless the proposed optimization procedure will remain the same.

The optimization follows the procedure illustrated in Fig. 2, except that the performance is not computed but measured experimentally. Specifically, for each new point $\mathbf{x}_n$, the sail is remotely trimmed to the corresponding L_{car} and L_{sheet} values and, after the flow has stabilized, the measured aerodynamic forces are averaged over a 30 s period to filter-out noisy fluctuations. The averaged forces are subsequently normalized to estimate the experimental performance denoted $\mathcal{P}^{\text{exp}}(\mathbf{x}_n)$. In these experiments, the optimization procedure is considered converged when two successive new trimming points $\mathbf{x}_n$ are within a distance less than 1 mm, a tolerance corresponding to the precision on the trimming parameters that can be applied experimentally.

In the rest of the section, we apply the optimization procedure in two different settings in order to demonstrate the effectiveness of the GP-based optimization on noisy experimental performance evaluations. To this end, we first assume in Section 3.2 a perfect performance measure, enforcing a measurement noise $\sigma_\epsilon = 0$ (noise-free situation) in the GP construct, so the GP model exactly interpolates the experimental points. This noise-free assumption is subsequently compared in Section 3.3 to the case where the GP construction uses a finite measurement noise estimated from the experimental observations.

3.2. Experimental optimization with $\sigma_\epsilon = 0$

In a first series of experiments, the optimization is carried out using GPs constructed fixing $\sigma_\epsilon = 0$. This situation corresponds to perfect performance measure assumption for the points in the initial sample set and subsequent optima of the AEI merit functions. As the result of setting $\sigma_\epsilon = 0$, the posterior mean of the GPs interpolates the experimental values. This is illustrated in Fig. 4 which depicts the GP model of $\mathcal{P}^{\text{exp}}(\mathbf{x})$ after iteration 36 of the optimization procedure. Fig. 4(a) shows the color contours of the posterior mean as a function of the trimming parameters L_{car} and L_{sheet} over the optimization domain, and Fig. 4(b) shows the corresponding standard deviation. Also shown using black dots are the 36 sample points of the initial set and successive AEI optima, that are used to construct the GP model. A large dispersion of the sample points is reported, the points being scattered over all the domain Ω . This dispersion denotes that the successive AEI merit function optima are not converging. Instead, the algorithm continues to explore Ω without discovering a particular sub-domain more likely to contain the true global optimum. This is confirmed by inspecting the posterior mean of the GP model, in Fig. 4(a), which presents at least 2 local minima, but without significantly attracting sample

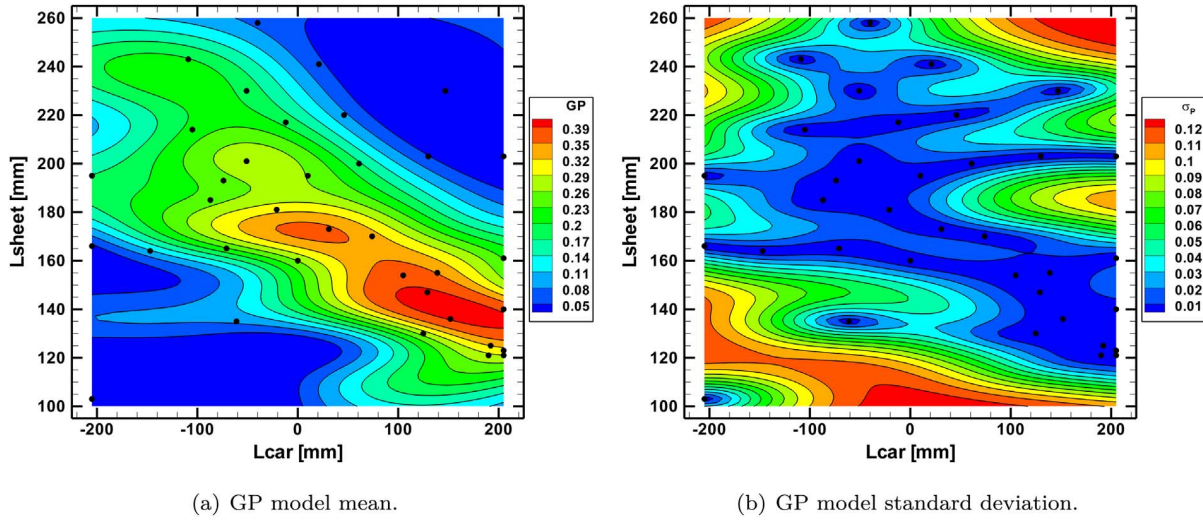


Fig. 4. GP model of $\mathcal{P}^{\text{exp}}(\mathbf{x})$ at iteration 36 and assuming $\sigma_\epsilon = 0$.

points. This behavior can be explained from the standard deviation field reported in Fig. 4(b); it has a zero value at the sample points (consistently with enforcing $\sigma_\epsilon = 0$) but quickly increases when moving away from these points: this structure dominates the AEI merit function and leads the optimization procedure to propose new points $\mathbf{x}_{n+1}$ in less explored areas because of the high variance of the GP model. In addition to distracting the iterates from promising areas to favor the exploration of highly uncertain regions, the inappropriately high confidence (low standard deviation) of the posterior mean in the immediate neighborhood of the sample points induces spurious local minima, due to the interpolatory nature of the mean. This behavior highlights the fact that the GP model fits the experimental noise.

Fitting the experimental noise is highly prejudicial in the present example, because the error in the performance is significant. The effect can be appreciated in Fig. 5 which shows the sequence of experimental performance measurements $\mathcal{P}^{\text{exp}}(\mathbf{x}_n)$ as a function of the iteration index (the first 10 iterations correspond to the initial sampling and are not actual iterations of the optimization algorithm). The plot shows that the sequence is not converging but sustains large fluctuations from an iterate to another. The magnitude of the fluctuations have remained essentially the same as for the initial random sampling, denoting the absence of improvement of the proposed optimum performance with the iterations. Clearly, the over-fitting of the noisy measurement $\mathcal{P}^{\text{exp}}$ is responsible for the failure of the present approach, which keeps exploring the domain without converging to the maximum underlying performance.

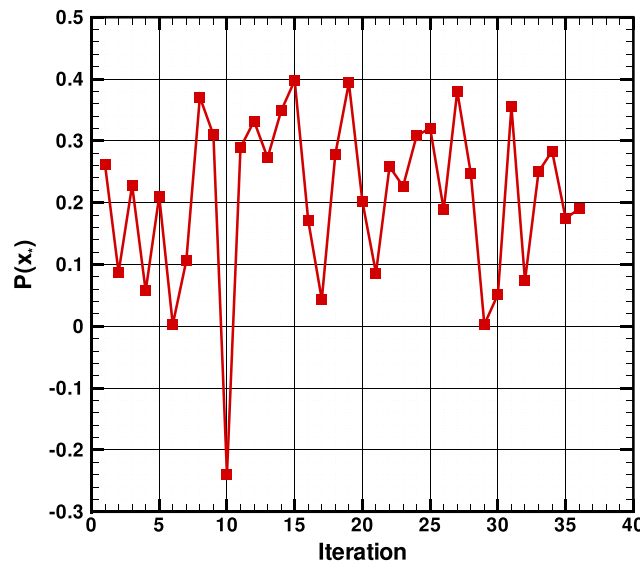


Fig. 5. Sequence of measured $\mathcal{P}^{\text{exp}}(\mathbf{x}_n)$ as a function of the iteration index. Case of $\sigma_\epsilon = 0$.

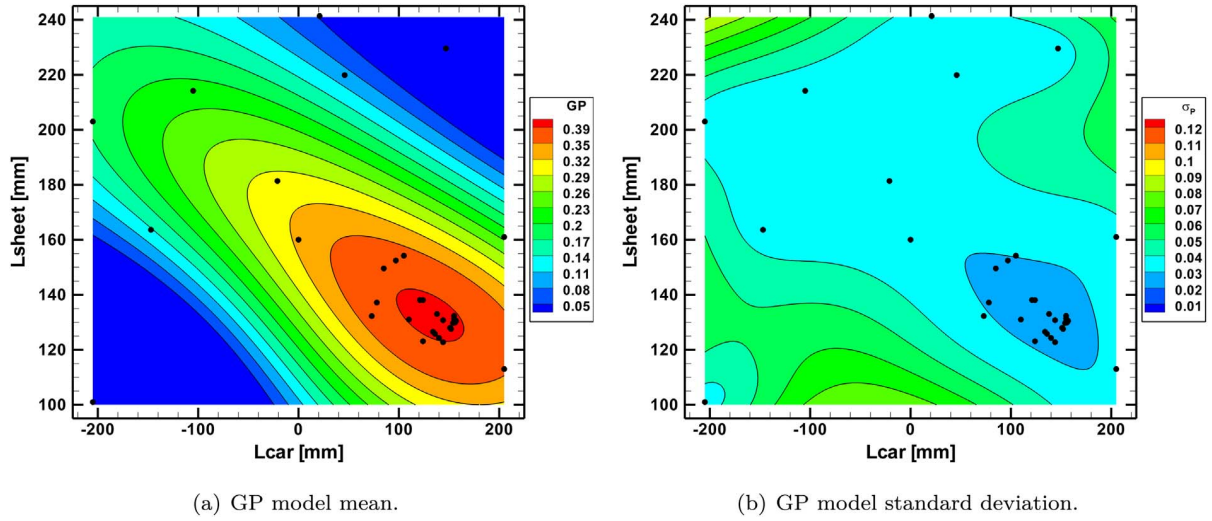


Fig. 6. GP model of $\mathcal{P}^{\text{exp}}(\mathbf{x})$ at iteration 33. Case of $\sigma_\epsilon = 0.027$.

3.3. Experimental optimization with finite noise

We next repeat the previous experiment, starting from the same initial sample set, but using a fixed finite measurement noise in the construction of the GPs. In this experiment, the value of σ_ϵ is not determined as part of the GP construction optimizing the log-marginal likelihood in (20), but is fixed to $\sigma_\epsilon = 0.027$. This value for σ_ϵ was experimentally estimated from repeated measurements at different trimmings points. Other than using a finite fixed non-zero value for σ_ϵ , the optimization procedure is the same as previously. However, the initial GP constructed from the initial sample set being different, the sequence of optima proposed is entirely different. In particular, the optimization reaches the stopping criterion at iteration 33.

Fig. 6 depicts the GP at convergence and should be contrasted with the previous case shown in Fig. 4. In particular, we observe that the posterior mean is much smoother than previously, with a single well-defined global optimum in Ω . In addition, it is seen that the sample points cluster in the neighborhood of the optimum, in sharp contrast to the previous case with $\sigma_\epsilon = 0$. The optimal trimming at convergence, $\hat{\mathbf{x}}_n$, is found for $L_{\text{sheet}} = 133$ mm and $L_{\text{car}} = 138$ mm corresponding to a predicted (GP) performance $\hat{\mathcal{P}}_n^{\text{exp}}(\hat{\mathbf{x}}_n) = 0.397$. The standard deviation of the GP model, shown in Fig. 6(b), is minimal in the neighborhood of the optimum, but remains finite and assumes values $\gtrsim \sigma_\epsilon$. Other regions of Ω far from the optimum are not explored by the optimization process, although the variance can be large.

The convergence of the optimization procedure can also be appreciated from the plot of Fig. 7, to be contrasted with Fig. 5. The

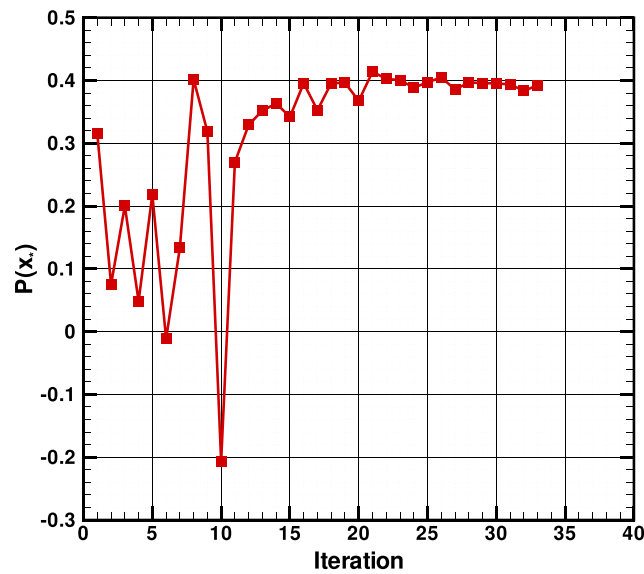


Fig. 7. Sequence of measured $\mathcal{P}^{\text{exp}}(\mathbf{x}_n)$ as a function of the iteration index. Case of $\sigma_\epsilon = 0.027$.

improvement in the measured performance, after the first 10 initial random sample points, is now clear, and the magnitude of the fluctuations quickly decreases and saturates at a low level with small amplitude $\approx \sigma_e$. Experimental imperfections can be reasonably deemed responsible for the remaining fluctuations, rather than an ineffective selection strategy of the successive optima.

The two experiments reported in this section have underlined the importance of properly treating the measurement noise. For the settings of the experimental test case, not accounting for the noise results in a complete failure of optimization procedure. On the contrary, when the experimental noise is accounted for, the GP-based optimization is robust, owing to a suitable AEI merit function able to disregard non-interesting areas of Ω , even when they support large prediction variance. Further, the GP model with correct noise level is able to smooth out the measurement noise, allowing to find the optimum in few iterations only.

4. Numerical optimization

The results reported in the previous section demonstrated the effectiveness of the GP-Based optimization procedure in determining the optimal trimming parameters in just a limited number of iterations, provided the experimental noise is appropriately accounted for. In this section, the same optimization procedure is used, but with the performance at trimming points estimated numerically. Our objective is to assess the capabilities of the optimization method applied to a coupled FSI solver, and compare the computed optimum with the experimental one. To this end, a numerical model of the wind tunnel experiment has been set up.

4.1. FSI solver

We assume that the FSI problem has a steady solution for all values of the trimming parameters in the optimization range, such that a quasi-steady approach can be used to couple the structural and fluid solvers. These solvers are now briefly described.

4.1.1. Structural solver

We rely on the ARA software developed by K-EPSILON for the structural model of the sail. The sail model involves Timoshenko beams (mast, boom, battens), cables (sheet) and Constant Strain Triangles (CST) membrane elements (cloth) of various types in large displacement formulation (Augier, 2012). A wrinkle model is added to capture the local sail deformations with a reasonable number of membrane elements (Katz and Plotkin, 2001). See Durand (2012) or Nakashino and Natori (2005) for further details on CST elements and the wrinkle model. The nonlinear solution is obtained by means of a Newton method with Aitken relaxation. The geometry of the model is imported from the SAILPACK software developed by BSG Developments, to best fit the experimental geometry. The mechanical characteristics of the mast, boom, and sail fabrics were measured during the experimental campaign and used in the numerical model. The number of membrane elements is a key parameter to control the convergence of the sail geometry and of the whole FSI solution. Fig. 8 shows 3 meshes of the sail with different number of membrane elements. The first mesh with ≈ 100 elements in Fig. 8(a) is quite coarse and is compared with an intermediate and fine meshes having $\approx 2,700$ and $\approx 9,000$ elements respectively (shown in Fig. 8(b) and Fig. (c)). Solving the complete FSI problem with these 3 meshes, we found that the coarse mesh yield a solution far from being converged. On the contrary the two most refined meshes yield fluid loads in excellent agreement with less than 0.5% discrepancies. To minimize the computational cost, and given the other sources of error, the structural mesh with $\approx 2,700$ membrane elements in Fig. 8(b) is used in the following.

The resulting structural model is illustrated in Fig. 9, where plotted are the membrane elements in Fig. 9(a) and the fabric stiffness contours in Fig. 9(b). The higher fabric stiffness at the clew, head and tack of the sail is clearly visible.

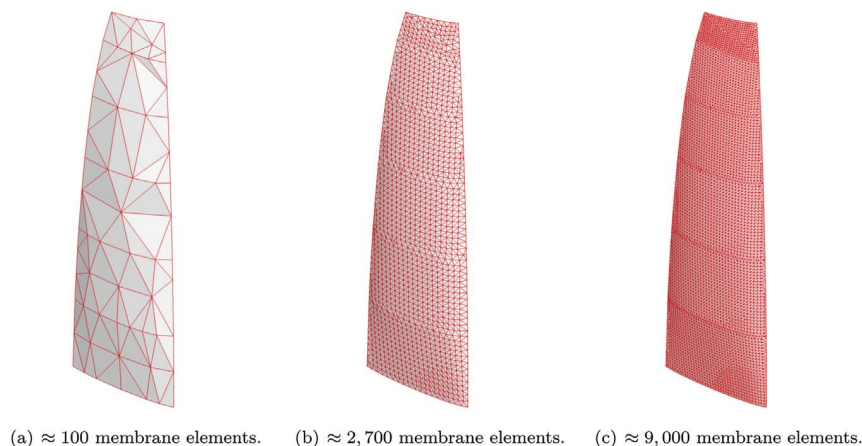


Fig. 8. Numerical meshes with different number of membrane elements.

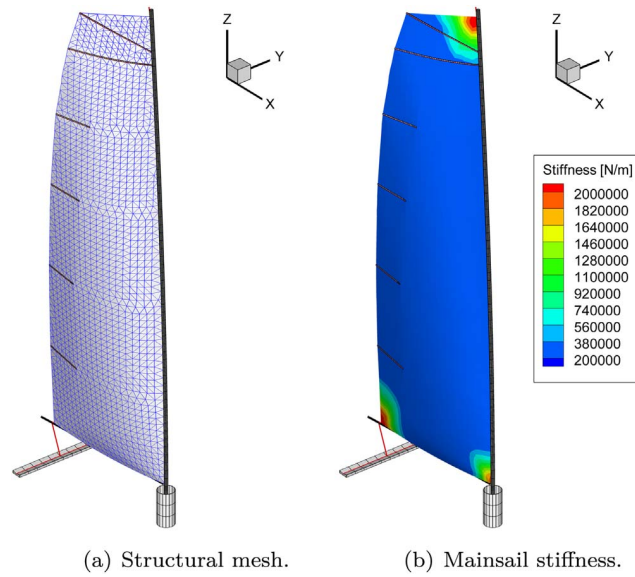


Fig. 9. Numerical model of the sail.

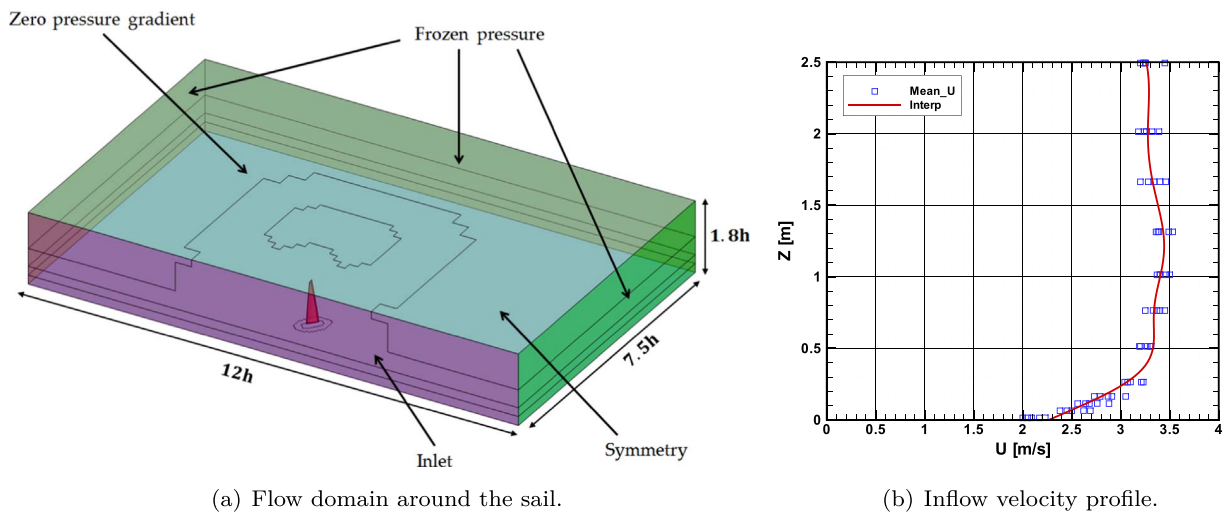


Fig. 10. Computational fluid domain and boundary conditions.

4.1.2. Flow solvers

The structural solver has been coupled with two different flow solvers.

First, it is coupled with the steady and unsteady inviscid flow solver AVANTI (Roux et al., 2002, 2008). The inviscid solver is based on the Vortex Lattice Method (VLM) (Katz and Plotkin, 2001) where the sail surface is discretized in rectangular panels supporting doublets distributions. The Kutta condition is enforced at the trailing edge to generate vorticity particles that model the wake and its nonlinear dynamics. This inviscid model assumes slip conditions over all the sail surface and therefore cannot account for detached flow situations. Its domain of validity is then limited to low-camber sails trimmings and a moderate angle of attacks. The steady inviscid solution is identified with the limit as $t \rightarrow \infty$ of the unsteady solutions.

The structural solver is also coupled with the parallel (MPI) incompressible viscous flow solver ISIS-CFD (of FINETM/Marine). This software solves the Unsteady Reynolds-Average Navier-Stokes Equations (URANS) with finite volume discretizations, accommodating both structured and unstructured meshes. We considered a parallelepiped computational domain enclosing the sail, with extension $7.5 \times 12 \times 1.8 h^3$ (in the X , Y and Z directions), following the numerical tests in Viola et al. (2013). The boundary conditions applied on the different faces of the computational domain are schematically illustrated in Fig. 10(a). The velocity profile at the inlet boundary is set to the profile measured experimentally in the wind-tunnel, which is shown in Fig. 10(b). The fluid domain is meshed to fit the sail geometries using HEXPRESSTM, a semi-automated mesh generator. Note that the mast is not represented in

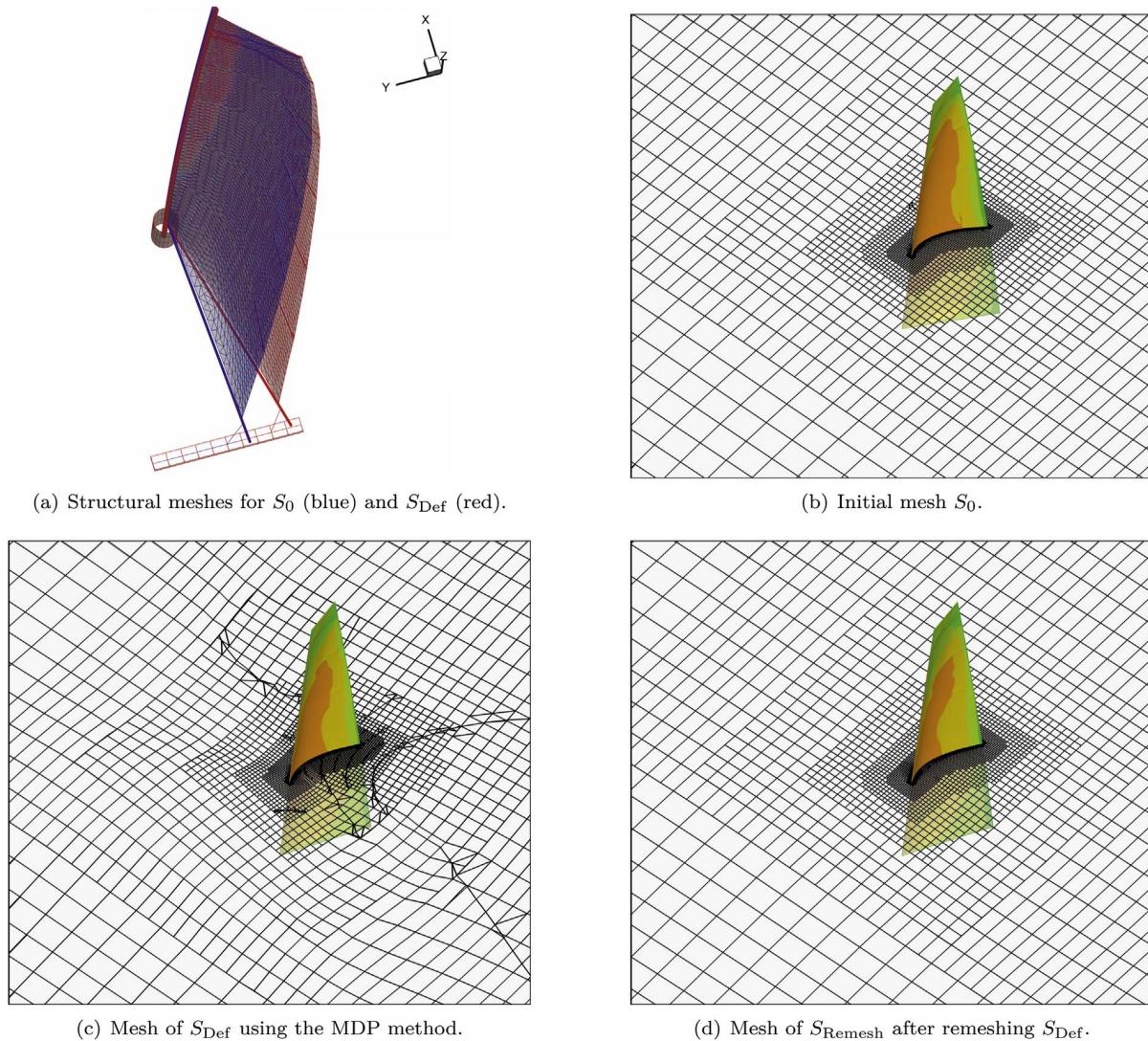


Fig. 11. Structural and fluid meshes for different trimming parameters and mesh generation methods (see text for more details).

the fluid model. Regarding the turbulence modeling, we used the two equations shear stress transport (SST) $k - \omega$ model (Menter et al., 2003) with wall function boundary conditions (Kalitzin et al., 2005).

4.1.3. Coupled problem

The FSI problems are solved coupling the structural and flow solvers through a quasi-monolithic algorithm (Durand, 2012), which is an implicit coupling procedure adapted to partitioned solvers. Briefly, the resolution of the structural problem is nested inside the iterations on the nonlinear steady flow solution. This approach preserves the convergence and stability properties of the monolithic approach. As the sail sustains large deformations, the flow solver must account for the change of the boundary geometry. This is easily achieved in the inviscid solver where only the sail surface is discretized. This is less trivial and computationally more demanding in the case of the viscous solver, which requires a volume mesh. In this work we use the mesh deformation propagation (MDP) method (Durand et al., 2014), proceeding from the sail boundary toward the inside of the fluid domain. More details on the solvers and the coupling algorithms can be found in Durand (2012), Roux et al. (2008), Roux et al. (2002).

4.2. Numerical validation

The inviscid and viscous FSI solvers have been extensively validated at the Naval Academy Research Institute with many experimental comparisons, see for instance Augier et al. (2012). A specific convergence analysis was conducted on the present problem to determine the discretization parameters of the fluid and structural problems needed to correctly capture the physics of

Table 1Comparison of the computed fluid loads for the Mesh Deformation Propagation method (S_{Def}) and the corresponding remeshed solution (S_{Remesh}).

	C_X	C_Y	$T_{sheet}[N]$
S_{Def}	0.476	−0.841	12.6
S_{Remesh}	0.477	−0.847	12.5
Differences	0.3%	0.7%	0.8%

Table 2

Comparison of the fluid loads computed using 3 different meshes with increasing resolution and their corresponding experimental values. Computational case corresponding to the experimental optimal trimming point.

	C_X	C_Y	$T_{sheet}[N]$	N_{cell}
Experimental	0.497	−1.026	14.9	–
M_0	0.476	−0.844	14.5	1.5 M
M_1	0.467	−0.831	14.3	3.4 M
M_2	0.483	−0.849	14.6	4.3 M

the FSI models, while maintaining an affordable computational cost. Achieving minimal computational time is indeed crucial to enable the optimization of the trimming parameters. A comparison with experimental measurements is also provided to assess the predictive capability of the (viscous flow) FSI solver.

4.2.1. Influence of mesh deformations

In these numerical tests, we analyze the impact of the fluid mesh generation on the computed FSI solutions. More precisely, we compare two solutions for trimming parameters $(L_{sheet}, L_{car}) = (123, 124)$ mm. The first one uses the MDP method (Durand et al., 2014), starting from the solution mesh at trimming parameters $(L_{sheet}, L_{car}) = (160, 0)$ mm, called S_0 . This first solution is called S_{Def} ; it requires a large deformation of the initial geometry of S_0 to adapt the flying shape at the new trimming parameters. The difference between the two flying shapes can be appreciated from Fig. 11(a), and the mesh deformation can be assessed comparing Figs. 11(b) and (c). Subsequently, we build a new mesh around the geometry of S_{Def} with ≈ 1.5 million finite volumes, and restart the FSI solver to accommodate changes in the fluid loads induced by the new mesh. This leads to a second solution over a much less deformed mesh, as can be seen in Fig. 11(d); we call this solution S_{Remesh} . We only report the differences between S_{Def} and S_{Remesh} in terms of fluid loads as the two flying shapes are almost indistinguishable. Values of C_X , C_Y and the sheet tensions T_{sheet} are reported in Table 1 for the two solutions. It is seen that they agree very well with less than 1% differences in the fluid loads, demonstrating the robustness of the MDP method that remains accurate even for these large changes in the sail geometry and mesh deformation. We shall exploit this robustness of the MDP method and avoid systematically iterating on the initial mesh at every new optimal point proposed when optimizing for the trimming parameters. Instead we shall perform a remeshing of the fluid domain only in the case of extreme mesh deformations.

4.2.2. Experimental comparison

The predictive capability of the viscous numerical FSI solver is also assessed by comparison with experimental measurements. In Table 2 we compare to the experimental values the computed fluid loads for 3 different meshes of the fluid domain having increasing resolution. The computations correspond to the experimental optimal trimming points $(L_{sheet}, L_{car}) = (133, 138)$ mm. The three meshes, M_0 , M_1 and M_2 have ≈ 1.5 , ≈ 3.4 and ≈ 4.3 million cells respectively. They were generated with the remeshing procedure discussed above. We can see that the numerical loads are close to each other, denoting again the convergence of the computation with respect to the fluid mesh size. The thrust force coefficient C_X and sheet tension T_{sheet} are also in excellent agreement with their experimental values. The magnitude of the side force coefficient C_Y , on the contrary, appears to be numerically under-estimated by roughly 15%. This discrepancy level is not surprising and can be explained by both the experimental uncertainties and the numerical modeling errors (see below). Consequently, the fluid mesh with around 1.5 million cells is selected in the following as it appears to provide sufficiently converged numerical fluid loads, given the discrepancies with the experimental loads. Specifically, our numerical tests show that the discretization error can not explain the discrepancy between numerical and experimental estimates. We shall discuss and investigate further the causes of the remaining discrepancies, namely the model errors, later in in Section 5.

The comparison with the experiment measurements can also be made for the geometries (flying shapes). For instance, Fig. 12(a) and (b) report the evolutions of the sail's twist angle and camber as functions of the height Z above the bottom boundary (the bottom of the sail is at $Z=0.2$ m). These two plots stress the convergence of the numerical prediction with respect to the mesh size, as no clear differences between the 3 meshes are reported. The twist distribution is also in excellent agreement with the measured one. On the contrary, differences between numerical and experimental values are not negligible for the camber distribution, with a maximum difference as large as $\approx 4.5\%$ of the chord length around 1.5 m height. Specifically, the camber of the sail appears to be numerically overestimated in the middle part of the sail. Regarding the top experimental point, we remark that it is the most sensitive to measurement errors, and the experimental camber is likely to be overestimated at this point. We however report this point for the

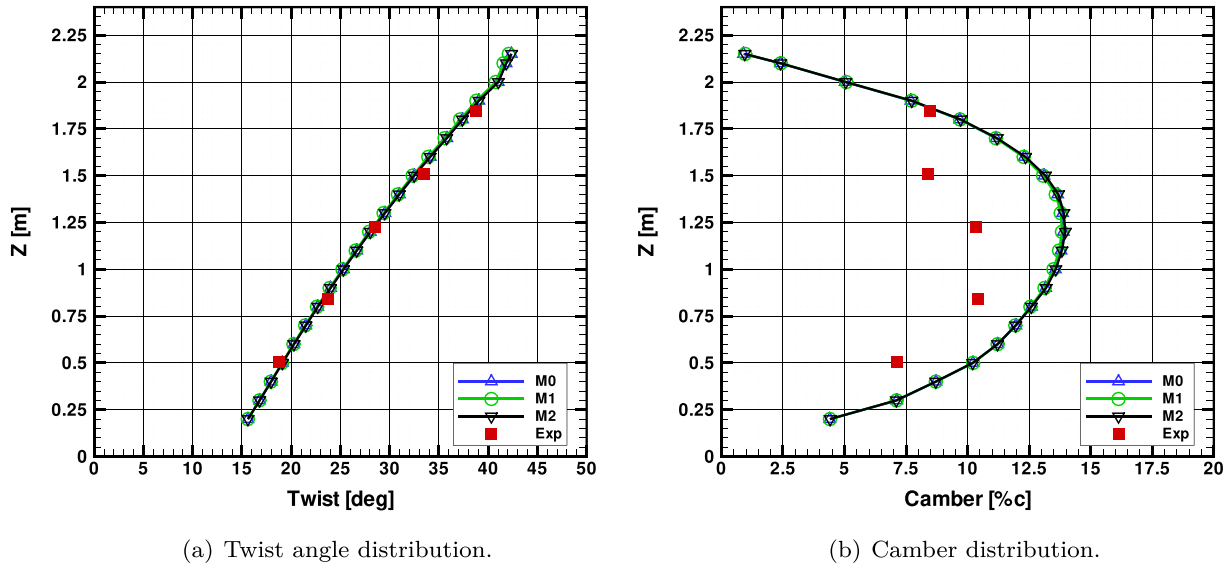


Fig. 12. Comparison of the experimental and numerical twist and camber distributions at the optimal experimental trimming points. Three numerical meshes $M_{0,1,2}$ are tested with size reported in Table 2.

completeness of the comparison.

To better appreciate the differences between the numerical and experimental results, we provide in Fig. 13 a visual comparison of the two flying shapes. In the experimental case shown in Fig. 13(a), a diagonal wrinkle of the sail is clearly visible. This is in contrast to the corresponding numerical flying shape, shown in 13(b), which is smooth and presents no such wrinkle. The wrinkle can explain the differences in the camber distribution (Fig. 12(b)), as wrinkles generally tend to reduce the sail camber. The absence of wrinkle in the numerical solution could be due to an incorrect prescription of the tension in the two top battens (horizontal beams) of the sail, which is difficult to measure, and the inability of the numerical model to capture these types of wrinkles. Another important possible source of discrepancies between the flying shapes is the boundary conditions and confinement effects, which are known to have a significant impact on the computed aerodynamic forces (Viola et al., 2013).

Based on this numerical experiment, we use in the following $\approx 2,700$ membrane elements for the structural mesh of the sail, and meshing parameters such that the fluid mesh size has ≈ 1.5 million finite volumes. With these discretization parameters, well converged viscous FSI solutions are obtained in a reasonable time of ≈ 5 h on a 64 CPUs cluster per computation.

4.3. Numerical optimization

The results of the numerical sail trimming optimization are now presented. The optimization considers again the performance function defined in (25). The optimization procedure is the same as for the experimental case, except that $\mathcal{P}(\mathbf{x})$ is numerically

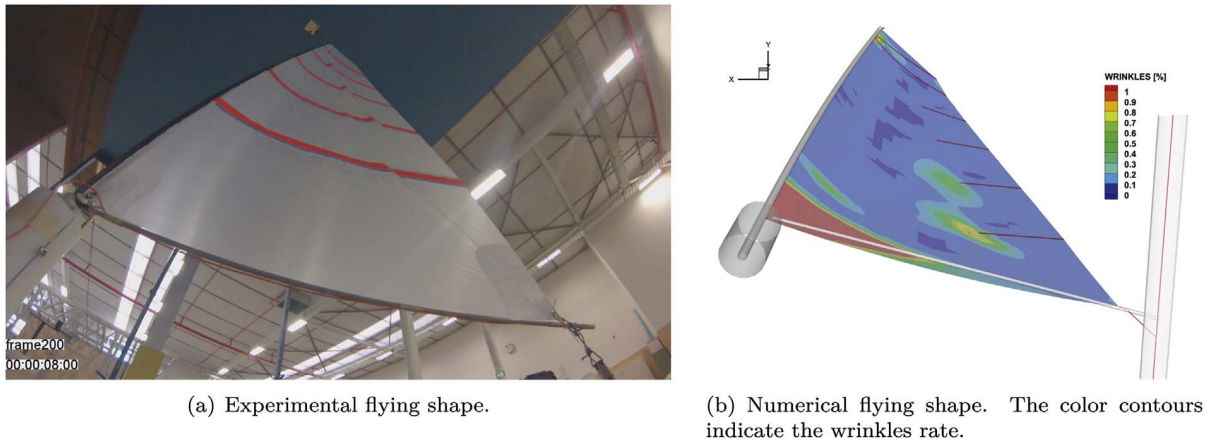


Fig. 13. Comparison of the experimental and numerical (viscous model) flying shapes at trimming parameters $L_{\text{sheet}} = 133$ and $L_{\text{car}} = 138$.

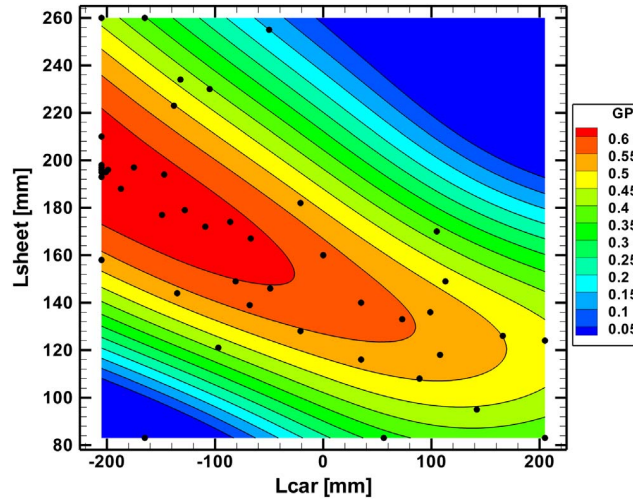


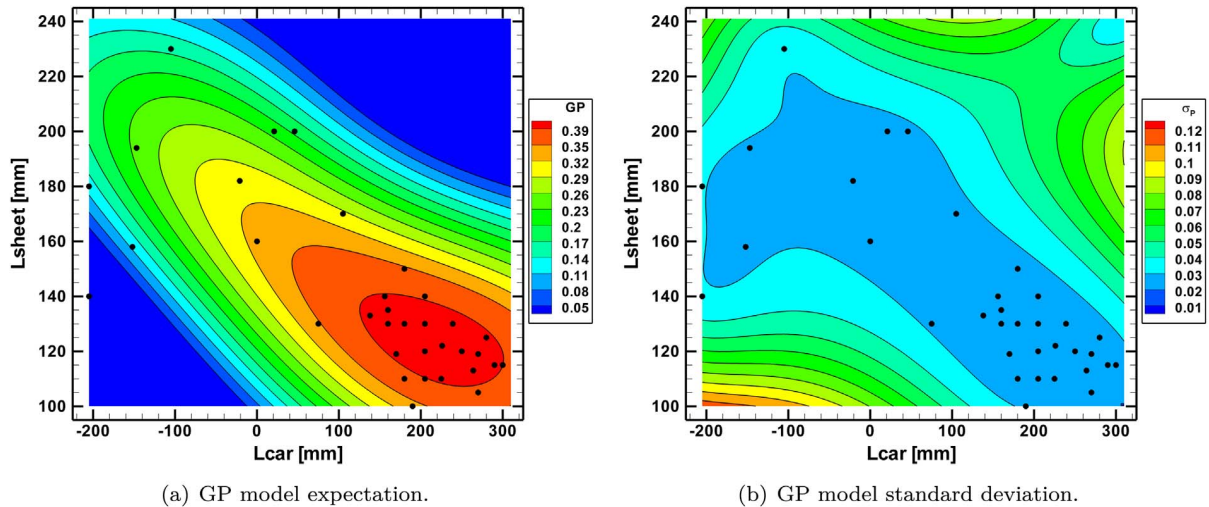
Fig. 14. Predictive mean $\hat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ at iteration 38 (inviscid model, $\sigma_\epsilon = 0.006$).

evaluated, solving the FSI problem at the proposed trimming points $\mathbf{x}_i$. We shall denote $\mathcal{P}^{\text{num}}(\mathbf{x})$ the numerical performance. In addition, the “noise” parameter σ_ϵ of the Gaussian process is not fixed to an a priori value, as in Section 3, but is inferred from the observations (see (20)).

4.3.1. Inviscid flow solver

The optimization is first applied on the structural solver coupled to the inviscid flow solver. The advantages of the inviscid model are its low computation time, compared to the viscous solver, and the simplicity of dealing with the sail deformations (no volume mesh needs be generated).

Fig. 14 shows the predictive mean $\hat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ at convergence, which is achieved at iterations 38 (it must be remembered that the first 10 iterates are selected randomly and are not actual iterations of the procedure). This predictive mean should be compared with its experimental counter-part depicted in Fig. 6(a). It can be seen that the inviscid and experimental predictive mean differ qualitatively, both in their magnitudes (note the change in the color range) and in the location of their respective optimum. Specifically, though $\hat{\mathcal{P}}_n^{\text{exp}}(\mathbf{x})$ and $\hat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ are both smooth with a unique maximum, the inviscid optimum corresponds to a trimming with L_{car} fully trimmed to the windward side and L_{sheet} quite eased, while the car is trimmed to the leeward side and the sheet is tightened in the experimental optimum. The optimal inviscid trimming corresponds to an unrealistically cambered sail. Such trimming is favored because the inviscid model cannot account for detached flow situations, yielding non-physical solutions with large overestimation of the performance in this regime. This example illustrates the importance of using a physically realistic predictive model that remains valid over the whole parameter domain. Otherwise, the numerical optimization procedure can produce an unphysical solution, even



(a) GP model expectation.

(b) GP model standard deviation.

Fig. 15. GP model of $\mathcal{P}^{\text{num}}(\mathbf{x})$ at iteration 34 (viscous model, $\sigma_\epsilon = 0.022$).

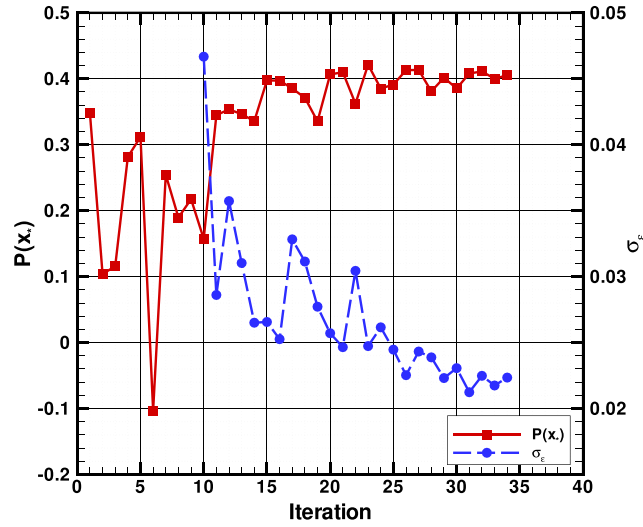


Fig. 16. Sequence of computed $\mathcal{P}^{\text{num}}(\mathbf{x}_n)$ as a function of the iteration index (in red, left axis), and corresponding inferred σ_ϵ (in blue, right axis). Case of the viscous model.

if the model is actually fairly accurate in the neighborhood of the physical optimum. In the present case, the problem could be remedied by means of an ad hoc penalization of the performance function; for instance one could add a righting moment penalty term or artificially decrease the performance for trimmings likely to produce a detached flow (Aubin et al., 2016).

4.3.2. Viscous flow solver

The optimization is now applied coupling the structural solver with the viscous solver. For reasons that will be explained soon, the parameter domain is extended to $L_{\text{car}} \in [-100, 310]$ (mm), while the range of L_{sheet} is kept as before. However, the optimization procedure is still initialized with the same LHS points as previously.

Fig. 15 shows the first two moments of the GP model of $\mathcal{P}^{\text{num}}(\mathbf{x})$ at the convergence of the optimization procedure, which is achieved at iteration 34. Note that the number of iterations to converge is very close to the experimental case (with noise). The predictive mean $\hat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ reported in Fig. 15(a) is again smooth with a unique optimum. Further, for $L_{\text{car}} \in [-210, 150]$ mm $\hat{\mathcal{P}}_n^{\text{num}}$ and $\mathcal{P}^{\text{exp}}$ are in good agreement (see Fig. 6(a)). However, the hill of $\hat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$ containing the numerical optimum has a larger extension and is flatter compared to the experimental case. Consequently, the sequence of proposed numerical optima converges to a value of L_{car} larger than for the experimental case, in fact slightly outside of the range of the experimental test. This was the motivation for extending the L_{car} range: to encompass the numerical optimum. The standard deviation of the GP model of $\mathcal{P}^{\text{num}}(\mathbf{x})$, shown in Fig. 15(b), has a structure comparable in shape and magnitude to the experimental one in Fig. 6(b), at least in the areas that were jointly sampled areas. Also, the inferred $\sigma_\epsilon = 0.022$ is not far from the experimental value, even though they have very different origins and account for totally different effects (experimental variability and numerical errors).

The convergence of the viscous optimization procedure is reported in Fig. 16, where the computed $\mathcal{P}^{\text{num}}(\mathbf{x}_n)$ is plotted as a function of the iteration index. Again, the first ten values correspond to the initial sampling of the parameter domain. After these first points, the computed performance improves and eventually stabilizes with slowly decaying fluctuations after iteration 20. This behavior is similar to the experimental results reported in Fig. 7. Fig. 16 also provides the evolution with the iteration index of the inferred σ_ϵ . It is seen that σ_ϵ decreases (on average) as the optimization proceeds and asymptotically converges to a finite value $\sigma_\epsilon \approx 0.022$ (read on the right-hand axis). The fact that σ_ϵ does not go to zero asymptotically indicates that the GP construction is informed of the numerical errors and does not attempt to fit exactly the computed performances. In other words, as for the experimental case, the GP-based optimization is able to smooth out part of the numerical errors inherent, for example, to the mesh deformations, the approximation of steady states or the convergence of nonlinear solvers.

To complete the comparison of the experimental and numerical solutions, we report in Table 3 the optimal trimming parameters and the corresponding optima of $\hat{\mathcal{P}}_n^{\text{num}}$ and $\hat{\mathcal{P}}_n^{\text{exp}}$ at convergence. We observe that the experimental and numerical solution are

Table 3

Comparison of the experimental and numerical optima: location, predicted optimal performance (see (21)), thrust and side force coefficients, and sheet tension.

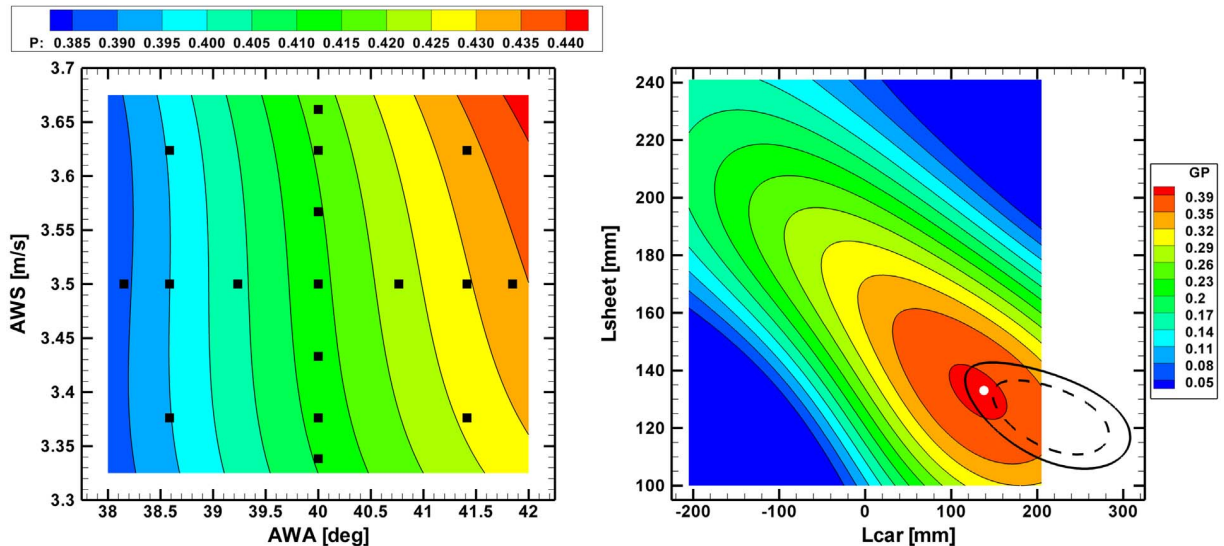
	L_{sheet} [mm]	L_{car} [mm]	$\hat{\mathcal{P}}_n(\mathbf{x}_{\text{opt}})$	C_X	$-C_Y$	T_{sheet} [N]
Exp.	133	138	0.397 ± 0.027	0.497 ± 0.012	1.026 ± 0.015	14.9 ± 0.02
Num.	122	226	0.413 ± 0.024	0.497	0.803	20.1

significantly different in terms of optimum locations, particularly for L_{car} : 138 mm in the experimental case and 226 mm in the numerical case. In contrast, the predicted performances in the fourth column of Table 3 are in much better agreement. The flatness of the performance around the experimental and numerical optima can explain the resulting large distance in terms of parameters: changing L_{car} around its optimal value weakens the numerical performance and, conversely, small numerical and modeling errors can drastically affect the optimal value of L_{car} . This claim is further confirmed by inspecting the fluid loads reported in the last column of Table 3. The loads reveal that the numerical and experimental thrust coefficients are in exact agreement, such that the side force coefficient explains all the difference in the performance. As mentioned previously, this is not surprising as the side force is known to be more sensitive to modeling errors than the thrust force. Finally, the last column of Table 3 compares the tensions in the sheet for the two optima, which significantly differ from one would have expected from the differences in the car trimming. In the next section we further refine the analysis of the different source of numerical and modeling errors and their impact on the predicted optimal trimmings.

5. Optimum uncertainty quantification

The comparison of the experimental and numerical optimizations has shown differences in the optimal trimming parameters that are consistent with the discrepancies between the measured and computed fluid forces. In addition, the numerical discretization parameters of the FSI solvers were selected as to ensure a sufficiently small numerical error, such that it can only be responsible for a limited contribution to the discrepancies between numerical and experimental optima. In fact, model errors and experimental imperfections are needed to explain most of the differences reported above. In a effort toward investigating model errors, we follow an uncertainty quantification approach in this Section, we propagate some uncertainties in the inflow conditions through the numerical optimization problem. Specifically, we are concerned with the impact of the model uncertainties caused by an incomplete knowledge of the actual wind-tunnel experiment, on the outcome of the optimization procedure. Two points of view will be considered. First, in Section 5.1, we analyze the impact of some model uncertainties on the predicted performance at the numerical optimal trimming determined in the previous section. Second, in Section 5.2, we directly estimate the uncertainty on the numerical optimal trimming due to the model uncertainties.

To perform these uncertainty analyses, we need to select the most relevant experimental uncertainties and define their variability range. To this end, we rely on expertise (from the staff of the Auckland wind-tunnel as well as our own appreciation) to review the different sources of experimental uncertainties likely to have a noticeable impact on the solution of the numerical model and on the optimization problem. The main sources identified were the measured apparent wind angle (AWA), apparent wind speed (AWS) and mechanical properties of the sail and spar elements. Regarding the mechanical properties of the mast and boom, as well as the stiffnesses of the sail fabrics, measured before and after the experimental tests were found to be quite accurate and also to have a small impact on the numerical solution. Assuming an overly large 2% range of variation for these parameters, no significant effect was observed in the predicted performance with consequently a negligible impact on the optimal solution. The tensions in the 2 top battens of the sail were identified as another potential source of uncertainties, that could significantly affect the numerical prediction, in particular through the direct modification of the sail camber and the creation-resorbing of wrinkles. Unfortunately, the batten



(a) Color contours of the PC approximation of the dependences of $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$ with the AWA and AWS. The dark points indicate the sparse grid points supporting the PC approximation.

(b) Color contours of $\hat{\mathcal{P}}_n^{\text{exp}}(\mathbf{x})$ superposed with the boundaries of the uncertainty sets $X(\sigma_{\mathcal{P}})$ (dashed line) and $X(\delta_{\text{max}})$ (solid line) and experimental optimum (white point).

Fig. 17. Effects of uncertainties in the AWA and AWS.

Table 4First and total order sensitivity indices of $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$, associated to AWA and AWS.

	AWA	AWS
First order	0.9776	0.016
Total order	0.9838	0.022

tensions were not systematically measured during the experimental campaign, such that determining a realistic variability range is not possible and we have preferred to disregard these model uncertainties in the present work. We are then left with the two main characteristics of the wind-tunnel inflow: the AWA and AWS. Based on experimental measurements and expertise, we have set the uncertainty ranges at 40 ± 2 degrees for the AWA and at 3.5 ± 0.175 m/s for the AWS. We further assume uniform distributions for the AWA and AWS which are considered independent, that is the least informative distribution given the variability range.

5.1. Uncertainty in the optimal performance

In this section we quantify the uncertainty, due to the uncertain AWA and AWS, in the numerical performance prediction at the optimal trimming point approximated by $\hat{\mathbf{x}}_n$ given by (21), that is $(L_{\text{sheet}}, L_{\text{car}}) = (122, 226)$ mm. To this end, $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$ is seen as a function of AWA and AWS which approximated using a Polynomial Chaos expansion (Ghanem and Spanos, 1991; Le Maître and Knio, 2010). The expansion coefficients are obtained using the non-intrusive Sparse Pseudo Spectral Projection (SPSP) method (Constantine et al., 2012; Conrad and Marzouk, 2013; Winokur et al., 2016) which requires the resolution of the FSI model for selected AWA and AWS values located on a sparse grid. An isotropic PC expansion with maximal degree 4 was found necessary to accurately approximate the dependences, corresponding to a sparse grid having a total number of 17 grid points.

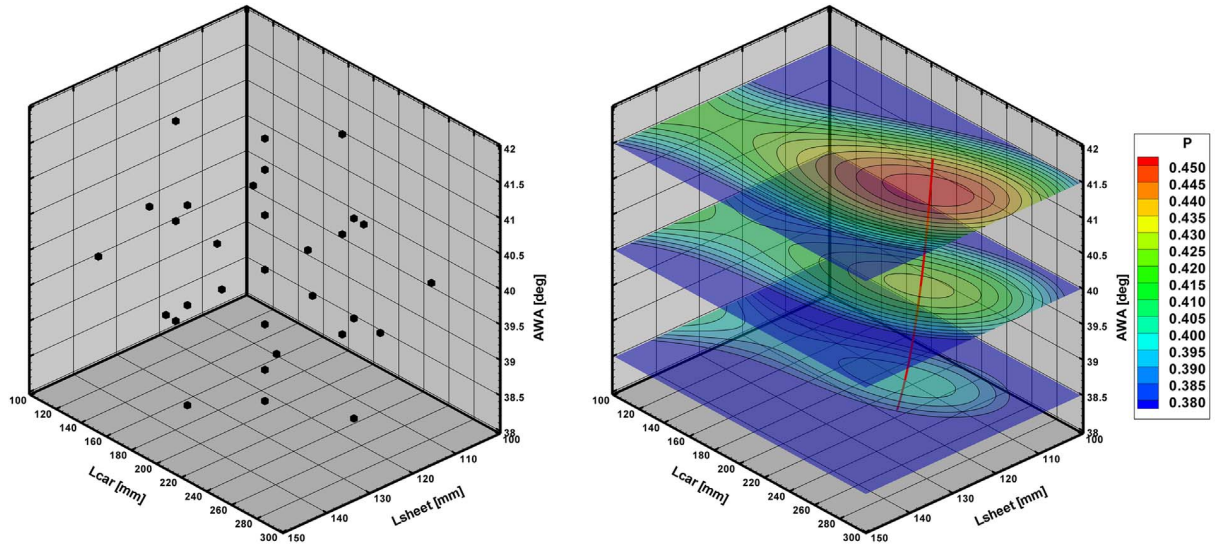
Fig. 17(a) shows the PC approximation of $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$ as a function of AWA and AWS in the uncertainty range. The sparse grid points involved in the PC construction are also shown using black squares. The performance is seen to have a low dependence on the wind speed AWS over the considered uncertainty range, especially at the lowest wind angle. The surface plot also indicated that the performance monotonically depends on the AWA, and this dependence is stronger at the highest wind speed.

The influence of AWA and AWS can be quantified using global sensitivity indices, also known as the Sobol indices (Sobol, 1993; Saltelli et al., 2004), which can be easily computed from the PC approximation (Crestaux et al., 2009; Sudret, 2008). First and total sensitivity indices of the predicted performance are reported in Table 4. The first order indices measure the fractions of the variance due to the respective factor (AWA or AWS) without any interaction with other factors. On the contrary, the total indices account for the effect of the factor and all its interactions with other factor. The table confirms the low influence of AWS compared to AWA, with a total sensitivity index barely exceeding 2%. Conversely, the AWA variability is responsible for 98% of the optimal prediction variability, and the two factors are weakly interacting. The low influence of the wind speed can be explained by the normalized form of the performance and the relative stiffness of the sail system that does not undergo significant elastic deformations over the AWS uncertainty range.

The PC approximation can be averaged over the uncertainty domain to give $\bar{\mathcal{P}}^{\text{num}}(\hat{\mathbf{x}}_n) = 0.413$; This averaged performance agrees up to the third significant digit with the nominal predicted performance $\hat{\mathcal{P}}_n^{\text{num}}(\hat{\mathbf{x}}_n)$ (that is for AWA=40 degrees and AWS=3.5 m/s). Then similarly, the standard deviation over the uncertainty range can be estimated from the PC expansion; we find $\sigma_p = 0.014$, corresponding to a low coefficient of variation (ratio of mean and standard deviation) less than 3.5%. Having characterized the uncertainty in the performance at the nominal optimum point $\hat{\mathbf{x}}_n$, we would like to deduce an uncertainty in the optimum itself. In Section 5.2 we shall consider the resolution of all the optimization problems with uncertain AWA and AWS, to assess the uncertainty in $\hat{\mathbf{x}}_n$. Here, we adopt a simpler characterization based on the distance to the nominal optimal performance $\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$; we define the family of uncertainty sets $X(\delta) \subset \Omega$, parametrized by $\delta \geq 0$, as follows:

$$X(\delta) = \{\mathbf{x} \in \Omega; \mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n) - \mathcal{P}^{\text{num}}(\mathbf{x}) < \delta\}.$$

In words, $X(\delta)$ is the set of trimmings points whose nominal performance is in a distance less than δ to the optimal nominal performance. In particular, $X(\sigma_p)$ is the set of trimmings points that are within a distance less than one AWA and AWS-induced standard deviation. To compute this set, we substitute the numerical performance $\mathcal{P}^{\text{num}}(\mathbf{x})$ with $\hat{\mathcal{P}}_n^{\text{num}}(\mathbf{x})$. The boundary of the set $X(\sigma_p)$ is plotted in Fig. 17(b) using a dotted line. It is superposed to the color contours of the experimental performance $\hat{\mathcal{P}}_n^{\text{exp}}(\mathbf{x})$. We see that the experimental optimum is not far from being included in $X(\sigma_p)$. Using $\delta = k\sigma_p$ with $k > 1$ in the definition of $X(\delta)$ defines larger and larger sets of parameters. However, k can be made arbitrarily large, in the present case, because of the finite variability of the performance caused by the AWA and AWS. Instead we consider for largest set $X(\delta_{\text{max}})$ the set defined by the upperbound distance $\delta_{\text{max}} = \bar{\mathcal{P}}^{\text{num}}(\hat{\mathbf{x}}_n) - \mathcal{P}_{\text{min}}^{\text{num}}$ where $\mathcal{P}_{\text{min}}^{\text{num}}$ is the minimal performance over the AWA and AWS uncertainty range. The boundary of $X(\delta_{\text{max}})$ is also shown in Fig. 17(b) using a solid line. It is seen that the set $X(\delta_{\text{max}})$ contains the experimental optimum. This result suggests that the discrepancy between the experimental and numerical optima can be partly explained by the experimental and model uncertainties, though other sources of modeling error not considered in this work could also have a significant impact.



(a) Sparse grid for the PC approximation of the numerical performance in the L_{sheet} , L_{car} and AWA domain.

(b) Color contours of the PC approximation of the performance for 3 values of the AWA: 38.5, 40 and 41.5 degrees. Also shown is the optimal trimming as a function of the AWA (red line).

Fig. 18. Effect of the AWA on the performance and optimal trimming.

5.2. Uncertainty on optimal trimming

The analyses presented in Section 5.1 demonstrate the effect of uncertainties in the AWA and AWS on the performances. In this section, we directly examine the uncertainty in the optimal trimming and its robustness. Following the previous findings on the dominant influence of the AWA relative to the AWS, we restrict ourselves to this unique uncertainty source in the following, keeping the same uncertainty range as before. Our objective is to characterize the dependence with the AWA of the numerical optimum, that is to determine $\hat{\mathbf{x}}_n(\text{AWA})$ and subsequently propose a robust optimum. To this end, we start by generating a PC expansion to approximate the dependences of the numerical performance with respect to the trimming parameters and the AWA. We again rely on the SPSP method assuming a uniform distribution of L_{car} , L_{sheet} and AWA in the domain $[100, 150] \times [100, 300] \times [38, 42]$. Note that the trimming domain has been restricted to a small portion of Ω around the nominal numerical optimum. The isotropic sparse grid used to construct the PC expansion of the performance is shown in Fig. 18(a); it has 31 points, requiring as many resolutions of the FSI problem, and leading to a polynomial approximation with maximum partial degree of 3. The resulting PC approximation is illustrated in Fig. 18(b), where shown are the color contours of the numerical performance in the $(L_{\text{car}}, L_{\text{sheet}})$ domain for three values of AWA = 38.5, 40.0 and 41.5 deg. The contours for AWA = 40 deg should be compared with the predicted mean $\hat{\mathcal{P}}_n^{\text{opt}}$ of the GP model shown in Fig. 15(a). Differences between the PC and GP models are visible, with in particular the presence of a second local maximum along the domain boundary. However, increasing accuracy of the PC expansion would require significantly more FSI solves, and the PC approximation is believed accurate enough for the uncertainty characterization of the optimum.

For the characterization, we rely on the determination of the optimal trimming solution $\hat{\mathbf{x}}_n(\text{AWA})$, for AWA $\in [38, 42]$. For a given value of the AWA, the optimum is determined by a simple Newton optimization method, exploiting the polynomial form of the PC approximation. The function $\hat{\mathbf{x}}_n(\text{AWA})$ is shown as a red line in Fig. 18(b). We can see that the optimal trimming corresponds to a car moved leeward and a sheet progressively eased as the AWA increases.

The uncertainty in the optimum $\hat{\mathbf{x}}_n$ can be appreciated from the values collected in Table 5, which reports the (AWA) averages

Table 5

AWA averages and standard deviations of the numerical optimum trimmings and optimal performance. Quantities are based of the PC expansion of the numerical performance.

	L_{sheet} [mm]	L_{car} [mm]	$\mathcal{P}^{\text{num}}(\hat{\mathbf{x}}_n)$ (PC approx.)
Mean	120	238	0.427
Standard deviation	1	13	0.025

Table 6

Robust optimal trimming based on (AWA)-averaged numerical performance. Also reported is the mean performance with $\pm$ standard deviation. Quantities are based on the PC expansion of the numerical.

L_{sheet} [mm]	L_{car} [mm]	Pred. $\mathcal{P}$
120	240	0.426 ± 0.026

and standard deviations of the trimming parameters and optimal performance. The AWA uncertainty is seen to have a greater influence on the optimal car trimming (13 mm standard deviation) than on the sheet trimming (1 mm standard deviation). Note that the averaged trimming point is close to the nominal numerical solution but remains quite different from the experimental one. In addition, the coefficient of variation of the optimal performance is seen to be slightly less than 6%, a significantly high value given the AWA range. This behavior is consistent with the findings of the previous of Section 5.1.

Finally, Table 6 presents the optimal trimming for the AWA-averaged numerical performance. It corresponds to the best (most robust) fixed trimming solution that one would select knowing that the AWA is uncertain, to yield the best performance on average. The standard deviation with respect to the AWA of this robust optimal solution is also provided. It is remarkable that the robust trimming parameters are close to the averaged trimming solution reported in Table 5 and to the optimal solution for the nominal AWA value. However, the robust and nominal trimmings yield different averaged performances, with an increase from 0.413, using the optimum for the nominal AWA, to 0.426 for the robust optimum solution.

6. Conclusion

We have proposed to use a Gaussian process model based optimization procedure to enable the optimization of trimming parameters in complex nonlinear sail systems. The approach was demonstrated on experimental optimization of a sail trimming in the Yacht Research Unit wind tunnel. These experiments have demonstrated the validity of the approach, which converges quickly on this problem, without exploring uninteresting trimming values. It was also shown that the GP model approach is very robust against experimental noise in the performance measurements.

To assess the validity and relevance of the experimental optimization, a comparison with a purely numerical optimization has been proposed. To this end, detailed numerical models of the wind tunnel experiment and sail have been established, considering the resolution of the nonlinear Fluid-Structure Interactions problems. An inviscid and a viscous flow solvers have been compared to highlight the importance of using a model physically correct over the whole optimization domain. The most complex, and physically correct, numerical model is based on a turbulent flow solver coupled with a nonlinear elastic solver, with a mesh deformation method to handle the changes in the sail geometries. This FSI solver can be considered representative of the state of the art for this kind of simulations. Discretization parameters were carefully selected to ensure sufficient convergence while maintaining a computational cost low enough to enable the numerical optimization, and the FSI solution was validated using several experimental measurements.

Performing the optimization on the numerical model confirmed the effectiveness of the GP-based approach. The convergence on the numerical optimum was achieved in a number of iterations comparable to the experimental case. In fact, we have shown that, using a limited set of computations, the GP model is able to accurately reconstruct the computed performance. Comparing the experimental and numerical optima, we found that while the corresponding performances and fluid forces are in excellent agreement, the optimal trimming parameters differ significantly, in particular the trimming of the car which is set more leeward in the numerical case than in the experimental one. Several modeling imperfections can explain these differences and it can be reasonably claimed that they are consistent with the current predictive capabilities of state of the art FSI solvers. In any cases, the GP-based optimization is not responsible for these differences, since it does not introduce noticeable errors, and only a better numerical modeling of the experimental set-up would help reducing the observed discrepancies. Possible avenues in this direction have been discussed such as improved batten models, better boundary conditions for the flow, accounting for mast/flow interaction or more advanced turbulence model. We also focused on the influence of experimental uncertainties on the optimal performance and trimming, focusing principally on the apparent wind angle. We have shown that the uncertainty in the AWA induces an uncertainty in the performance which is consistent with the distance between the experimental and numerical optimal trimmings, and the important effect of the AWA on the trimming of the car. We finally were interested in considering uncertainties in the definition of the optimum and introducing robust optimization.

Though the results reported in the paper have demonstrated the efficiency of the procedure, in terms of number of performance estimations, the method remains to be tested on problems involving larger sets of trimming parameters, for instance an optimization of a realistic full rig system. This is currently investigated numerically. In terms of future improvements of the optimization procedure, current developments concern the treatment of constrained problems and the incorporation of a prior belief on the optimal trimming parameters. We are also planning to introduce multi-fidelity concepts (de Baar et al., 2015) to further reduce the computational cost of optimizing complex nonlinear fluid-structure interaction problems.

Acknowledgments

The authors wish to acknowledge the researchers and staffs of the Yacht Research Unit for their welcoming and support during the experimental campaign at the wind tunnel. Contributions of INCIDENCE SAILS and NUMECA companies are also acknowledged. This work was partially funded by the European Union's Seventh Program for research, technological development and demonstration under grant agreement No PIRSES-GA-2012-318924, and from the Royal Society of New Zealand for the UK-France-NZ collaboration project SAILING FLUIDS (see www.sailingfluids.org). This work was supported by the “Laboratoire d'Excellence” LabexMER (ANR-10-LABX-19) and co-funded by a grant from the French government under the program “Investissements d'Avenir”. We also acknowledge the anonymous reviewers for their insightful comments and suggestions to improve the present paper.

References

- Aghajari, N., Schäfer, M., 2015. Efficient shape optimization for fluid-structure interaction problems. *J. Fluids Struct.* 57, 298–313.
- Aubin, N., Augier, B., Bot, P., Hauville, F., Floch, R., 2016. Inviscid approach for upwind sails aerodynamics. how far can we go? *J. Wind Eng. Ind. Aerodyn.* 155, 208–215. <http://dx.doi.org/10.1016/j.jweia.2016.06.005>, (ISSN 0167-6105, URL (<http://www.sciencedirect.com/science/article/pii/S016761051530177X>)).
- Augier, B., Bot, P., Hauville, F., Durand, M., 2012. Experimental validation of unsteady models for fluid structure interaction: application to yacht sails and rigs. *J. Wind Eng. Ind. Aerodyn.* 101, 53–66.
- Augier, B., 2012. Etudes expérimentales de l'interaction fluide-structure sur surface souple: application aux voiles de bateaux (Ph.D. thesis), Université de Bretagne Occidentale.
- Bäck, T., Schwefel, H.P., 1993. An overview of evolutionary algorithms for parameter optimization. *Evolut. Comput.* 1 (1), 1–23.
- Barnett, S., Barnett, S., 1979. *Matrix Methods for Engineers and Scientists*. McGraw-Hill.
- Chandrashekar, P., Duvigneau, R., 2009. Study of some strategies for global optimization using Gaussian process models with application to aerodynamic design, Tech. Rep. RR-6964, INRIA.
- Chapin, V.G., Neyhousser, R., Dulliand, G., Chassaing, P., 2008. Design optimization of interacting sails through viscous CFD. In: *INNOVSail, Innovation in high performance sailing Yacht*, Lorient.
- Conrad, P.R., Marzouk, Y.M., 2013. Adaptive smolyak pseudospectral approximations. *SIAM J. Sci. Comput.* 35 (6), A2643–A2670.
- Constantine, P.G., Eldred, M.S., Phipps, E.T., 2012. Sparse pseudospectral approximation method. *Comput. Methods Appl. Mech. Eng.* (229–232), 1–12.
- Crestaux, T., Le Maître, O., Martinez, J.M., 2009. Polynomial chaos expansion for sensitivity analysis. *Reliab. Eng. Syst. Saf.* 94 (7), 1161–1172.
- de Baar, J., Roberts, S., Dwight, R., Mallol, B., 2015. Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging. *Comput. Fluids* 123, 185–201.
- Degroote, J., Couckuyt, I., Vierendeels, J., Segers, P., Dhaene, T., 2012. Inverse modelling of an aneurysm's stiffness using surrogate-based optimization and fluid-structure interaction simulations. *Struct. Multidiscip. Optim.* 46 (3), 457–469.
- di Pierro, F., Khu, S.-T., Savić, D., Berardi, L., 2009. Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms. *Environ. Model. Softw.* 24 (2), 202–213. <http://dx.doi.org/10.1016/j.envsoft.2008.06.008>, (ISSN 1364-8152, URL (<http://www.sciencedirect.com/science/article/pii/S1364815208001084>)).
- Durand, M., Leroyer, A., Lothodé, C., Visonneau, M., Floch, R., Guillaume, L., 2014. FSI investigation on stability of downwind sails with an automatic dynamic trimming. *Ocean Eng.* 90, 129–139.
- Durand, M., Interaction fluide-structure souple et légère, 2012. application aux voiliers (Ph.D. thesis), Ecole Centrale de Nantes.
- Duvigneau, R., Chandrashekar, P., 2012. Kriging-based optimization applied to flow control. *Int. J. Numer. Methods Fluids* 69 (11), 1701–1714.
- Espinete, A.J., Shoemaker, C.A., 2013. Comparison of optimization algorithms for parameter estimation of multi-phase flow models with application to geological carbon sequestration. *Adv. Water Resour.* 54, 133–148. <http://dx.doi.org/10.1016/j.advwatres.2013.01.003>, (ISSN 0309-1708, URL (<http://www.sciencedirect.com/science/article/pii/S0309170813000122>)).
- Flay, R.G.J., 1996. A twisted flow wind tunnel for testing yacht sails. *J. Wind Eng. Ind. Aerodyn.* 63 (1–3), 171–182, (ISSN 0167-6105, special issue on sail aerodynamics.).
- Ghanem, R.G., Spanos, P.D., 1991. *Stochastic Finite Elements: A Spectral Approach*. Springer Verlag.
- Gibbs, M.N., 1997. Bayesian Gaussian Processes for Classification and Regression (Ph.D. thesis), University of Cambridge.
- Ginsbourger, D., Le Riche, R., Carraro, L., 2010. Computational Intelligence in Expensive Optimization Problems, chap. Kriging Is Well-Suited to Parallelize Optimization. Springer Berlin Heidelberg, Berlin, Heidelberg, 131–162.
- Glaz, B., Friedmann, P.P., Liu, L., 2009. Helicopter vibration reduction throughout the entire flight envelope using surrogate-based optimization. *J. Am. Helicopter Soc.* 54 (1), (12007–12007).
- Hansen, N., Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution. *Strateg. Evolut. Comput.* 9 (2), 159–195.
- Hansen, H., Jackson, P.S., Hochkirch, K., 2003. Real-time velocity prediction program for wind tunnel testing of sailing yachts. In: *Proceedings of The Modern Yacht*, Southampton, UK.
- Hansen, N., 2006. The CMA Evolution Strategy: A Comparing Review. In: *Towards a new evolutionary computation*, Springer, pp. 75–102.
- Huang, D., Allen, T.T., Notz, W.I., Zeng, N., 2006. Global optimization of stochastic black-box systems via sequential kriging meta-models. *J. Glob. Optim.* 34 (3), 441–466.
- Huetz, L., Guillerme, P.E., 2014. Database building and statistical methods to predict sailing yacht hydrodynamics. *Ocean Eng.* 90, 21–33.
- Jeong, S., Murayama, M., Yamamoto, K., 2005. Efficient optimization design method using kriging model. *J. Aircr.* 42 (2), 413–420.
- Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* 13 (4), 455–492.
- Jones, D.R., 2001. A taxonomy of global optimization methods based on response surfaces. *J. Glob. Optim.* 21 (4), 345–383.
- Kalitzin, G., Medic, G., Iaccarino, G., Durbin, P., 2005. Near-wall behavior of RANS turbulence models and implications for wall functions. *J. Comput. Phys.* 204 (1), 265–291.
- Katz, J., Plotkin, A., 2001. *Low-Speed Aerodynamics second edn.*. Cambridge University Press.
- Kleijnen, J.P., 2009. Kriging metamodeling in simulation: a review. *Eur. J. Oper. Res.* 192 (3), 707–716. <http://dx.doi.org/10.1016/j.ejor.2007.10.013>, (ISSN 0377-2217 URL (<http://www.sciencedirect.com/science/article/pii/S0377221707010090>)).
- Korpus, R., 2007. Performance Prediction without Empiricism: A RANS-Based VPP and Design Optimization Capability. In: *Proceedings of the 18th Chesapeake Sailing Yacht Symposium, SNAME*.
- Laurenceau, J., Sagaut, P., 2008. Efficient response surfaces of aerodynamic functions with kriging and cokriging. *AIAA J.* 46 (2), 498–507.
- Le Maître, O.P., Knio, O.M., 2010. *Spectral Methods for Uncertainty Quantification*, Scientific Computation. Springer, New York, NY.
- Le Pelley, D.J., Modral, O., 2008. V-Spars: A Combined Sail and Rig Shape Recognition System Using Imaging Techniques. In: *Proceedings of the 3rd High Performance Yacht Design Conference Auckland, New Zealand, Dec, 2–4*.
- Liu, J., Song, W.-P., Han, Z.-H., Zhang, Y., 2016. Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models. *Struct. Multidiscip. Optim.*. <http://dx.doi.org/10.1007/s00158-016-1546-7>, (1–19 ISSN 1615-1488, URL (<http://dx.doi.org/10.1007/s00158-016-1546-7>)).
- McKay, M.D., Beckman, R.J., Conover, W.J., 2000. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 42 (1), 55–61.

- Menotti, W., Durand, M., Gross, D., Roux, Y., Glehen, D., Dorez, L., 2013. An unsteady FSI investigation into the cause of the dismating of the Volvo 70 Groupama 4. In: INNOVSail, Innovation in high performance sailing Yacht, Lorient, 197.
- Menter, F.R., Kuntz, M., Langtry, R., 2003. Turbulence heat and mass transfer. Ten years Ind. Exp. SST Turbul. Model 4, 625–632.
- Nakashino, K., Natori, M.C., 2005. Efficient modification scheme of stress-strain tensor for wrinkled membranes. AIAA J. 43 (1), 206–215.
- Nelder, J., Mead, R., 1965. A simplex method for function minimization. Comput. J. 7 (4), 208–313.
- Oossanen, P.V., 1993. Predicting the speed of sailing yachts. SNAME 101, 337–397.
- Picheny, V., Wagner, T., Ginsbourger, D., 2013. A benchmark of kriging-based infill criteria for noisy optimization. Struct. Multidiscip. Optim. 48 (3), 607–626. <http://dx.doi.org/10.1007/s00158-013-0919-4>, (ISSN 1615-1488, URL (<http://dx.doi.org/10.1007/s00158-013-0919-4>)).
- Ranzenbach, R., Armitage, D., Carrau, A., 2013. Mainsail Planform Optimization for IRC 52 Using Fluid Structure Interaction. In: Proceedings of the 21st Chesapeake Sailing Yacht Symposium, SNAME.
- Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. MIT Press.
- Rousselon, N., 2008. Optimization for Sail Design. In: modeFRONTIER Conference
- Roux, Y., Huberson, S., Hauville, F., Boin, J.P., Guilbaud, M., Malick, B., 2002. Yacht performance prediction : Towards a numerical VPP. In: Proceedings of the 1st High Performance Yacht Design Conference, 4–6 December, Auckland
- Roux, Y., Durand, M., Leroy, A., Queutey, P., Visonneau, M., Raymond, J., Finot, J.M., Hauville, F., Purwanto, A., 2008. Strongly coupled VPP and CFD RANSE code for sailing yacht performance prediction. In: Proceedings of the 3rd High Performance Yacht Design Conference, 2–4 December, Auckland, 215–225.
- Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. John Wiley & Sons.
- Saul'ev, V.K., Samoilova, I.L., 1975. Approximation methods for the unconstrained optimization of functions of several variables. J. Sov. Math. 4 (6), 681–705.
- Simpson, T., Poplinski, J., Koch, N.P., Allen, J., 2001. Metamodels for computer-based engineering design: survey and recommendations. Eng. Comput. 17 (2), 129–150. <http://dx.doi.org/10.1007/PL00007198>, (ISSN 1435-5663, URL (<http://dx.doi.org/10.1007/PL00007198>)).
- Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Math. Model. Comput. Exp. 1, 407–414.
- Stein, M.L., 2012. Interpolation of Spatial Data - Some Theory for Kriging. Springer Science & Business Media.
- Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. 93 (7), 964–979.
- Trimarchi, D., 2012. Analysis of downwind sail structures using non-linear shell finite elements (Ph.D. thesis), University of Southampton.
- Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer-Verlag New York, Inc., New York, NY, USA, (ISBN 0-387-94559-8).
- Viola, I.M., Bot, P., Riotte, M., 2013. Upwind sail aerodynamics: a RANS numerical investigation validated with wind tunnel pressure measurements. Int. J. Heat. Fluid Flow. 39, 90–101.
- Von Mises, R., 1964. Mathematical Theory of Probability and Statistics. Academic Press.
- Winokur, J., Kim, D., Bisetti, F., Le Maître, O.P., Knio, O.M., 2016. Sparse pseudo spectral projection methods with directional adaptation for uncertainty quantification. J. Sci. Comput., 1–28.

7.3 Conclusion

L'application de l'optimisation par méta-modèles Gaussiens aux paramètres de réglages d'une grand-voile est proposée dans cette contribution. Cela est premièrement effectué à l'aide d'un modèle expérimental dont les mesures sont réalisées dans la soufflerie du YRU à Auckland. La robustesse de la méthode EGO est validée sur ce problème physique réel et complexe, comprenant des observations bruitées.

Une seconde optimisation par méta-modèles est effectuée à partir d'un modèle numérique construit sur la base du modèle expérimental. Cette optimisation nécessite la résolution de calculs d'interactions fluide-structure comprenant des déformations de maillages non négligeables pour chaque jeu de réglages de la grand-voile. Les comparaisons de la réponse simulée par rapport à celle mesurée montrent de bons accords sur le domaine de variabilité des paramètres d'optimisation. Des différences sur la localisation de l'optimum, en particulier pour la position du chariot de grand-voile, entre les modèles numérique et expérimental sont aussi présentées. Ces différences proviennent principalement des incertitudes inhérentes aux mesures expérimentales ainsi qu'aux simplifications du modèle numérique.

Enfin, une analyse de sensibilité permet d'identifier les principales sources d'incertitudes expérimentales sur la fonction objective. En particulier, l'incertitude sur l'angle du vent apparent introduit une incertitude sur la fonction performance qui rend les différences de localisation des optimums numérique et expérimental consistantes. La robustesse des réglages optimaux de la grand-voile vis-à-vis des incertitudes sur l'angle du vent apparent est aussi étudiée.

Pour conclure, ces premiers travaux de thèse permettent de valider les développements et les implémentations des méthodes d'optimisations par méta-modèles Gaussiens.

Optimisation en présence de points non évaluables

Sommaire

8.1	Introduction	147
8.2	Publication dans « <i>Structural and Multidisciplinary Optimization</i> »	149
8.3	Conclusion	170

8.1 Introduction

Le Chapitre 8 présente la deuxième contribution de cette thèse qui concerne l'optimisation par méta-modèles Gaussiens en présence de points non évaluables. La définition des points non évaluables est donnée dans le Chapitre 5 à la Section 5.1. L'efficacité de la méthode EGO-LS-SVM décrite dans la Section 5.3 est présentée dans ce Chapitre 8.

La difficulté des points non évaluables dans l'application de l'EGO apparaît généralement lorsque l'on cherche à résoudre des problèmes d'optimisation réels, dans le sens « non analytiques », ayant une dimension $d \gg 1$. L'idée d'utiliser un modèle de classification est apparue dès les premières tentatives de développements. Les diagrammes de « Voronoï » (Aurenhammer, 1991) ou la méthode des « k -nearest neighbors (K-NN) » (Keller et al., 1985) sont deux approches qui s'implémentent facilement et qui ont été initialement testées avant la classification LS-SVM. La Figure 8.1 présente un exemple d'optimisation de la fonction Branin où chacun des trois optimums est considéré comme un point non évaluable (carrés rouges). Les méta-modèles de la prédiction $\hat{y}(\mathbf{x})$ sont tracés pour les itérations 1, 10 et 30 dans les cas de classifications par diagrammes de Voronoï et par la méthode K-NN. Les aires blanches correspondent aux zones classées comme non évaluables. On remarque que la classification par diagrammes de Voronoï est moins efficace que la méthode K-NN car la diminution des aires non évaluables autour des optimaux est plus lente. En pratique, ces deux méthodes

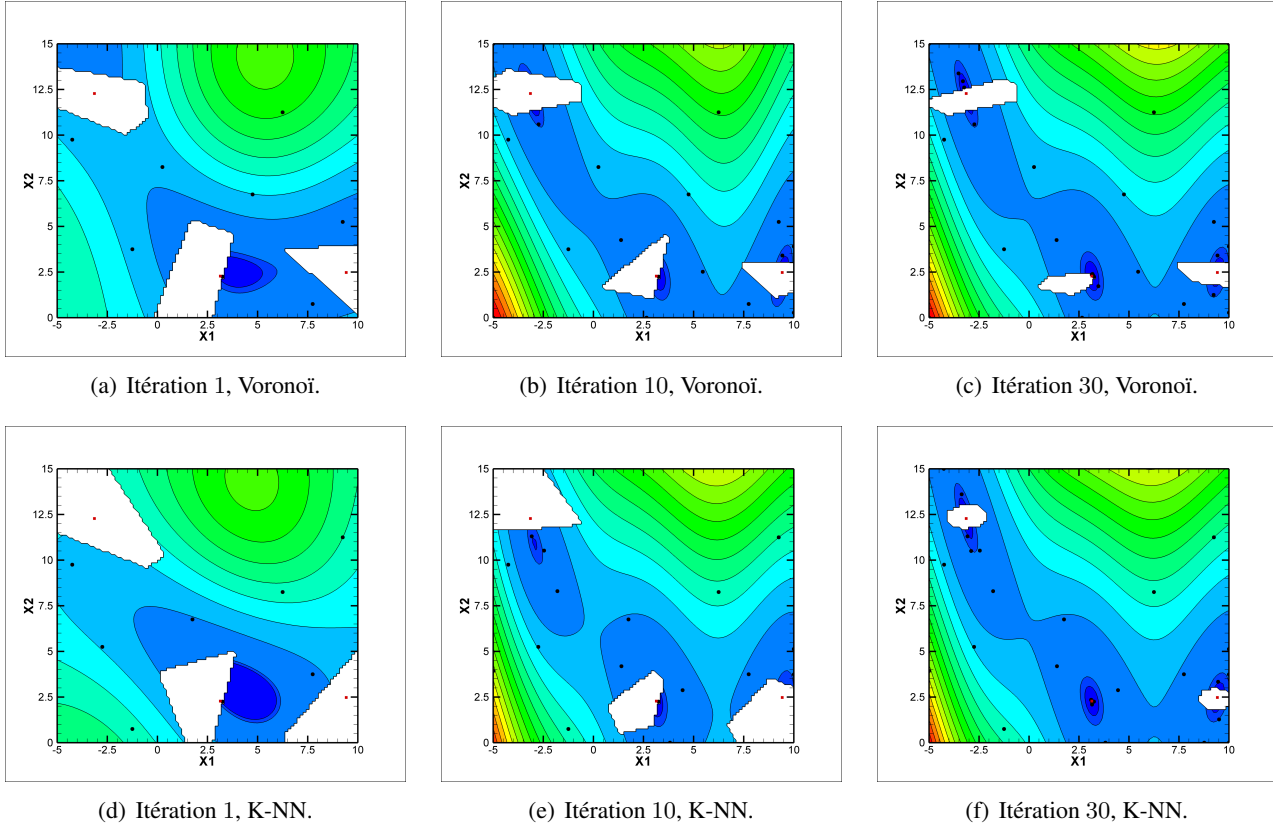


Figure 8.1 – Premiers essais de classifications des points non évaluables. Méta-modèles de la prédiction $\hat{y}(\mathbf{x})$ comprenant les zones classées comme non évaluables (zones blanches).

ont l'inconvénient de devenir coûteuses lorsque le nombre d'observations (évaluables + non évaluables) est important. De plus, leur définition sous la forme d'une probabilité d'appartenance à la classe évaluable n'est pas immédiate, même si la méthode K-NN dans sa version pondérée peut s'en rapprocher. L'utilisation du classifieur LS-SVM et sa version probabiliste donnée en (5.3.32) est alors évidente car bien plus efficace.

La validation de la méthode EGO-LS-SVM développée pendant cette thèse est illustrée dans la Section 8.2 ci-après, sous la forme de la publication (Sacher et al., 2018b). Ces travaux sont également présentés dans (Sacher et al., 2017b). Cette contribution est composée de différents problèmes d'optimisation analytiques comprenant des points non évaluables et présentant des niveaux de difficulté et des propriétés variés. L'efficacité de la méthode EGO-LS-SVM est comparée à la méthode CMA-ES (Hansen, 2006b). Un cas concret d'optimisation sous contraintes comprenant des points non évaluables est également illustré dans la dernière partie de cette publication. Cela concerne l'optimisation du design d'un hydrofoil flexible composé de paramètres géométriques et structuraux (voir le Chapitre 9). L'objectif est de minimiser l'effort de traînée de l'hydrofoil à différents points de fonctionnement (vitesses) du voilier, tout en respectant des contraintes d'inégalité afin d'assurer un régime d'écoulement non cavitant. Dans ce dernier problème, la méthode EGO-LS-SVM est confrontée à la méthode CMA-ES et la méthode de modélisation des contraintes introduite à la Section 5.2. Les résultats portant sur les performances hydrodynamiques de l'hydrofoil flexible sont détaillés dans le Chapitre 9.

8.2 Publication dans « *Structural and Multidisciplinary Optimization* »

Structural and Multidisciplinary Optimization
<https://doi.org/10.1007/s00158-018-1981-8>

RESEARCH PAPER



A classification approach to efficient global optimization in presence of non-computable domains

Matthieu Sacher¹ · Régis Duvigneau² · Olivier Le Maître³ · Mathieu Durand^{4,5,6} · Élisabeth Berrini^{2,7} · Frédéric Hauville¹ · Jacques-André Astolfi¹

Received: 26 October 2017 / Revised: 26 February 2018 / Accepted: 26 March 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Gaussian-Process based optimization methods have become very popular in recent years for the global optimization of complex systems with high computational costs. These methods rely on the sequential construction of a statistical surrogate model, using a training set of computed objective function values, which is refined according to a prescribed infilling strategy. However, this sequential optimization procedure can stop prematurely if the objective function cannot be computed at a proposed point. Such a situation can occur when the search space encompasses design points corresponding to an unphysical configuration, an ill-posed problem, or a non-computable problem due to the limitation of numerical solvers. To avoid such a premature stop in the optimization procedure, we propose to use a classification model to learn non-computable areas and to adapt the infilling strategy accordingly. Specifically, the proposed method splits the training set into two subsets composed of computable and non-computable points. A surrogate model for the objective function is built using the training set of computable points, only, whereas a probabilistic classification model is built using the union of the computable and non-computable training sets. The classifier is then incorporated in the surrogate-based optimization procedure to avoid proposing new points in the non-computable domain while improving the classification uncertainty if needed. The method has the advantage to automatically adapt both the surrogate of the objective function and the classifier during the iterative optimization process. Therefore, non-computable areas do not need to be *a priori* known. The proposed method is applied to several analytical problems presenting different types of difficulty, and to the optimization of a fully nonlinear fluid-structure interaction system. The latter problem concerns the drag minimization of a flexible hydrofoil with cavitation constraints. The efficiency of the proposed method compared favorably to a reference evolutionary algorithm, except for situations where the feasible domain is a small portion of the design space.

Keywords Global optimization · Gaussian process model · Constrained optimization · Classification

Responsible Editor: Raphael Haftka

✉ Matthieu Sacher
matthieu.sacher@ecole-navale.fr

¹ Naval Academy Research Institute - IRENAV, Brest, France

² INRIA - CNRS, Université Côte d'Azur, Valbonne, France

³ LIMSI - CNRS, Orsay, France

⁴ K-Epsilon, Valbonne, France

⁵ Groupama Team France, Lorient, France

⁶ Sirli Innovations, Biot, France

⁷ MyCFD, Golfe-Juan, France

1 Introduction

The use of surrogate models is a classical approach to reduce the computational burden related to the optimization of complex systems (Simpson et al. 2001). Polynomial models were first used as surrogates, thanks to their ease of construction. More recently, sophisticated surrogate models such as Gaussian processes (GP) (Kleijnen 2009) or Support Vector Machines (SVM) (Vapnik 1995) have emerged. Gaussian Process (GP) modeling is a statistical method to approximate functions from a finite set of observations, possibly noisy, at arbitrary points. Using the observations, it

explicitly updates the prior mean and covariance functions into their posterior counterparts, minimizing the mean-squared-error of the estimator. In the case of Gaussian observation noise, the estimator corresponds to the optimal Bayesian posterior. Contrary to alternative regression-type approaches, GP modeling does not require to prescribe explicitly a basis for the approximation space. GP surrogates have been applied to several problems such as uncertainty quantification (Marrel et al. 2009; Wang et al. 2013), binary classification (Nickisch and Rasmussen 2008), multi-levels system (Liu et al. 2016), multi-fidelity surrogate (Park et al. 2017) or clustering-based space exploration (Dong et al. 2017).

GP models have also been found especially appealing for optimization, in the framework of the Efficient Global Optimization (EGO) (Jones et al. 1998) method, because their statistical nature allows to provide both a prediction of the objective function, in terms of model mean, and an error estimate, in terms of model variance. EGO strategies have been used in several engineering applications, such as aerodynamic drag reduction of transonic wings (Liu et al. 2017), vibration reduction for rotating aircrafts (Glaz et al. 2008; 2009), optimization of FSI problems (Aghajari and Schäfer 2015) and sail trimming optimization (Sacher et al. 2017). The EGO efficiency has been demonstrated for the optimization of complex systems with costly objective function evaluations (Simpson et al. 2001).

In all these problems, a statistical criterion accounting for the Expected Improvement (EI) (Picheny et al. 2013), usually referred to as *merit function*, is used to select a new (or multiple (Li et al. 2016)) design point, at which the objective function should be evaluated. This new evaluation is used to improve the accuracy of the surrogate. However, in some situations, the analysis may not be possible for the newly selected design point. In the context of simulation-based optimization, such situations can be related for instance to the existence of unphysical configurations, ill-posed problems, or the lack of numerical robustness (weak convergence of the solvers, non-convergence to a steady solution, unstable computation, poor mesh quality for extreme configuration, etc). The origin of such failures is dependent on the application domain and an example will be provided below in the context of hydrodynamic design. Note that the automatic detection of non-computable points is a real issue in practical applications. Indeed, the origins of the failures can be very diverse and their detection is therefore not obvious. In some cases, the failure can be relatively easy to detect by monitoring the evolution of residuals or any variable of interest (e.g. forces). In other cases, the computational time might be a way to identify the simulations that are not converging. The use of clustering approaches (Jain et al. 1999) may also be an alternative to detect such computational failures. All these situations will be referred generically as cases of

non-computable design points. The occurrence of such non-computable points is critical to sequential EGO method because it leads to a premature halt of the optimization process, due to the impossibility to update the surrogate model. A possible way to overcome this problem would be to interpolate/extrapolate the non-computable objective function values, using directly the current GP model or more sophisticated methods (Zhang et al. 2017). However, the accuracy of these values and their impact on the convergence of the optimization procedure are questionable. Moreover, such methods are not always satisfactory since we may not be interested in non-computable points that correspond to unphysical designs.

Unphysical or ill-posed situations typically arise from the definition of the design space that is too large and includes non-feasible situations. Similarly, the presence of design points that are not computable because of numerical issues should be traced back to a design space that is too rich and contains extreme situations that should not be considered as potential candidates for the optimum. Otherwise, it implies that the numerical solvers considered for the evaluation of the objective function are simply not suitable for the optimization task. However, even though the numerically critical areas of the design space may not be close to the sought optimum, the EGO eventually visit them because of the infilling strategy. A naive fix to the encounter of a non-computable point consists in associating to it a prescribed high (resp. low) value for the objective function to be minimized (resp. maximized). This avoids halting the EGO procedure and prevents the future exploration of the non-computable point neighborhood. Unfortunately, the selection of an appropriate value for the non-computable points is not a trivial task with, in practice, detrimental consequences such as a significant loss of accuracy on the whole GP model approximating the objective function and a possibly strong deterioration of the EGO convergence. Alternative remedies are a better definition of the design space, to explicitly remove design points leading to unphysical or ill-posed problems, and the improvement of the numerical robustness through, e.g., the tuning of solvers parameters. Unfortunately, depending on the situation these two tasks can be extremely cumbersome, if possible at all.

In this work, we propose to combine a classical EGO method and a classification approach in order to deal with the possible existence of non-computable design areas. In this approach, the subset of the design space corresponding to computable points is progressively determined, by the Least-Squares Support Vector Machine (LS-SVM) classification method (Suykens and Vandewalle 1999), on the basis of previously classified computable / non-computable points. The SVM classification method was first proposed in the EGO context by Basudhar et al. (2012) for the treatment of discontinuous or binary constraints. We extend this idea to the case of non-computable design

areas, with the suggestion of using LS-SVM instead of SVM models. Indeed, LS-SVMs are computable at a lower cost than SVMs since the SVM formulates the classifier as a quadratic programming problem while LS-SVM deals with a set of linear equations. Specifically, each iteration step in the iterative method for solving the SVM consists of a linear system to be solved which has the same complexity as one single LS-SVM. Moreover, benchmarks in van Gestel et al. (2004) have shown that LS-SVMs are performing consistently very well in comparison with many other methods. The LS-SVM classifier provides a probability that a new point is computable. We incorporate this probabilistic information in the definition of the merit function when selecting the new design point to be analyzed. The merit function also involves a classical EI criterion, based on the GP model of the objective function constructed on the subset of computable design points. Therefore, the key advantages of the proposed method, called EGO-LS-SVM hereafter, are the following: it does not halt when a non-computable point is proposed; it does not require an explicit definition of the computable domain, but rather it constructs it; the GP model of the objective function uses only computable values and so is not polluted by arbitrary values at non-computable points. Finally, the EGO-LS-SVM method can handle optimization problems with inequality constraints and non-computable points, in a unified fashion employing the same classification strategy.

The present paper is organized as follows. The EGO-LS-SVM method is progressively introduced in Section 2, starting with the classical EGO method, presenting next the LS-SVM classification method to handle non-computable domains and inequality constraints, and finally deriving several merit functions. The efficiency of the EGO-LS-SVM method is first assessed on a series of analytical optimization problems in Section 3. A comparison with the reference evolutionary algorithm CMA-ES (Hansen 2006) is also provided. The optimization of a fully nonlinear Fluid-Structure Interaction (FSI) system involving non-computable domains is subsequently considered in Section 4. It consists in the minimization of the hydrodynamic drag of a two-dimensional flexible hydrofoil, with non-cavitation constraints, for a 5-dimensional design space defining the unloaded geometry and mechanical properties of the hydrofoil. Finally, the conclusions of the work are proposed in Section 5.

2 GP-based constrained optimization

Our objective is to estimate the solution $\mathbf{x}_{\text{opt}}$ of the following generic optimization problem

$$\mathbf{x}_{\text{opt}} = \arg \min_{\mathbf{x} \in \Omega} f(\mathbf{x}), \quad \text{s.t. } \mathbf{Q}(\mathbf{x}) \geq \mathbf{0}, \quad (1)$$

where $\mathbf{x}$ is the vector of design parameters, $\Omega \subset \mathbb{R}^d$ is the optimization domain, $f : \Omega \mapsto \mathbb{R}$ the objective function and $\mathbf{Q} : \mathbf{x} \mapsto \mathbb{R}^m$ is the vector of (non-linear) constraints. Even in the unconstrained case, $m = 0$, finding the global optimum of f can be very costly, in particular when its evaluation is numerically expensive. The use of surrogate models in place of f is then a classical approach to reduce the computational burden related to the optimization of complex systems (Simpson et al. 2001). In this work, we consider the use of Gaussian processes (GP) (Kleijnen 2009) which, owing to their statistical nature, provide both a prediction of the objective function and a measure of the uncertainty (variance) in the prediction. These features are appealing in optimization, as they can be used to derive rigorous optimization strategies based on the maximization of the Expected Improvement (EI) criterion. These methods are globally referred to as the Efficient Global Optimization (EGO) (Jones et al. 1998) methods. A summary of the construction of the GP model for f is provided in Section 2.1, as well as the resulting optimization strategy in the unconstrained case. The classification approach to account for constraints and non-computable subsets of points is considered in Section 2.2.

2.1 Unconstrained EGO using GP

Consider a set of n training inputs points $\mathbf{X}_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, each associated to a noisy observation y_i of the objective function f . It is assumed that $y_i = f(\mathbf{x}_i) + \varepsilon_i$, where the ε_i are Gaussian measurement noises, assumed for simplicity independent and identically distributed with variance σ_ε^2 . The GP construction considers that $f(\mathbf{x})$ is a realization of a zero-mean multivariate Gaussian process with covariance function C_f . In this work, we consider the multidimensional squared exponential covariance functions defined by

$$C_f(\mathbf{x}, \mathbf{x}'; \Theta) \doteq \theta_1 \prod_{i=1}^d \exp\left(-\frac{(x_i - x'_i)^2}{2l_i^2}\right) + \theta_2, \quad (2)$$

where $\Theta = \{\theta_1, \theta_2, l_1, l_2, \dots, l_d\}$ is a vector of hyperparameters to be inferred from the observations. Denoting $\mathbf{C}(\Theta) \in \mathbb{R}^{n \times n}$ the covariance matrix with entries $C_{i,j}(\Theta) \doteq C_f(\mathbf{x}_i, \mathbf{x}_j; \Theta)$, $1 \leq i, j \leq n$, the joint Gaussian distribution of the noisy observation vector, $\mathbf{Y}_n = (y_1, \dots, y_n)^T$, and the predicted observation $y(\mathbf{x})$ is given by

$$\begin{pmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{C} + \sigma_\varepsilon^2 \mathbf{I} & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^T(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\varepsilon^2 \end{bmatrix}\right). \quad (3)$$

In (3) we have denoted $\kappa(\mathbf{x}) \doteq C_f(\mathbf{x}, \mathbf{x}; \Theta)$, $\mathbf{k}(\mathbf{x}) \doteq (C_f(\mathbf{x}, \mathbf{x}_1; \Theta) \cdots C_f(\mathbf{x}, \mathbf{x}_n; \Theta))^T$ and $\mathbf{I}$ the identity matrix of $\mathbb{R}^n$. From the conditional rules of joint Gaussian distributions (Rasmussen and Williams 2006), the best

prediction $\hat{f}(\mathbf{x})$ of $f(\mathbf{x})$, i.e. the mean of y , and the prediction variance $\hat{\sigma}_f^2(\mathbf{x})$ are given by

$$\hat{f}(\mathbf{x}) = \mathbf{k}^T(\mathbf{x}) \left(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I} \right)^{-1} \mathbf{Y}_n, \quad (4)$$

$$\hat{\sigma}_f^2(\mathbf{x}) = \kappa(\mathbf{x}) + \sigma_\epsilon^2 - \mathbf{k}^T(\mathbf{x}) \left(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I} \right)^{-1} \mathbf{k}(\mathbf{x}). \quad (5)$$

The hyper-parameters Θ and noise variance σ_ϵ^2 can be determined by maximizing the log-marginal likelihood (Rasmussen and Williams 2006) using the ‘‘Covariance Matrix Adaptation Evolution Strategies’’ (CMA-ES) algorithm (Hansen 2006). This powerful evolutionary based optimizer is an improvement of the original $(\mu/\rho^+ \lambda)$ -ES variants (Hansen et al. 2015), thanks to the covariance matrix adaptation property. Indeed, the CMA-ES allows generations of new individuals that are possibly not distributed along the principal axis of the covariance matrix. In the present work, we apply this evolutionary algorithm by using the recommended parameters values (step-size $\sigma = 0.5$, population size $\lambda = 4 + \lfloor 3 \log d \rfloor$, number of parents for recombination $\mu = \frac{\lambda}{2}$, etc.) of Hansen (2006). A complete tutorial of the CMA-ES algorithm, including a source code and best uses, is available in Hansen (2016). More details on GP meta-models can also be found in Rasmussen and Williams (2006).

Let $\hat{\mathbf{x}}_n$ be the optimum of $\hat{f}(\mathbf{x})$. It is expected that $\hat{\mathbf{x}}_n \approx \mathbf{x}_{\text{opt}}$ if the approximation error $\hat{f} - f$ is small enough. The advantage of minimizing $\hat{f}$ instead of f is that GP models are usually inexpensive to evaluate compared to the original objective function. To control the error in the approximation, one proceeds sequentially by adding progressively new points in the area of interest. A deterministic optimization procedure would choose the next point $\mathbf{x}_{n+1}$ as the optimal point of $\hat{f}$. However, the GP model provides probabilistic information that can be exploited to propose more robust strategies based on merit functions, which combine the prediction and its variance. In this work, we use the Augmented Expected Improvement (AEI) merit function (Huang et al. 2006), which estimates the expected progress in the cost, taking into account the noise in the observed values and the prediction variance:

$$AEI(\mathbf{x}) = EI(\mathbf{x}) \left(1 - \frac{\sigma_\epsilon}{\sqrt{\hat{\sigma}_f^2(\mathbf{x}) + \sigma_\epsilon^2}} \right), \quad (6)$$

where the Expected Improvement $EI(\mathbf{x})$ is defined by

$$EI(\mathbf{x}) = \hat{\sigma}_f(\mathbf{x}) [u(\mathbf{x})\Phi(u(\mathbf{x})) + \phi(u(\mathbf{x}))], \quad (7)$$

$$u(\mathbf{x}) = \frac{\hat{f}(\mathbf{x}_{*,n}) - \hat{f}(\mathbf{x})}{\hat{\sigma}_f(\mathbf{x})}, \quad (8)$$

with Φ and ϕ the cumulative and density functions of the standard Gaussian distribution, and $\mathbf{x}_{*,n} \in \mathbf{X}_n$ is the current effective best solution (see Huang et al. (2006)).

The optimum $\mathbf{x}_{n+1}$ of the AEI is added to $\mathbf{X}_n$, and f is evaluated at the new point providing y_{n+1} . Setting $n \leftarrow n + 1$, a new iteration can start updating the GP model. Overall, each iteration requires one computation of the cost and the resolution of two optimization problems: a first one for the hyper-parameters of the GP model, and a second one to find the AEI optimum. The iterations of the GP-based optimization problem are continued until a stopping criterion is satisfied or the resources allocated to the optimization have been exhausted. Obviously, if f is not computable for the proposed $\mathbf{x}_{n+1}$, the optimization process is stopped prematurely, because $\mathbf{x}_{n+1}$ and y_{n+1} cannot be added to $\mathbf{X}_n$ and $\mathbf{Y}_n$.

2.2 Classification method

As underlined in the introduction, the proposed approach based on a classification method is closely related to the handling of inequality constraints in the optimization problem (1). Indeed, the non-computable simulations are usually located in some specific regions of the design space. Therefore, the presence of non-computable points can be managed as inequality constraints if these regions of failures can be identified thanks to a probabilistic formulation. In this framework, a failure probability would be used within the optimization process to avoid the non-computable simulations.

EGO methods with inequality constraints were considered by Schonlau (1997). The key idea followed in this reference is to rely on m additional surrogates to estimate the constraints $Q_i(\mathbf{x})$. For Gaussian Process models, one can easily determine the probability $P_i(Q_i \geq 0|\mathbf{x})$ that the constraint Q_i is satisfied at $\mathbf{x}$. Assuming the independence of the constraints probability, the consolidated probability $P(\mathbf{Q} \geq \mathbf{0}|\mathbf{x}) = \prod_{i=1}^m P_i(Q_i \geq 0|\mathbf{x})$ is obtained. This probability is used to modify the unconstrained AEI criterion and favor regions feasible with a probability:

$$AEIQ(\mathbf{x}) = AEI(\mathbf{x})P(\mathbf{Q} \geq \mathbf{0}|\mathbf{x}). \quad (9)$$

Although effective in many problems, this GP modeling of the constraints faces several limits. First, its computational cost increases with the number m of constraints and can be an issue for problems with large m . Second, the approximation by GP models assumes a sufficient smoothness of the Q_i . This rules out the case of binary constraints (feasible/infeasible). Another situation, non-amenable to the GP approximations of the constraints, is when the search domain Ω possesses a sub-domain Ω_χ over which f is non-computable, i.e. undefined, but there is no explicit expression or a priori knowledge of this sub-domain. As a result, Ω_χ cannot be expressed in terms of inequality constraints.

A classification approach to efficient global optimization in presence of non-computable domains

Classification methods recently proposed by Basudhar et al. (2012) are better suited to deal with discontinuous and binary constraints in GP-based optimization procedures. Therefore, we aim at using this approach to define a unique framework to deal with inequality constraints and non-explicit feasible domain definition in presence of a non-computable domain $\Omega_\chi \subset \Omega$. In essence, classification is based on the reformulation of problem (1) into the unconstrained form

$$\min_{\mathbf{x} \in \Omega_{\text{adm}}} f(\mathbf{x}), \quad \Omega_{\text{adm}} := \Omega_Q \setminus (\Omega_\chi \cap \Omega_Q), \quad (10)$$

where $\Omega_Q := \{\mathbf{x} \in \Omega, Q_{i=1,\dots,m}(\mathbf{x}) \geq 0\}$ is the subset of Ω satisfying the inequality constraints. As the admissible set Ω_{adm} has no explicit form, in general, we rely on a binary classifier with two classes C^+ and C^- over Ω , corresponding to the admissible ($\mathbf{x}$ satisfies all constraints and $f(\mathbf{x})$ is computable) and non-admissible ($\mathbf{x}$ does not satisfy all constraints or $f(\mathbf{x})$ is not computable) domains respectively. To construct this classifier, each $\mathbf{x}_i$ of $\mathbf{X}_n$ is equipped with a value $z_i = \pm 1$ depending on its membership $C^\pm$. To predict the class of a new point $\mathbf{x}$ we introduce a classification function $h : \mathbf{x} \in \Omega \rightarrow \mathbb{R}$, such that $z(\mathbf{x}) = \text{sign}h(\mathbf{x})$. A Least-Squares Support Vector Machine (LS-SVM) (Suykens and Vandewalle 1999) is used to construct h . The LS-SVM method extends the original Support Vector Machine (SVM) classifier (Vapnik 1995) to quadratic penalization, resulting in a linear (but non-sparse) system to be solved (see below). Further details and discussion on SVM and LS-SVM methods can be found in Cawley (2006).

The LS-SVM method (Suykens and Vandewalle 1999) is a linear classifier, for C^+ and C^- , in a feature space induced by the transformation $\phi : \Omega \rightarrow \mathcal{F}$:

$$h(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b. \quad (11)$$

Here, $\mathbf{w}$ and $\phi(\mathbf{x})$ are the weights and features vectors, while $b \in \mathbb{R}$ is a constant. The feature space $\mathcal{F}$ is generated by a reproducing kernel $r : \Omega \times \Omega \rightarrow \mathbb{R}$ representing the inner product between images in $\mathcal{F}$ of vectors: $r(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x}) \cdot \phi(\mathbf{x}')$. We use the classical Gaussian kernel,

$$r(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\lambda^2}\right), \quad (12)$$

with scale factor $\lambda \in \mathbb{R}$ to be adjusted. The LS-SVM parameters $(\mathbf{w}, b)$ satisfy the primal constrained optimization problem

$$\begin{aligned} \min_{\mathbf{w}, b, \mathbf{e}} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2, \\ \text{s.t.} \quad & z_i = \mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, n, \end{aligned} \quad (13)$$

with trade-off parameter $\gamma \in \mathbb{R}^+$ (to be fixed) and relaxation variables e_i allowing for miss-classification. The Lagrangian of this optimal problem is

$$\mathcal{L}(\mathbf{w}, b, \alpha_n, \mathbf{e}) = \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2 - \sum_{i=1}^n \alpha_i \left(\mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i - z_i \right), \quad (14)$$

where the $\alpha_i \in \mathbb{R}$ are the Lagrange multipliers of the constraints. Denoting $\mathbf{Z}_n = (z_1 \dots z_n)^T$, the optimality conditions of the Lagrangian (15)

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = \mathbf{0} \rightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i \phi(\mathbf{x}_i), \\ \frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i = 0, \\ \frac{\partial \mathcal{L}}{\partial \alpha_n} = \mathbf{0} \rightarrow \alpha_i = \gamma e_i, \quad i = 1, \dots, n, \\ \frac{\partial \mathcal{L}}{\partial \mathbf{e}} = \mathbf{0} \rightarrow \mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i - z_i = 0, \quad i = 1, \dots, n, \end{cases} \quad (15)$$

are used to derive a linear system for the dual model parameters b and $\alpha_n = (\alpha_1 \dots \alpha_n)^T$,

$$\begin{bmatrix} \mathbf{R} + \gamma^{-1} \mathbf{I} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \alpha_n \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_n \\ 0 \end{bmatrix}, \quad (16)$$

where $\mathbf{R} \in \mathbb{R}^{n \times n}$ is the kernel matrix with $R_{i,j} = r(\mathbf{x}_i, \mathbf{x}_j)$. This system is solved via a Cholesky factorization (Cawley 2006), and h in (11) is expressed in terms of dual model parameters to obtain

$$h(\mathbf{x}) = \sum_{i=1}^n \alpha_i r(\mathbf{x}_i, \mathbf{x}) + b. \quad (17)$$

The LS-SVM classifier depends on two parameters, γ and λ , to be determined when solving (16). This step is called *model selection* and it is usually performed with cross-validation approaches (Stone 1974). The most common method is the k -fold cross-validation, in which the training set is split into k disjoint subsets. Then, the LS-SVM model is sequentially trained on $k - 1$ subsets and the remaining subset is used to evaluate the classification efficiency by computing the error rate. The global k -fold estimate of the error rate is the average of the k error rates. The extreme case where $k = n$ is called the Leave-One-Out (LOO) cross-validation and it allows an almost unbiased estimate of the error rate (Vapnik and Vapnik 1998). Unfortunately, the LOO computation is too expensive to be applied on common training sets. However, in the specific case of LS-SVM classifiers, it was shown (Cawley and Talbot 2003, 2004, 2007) that exact LOO can be performed in only $\mathcal{O}(n)$ operations. In this work, we rely on this efficient LOO procedure to determine γ and λ that

minimize the Predicted Residual Sum-of-Squares (PRESS) criterion (Allen 1974),

$$\text{PRESS} = \sum_{i=1}^n \left(z_i - \hat{z}_i^{(-i)} \right)^2, \quad (18)$$

where $\hat{z}_i^{(-i)}$ is the predicted output $z(\mathbf{x}_i)$ of the LS-SVM, when the i -th training point $(\mathbf{x}_i, z_i)$ is disregarded from the construction of the classifier. A simple expression for the predicted residuals is proposed in Cawley and Talbot (2007),

$$z_i - \hat{z}_i^{(-i)} = \frac{\alpha_i}{D_{ii}^{-1}}, \quad i = 1, \dots, n, \quad (19)$$

where the D_{ii}^{-1} are the diagonal entries of the inverse of $\mathbf{D}$, the matrix of system (16). An explicit expression is also provided in Cawley and Talbot (2007) to compute D_{ii}^{-1} during the factorization of $\mathbf{D}$. The minimization of the PRESS (18) for λ and γ is performed with the CMA-ES algorithm (Hansen 2006) in all the examples presented below.

The LS-SVM binary classifier is finally extended to a probabilistic classification, relating h in (17) to the probability of the class C^+ , denoted $P(C^+|\mathbf{x})$. A comparison of several probability models for the LS-SVM classification is provided in Van Calster et al. (2007). We use the sigmoid function (Platt 1999) and expressing the probability of C^+ as

$$P(C^+|\mathbf{x}) = \frac{1}{1 + \exp(Ah(\mathbf{x}) + B)}. \quad (20)$$

The parameters A and B of the sigmoid are determined by minimizing the probability of misclassification, see Platt (1999) and Lin et al. (2007). In practice, the probability P goes to 1 (resp. 0) as the classifier is certain that $\mathbf{x} \in \Omega_{\text{adm}}$ and belongs to C^+ ($\mathbf{x} \notin \Omega_{\text{adm}}$ and belongs to C^-), while a value of $P = 1/2$ denotes a complete uncertainty in the classification. This can occur when $\mathbf{x}$ is far from any observations in $\mathbf{X}_n$ or close to the interface between the two classes.

2.3 EGO-LS-SVM method

2.3.1 Extended merit functions

We now return to the optimization problem and introduce the EGO-based method that we call EGO-LS-SVM for it is based on the LS-SVM classifier. The EGO-LS-SVM method uses the probability $P(C^+|\mathbf{x})$ to derive a merit function from the AEI, in order to select a new point $\mathbf{x}_{n+1} \in \Omega$. Following (Schonlau 1997; Basudhar et al. 2012), the selected point should present high expected improvement (relatively to another point) of the objective function value and a high probability of belonging to Ω_{adm} .

These considerations lead us to propose the following sequential infilling strategy, selecting alternatively one of the following definitions for $\mathbf{x}_{n+1}$ in the EGO-LS-SVM method:

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \Omega} AEI(\mathbf{x}) P(C^+|\mathbf{x}), \quad (21)$$

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \Omega} AEI(\mathbf{x}) \quad \text{s.t.} \quad P(C^+|\mathbf{x}) \geq \rho, \quad (22)$$

$$\mathbf{x}_{n+1} = \arg \min_{\mathbf{x} \in \Omega} \hat{f}(\mathbf{x}) \quad \text{s.t.} \quad P(C^+|\mathbf{x}) \geq \rho, \quad (23)$$

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \Omega} [AEI(\mathbf{x}) P(C^+|\mathbf{x}) (1 - P(C^+|\mathbf{x}))]. \quad (24)$$

The first definition in (21) corresponds to the extension of the AEI favoring points with high chance of feasibility. The second expression (22) maximizes the original AEI definition by enforcing a minimal probability ρ of feasibility; we use $\rho = 0.5$ in all subsequent computations. In (23), the predicted objective function is directly minimized over the approximated feasible domain defined by $\{\mathbf{x} \in \Omega, P(C^+|\mathbf{x}) > \rho\}$. Finally (24) combines the two classes probabilities to favor areas where the classification is the most uncertain ($P \sim 0.5$), to improve the exploration along the estimated feasible domain boundaries.

Finally the new points $\mathbf{x}_{n+1}$ are determined using CMA-ES algorithms without (Hansen 2006) or with constraints (Arnold and Hansen 2012), depending on the considered definition. Whence the new point has been selected, the constraints Q_i and the objective function f are evaluated at the selected point. If $\mathbf{x}_{n+1} \notin \Omega_{\chi}$, that is $f(\mathbf{x}_{n+1})$ is computable we add $f(\mathbf{x}_{n+1})$ to the training set and update the GP model of f . Otherwise the new point is simply considered as a missing data and disregarded in the subsequent constructions of the GP model of f . In any cases, we set $z_{n+1} = +1$ if $f(\mathbf{x}_{n+1})$ is computable and $Q(\mathbf{x}_{n+1}) \leq 0$, and $z_{n+1} = -1$ otherwise. Therefore, the GP model of f and the LS-SVM classification may involve different numbers of observations if some of them fall in Ω_{χ} . Note that if $\mathbf{x}_{n+1} \in \Omega_{\chi}$, we have $\hat{f}_{n+1} = \hat{f}$, but because of the update of $P(C^+|\mathbf{x})$ we shall have in general $\mathbf{x}_{n+2} \neq \mathbf{x}_{n+1}$. Consequently, the EGO-LS-SVM method will improve the estimation of Ω_{adm} till it proposes a new point $\mathbf{x}_{n+1} \in \Omega_{\text{adm}}$ to update $\hat{f}$.

Compare to the similar classification based approach proposed in Basudhar et al. (2012), we remark that our approach is based on the original sigmoid probability in (34) such that $P(C^+|\mathbf{x})$ remains differentiable, facilitating the search of $\mathbf{x}_{n+1}$. Second, a single point is added and have to be evaluated per EGO iteration, where the approach in Basudhar et al. (2012) enriches the training set with $d + 1$ new points at each iteration. The computational cost to perform these $d + 1$ evaluations of f is expected to be too important in many application and is not necessary for the LS-SVM classification procedure. This is demonstrated

on a numerical example later, in Section 3.5, where we discuss further the differences and respective efficiencies of the method of Basudhar et al. (2012) and our.

The work-flow of one iteration step of the EGO-LS-SVM method is summarized in the Algorithm 1 below. In the full optimization process, the Algorithm 1 is called iteratively until a convergence is reached or a maximum workload has been reached. The lines 1 and 2 correspond to the GP surrogate construction of the objective function f . This allows computing the mean and variance predictions given in (4) and (5). Lines 3 to 5 concern the LS-SVM classification model with the classification probability in (20) as output. To do this, a first optimization problem is solved for the parameters (γ, λ) (line 3). Then, the system (16) is solved, at line 4, to determine (α_n, b) in (17). Finally, the coefficients A and B in (20) are computed by Newton's method (Lin et al. 2007). To the end, the new point $\mathbf{x}_{n+1}$ is determined at line 6 by solving a last optimization problem for one of the merit functions in (21)–(24).

Algorithm 1 Work-flow of one EGO-LS-SVM iteration.

Input: Sample set points $\mathbf{X}_n$, objective function values $\mathbf{Y}_n$, class labels $\mathbf{Z}_n$

Output: New selected point $\mathbf{x}_{n+1}$

- ```

!Build GP surrogate!
1 Find $(\Theta, \sigma_\epsilon^2)$ maximizing the log-likelihood !Using
 CMA-ES
2 Factorize $(\mathbf{C}(\Theta) + \sigma_\epsilon^2 \mathbf{I})$!With LU decomposition
 !Build LS-SVM model!
3 Find (γ, λ) minimizing the PRESS criterion (18)
 !Using (19) LOO expression and CMA-ES
4 Compute (α_n, b) solving system (16) !Using
 Cholesky factorization
5 Find (A, B) minimizing the misclassification
 probability !Using Newton's method
 !Merit function!
6 Find $\mathbf{x}_{n+1}$ the extended merit optimum !Using one
 of equations (21)–(24) and CMA-ES

```
- 

### 2.3.2 Convergence assessment

In the following sections, we are assessing and comparing the efficiency of the proposed EGO-LS-SVM method on several problems. To this end, we need to monitor the convergence of the sequence of optimization iteration. The convergence can be characterized in different ways. For instance one may report the best point in the training set,  $\mathbf{x}_{\text{best}} \in \mathbf{X}_n$ , corresponding to the smallest objective function value  $f_{\text{best}} \doteq f(\mathbf{x}_{\text{best}})$ . Note that  $f_{\text{best}}$  is an upper bound of the true optimal value, and that  $f_{\text{best}}$  can only improve as the optimization iterations proceeds. Alternatively, in the

analytical tests presented in the next section, where the objective function is easily evaluated, one can determine at each iteration the approximated optimal point  $\mathbf{x}_{\text{opt}}^*$  and objective function value  $f_{\text{opt}}^* \doteq f(\mathbf{x}_{\text{opt}}^*)$  associated to the current state of knowledge of  $f$  and computable domain  $P(C^+|\mathbf{x})$ . Specifically, one can solve for  $\mathbf{x}_{\text{opt}}^*$  the following constrained optimization problem

$$\mathbf{x}_{\text{opt}}^* = \arg \min_{\mathbf{x} \in \Omega} \hat{f}(\mathbf{x}) \quad \text{s.t.} \quad P(C^+|\mathbf{x}) \geq 0.5, \quad (25)$$

and subsequently evaluate  $f_{\text{opt}}^* = f(\mathbf{x}_{\text{opt}}^*)$ . It is noted, however, that  $\mathbf{x}_{\text{opt}}^*$  may not be a computable point since the classifier may have a non-zero misclassification rate. Further, when the exact solution of the optimization problem is known, one can report the distance to current estimate  $\mathbf{x}_{\text{opt}}^*$  to the optimum true optimum  $\mathbf{x}_{\text{opt}}$ , in terms of Euclidian distance  $\|\mathbf{x}_{\text{opt}} - \mathbf{x}_{\text{opt}}^*\|$  or objective function values  $f(\mathbf{x}_{\text{opt}}^*) - f(\mathbf{x}_{\text{opt}}) \geq 0$ .

### 2.3.3 Illustration of the infilling strategy

Before applying the EGO-LS-SVM method to generic problems, we first illustrate the properties of the different merit functions in the criteria (21)–(24), and the benefit of applying them in a sequential fashion. To this end, we consider the one-dimensional domain  $\Omega = [0, 1]$  and the *Ricker Wavelet* objective function  $f(x) = (1 - 2\pi^2 x^2) \exp(-\pi^2 x^2)$ . We define the set of non-computable points through

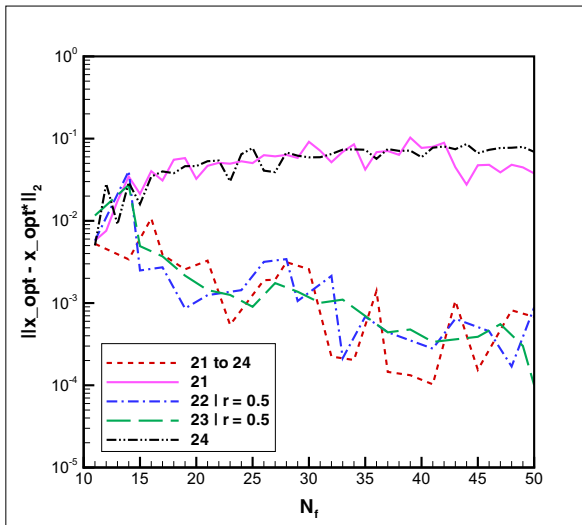
$$\Omega_\chi = \{x \in \Omega : |x - x_{\min}| < 0.2\}, \quad (26)$$

where  $x_{\min} = \frac{\sqrt{3/2}}{\pi}$  is the minimum of  $f$  in  $\Omega$ . Letting  $\Omega_{\text{adm}} = \Omega \setminus \Omega_\chi$ , the optimization problem becomes

$$\min_{x \in \Omega_{\text{adm}}} f(x) = (1 - 2\pi^2 x^2) \exp(-\pi^2 x^2). \quad (27)$$

The optimum of (27) is found on the boundary of  $\Omega_{\text{adm}}$  at  $x_{\text{opt}} = x_{\min} + 0.2$ . We solve this problem for several infilling strategies, with an initial sample set with size  $n_{\text{LHS}} = 10$  observations, and fix the maximum number of calls to  $f$  to  $N_f = 50$  (thus 40 optimization iterations are performed). Figure 1 compares the distances to the optimum as a function of  $N_f$  for the sequential use of all the merit functions (21)–(24), or using systematically only one of them. It is seen that using only (21) (pink curve) or (24) (black curve) to drive the procedure does not yield satisfactory results on this example. In contrast, using (22) alone (blue curve) or (23) alone (green curve) yields a convergence similar to that obtained using the four merit functions sequentially (red curve). These behaviors can be explained on this simple problem: because the criteria (22) and (23) produce points closer to the boundary than the unconstrained merit functions in (21) and (24), they are more efficient to recover an optimum located





**Fig. 1** Convergence with the number of calls  $N_f$  of the distance to optimum,  $\|\mathbf{x}_{\text{opt}} - \mathbf{x}_{\text{opt}}^*\|_2$ . Values are only reported for computable optimum  $\mathbf{x}_{\text{opt}}^*$  solution of (25). Case of problem (27)

on the boundary. Nevertheless, using sequentially all the merit functions (21)–(24) does not impact negatively the convergence. We thus advise for the alternative use of all the four merit functions in a sequential infilling strategy. This setting favors both the exploration of promising areas (with large AEI) of the design space and of the most uncertain boundaries of the computable domain ( $P \sim 0.5$ ).

Finally, we report in Fig. 2 the convergence of EGO-LS-SVM when using the merit functions (22) alone (left plot), or (23) alone (right plot), and for different values of the threshold probability  $\rho = 0.2, 0.5$  and  $0.8$ . Again, the curves depict the distance to the optimum as a function of the total number  $N_f$  of calls to  $f$ . These plots show that the value of  $\rho$  seems to have a moderate effect on the convergence in the present example. In addition, using  $\rho = 0.5$  appears to be the best choice for the two merit functions. This value will be considered as the default value in the following experiments.

### 3 Analytical test problems

In this section, we illustrate the accuracy and the computational efficiency of the EGO-LS-SVM approach by solving few optimization problems defined analytically. We consider problems with objective functions that are not computable over a subset  $\Omega_\chi$  of the full design space  $\Omega$ . The classification approach is used to estimate the admissible (i.e. computable) domain  $\Omega_{\text{adm}} := \Omega \setminus \Omega_\chi$ , and the GP model of  $f$  is constructed using only evaluations at points proposed in  $\Omega_{\text{adm}}$ . For reference and comparison purposes,

we rely on the evolutionary algorithm CMA-ES (Hansen 2006) to solve the different problems, employing a high penalization value  $f(\mathbf{x}) = f_\infty$  for the individual  $\mathbf{x} \notin \Omega_{\text{adm}}$ . This penalization  $f_\infty$  and other parameters of the CMA-ES method (population size, cross-over parameter...) are carefully set for each optimization problem, following the recommendations (Hansen 2006).

The selected test problems are presented in the following. A simple two-dimensional illustrative example is firstly proposed in Section 3.1; a multi-dimensional optimization problem is considered in Section 3.2, which includes optimizations in increasing dimensions and calculations with different initial sample set sizes; Section 3.3 investigate the behavior of the EGO-LS-SVM method in the case of a multi-dimensional optimization problem with complex domain  $\Omega_\chi$  corresponding to a large portion of  $\Omega$  (case of an over-constrained problem); Section 3.4 proposes the minimization of a problem where the domain  $\Omega_{\text{adm}}$  and objective function  $f$  are defined probabilistically; finally Section 3.5 concerns the efficiency comparison of the original SVM based approach (Basudhar et al. 2012) to our new EGO-LS-SVM method.

#### 3.1 Two-dimensional illustrative example

We consider the two-dimensional domain  $\Omega = [0, 4]^2$  and the simple convex objective function  $f(\mathbf{x}) = x_1^2 + x_2^2$ . We define the set of non-computable points through

$$\Omega_\chi = \{\mathbf{x} \in \Omega : c_1(\mathbf{x}) < 0, c_2(\mathbf{x}) < 0\}, \quad (28)$$

where  $c_1(\mathbf{x}) = 0.25x_1^2 + 0.75x_2^2 - 1$  and  $c_2(\mathbf{x}) = 0.75x_1^2 + 0.25x_2^2 - 1$ . Letting  $\Omega_{\text{adm}} = \Omega \setminus \Omega_\chi$ , the optimization problem becomes

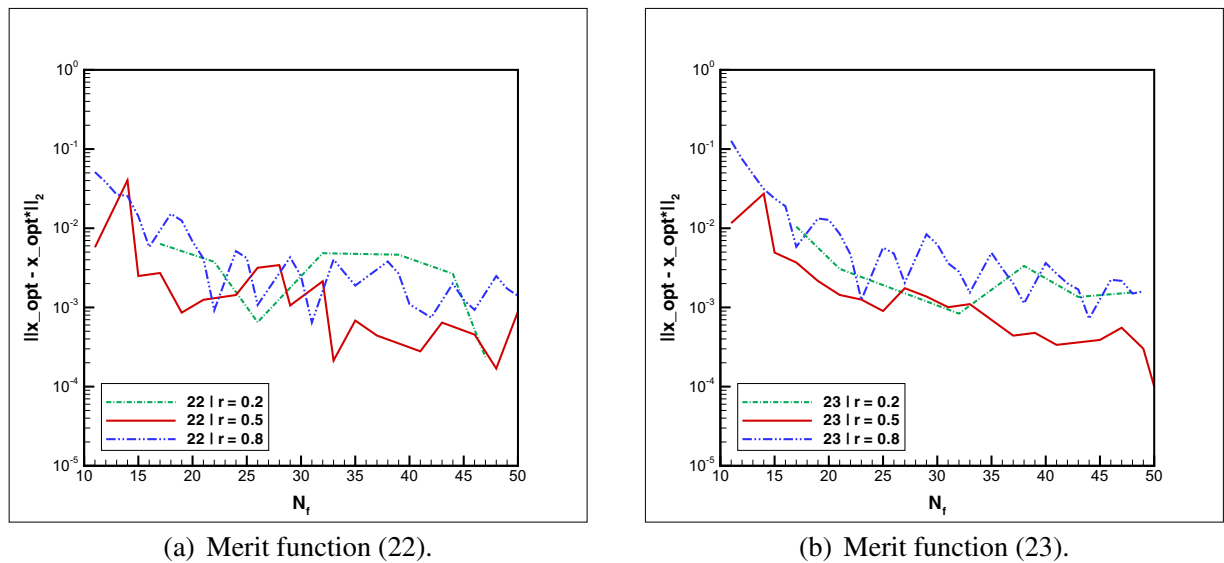
$$\min_{\mathbf{x} \in \Omega_{\text{adm}}} f(\mathbf{x}) = x_1^2 + x_2^2. \quad (29)$$

The exact optimum of the problem is found to be on the boundary of  $\Omega_{\text{adm}}$  at  $\mathbf{x}_{\text{opt}} = (1, 1)$ .

The EGO-LS-SVM algorithm is initialized using a Latin Hypercube Sampling (LHS) set of 15 points drawn randomly in  $\Omega$ . The optimization iterations are repeated till the AEI merit function reaches a value less than  $10^{-3}$ . In the CMA-ES method, the optimization is carried-out until the estimated optimum is in  $\Omega_{\text{adm}}$  with a value of  $f$  less than the EGO-LS-SVM solution. As mentioned before, a large penalization  $f_\infty = 100$  is applied to points not belonging to  $\Omega_{\text{adm}}$  in the CMA-ES algorithm.

Figure 3 shows the convergence of the error (red squares and blue triangles, left axis) with the cumulated number of calls to (attempts to evaluate)  $f$ , denoted  $N_f$ . Note that  $N_f$  is also accounting for calls to  $f$  with  $\mathbf{x} \notin \Omega_{\text{adm}}$ . The error on the solution is defined as the Euclidian distance between the current estimate  $\hat{\mathbf{x}}_{\text{opt}}$  of the optimum and the exact one,

A classification approach to efficient global optimization in presence of non-computable domains

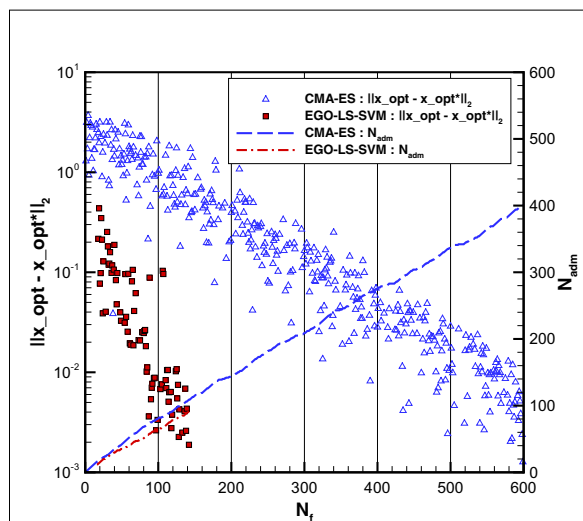


**Fig. 2** Convergence with the number of calls  $N_f$  of the distance to optimum,  $\|x_{\text{opt}} - x_{\text{opt}}^*\|_2$ , for the infilling strategy based on single merit function (22) (left) or (23) (right), and for different threshold probabilities  $\rho$ . Values are only reported for computable optimum  $x_{\text{opt}}^*$  solution of (25). Case of problem (27)

$x_{\text{opt}}$ . For the EGO-LS-SVM method,  $\hat{x}_{\text{opt}}$  is the solution of (29) where  $f$  is substituted with its GP model. For the CMA-ES approach, the distance to the exact optimum of each member of the population in  $\Omega_{\text{adm}}$  is reported. The number  $N_{\text{adm}}$  of computable calls to  $f$  is also reported for both methods, using red and blue lines (right axis) in the same Figure. The EGO-LS-SVM (squares and red line)

terminated at  $N_f = 142$ , with a number of computable calls equal to 91, such that roughly 50 points have been proposed in the non-feasible region  $\Omega_{\chi}$ . The evolutionary method needs 599 calls to  $f$  before it finds a better optimum than the EGO-LS-SVM one, among which 401 were for computable points. This result shows that, in this example, the EGO-LS-SVM method is able to properly estimate the feasible domain  $\Omega_{\text{adm}}$  in the neighborhood of the optimum, and converges to this optimum at a significantly faster rate than the CMA-ES approach.

To better appreciate the effectiveness of the classification approach, contours of  $P(C^+|\mathbf{x})$  are reported in Fig. 4 at  $N_f = 50$  and 100. Also shown in the plots are the iso-value corresponding to  $P(C^+|\mathbf{x}) = 0.5$  (dashed white line), proposed computable points (white squares) and non-computable points (black triangles). After 50 calls to  $f$ , the classification has well discovered  $\Omega_{\text{adm}}$  and the level  $P(C^+|\mathbf{x}) = 0.5$  is not far from the actual boundary of  $\Omega_{\text{adm}}$ . After 100 calls to  $f$ , a large fraction of the proposed points are concentrated around the optimum and the contour  $P(C^+|\mathbf{x}) = 0.5$  is not following anymore the boundary of  $\Omega_{\text{adm}}$ , except in the immediate neighbor of the optimum. Farther from the optimum in areas less explored, the classification probability quickly returns to  $P(C^+|\mathbf{x}) \approx 0.5$  as we move away from observed points. This behavior is due to the classification parameters  $\lambda$  and  $\gamma$  (see Section 2.2) that are adjusted to capture the steep transition in  $P(C^+|\mathbf{x})$  at the interface between  $\Omega_{\text{adm}}$  and  $\Omega_{\chi}$ , where the optimum point  $x_{\text{opt}}$  is located. This behavior is desirable because it allows considering the exploration of



**Fig. 3** Distance to the optimum ( $\|x_{\text{opt}} - x_{\text{opt}}^*\|_2$ ) and cumulated number of computable calls to  $f$  ( $N_{\text{adm}}$ ) as functions of the total number of calls  $N_f$  and for the EGO-LS-SVM and CMA-ES optimization procedures. Distance is reported only for computable  $x_{\text{opt}}^*$  solution of (25). Case of problem (29)

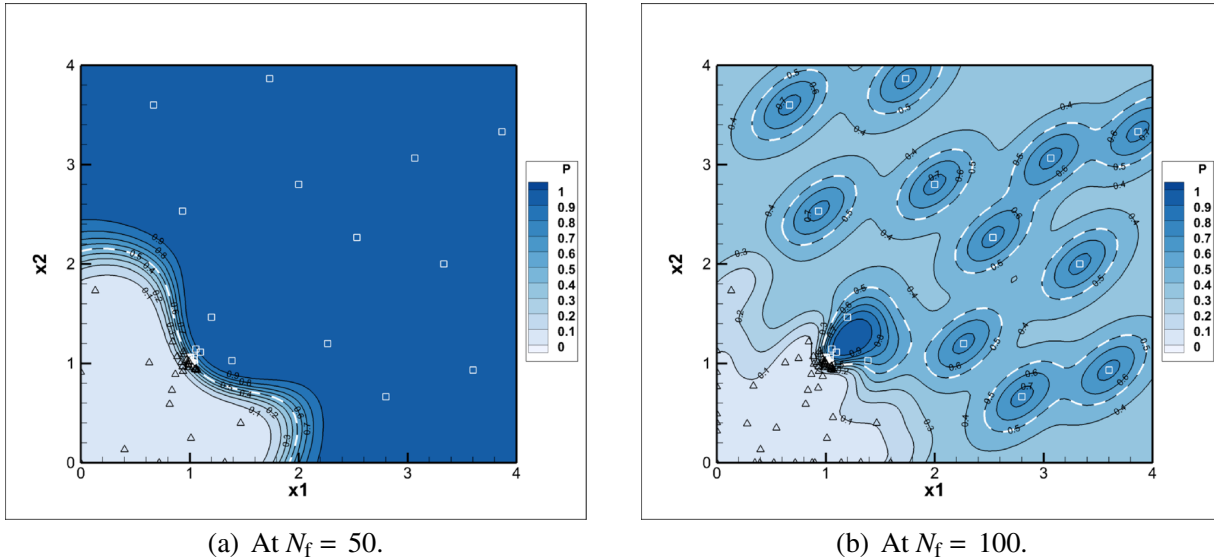


Fig. 4 Contours of the classification probability  $P(C^+|\mathbf{x})$  after 50 and 100 calls. Case of problem (29)

areas that were initially classified as unfeasible with high probability if they exhibit large merit function AEI. Indeed, the plot also reveals several calls to  $f$  for points in  $\Omega_\chi$ , far from the optimum, that confirm non-feasibility and balance large merit function values (recall that the unconstrained optimum is at the origin).

### 3.2 Higher dimensional problems

#### 3.2.1 Influence of dimensionality

To investigate the behavior of the EGO-LS-SVM method as the dimensionality of the search space  $\Omega$  is increasing, we propose the following problem where no function value is provided for points not belonging to the feasible domain. Let  $d > 1$  be the dimensionality of  $\Omega := [-1, 1]^d$ . We denote  $B_d$  the hypersphere center on  $\mathbf{x}_c = (0.1 \dots 0.1)$  with radius  $R_d = \sqrt{0.05(d-1)}$  and set  $\Omega_\chi := B_d$ . The optimization problem is then defined by

$$\min_{\mathbf{x} \in \Omega \setminus B_d} f(\mathbf{x}) = \sum_{i=1}^d x_i^2. \quad (30)$$

As for the previous problem, the optimal solution is found on the boundary of  $\Omega_\chi$ . The initialization of the EGO-LS-SVM method uses LHS set having increasing size with  $d$ . Specifically the tests shown below correspond to  $d = 2, 5, 10$  and  $20$  and respective LHS sizes  $n_{\text{LHS}} = 15, 40, 100$  and  $200$ . The stopping criteria for the EGO-LS-SVM and CMA-ES methods are identical to the previous one in Section 3.1.

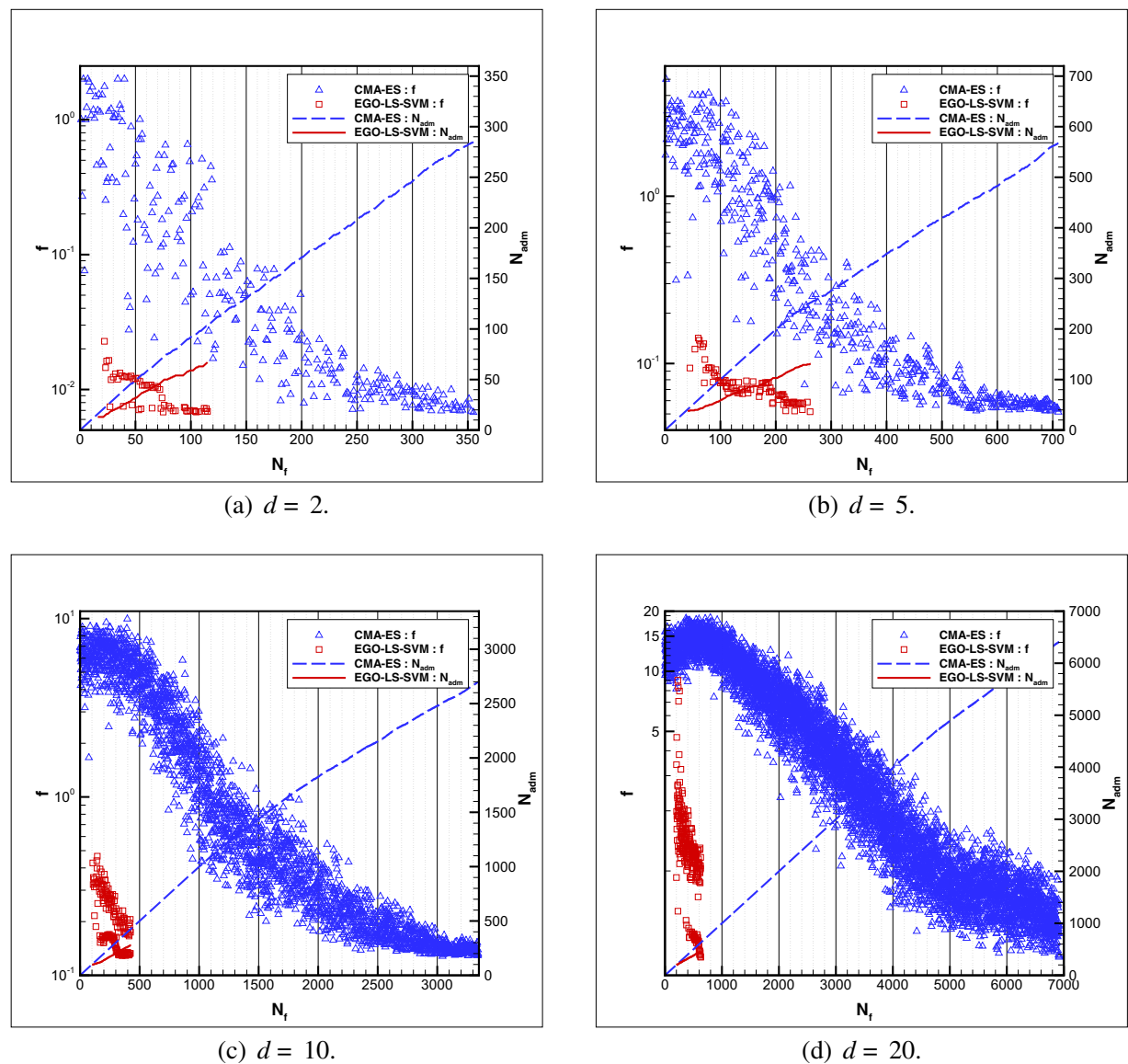
Figure 5 shows the convergences for the different dimensionality  $d$  tested. The plot shows the sequence of values  $f(\mathbf{x})$  at computable proposed points  $\mathbf{x}$  as a function of  $N_f$ . The evolutions of the cumulated number of computable calls  $N_{\text{adm}}$  are also reported.

For all the dimensionality  $d$  tested, the CMA-ES method needs more function calls than the EGO-LS-SVM method and the ratio of numbers of calls increases with  $d$ . In other words, the efficiency of the EGO-LS-SVM method relatively to the CMA-ES method is increasing with  $d$ . It is remarked that for the same number of calls to  $f$ , the CMA-ES is proposing more points in  $\Omega_{\text{adm}}$  than the EGO-LS-SVM method, but the latter is more effective at proposing points with low  $f$  value. For the EGO-LS-SVM method, we also note the presence of distinct branches of function values. These branches reflect the competing trends involved in the selection of the new optimum candidates  $\mathbf{x}_{n+1}$ , because of the alternated use of conservative or aggressive strategies specified by (23) and (24).

#### 3.2.2 Influence of initial sample set size

This section aims at assessing the impact of the initial sample set size on the convergence of the optimization procedure. Indeed, the initial design of experiments can be crucial in the optimization process and may significantly affect the EGO procedure. To study the robustness of the EGO-LS-SVM method, we solve the problem (30), with  $d = 5$  and three different initial sample set sizes:  $n_{\text{LHS}} = 10, 50$  and  $100$ . For these three cases, we perform 300 EGO iterations.

A classification approach to efficient global optimization in presence of non-computable domains



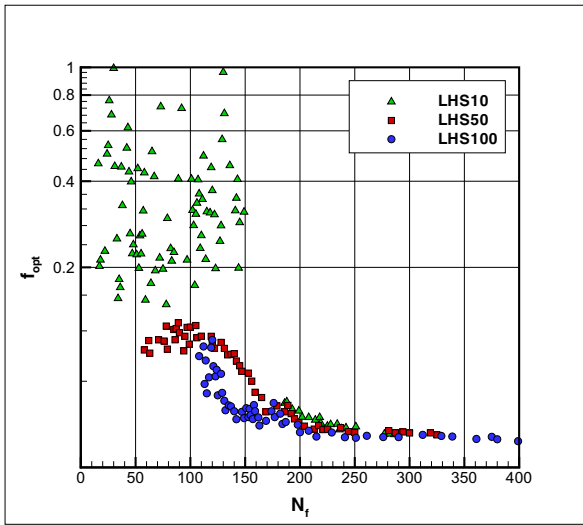
**Fig. 5** Sequence of computable values of  $f(\mathbf{x}_n)$  and cumulated number of computable calls  $N_{\text{adm}}$  as functions of  $N_f$  and for the EGO-LS-SVM and CMA-ES optimization procedures. Case of problem (30) with different dimensionality  $d$  as indicated

Figure 6 shows the convergences for the different sample set sizes  $n_{\text{LHS}}$ . The plot shows the sequence of objective function values  $f$  at computable points as a function of  $N_f$  (also counting the initial sample set generation). We see that for the  $n_{\text{LHS}} = 50$  and 100 cases the algorithm immediately proposed new points that improve the function values of  $f$  and accordingly converge quickly within  $\approx 200$  iterations. The case of only  $n_{\text{LHS}} = 10$  initial samples exhibits a very different behavior with an initial exploration phase with selected points having high objective function values, up to  $N_f \approx 150$ , after which the optimization quickly converges. Overall, the convergence requires slightly more

iterations for  $n_{\text{LHS}} = 10$  than for the larger initial sample sets. Nevertheless, this example highlights the robustness of the EGO-LS-SVM method since the procedure converges to the optimum even when a coarse initial sample set (10 points in 5 dimension) is employed.

### 3.3 Over-constrained problem

The following problem is designed to illustrate the limits of the EGO-LS-SVM method in the case where the feasible domain  $\Omega_{\text{adm}}$  is hard to learn and explore, because of a complex structure. This problem which was proposed



**Fig. 6** Sequences of optimum computable values of  $f_{\text{opt}}$ , as a function of  $N_f$ , for different initial sample set sizes  $n_{\text{LHS}}$ . Case of problem (30) with  $d = 5$

in Tenne and Goh (2010) is then said over-constrained. The domain  $\Omega_{\text{adm}}$  is here defined as a subset of  $\Omega = [-10, 10]^d$ , for  $d = 7$ , through

$$\Omega_{\text{adm}} = \{\mathbf{x} \in \Omega, c_{i=1,\dots,4}(\mathbf{x}) \leq 0\}, \quad (31)$$

where

$$\begin{aligned} c_1(\mathbf{x}) &= 2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 - 127, \\ c_2(\mathbf{x}) &= 7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 - 282, \\ c_3(\mathbf{x}) &= 23x_1 + x_2^2 + 6x_6^2 - 8x_7 - 196, \\ c_4(\mathbf{x}) &= 4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7. \end{aligned}$$

The optimization problem finally writes as

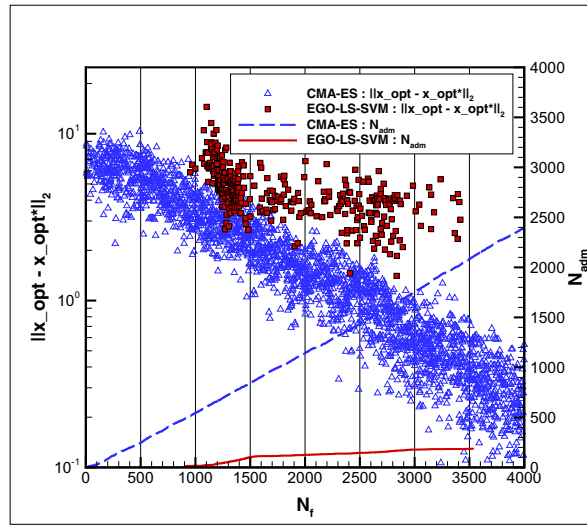
$$\begin{aligned} \min_{\mathbf{x} \in \Omega_{\text{adm}}} f(\mathbf{x}) &= (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 \\ &\quad + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7. \end{aligned} \quad (32)$$

The global optimum is located at

$$\mathbf{x}_{\text{opt}} = (2.330499, 1.951372, -0.4775414, 4.365726, -0.624487, 1.038131, 1.594227).$$

As before, only the proposed points in  $\Omega_{\text{adm}}$  are evaluated and used to approximate  $f$ , while the classification proceeds with all the points, computable or not. Therefore, a large LHS set of 800 points drawn uniformly in  $\Omega$  is needed to ensure that the initial GP model of  $f$  uses sufficiently many points. In the experiment presented below, only 7 of the 800 LHS points were drawn in  $\Omega_{\text{adm}}$ , a low fraction illustrating the complexity of the constrained problem.

Figure 7 shows the evolution of the Euclidean distances to the exact optimum,  $\|\mathbf{x}_{\text{opt}} - \hat{\mathbf{x}}_{\text{opt}}\|_2$  as a function of  $N_f$ .



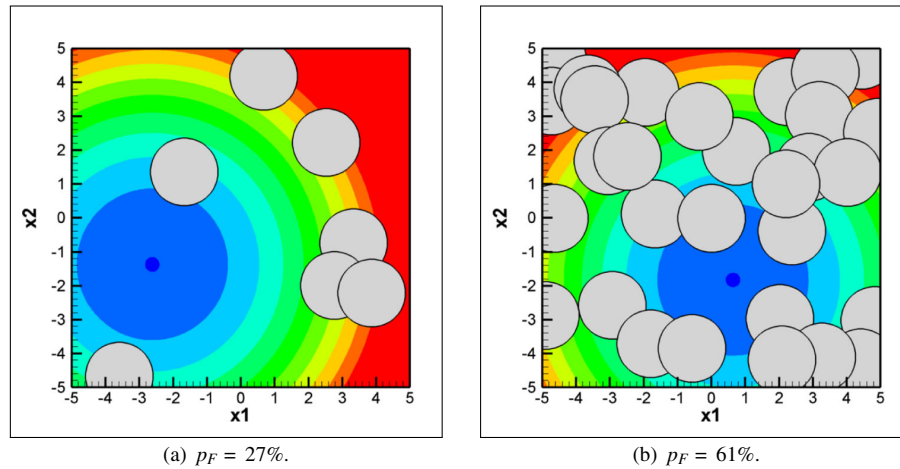
**Fig. 7** Distance to the optimum  $\|\mathbf{x}_{\text{opt}} - \mathbf{x}_{\text{opt}}^*\|_2$  and cumulated number of computable calls to  $f$   $N_{\text{adm}}$  as functions of  $N_f$  and for the EGO-LS-SVM and CMA-ES optimization procedures. Case of problem (32)

The EGO-LS-SVM method progresses during the first 500 calls to  $f$  (EGO iterations), but subsequently stalls. Note that when the optimization procedure is halted, only 187 evaluations have been generated for a total of 3541 calls to  $f$ . The low ratio of computable proposed points  $\mathbf{x}_n$  indicated that for this example the EGO-LS-SVM method is not able to properly learn the domain  $\Omega_{\text{adm}}$ . On the contrary, the CMA-ES method performs satisfactorily on this problem and appears to be much more robust against the complexity of  $\Omega_{\text{adm}}$ . Nevertheless, it should be underlined that this problem corresponds to an extreme situation where the objective function is not computable over most of the search domain  $\Omega$ .

### 3.4 Problem with random computable domain

To complete the series of tests on analytically defined problems, we propose a problem for which the domain of non-computable function  $\Omega_{\chi}$  and the objective function are defined randomly. We set  $\Omega = [-5, 5]^2$  and generate randomly  $\Omega_{\chi}$  using the following procedure. We first generate randomly 100 points  $\mathbf{x}_{i=1,\dots,100}^c \in \Omega$ . Then, for each point  $\mathbf{x}_i^c$  we decide with a probability  $p_{\chi} \in [0, 1]$  if  $f$  is non-computable in the neighborhood  $\|\mathbf{x} - \mathbf{x}_i^c\|_2 \leq 1$ . We then define  $\Omega_{\chi}$  as the intersection of  $\Omega$  with the non-computable neighborhoods. For this construction, the probability  $P_{\chi}$  that  $f$  is not computable at a given  $\mathbf{x} \in \Omega$  is approximated from the cumulative distribution function of the binomial law  $B(100, \frac{\pi}{|\Omega|} p_{\chi})$ , where  $|\Omega|$  is the volume of  $\Omega$ . Figure 8 presents two random realizations of  $\Omega_{\chi}$  (gray





**Fig. 8** Realizations of the non-computable domain  $\Omega_\chi$  (gray circles) for probabilities  $P_\chi = 0.21$  and  $0.61$ . Also shown are color contours of  $f$

areas) for a probability  $P_\chi = 27\%$  and  $61\%$  in the left and right plot respectively.

The optimization problem is finally defined as

$$\min_{\mathbf{x} \in \Omega \setminus \Omega_\chi} f(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_{\text{opt}}\|_2^2, \quad (33)$$

where the solution  $\mathbf{x}_{\text{opt}}$  is drawn at random in  $\Omega \setminus \Omega_\chi$ . In the computations, the EGO-LS-SVM method is initialized with an LHS set of  $n_{\text{LHS}} = 10$  points in  $\Omega$  and the optimization procedure is stopped at  $N_f = 50$ .

We solve the optimization problem (33) for several probabilities  $P_\chi$ . For every value of  $P_\chi$  several realizations of  $\Omega_\chi$  and  $\mathbf{x}_{\text{opt}}$  are generated to estimate statistics of the solution. Figure 9 shows the convergence of the optimization procedure with  $N_f$  (optimization iteration) and for increasing values of  $P_\chi$ . Figure 9a–c correspond to increasing values of  $P_\chi$  and reports the median (blue line) and first and third quartiles (red lines) of the distance to the optimum. Figure 9d compares the convergence of the median distance to the optimum, for several values of  $P_\chi$ . The plots show that as  $P_\chi$  increases from 0, the convergence of the median and first quartile is only mildly affected. In contrast, the third quartile is significantly impacted, with an  $O(1)$  distance to the optimum when  $P_\chi = 61\%$ . This behavior indicates that the EGO-LS-SVM method is quite robust and successful in finding the optimum, except when  $P_\chi$  becomes too large with computable domains  $\Omega_{\text{adm}}$  with a low ratio of  $|\Omega_{\text{adm}}|/|\Omega|$  and complicated structures. In these situations, the EGO-LS-SVM method is not making progress and becomes inefficient at proposing new points  $\in \Omega_{\text{adm}}$  to improve the model of  $f$ . As a result, the approach is either successful or completely failing, with a significantly skewed distribution of the distance to the optimum as evidenced by the distance between the first and third quartiles to the median. However, the results reported in Fig. 9d show that the EGO-LS-SVM method is

quite robust since for probabilities as high as  $P_\chi = 79\%$  the approach is still converging for at least half of the generated problems.

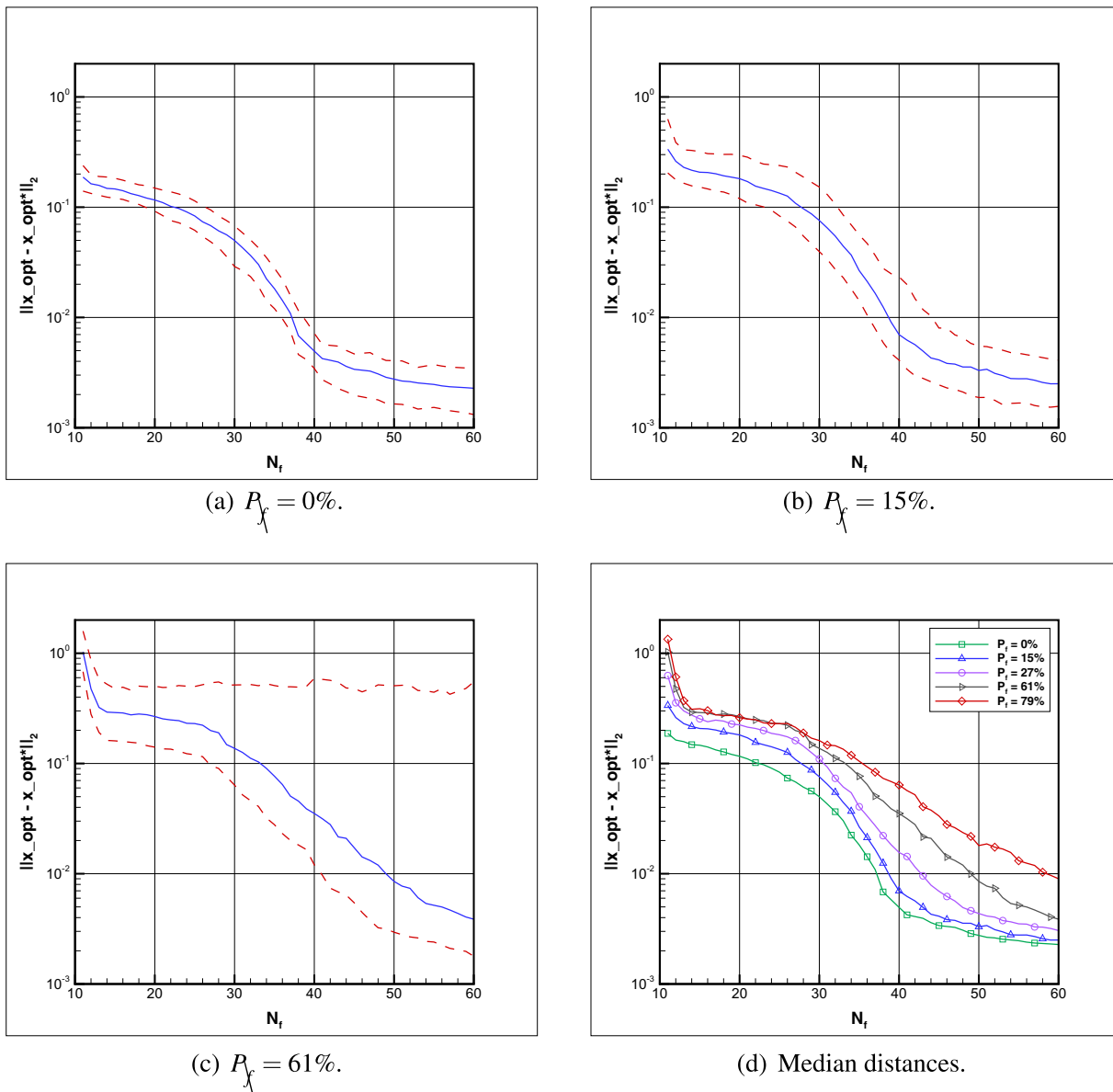
### 3.5 Comparison with EGO-SVM( $d+1$ )

A classification approach was proposed in Basudhar et al. (2012) for the treatment of constraints, and it can be adapted to the case of non-computable domains. This section aims at contrasting this approach with ours, and at highlighting the improved efficiency of our approach.

As mentioned before, the method proposed in Basudhar et al. (2012) uses a classifier to predict the feasibility of new points (i.e. whether constraints are satisfied), and so can be used in the same way to predict if  $f$  is computable. Compared to our approach, the method in Basudhar et al. (2012) differs on four main points. First, it uses an SVM method to determine the classification function  $h(\mathbf{x})$ , in a polynomial space, where we use an LS-SVM method in general Reproducing Kernel Hilbert spaces. The LS-SVM has thus generally better approximation properties when the SVM method selects a subset of points to construct  $h$  with possibly a reduced computational cost. Second, using a reduced training set in SVM calls for a correction of the classification probability to ensure that disregarded training samples are correctly classified. In Basudhar et al. (2012) the following correction of the sigmoid function (20) was considered,

$$P(C^+|\mathbf{x}) = \frac{1}{1 + \exp\left(Ah(\mathbf{x}) + B\left(\frac{d^-(\mathbf{x})}{d^+(\mathbf{x})+\delta} - \frac{d^+(\mathbf{x})}{d^-(\mathbf{x})+\delta}\right)\right)}, \quad (34)$$

where  $d^+(\mathbf{x})$  (resp.  $d^-(\mathbf{x})$ ) is the distances of  $\mathbf{x}$  to the closest point in the training set belonging to  $C^+$  (resp.  $C^-$ ), and  $\delta = 10^{-10}$  is used to avoid numerical overflows.



**Fig. 9** Statistic (median, 25% and 75% quartiles) of the distance to the optimum  $\|x_{\text{opt}} - x_{\text{opt}}^*\|$  as a function of  $N_f$  and for several failure probabilities  $P_{\chi}$

This modification is introduced to account for the fact that the class of the training samples is exactly known. Third, the method in Basudhar et al. (2012) considers a single merit function based on the expected improvement (EI). Forth and finally, the infilling strategy in Basudhar et al. (2012) adds not only the point  $x_{n+1}$  classified in  $C^+$  and maximizing the EI but also  $d$  additional new points in its neighborhood to maintain the isotropy of the training set (a requirement made necessary by the corrected probability in (34)). In contrast, our approach uses the simpler and

smoother sigmoid function, alternates between the criteria in (21)–(24) to accommodate different situations and enrich the training set by a single element at each iteration of the iterative procedure. As subtle as they may appear, these differences greatly impact the efficiency of the optimization procedures.

To distinguish between the impacts of the classifier construction methods and of the infilling strategies (classification probability, selection criteria for  $x_n$ ), we considered the following methods:

- EGO-LS-SVM : the approach proposed in this paper, with LS-SVM construction of the classification function  $h(\mathbf{x})$ , sigmoid classification probability definition, sequential selection using criteria (21)–(24).
- EGO-SVM( $d+1$ ) : the approach proposed (Basudhar et al. 2012), with SVM construction of  $h(\mathbf{x})$ , modified sigmoid classification probability definition in (34), EI-only selection criteria and  $d$ -points neighborhood exploration.
- EGO-SVM : the same as EGO-LS-SVM but with a SVM construction of  $h(\mathbf{x})$ .
- EGO-LS-SVM( $d+1$ ) : the same as EGO-SVM( $d+1$ ) but with a LS-SVM construction of  $h(\mathbf{x})$ .

These methods are tested on problem (30), fixing  $d = 5$ . The results of these experiments are summarized in Fig. 10. Note that the SVM and LS-SVM methods both use kernel-based constructions. The plots report the sequence of computable function values  $f(\mathbf{x}_n)$  as a function of the number of function calls  $N_f$ . In the case of the EGO-SVM( $d+1$ ) and EGO-LS-SVM( $d+1$ ), the function values are reported only for the newly selected point  $\mathbf{x}_{n+1}$ , if computable, and not for its  $d$  additional companion points. In addition, the convergence of the best feasible function value  $f_{\text{best}}$  is reported, using lines, and considers all evaluations of  $f$  (including the eventual  $d$  additional points).

Figure 10a compares the LS-SVM methods for an initial sample set with  $n_{\text{LHS}} = 10$  points. It is seen that our approach EGO-LS-SVM explores the design space during about 100 function calls and converges to the computable optimum in about 250 calls. In contrast, the EGO-LS-SVM( $d+1$ ) method that uses the probability function and infilling strategy of Basudhar et al. (2012) is seen to have not converged even after 1,800 function evaluations, being unable to focus on the correct area. Figure 10b shows the case of the SVM methods using again  $n_{\text{LHS}} = 10$  samples. The plot shows that the EGO-SVM method converges essentially as fast as for our proposed EGO-LS-SVM, albeit with a somehow longer exploration phase, as one would have expected for using the SVM in place of the LS-SVM model. For the EGO-SVM( $d+1$ ) method, corresponding to the original method proposed in Basudhar et al. (2012), we observe that it is now converging too, but it requires a few times more function evaluations to achieve a comparable best value  $f_{\text{best}}$  compared to our infilling strategy. This result highlights the computational overhead due to the need of adding  $d$  additional points to maintain the isotropy of the sample points. Figure 10c (resp. Fig. 10d) corresponds to the same methods as in Fig. 10a (resp. Fig. 10b) but now using an initial sample set with  $n_{\text{LHS}} = 100$  points. It shows that our infilling method does not improve significantly when

increasing  $n_{\text{LHS}}$ , or conversely demonstrating once more the robustness and interest of alternating between merit criteria.

We finally remark that the construction of the SVM classifier is much more time consuming than the LS-SVM one, in the presented experiments using a kernel-based approach. This can be appreciated from the computational times reported Table 1 for different training set sizes. The differences in the computational times are mainly due to *model selection* step in the SVM classifier, which involves the resolution of a quadratic problem (using an SMO method (Platt 1999)). The higher computational cost of the SVM classifier construction further pleads for using the LS-SVM.

## 4 Application to flexible hydrofoil optimization

### 4.1 Optimization problem and solvers

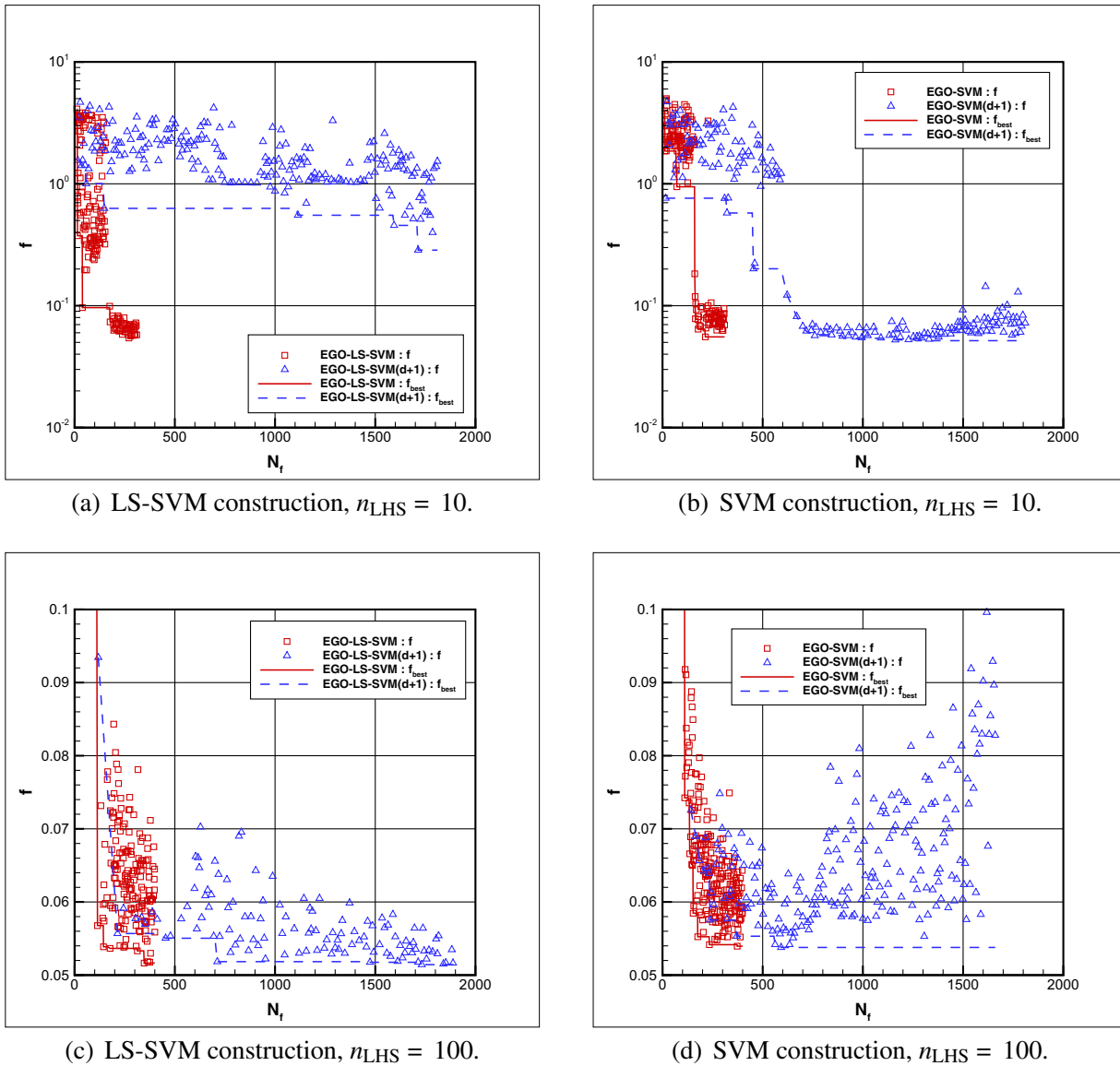
We now apply the EGO-LS-SVM on a realistic problem corresponding to the optimization of a flexible hydrofoil. The objective of the optimization is to minimize the hydrofoil drag force at selected conditions (forward speed and lifting force) while ensuring non-cavitating flows. The complete description of the optimization problem and the hydrodynamical analysis of the optimized hydrofoil are not provided in this paper, which focuses on the method; interested readers can refer to Sacher et al. (2017). The original optimization problem proposed in Sacher et al. (2017) involved 11 design variables describing the shape and elastic characteristics of the foil flexible trailing edge, as schematically illustrated in Fig. 11. In the present work, the number of design variables is reduced to 5 in order to maintain reasonable computational times and to allow extensive comparisons of the optimization methods. Among the retained optimization variables, 4 concern the trailing edge camber geometry at rest ( $p_{1,\dots,4}$ ), while the last one modifies the Young modulus  $E_{\text{Beam}}$  of the flexible part. We set  $\Omega = [-0.1, 1.5] \times [-0.5, 0.2] \times [-0.3, 0.5] \times [-0.3, 0.3] \times [0.1, 1.2]$ .

As mentioned before, the optimization concerns the minimization, for 4 conditions, of the hydrofoil drag coefficients  $C_{D_{i=1,\dots,4}}$ . Each condition corresponds to a prescribed forward speed and hydrodynamic lift force developed by the hydrofoil. The optimization problem is written as

$$\min_{\mathbf{x} \in \Omega_{\text{adm}}} f(\mathbf{x}) = \sum_{i=1}^4 w_i C_{D_i}(\mathbf{x}), \quad (35)$$

where the  $w_i > 0$  are prescribed weighting coefficients fixing the relative importance of the 4 conditions. Constrains





**Fig. 10** Sequence of computable objective function values  $f(\mathbf{x}_n)$  (symbols) and convergence of the best value  $f_{\text{best}}$  (lines) as functions of the total number of calls  $N_f$  for different classifiers constructions, infilling strategies and initial sample set sizes. Case of problem (30) with  $d = 5$

$Q_i(\mathbf{x})$  are introduced to prevent the cavitation of the flow for 4 other conditions. The admissible and computable domain is then defined as the intersection of the set of design points

$\mathbf{x}$  satisfying the cavitation constraints, with  $\Omega \setminus \Omega_\chi$  the set of  $\mathbf{x}$  having computable function  $f$ :

$$\Omega_{\text{adm}} = \Omega \cap \{\mathbf{x}, Q_{i=1,\dots,4}(\mathbf{x}) \leq 0\} \setminus \Omega_\chi, \quad (36)$$

Indeed, parts of the search domain  $\Omega$  corresponds to non-computable points, because of the lack of robustness of the simulation tool. As an example, Fig. 12 shows a hydrofoil corresponding to a non-computable point of the search space, at which the flow solver is not able to determine the hydrodynamic forces. Although it may be possible to finely tune the solvers to improve their robustness, automating the tuning process can be a difficult task. In addition,

**Table 1** Measured computational times (in seconds) for building the LS-SVM and SVM classifiers on training set with different sizes

| Training set size      | 50  | 100 | 150  | 200  |
|------------------------|-----|-----|------|------|
| CPU time (s) of LS-SVM | 0.2 | 0.4 | 1.1  | 2.1  |
| CPU time (s) of SVM    | 0.9 | 3.7 | 15.5 | 30.9 |

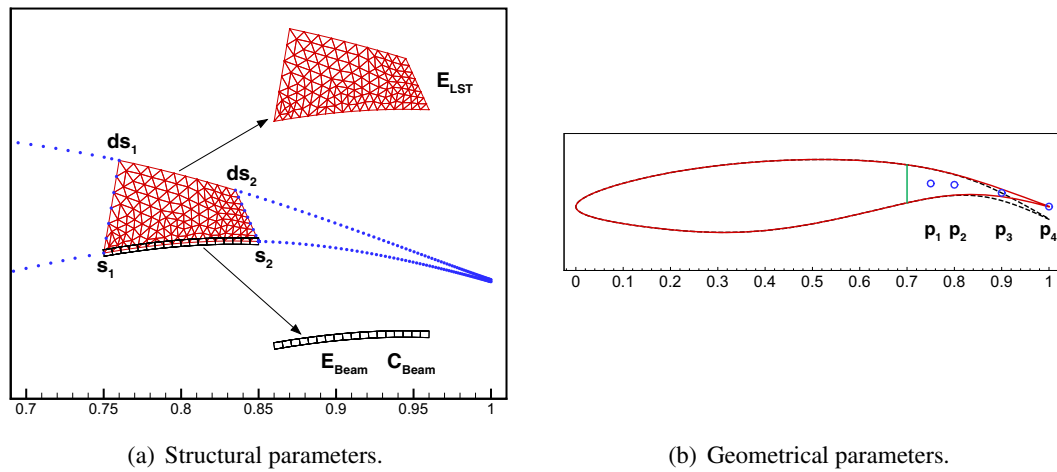


Fig. 11 Illustration of hydrofoil parameters

these non-computable configurations often correspond to uninteresting or even unphysical design points, such that the direct classification of these non-computable areas has been found very effective.

The explicit constraints  $Q_i$  are expressed as

$$Q_i(\mathbf{x}) = -C_{p_i}(\mathbf{x}) - \lambda_i, \quad \lambda_i = \frac{\bar{p}_i - p_v}{\frac{1}{2}\rho U_i^2}, \quad (37)$$

where  $\rho$  is the fluid density,  $p_v$  the saturated vapor pressure,  $C_{p_i}$  the minimal pressure coefficient,  $\lambda_i$  the cavitation number, and  $U_i$  and  $\bar{p}_i$  are the reference velocity and pressure for the  $i$ -th condition. The minimum of the pressure coefficient is defined by  $C_{p_i}(\mathbf{x}) := \frac{p_i^-(\mathbf{x}) - \bar{p}_i}{0.5\rho U_i^2}$ , where  $p_i^-$  is the lowest pressure over the hydrofoil surface. The condition  $Q_i \leq 0$  expresses the constraint that the minimum of the pressure around the hydrofoil should remain higher than the vapor pressure:  $p_i^-(\mathbf{x}) \geq p_v$ .

Given a value  $\mathbf{x}$  of the design variable, the evaluation of the objective function  $f$  and the constraints  $Q_i$  requires the resolution of 8 nonlinear fluid-structure interaction problems. For the flow, we rely on a static vortex lattice method with viscous boundary layer equations, using the solver XFOIL (Drela 1989; Morgado et al. 2016), while a nonlinear elasticity solver (Durand et al. 2014) is used for the elastic deformations of the foil (Durand et al. 2014),

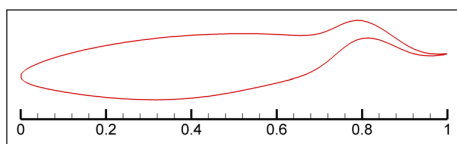


Fig. 12 Configuration of the hydrofoil leading to a non computable case

modeling the elastic trailing edge with 2D Linear Strain Triangles (LST) (Pedersen 1973) and Timoshenko beam elements (see Fig. 11a). The nonlinear equilibrium FSI solutions are computed by a Newton method with Aitken relaxation. We mention that the conditions call for the enforcement of a prescribed lift force. This is achieved in XFOIL by adjusting the Angle of Attack (AoA) of the hydrofoil.

## 4.2 Optimization results

In the following, we consider two EGO approaches to handle the constraints related to the cavitation. First, the proposed EGO-LS-SVM method is used to estimate directly  $\Omega_{\text{adm}}$  from the classification of previous points  $\mathbf{x}$ . In other words, the LS-SVM method is employed to predict both the satisfaction of the constraints  $Q_i(\mathbf{x}) \leq 0$  and the computability of future points. For the second method, the LS-SVM classification is used only to predict the computability, while the functionals  $Q_i$  of the cavitation constraints are approximated by individual GP models following the EGO method presented in Schonlau (1997). We call this second approach the EGO-GPC. These two approaches are also compared to the reference CMA-ES algorithm (Hansen 2006) with a penalty value  $f_\infty = 100$  for the non-computable points. In all computations, the two EGO methods are initialized using an LHS set of size 100. For this example, the EGO optimizations are continued until a maximum computational budget of 36 hours is exhausted. Both the resolutions of the non-linear FSI problems, the construction of the GP models and the search for the maximizer of the AEI criteria are included in the computational time. Note that on a classical workstation it takes approximatively  $T_f = 75\text{s}$  to decide the admissibility of a point and evaluate  $f$  at a given

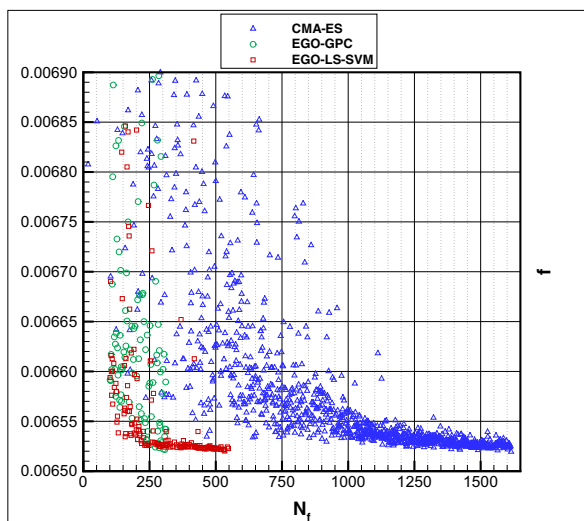
computable  $\mathbf{x}$ , solving the FSI problems corresponding to the prescribed 8 conditions.

Figure 13 shows the evolution of the successive values of  $f$  with the number of calls to  $f$ . It is seen that the EGO-LS-SVM and EGO-GPC methods require a significantly lower number of function evaluations than CMA-ES to reach the optimum neighborhood. Specifically, the CMA-ES optimization needs 1615 function evaluations, that is, twice as much, to eventually produce a better point than the EGO-LS-SVM optimum (stopped after 36h of computations). The two EGO methods have similar behavior, although the sequence of computed values of  $f$  produced the EGO-GPC method is slightly more dispersed than for the EGO-LS-SVM method. This difference can be explained by the additional evaluations needed to construct the GP models of the constraints in the EGO-GPC method.

If the two EGO methods behave similarly in term of the convergence with the number of calls to  $f$ , they do present significant differences in terms of computational times. This can be appreciated from the plots of Fig. 14. Figure 14a reports the cumulated computational times of the EGO-LS-SVM and EGO-GPC methods as functions of the optimization iteration. The computational times are split into the time dedicated to the FSI problems resolution, called FSI time, and the construction of the GP models and search of maximizer of the extended AEI, called EGO time. It is seen from the graph that the EGO time is dominant for the two approaches and increases with the iteration index at a polynomial rate  $r \gtrsim 2$ , because of the increasing size of the GP models construction. Furthermore, the EGO time of the EGO-GPC method is significantly

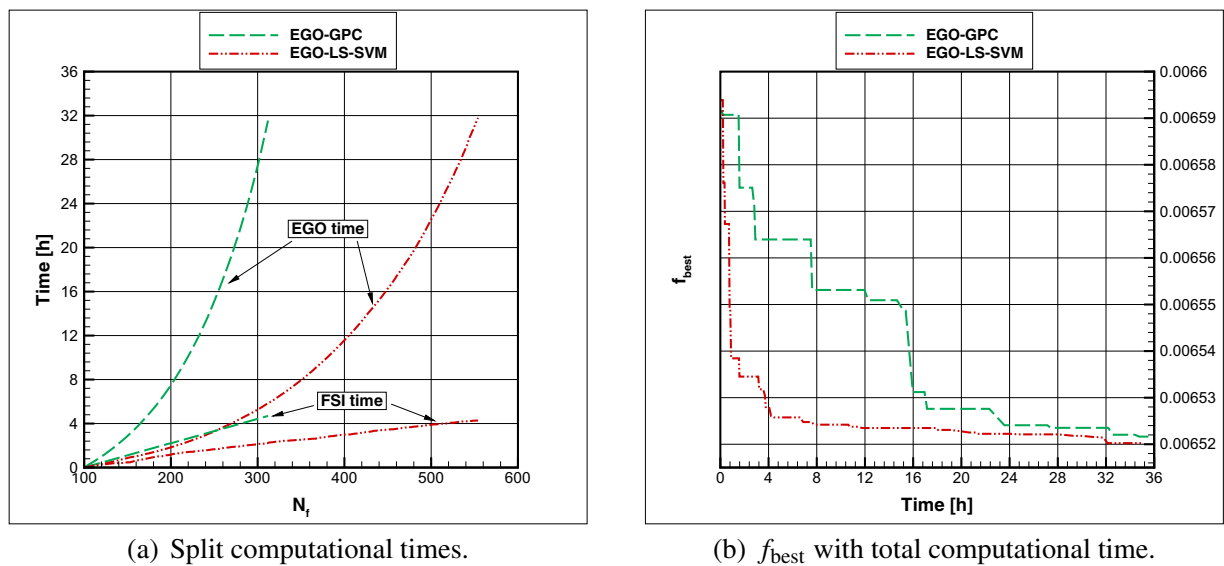
higher than that of the EGO-LS-SVM method, as one may have expected since more GP processes must be estimated. In contrast to the EGO time, the computational times spent on the FSI problems increase roughly at a linear rate, with the EGO-LS-SVM FSI time being half of the EGO-GPC FSI time. A closer inspection of the results reveals that, when the computational budget of 36 hours is reached, the EGO-GPC method called for 312 evaluations of  $f$ , including 8 non-computable points, while EGO-LS-SVM made up to 554 calls to  $f$ , among which 294 violated the constraints and 5 were not computable. These numbers explain the lower FSI time of the EGO-LS-SVM method, compared to the EGO-GPC method. Indeed, the full classification approach of the EGO-LS-SVM method allows for significant computational savings because once the constraint on cavitation for a certain condition has been found to be non-satisfied, there is no need to solve the remaining FSI problems for the other conditions and to compute  $f$ . For the EGO-GPC method, on the contrary, one must solve systematically all the FSI problems for the whole set of 8 (constraints and objective) conditions to improve the GP models of all the  $Q_i$  and  $f$ . To further illustrate this point, we compare in Fig. 14b the current best value of  $f$  observed, denoted  $\min(f)$ , as a function of the computational time, for the two EGO methods. The EGO-LS-SVM method is seen to be more efficient as it gives a better optimum than the EGO-GPC approach. Moreover, the convergence rate seems faster for the EGO-LS-SVM method: an excellent minimum within less than 0.1% of the terminal value of  $f$  is obtained in just 4 hours when for the GP method a comparable point is obtained after 17 hours. These differences are consistent with the fact that the number of EGO-LS-SVM iterations performed in a given computational time is essentially twice that of the EGO-GPC: although it is proposing a significant fraction of non-feasible points, the EGO-LS-SVM method better explores the design space and takes advantage of a cheaper evaluation cost for non-feasible points.

Nevertheless, it should be noted that the CMA-ES algorithm can be computationally cheaper than the EGO methods, depending on the computational budget and the problem considered. For instance, Fig. 15a shows the best admissible (and computable) value of  $f$  obtained as a function of the computational time for the two EGO methods and the CMA-ES procedure. It is seen that in the case of the hydrofoil optimization with the numerical models considered here, the CMA-ES always outperforms EGO-GPC, except for a point around 17 hours of computational time. Compared to the EGO-LS-SVM, the CMA-ES is significantly less efficient at the beginning of the optimization procedure but eventually do better as the computational time increases, with a better point found by the evolutionary approach after  $\approx 26$  hours



**Fig. 13** Sequence of computable function values  $f(\mathbf{x}_n)$  as a function of the number of calls  $N_f$  for the EGO-LS-SVM, EGO-GPC and CMA-ES methods. Case of problem (35)

A classification approach to efficient global optimization in presence of non-computable domains

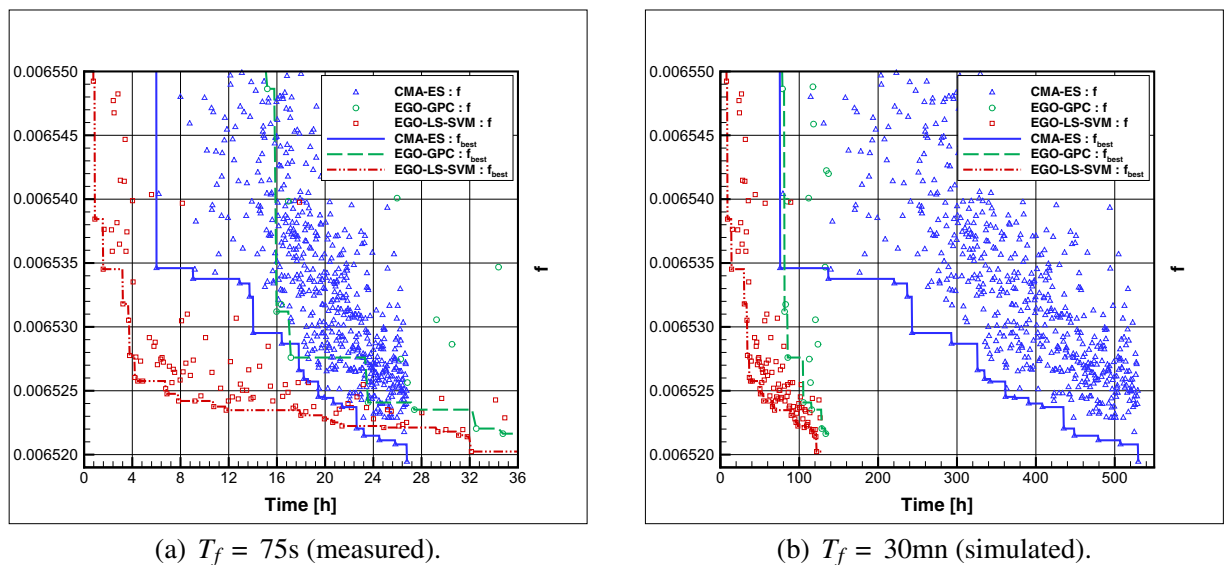


**Fig. 14** Left: comparison of the computational times of the EGO-LS-SVM and EGO-GPC methods, splitting the FSI solves (FSI time) and the GP models construction and AEI exploration (EGO time)

contributions. Right: convergence of the computable  $f_{best}$  value with the total computational time. Case of problem (35)

of calculation. Obviously, these results are for a single realization of the optimization procedures that are random in essence (through the selection / re-sampling of the population in CMA-ES and initial LHS sample in the EGO methods). A complete comparison should consider averages over multiple realizations, but the results reported in Fig. 15a are representative of the relative behavior of

the different methods. However, these relative efficiencies are highly dependent on the problem and the numerical model involved. In the present case, the important fraction of the computational time dedicated to the GP models construction is detrimental to the relative performance of the EGO methods. To illustrate this, we can estimate the computational times that would result if the resolution of



**Fig. 15** Evolution of the best computable objective function value  $f_{best}$  with the total computational time for the EGO-LS-SVM, EGO-GPC and CMA-ES methods and considering different evaluation times  $T_f$ . The sequence of computable function values  $f(\mathbf{x}_n)$  are also reported. Case of problem (35)

the FSI problems were much more demanding, so the EGO times become relatively less significant. This is made in Fig. 15b where we have considered the case of an evaluation of  $f$  requiring 30 minutes of computational time, instead of 75 seconds. In this situation, the CMA-ES optimization would need roughly 525 hours to find a better point than the one obtained in just 125 hours by the EGO-LS-SVM. Even the EGO-GPC method would become more competitive than the CMA-ES strategy for  $T_f$  large enough. In conclusion, the EGO-based methods are most suited to the case of costly models.

## 5 Conclusion

In this work, we have proposed the use of a classification approach in EGO procedures to deal with the existence of non-computable design points and inequality constraints. The key idea of the approach is to define the admissible sub-domain of the original design space consisting of points being computable and satisfying the problem constraints. The method, named EGO-LS-SVM, relies on an LS-SVM classifier to predict the admissibility of points in the global search domain. The LS-SVM classifier composes a Gaussian process model with a sigmoid function, leading to an admissibility probability. It is trained on the set of previously classified points. We combine the probabilistic characterization of the admissibility with the classical Augmented Expected Improvement (AEI) statistical criterion, to propose several merit functions for the selection of the next EGO point. These merit functions induce conservative or on the contrary aggressive strategies, with respect to admissibility when exploring the design space. In fact, we recommend to actually alternate between the merit functions to yield a robust but effective sequential infilling strategy, ensuring the intensive exploration of promising areas of the search domain and the boundaries of the admissible domain.

The EGO-LS-SVM method has been tested on several analytical problems presenting different types of difficulties. These tests have validated the method and revealed its robustness with respect to the number of design variables. In fact, it is shown that compared to standard evolutionary algorithms (CMA-ES) the relative efficiency of EGO-LS-SVM is increasing with the number of parameters in the range tested. However, the EGO-LS-SVM method can be challenged in extreme situations where the admissible domain consists of a small portion of the global design set with complex geometry. For these situations, the EGO-LS-SVM may not have found admissible points, stalling the convergence.

Finally, the method was tested on the constrained optimization of a fully nonlinear numerical fluid-structure

interaction system. Here we aimed at minimizing the drag of a flexible hydrofoil in several operating conditions, with constraints related to the cavitation. On this example, the EGO-LS-SVM method has been compared with the CMA-ES and EGO methods with GP models of the constraints. This example highlighted the interest of using EGO-based methods compared to the CMA-ES on these complex problems, and the advantages of treating the constraints and computable domain in a unified approach to the EGO-LS-SVM method.

Future works should focus on the current limits of the EGO-LS-SVM method. In particular, the initialization procedure must be improved in the situation where the computable domain is small compared to the whole search domain. One possible route in this direction could be to initiate the optimization with an evolutionary method, and subsequently switch to EGO-LS-SVM when sufficiently many points have been found in the admissible domain. Also, when the absolute probabilities of  $C^\pm$  are much unbalanced (small or large computable domain, compared to the search domain) the use of a Weighted-SVM approach (Yang et al. 2005) could help achieving better classification performances, in particular during the initial optimization iterations.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Simpson T, Poplinski J, Koch NP, Allen J (2001) Metamodels for computer-based engineering design: survey and recommendations. *Eng Comput* 17(2):129–150
- Kleijnen JP (2009) Kriging metamodeling in simulation: a review. *Eur J Oper Res* 192(3):707–716
- Vapnik V (1995) *The nature of statistical learning theory*. Springer, New York
- Marrel A, Iooss B, Laurent B, Roustant O (2009) Calculations of sobol indices for the gaussian process metamodel. *Reliab Eng Syst Saf* 94(3):742–751
- Wang P, Lu Z, Tang Z (2013) An application of the kriging method in global sensitivity analysis with parameter uncertainty. *Appl Math Model* 37(9):6543–6555
- Nickisch H, Rasmussen CE (2008) Approximations for binary gaussian process classification. *J Mach Learn Res* 9:2035–2078
- Liu Y, Shi Y, Zhou Q, Xiu R (2016) A sequential sampling strategy to improve the global fidelity of metamodels in multi-level system design. *Struct Multidiscip Optim* 53(6):1295–1313
- Park C, Haftka RT, Kim NH (2017) Remarks on multi-fidelity surrogates. *Struct Multidiscip Optim* 55(3):1029–1050
- Dong H, Song B, Wang P, Dong Z (2017) Surrogate-based optimization with clustering-based space exploration for expensive multimodal problems. *Struct Multidisc Optim* 57(4):1553–1577
- Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive Black-Box functions. *J Glob Optim* 13(4):455–492



- Liu J, Song WP, Han ZH, Zhang Y (2017) Efficient aerodynamic shape optimization of transonic wings using a parallel infilling strategy and surrogate models. *Struct Multidiscip Optim* 55(3):925–943
- Glaz B, Friedmann PP, Liu L (2008) Surrogate based optimization of helicopter rotor blades for vibration reduction in forward flight. *Struct Multidiscip Optim* 35(4):341–363
- Glaz B, Friedmann PP, Liu L (2009) Helicopter vibration reduction throughout the entire flight envelope using surrogate-based optimization. *J Amer Helicopter Soc* 54(1):12007
- Aghajari N, Schäfer M (2015) Efficient shape optimization for fluid–structure interaction problems. *J Fluids Struct* 57:298–313
- Sacher M, Hauville F, Duvigneau R, Maître OL, Aubin N, Durand M (2017) Efficient optimization procedure in non-linear fluid–structure interaction problem: application to mainsail trimming in upwind conditions. *J Fluids Struct* 69:209–231
- Picheny V, Wagner T, Ginsbourger D (2013) A benchmark of kriging-based infill criteria for noisy optimization. *Struct Multidiscip Optim* 48(3):607–626
- Li Z, Ruan S, Gu J, Wang X, Shen C (2016) Investigation on parallel algorithms in efficient global optimization based on multiple points infill criterion and domain decomposition. *Struct Multidiscip Optim* 54(4):747–773
- Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. *ACM computing surveys (CSUR)* 31(3):264–323
- Zhang Y, Park C, Kim NH, Haftka RT (2017) Function prediction at one inaccessible point using converging lines. *J Mech Des* 139(5):051402
- Suykens J, Vandewalle J (1999) Least squares support vector machine classifiers. *Neural Process Lett* 9(3):293–300
- Basudhar A, Dribusch C, Lacaze S, Missoum S (2012) Constrained efficient global optimization with support vector machines. *Struct Multidiscip Optim* 46(2):201–221
- van Gestel T, Suykens JA, Baesens B, Viaene S, Vanthienen J, Dedene G, de Moor B, Vandewalle J (2004) Benchmarking least squares support vector machine classifiers. *Mach Learn* 54(1):5–32
- Hansen N (2006) The CMA Evolution Strategy: a comparing review. In: *Towards a new evolutionary computation*. Springer, pp 75–102
- Rasmussen CE, Williams CKI (2006) *Gaussian processes for machine learning*. MIT Press, Cambridge
- Hansen N, Arnold DV, Auger A (2015) Evolution strategies. In: *Springer handbook of computational intelligence*. Springer, pp 871–898
- Hansen N (2016) The cma evolution strategy: a tutorial. [arXiv:1604.00772](https://arxiv.org/abs/1604.00772)
- Huang D, Allen TT, Notz WI, Zheng N (2006) Global optimization of stochastic Black-Box systems via sequential kriging meta-models. *J Global Optim* 34(3):441–466
- Schonlau M (1997) *Computer Experiments and Global Optimization*. PhD thesis, University of Waterloo, Waterloo, Ont., Canada, Canada. AAINQ22234
- Cawley GC (2006) Leave-one-out cross-validation based model selection criteria for weighted ls-svms. In: *2006 IEEE international joint conference on neural network proceedings. IEEE*, pp 1661–1668
- Stone M (1974) Cross-validated choice and assessment of statistical predictions. *J R Stat Soc Ser B Methodol* 36(2):111–147
- Vapnik VN, Vapnik V (1998) *Statistical learning theory*, vol 1. Wiley, New York
- Cawley GC, Talbot NL (2003) Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recogn* 36(11):2585–2592
- Cawley GC, Talbot NL (2004) Fast exact leave-one-out cross-validation of sparse least-squares support vector machines. *Neural Netw* 17(10):1467–1475
- Cawley GC, Talbot NL (2007) Preventing over-fitting during model selection via bayesian regularisation of the hyper-parameters. *J Mach Learn Res* 8:841–861
- Allen DM (1974) The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* 16(1):125–127
- Van Calster B, Luts J, Suykens JAK, Condous G, Bourne T, Timmerman D, Van Huffel S (2007) Comparing methods for multi-class probabilities in medical decision making using LS-SVMs and kernel logistic regression. In: *Artificial neural networks – ICANN 2007: 17th international conference, Porto, Portugal, September 9–13, 2007, Proceedings, part II*. Springer, Berlin, pp 139–148
- Platt JC (1999) Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advan Large Margin Classifiers* 10(3):61–74
- Lin HT, Lin CJ, Weng RC (2007) A note on platt’s probabilistic outputs for support vector machines. *Mach Learn* 68(3):267–276
- Arnold DV, Hansen N (2012) A (1+1)-CMA-ES for constrained optimisation. In: *Soule T, Moore JH (eds) GECCO, Philadelphia, United States, ACM. ACM Press*, pp 297–304
- Tenne Y, Goh C (2010) *Computational intelligence in expensive optimization problems. Adaptation, learning, and optimization*. Springer, Berlin
- Platt JCSchölkopf B, Burges CJC, Smola AJ (eds) (1999) *Fast training of support vector machines using sequential minimal optimization*. MIT Press, Cambridge
- Sacher M, Durand M, Berrini E, Hauville F, Duvigneau R, Le Maître O, Astolfi JA (2017) Flexible hydrofoil optimization for the 35th america’s cup with constrained ego method. In: *International Conference on Innovation in High Performance Sailing Yachts, Innov’Sail*, pp 193–205
- Drela M (1989) *XFOIL: an analysis and design system for low Reynolds number airfoils*. Springer, Berlin, pp 1–12
- Morgado J, Vizinho R, Silvestre M, Piscoa J (2016) {XFOIL} vs {CFD} performance predictions for high lift low reynolds number airfoils. *Aerosp Sci Technol* 52:207–214
- Durand M, Leroyer A, Lothodé C, Hauville F, Visonneau M, Floch R, Guillaume L (2014) FSI Investigation on stability of downwind sails with an automatic dynamic trimming. *Ocean Eng* 90:129–139
- Pedersen P (1973) Some properties of linear strain triangles and optimal finite element models. *Int J Numer Methods Eng* 7(4):415–429
- Yang X, Song Q, Cao A (2005) Weighted support vector machine for data classification. In: *Proceedings of the 2005 IEEE International Joint Conference on Neural Networks*, vol 2, pp 859–864

### 8.3 Conclusion

La validation de la méthode EGO-LS-SVM est présentée sous la forme d'une publication dans la Section 8.2. Dans cette contribution, différents problèmes analytiques comportant des points non évaluables sont résolus afin d'illustrer l'efficacité de la procédure. L'approche EGO-LS-SVM se montre plus performante que la méthode CMA-ES (Hansen, 2006b) dans tous les cas analytiques testés excepté dans un cas « extrême » que l'on considère sur-constraint. On compare également nos résultats à la formulation originale (Basudhar et al., 2012) qui utilise un modèle de classification SVM et une méthode d'enrichissement différente de notre proposition. La méthode EGO-LS-SVM est plus efficace en terme de vitesse de convergence ainsi qu'en coût de calcul réel sur le problème testé. Finalement, le problème d'optimisation géométrique et structurel de l'hydro-foil flexible permet d'illustrer l'intérêt d'utiliser la méthode EGO-LS-SVM dans un cas d'optimisation sous de multiples contraintes d'inégalité. En effet, la vision binaire des contraintes apportée par classification LS-SVM réduit significativement le coût global de l'optimisation.

Pour conclure, ces travaux de thèse permettent de valider la méthode EGO-LS-SVM sur des cas analytiques ayant des niveaux de difficulté et des propriétés variés.

## Design optimal sous contraintes d'un hydrofoil à paramètres multiples

### Sommaire

|            |                                                      |            |
|------------|------------------------------------------------------|------------|
| <b>9.1</b> | <b>Introduction</b>                                  | <b>171</b> |
| <b>9.2</b> | <b>Publication dans « <i>Ocean Engineering</i> »</b> | <b>174</b> |
| <b>9.3</b> | <b>Conclusion</b>                                    | <b>185</b> |

### 9.1 Introduction

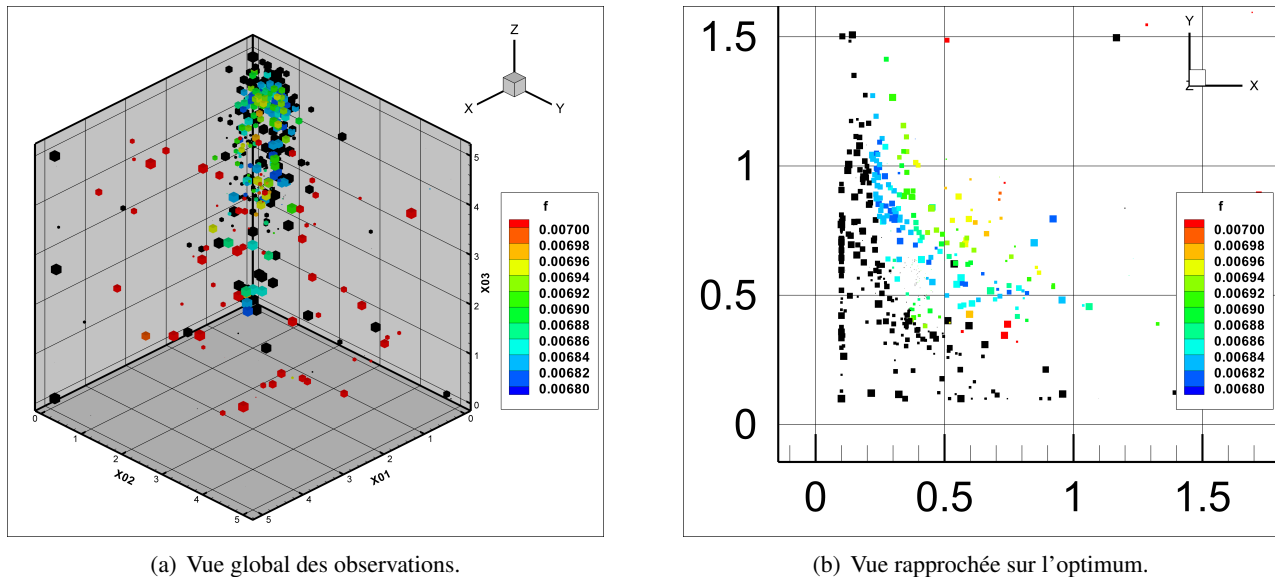
Le Chapitre 9 contient une application de l'optimisation par méta-modèles Gaussiens à un problème physique réel. Cette application concerne la définition optimale sous contraintes du design d'un hydrofoil paramétré par un ensemble de variables de natures variées. Les hydrofoils sont des corps portants utilisés dans l'eau. Leur définition et principe de fonctionnement sont analogues aux foils que l'on trouve dans l'aviation. Comme pour une aile d'avion, l'idée est de générer un effort vertical qui, dans le cas d'un hydrofoil monté sur un navire, soulagera la carène et par conséquent en réduira son effort de traînée. Le cas extrême est dit « volant », ou « foiling mode », lorsque la carène n'est plus en contact avec la surface libre, et la traînée du navire est alors uniquement due aux appendices. Le design des hydrofoils est toujours un challenge actuel qui date déjà de quelques années, comme le montre la thèse de [Bose \(1982\)](#).

La « Coupe de l'America » est l'une des plus anciennes compétitions sportives ; elle a été créée en 1851. L'édition 35 s'est déroulée en juin 2017 sur des catamarans appelés « AC50 ». Ces voiliers ont la particularité d'être équipés d'une aile rigide et d'hydrofoils. Le design des AC50 est limité par des règles données dans l'« AC Class Rule ». Dans cette 35<sup>ième</sup> édition, les hydrofoils sont les éléments les plus libres en terme de choix de design. Ils constituent donc une part importante du design global du voilier. Leur objectif principal est de



générer un effort de traînée minimal sur l'ensemble des vitesses possibles de l'AC50 tout en assurant un effort de portance suffisant pour permettre la propriété volante du voilier, en particulier aux faibles vitesses. Dans l'« AC Class Rule », la règle 11.11 précise : « *Daggerboard components shall be rigidly fixed to each other and the daggerboard shape shall not be adjusted while racing.* » Il n'est donc pas autorisé d'utiliser une solution « active » contrôlant la géométrie des l'hydrofoils. Cependant, une solution « passive » composée d'éléments rigides acceptants des déformations élastiques qui soient dépendantes de la vitesse respecterait la règle 11.11. C'est l'objet de l'application présentée dans ce Chapitre 9, où une section d'hydrofoil comportant une structure flexible est optimisée. Cette étude est menée en partenariat avec l'équipe de design « Groupama Team France » pour la 35<sup>ième</sup> Coupe de l'America.

L'optimisation de l'hydrofoil est sujette à différentes contraintes, telles que des limites sur les chargements de la structure, sur l'inertie de la section, ou sur la cavitation. Les paramètres à optimiser sont de nature géométrique et structurelle. En plus de la complexité de l'application en terme de calculs IFS, on a ici, de par la dimension élevée du problème d'optimisation, la présence de points non évaluables (voir Section 5.1). Un exemple illustrant la présence de points non évaluables dans une optimisation d'hydrofoil où  $d = 4$  est donné sur la Figure 9.1. Chaque symbole correspond à une observation explorée par l'EGO, les symboles noirs sont des points non évaluables et la quatrième dimension est représentée par la taille des symboles. Malgré la difficulté de représenter ces résultats, on observe qualitativement que la zone contenant l'optimum est proche de zone non évaluable. On voit de plus qu'il est difficile de définir un domaine *a priori* non évaluable, en particu-



**Figure 9.1** – Illustration d'observations évaluables et non évaluables lors d'une optimisation à 4 paramètres d'un hydrofoil flexible.

lier lorsque  $d \gg 1$  et avec des paramètres de natures variées, d'où l'utilité de la méthode EGO-LS-SVM (voir Section 5.3) pour cette application.

La présentation de l'optimisation est donnée dans la Section 9.2 par la publication (Sacher et al., 2018a). Les

résultats hydrodynamiques sont détaillés avec en particulier une comparaison d'un hydrofoil flexible optimal à un hydrofoil rigide optimisé pour les même conditions. Cela permet d'illustrer les performances apportées par la flexibilité de l'hydrofoil. Ces travaux sont également présentés dans (Sacher et al., 2017a).

## 9.2 Publication dans « *Ocean Engineering* »

Ocean Engineering 157 (2018) 62–72



Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: [www.elsevier.com/locate/oceaneng](http://www.elsevier.com/locate/oceaneng)



### Flexible hydrofoil optimization for the 35th America's Cup with constrained EGO method

Matthieu Sacher<sup>a,\*</sup>, Mathieu Durand<sup>b,c,d</sup>, Élisabeth Berrini<sup>e,f</sup>, Frédéric Hauville<sup>a</sup>, Régis Duvigneau<sup>e</sup>, Olivier Le Maître<sup>g</sup>, Jacques-André Astolfi<sup>a</sup>

<sup>a</sup> Naval Academy Research Institute - IRENAV CC600, 29240, BREST Cedex 9, France

<sup>b</sup> K-Epsilon, 1300 Route des Crêtes, 06560, Valbonne, France

<sup>c</sup> Groupama Team France, Lorient La Base - Batiment des défis, 56100, Lorient, France

<sup>d</sup> Sirli Innovations, 600 Chemin des Aspres, 06410, Biot, France

<sup>e</sup> Université Côte d'Azur, Inria, CNRS, LJAD, 2004 Route des Lucioles, 06902, Valbonne, France

<sup>f</sup> MyCFD, 29 Avenue des Frères Roustau, 06220, Golfe-Juan, France

<sup>g</sup> LIMSI - CNRS, Rue John Von Neumann, 91400, Orsay, France

#### ARTICLE INFO

##### Keywords:

Constrained optimization  
Gaussian process model  
Hydrofoil  
Cavitation  
Fluid-structure interaction

#### ABSTRACT

This paper investigates the use of constrained surrogate models to solve the multi-design optimization problem of a flexible hydrofoil. The surrogate-based optimization (EGO) substitutes the complex objective function of the problem by an easily evaluable model, constructed from a limited number of computations at carefully selected design points. Associated with ad-hoc statistical strategies to propose optimum candidates within the estimated feasible domain, EGO enables the resolution of complex optimization problems. In this work, we rely on Gaussian processes (GP) to model the objective function and adopt a probabilistic classification method to treat non-explicit inequality constraints and non-explicit representation of the feasible domain. This procedure is applied to the design of the shape and the elastic characteristics of a hydrofoil equipped with deformable elements providing flexibility to the trailing edge. The optimization concerns the minimization of the hydrofoil drag while ensuring a non-cavitating flow, at selected sailing conditions (boat speed and lifting force). The drag value and cavitation criterion are determined by solving a two-dimensional nonlinear fluid-structure interaction problem, based on a static vortex lattice method with viscous boundary layer equations, for the flow, and a nonlinear elasticity solver for the deformations of the elastic components of the foil. We compare the optimized flexible hydrofoil with a rigid foil geometrically optimized for the same sailing conditions. This comparison highlights the hydrodynamical advantages brought by the flexibility: a reduction of the drag over a large range of boat speeds, less susceptibility to cavitation and a smaller angle of attack tuning range.

#### 1. Introduction

Created in 1851, the America's Cup trophy is one of the oldest international sports competition. The race opposes two sailing boats, representing a defender and a challenger yacht clubs, that race one against the other to be the winner of the America's Cup match races. In each edition, specific rules are provided by the defender in accordance with the *Deed of Gift*, which define the format of the regatta, the location of the races and restrictions on the type of boats allowed to compete. When a challenging yacht club wins the match races, it becomes the defender and gains the stewardship for the next America's Cup edition. The 35th edition, which took place in June 2017, has been raced on so-called AC50s

catamarans with size fixed to 50 feet, equipped with hydrofoils and wing-sails. The hydrofoils are expected to be a key component of the AC50s performance, and the competing teams are dedicating a growing fraction of their design effort to their optimization in comparison to older yacht design approach (Richards et al., 2001). The optimization aims at designing hydrofoil with a minimal drag force over a whole range of boat speeds (BS), while ensuring a sufficient lift force for the boat to fly, especially at the lowest BS. The design of the hydrofoil is restricted by the AC Class Rule documents. In particular, the rule number 11.11 states: *Daggerboard components shall be rigidly fixed to each other and the daggerboard shape shall not be adjusted while racing*. One possible way to comply with this rule, while allowing for some adaptivity of the shape with the

\* Corresponding author.

E-mail address: [matthieu.sacher@ecole-navale.fr](mailto:matthieu.sacher@ecole-navale.fr) (M. Sacher).

<https://doi.org/10.1016/j.oceaneng.2018.03.047>

Received 7 September 2017; Received in revised form 6 March 2018; Accepted 13 March 2018

BS, is to design a hydrofoil composed of rigidly mounted components, undergoing large elastic deformations under the hydrodynamical loads and without crew adjustments (Hueber et al., 2017). For instance, a reduction of the hydrofoil camber, with increasing BS and hydrodynamic loads, can be obtained using a flexible trailing edge.

Drag reduction by shape optimization is an important research topic, with many applications in domains ranging from aeronautic to wind-energy recovery. Multiple studies have been proposed, based on physical models with different levels of complexity. For instance, the geometrical optimization of an airfoil with a two-dimensional steady flow model is considered in Ribeiro et al. (2012) when the unsteady flow situation is treated in Srinath and Mittal (2010). Complex optimizations for a range of Mach numbers of a two-dimensional airfoil have been performed by Li et al. (2002), and the robust optimization for uncertain flows and geometries is considered in Papadimitriou and Papadimitriou (2016). The numerical optimization of hydrofoils is to our knowledge much rarer, and the hydrofoil literature generally focuses on the stability (Inukai et al., 2001), wake energy reduction (Kandasamy et al., 2011), and free surface elevation (Duvigneau and Visonneau, 2004) aspects.

This paper reports the development and application, in partnership with the *Groupama Team France (GTF)* design team, of an Efficient Global Optimization (EGO) procedure tailored to the flexible hydrofoil optimization problem. We restrict ourselves to the optimization of a two-dimensional hydrofoil consisting of rigid forward and trailing parts, connected by a deformable element. The objective of the optimization is the reduction of the hydrodynamic drag at several BS and lift conditions, with cavitation constraints. Cavitation is indeed an important concern for the hydrofoil performance (Sedlar et al., 2016), but its numerical prediction remains a difficult problem, as shown for instance in Leroux et al. (2005), Coutier-Delgosha et al. (2007), Ducoin et al. (2009), Akcabay et al. (2014). These difficulties explain that cavitation aspects are usually not considered in hydrofoil shape optimization, unless when the objective is precisely to delay the cavitation, such as in WEI et al. (2015).

The optimization of flexible hydrofoils raises many challenges. First it requires multi-design variables to prescribe the foil geometry at rest (unloaded shape), on the one hand, and the elastic properties and the internal structure of the deformable parts, on the other hand. Second, the evaluation of the performance of a given design requires the resolution of a fully nonlinear fluid-structure interaction (FSI) problem, possibly combining different boat speeds, with a numerical cost that inherently limits the number of designs that can be evaluated. Finally, for a given design, the consideration of non-cavitating flow requires an appropriate treatment of constraints that are costly to evaluate. As a matter of fact, the numerical optimization requires dedicated techniques to circumvent the computational complexity of the problem. In the present work, the numerical modeling involves the coupling of the two-dimensional vortex lattice method solver XFOIL (Drela, 1989), to compute the flow, with the structural ARA software developed by K-Epsilon (Durand et al., 2014), to compute the hydrofoil deformations. Even for this two-dimensional modeling, the computational times prevent the direct optimization based on the FSI solution with standard methods.

We then rely on a surrogate model to reduce the computational burden of optimizing complex systems with costly objective function estimation (Simpson et al., 2001). Specifically, we use Gaussian processes (Kleijnen, 2009) surrogate models with Efficient Global Optimization (EGO) strategies (Jones et al., 1998), that have been previously applied to aerodynamic drag reduction (Jeong et al., 2005), vibration reduction for rotating aircrafts (Glaz et al., 2009), optimization of FSI problems (Aghajari and Schäfer, 2015) and sail trimming optimization (Sacher et al., 2017). The classification approach for discontinuous or binary constraints proposed in Basudhar et al. (2012) is extended to the treatment of the cavitation constraints in the optimization procedure. In our approach, The admissible domain is approximated using Least-Squares Support Vector Machine (LS-SVM) (Suykens and Vandewalle, 1999) regression from the previous observations. An original contribution of the work is a new probabilistic treatment of the constraints which is

designed to mitigate the risk of exploring unfeasible solutions (which yield a cavitating flow).

This optimization procedure used in the present work is briefly described in Section 2, while Section 3 discusses the formulation of the optimization problem for the flexible hydrofoil, including the definition of the optimization variables and a brief discussion of the numerical models. Section 4 presents the optimization results, emphasizing on impact of the flexibility of the performance of the optimal hydrofoil. In particular, these performances are contrasted with the case of a rigid hydrofoil optimized for the same conditions. We have tried to provide as much as possible quantitative results, but for confidentiality reasons related to the partnership with the GTF design team, some values can not be provided and only trends are reported. Finally, conclusions and recommendations for future development are given in Section 5.

## 2. GP-based constrained optimization

The objective is to estimate the solution  $\mathbf{x}_{\text{opt}}$  of the following abstract optimization problem

$$\min_{\mathbf{x} \in \Omega} \mathcal{P}(\mathbf{x}), \quad \text{s.t.} \quad \mathbf{Q}(\mathbf{x}) \geq 0,$$

where  $\mathbf{x} \in \mathbb{R}^d$  is the vector of control variables,  $\Omega$  the optimization domain,  $\mathcal{P} : \Omega \rightarrow \mathbb{R}$  the objective function and  $\mathbf{Q} : \mathbf{x} \rightarrow \mathbb{R}^m$  is the vector of constraints. Even in the unconstrained case,  $m = 0$ , finding the global optimum of  $\mathcal{P}$  can be very costly, in particular when its evaluation is numerically expensive. The use of surrogate models for  $\mathcal{P}$  is then a classical solution to reduce the computational burden of complex systems optimization (Simpson et al., 2001). We use Gaussian processes (GP) (Kleijnen, 2009) which, owing to their statistical nature, provide both a prediction of the objective function and a measure of the uncertainty (variance) in the prediction. These features are appealing in optimization, as they can be used to derive rigorous optimization strategies based on the maximization of the Expected Improvement (EI) criterion, leading to methods referred globally as the Efficient Global Optimization (EGO) (Jones et al., 1998) methods. In Section 2.1 we review the construction of the GP model for  $\mathcal{P}$  and the resulting optimization strategy in the unconstrained case. The constrained case is considered in Section 2.2.

### 2.1. Unconstrained EGO using GP

Consider a set of  $n$  training inputs points  $\mathbf{X}_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , each associated to a noisy observation  $y_i$  of the objective function. Specifically, it is assumed that  $y_i = \mathcal{P}(\mathbf{x}_i) + \varepsilon_i$ , where the  $\varepsilon_i$  are Gaussian measurement noises, assumed for simplicity independent and identically distributed with variance  $\sigma_\varepsilon^2$ . The GP construction considers that  $\mathcal{P}$  is a realization of a zero-mean multivariate Gaussian process with covariance function  $C_{\mathcal{P}}$ . In this work, we consider the multidimensional squared exponential covariance functions defined by

$$C_{\mathcal{P}}(\mathbf{x}, \mathbf{x}_{\text{opt}}; \boldsymbol{\Theta}) \doteq \theta_1 \prod_{i=1}^d \exp\left(-\frac{(x_i - x'_{\text{opt},i})^2}{2l_i^2}\right) + \theta_2, \quad (1)$$

where  $\boldsymbol{\Theta} = \{\theta_1, \theta_2, l_1, l_2, \dots, l_d\}$  is a vector of hyper-parameters to be inferred from the observations. From the conditional rules of joint Gaussian distributions (Rasmussen and Williams, 2006), the best prediction  $\widehat{\mathcal{P}}_n(\mathbf{x})$  of  $\mathcal{P}(\mathbf{x})$ , i.e. the mean of  $y$ , and the prediction variance  $\widehat{\sigma}_{\mathcal{P}}^2(\mathbf{x})$  are given by

$$\widehat{\mathcal{P}}_n(\mathbf{x}) = \mathbf{k}^T(\mathbf{x}) (\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\varepsilon^2 \mathbf{I})^{-1} \mathbf{Y}_n, \quad (2)$$

$$\widehat{\sigma}_{\mathcal{P}}^2(\mathbf{x}) = \kappa(\mathbf{x}) + \sigma_\varepsilon^2 - \mathbf{k}^T(\mathbf{x}) (\mathbf{C}(\boldsymbol{\Theta}) + \sigma_\varepsilon^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}). \quad (3)$$

where  $\mathbf{C}(\boldsymbol{\Theta}) \in \mathbb{R}^{n \times n}$  is the covariance matrix with entries  $C_{ij}(\boldsymbol{\Theta}) \doteq C_{\mathcal{P}}(\mathbf{x}_i, \mathbf{x}_j; \boldsymbol{\Theta})$ ,  $1 \leq i, j \leq n$ ,  $\kappa(\mathbf{x}) \doteq C_{\mathcal{P}}(\mathbf{x}, \mathbf{x}; \boldsymbol{\Theta})$ ,  $\mathbf{k}(\mathbf{x})$  is the covariance vector between

the observations in  $\mathbf{X}_n$  and  $\mathbf{x}$ , and  $\mathbf{I}$  is the identity matrix of  $\mathbb{R}^n$ . The hyper-parameters  $\Theta$  and noise variance  $\sigma_\varepsilon^2$  can be determined by maximizing the log-marginal likelihood (Rasmussen and Williams, 2006) using an evolution strategy algorithm (Hansen, 2006). More details on GP meta-models can be found in (Rasmussen and Williams, 2006).

Let  $\hat{\mathbf{x}}_n$  be the optimum of  $\hat{\mathcal{P}}_n(\mathbf{x})$ . It is expected that  $\hat{\mathbf{x}}_n \approx \mathbf{x}_{\text{opt}}$  if the approximation error  $\hat{\mathcal{P}}_n - \mathcal{P}$  is small enough. The advantage of minimizing  $\hat{\mathcal{P}}_n$  instead of  $\mathcal{P}$  is that GP models are usually inexpensive to evaluate compared to the original objective function. To control the error in the approximation, one proceeds sequentially by adding progressively new observation points in the area of interest. A deterministic optimization procedure would choose the next point  $\mathbf{x}_{n+1}$  as the optimal point of  $\hat{\mathcal{P}}_n$ . However, the GP model provides probabilistic information that can be exploited to propose more robust strategies based on merit functions, which combine the prediction and its variance. In this work, we use the Augmented Expected Improvement (AEI) merit function (Huang et al., 2006), which estimates the expected progress in the objective, taking into account the noise in the observed values and the prediction variance:

$$AEI(\mathbf{x}) = EI(\mathbf{x}) \left( 1 - \frac{\sigma_\varepsilon}{\sqrt{\hat{\sigma}_{\mathcal{P}}^2(\mathbf{x}) + \sigma_\varepsilon^2}} \right), \quad (4)$$

where the Expected Improvement  $EI(\mathbf{x})$  is defined by

$$EI(\mathbf{x}) = \hat{\sigma}_{\mathcal{P}}(\mathbf{x}) [\Phi(u(\mathbf{x})) + \phi(u(\mathbf{x}))], \quad (5)$$

$$u(\mathbf{x}) = \frac{\hat{\mathcal{P}}_n(\mathbf{x}_{n,n}) - \hat{\mathcal{P}}_n(\mathbf{x})}{\hat{\sigma}_{\mathcal{P}}(\mathbf{x})}, \quad (6)$$

with  $\Phi$  and  $\phi$  the cumulative and density functions of the standard Gaussian distribution, and  $\mathbf{x}_{n,n} \in \mathbf{X}_n$  is the current effective best solution (see Huang et al. (2006)). The optimum  $\mathbf{x}_{n+1}$  of the AEI is added to  $\mathbf{X}_n$ , and  $\mathcal{P}$  is evaluated at the new point providing  $y_{n+1}$ . Setting  $n \leftarrow n+1$ , a new iteration can start updating the GP model with the new observation. Overall, each iteration requires one computation of the objective and the resolution of two optimization problems: a first one for the hyper-parameters of the GP model, and a second one to find the AEI optimum. The iterations of the GP-based optimization problem are continued until a stopping criterion is satisfied or the resources allocated to the optimization have been exhausted.

## 2.2. EGO under constraints, a classification based approach

EGO methods with inequality constraints was considered in Schonlau (1997). The key idea was to rely on  $m$  additional surrogates to estimate the constraints  $Q_i(\mathbf{x})$ . For Gaussian Process models, one can easily determine the probability  $P_i(Q_i \geq 0|\mathbf{x})$  that a constraint  $Q_i$  is satisfied at  $\mathbf{x}$  and, assuming the independence of the constraints probability, to come up with the consolidated probability  $P(\mathbf{Q} \geq 0|\mathbf{x}) = \prod_{i=1}^m P_i(Q_i \geq 0|\mathbf{x})$ . This probability is used to modify the unconstrained AEI criterion and favor feasible regions:

$$AEIQ(\mathbf{x}) = AEI(\mathbf{x})P(\mathbf{Q} \geq 0|\mathbf{x}).$$

Although effective in many problems, this GP modeling of the constraints faces several limitations. First, its computational cost increases with the number  $m$  of constraints and can be an issue for problems with large  $m$ . Second, the approximation by GP models assumes a sufficient smoothness of the  $Q_i$ , which must be evaluable for almost every  $\mathbf{x} \in \Omega$ . This rules out the case of binary constraints (feasible/infeasible) and models having no solution for  $Q(\mathbf{x}) \neq 0$ . Finally, the approach is also limited to situations where the feasible domain has an explicit representation in terms of the constraints. The optimization of hydrofoils

reaches some of these limitations as it is difficult to express feasible geometries in terms of constraints on design parameters, while some values may lead to uncomputable solutions.

Classification methods recently proposed in Basudhar et al. (2012) are better suited to deal with discontinuous or binary constraints in a GP-based optimization procedure. The binary classifier considers two classes  $C^+$  and  $C^-$  over  $\Omega$ , corresponding to the feasible and unfeasible domains respectively. Each  $\mathbf{x}_i$  of  $\mathbf{X}_n$  is equipped with a value  $z_i = \pm 1$  depending on its membership  $C^\pm$ . To predict the class of a new point  $\mathbf{x}$  we introduce a classification function  $h: \mathbf{x} \in \Omega \rightarrow \mathbb{R}$ , such that  $z(\mathbf{x}) = \text{sign}h(\mathbf{x})$ . A Least-Squares Support Vector Machine (LS-SVM) (Suykens and Vandewalle, 1999) is used to construct  $h$ . The LS-SVM method extends the original Support Vector Machine (SVM) classifier Vapnik (1995) to quadratic penalization, resulting in a linear (but non-sparse) system to be solved (see below). Further details and discussion on SVM and LS-SVM methods can be found in Cawley (2006).

The LS-SVM method (Suykens and Vandewalle, 1999) is a linear classifier, for  $C^+$  and  $C^-$ , in a feature space induced by the transformation  $\phi: \Omega \rightarrow \mathcal{F}$ :

$$h(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b. \quad (7)$$

Here,  $\mathbf{w}$  and  $\phi(\mathbf{x})$  are the weights and features vectors, while  $b \in \mathbb{R}$  is a constant. The feature space  $\mathcal{F}$  is generated by a reproducing kernel  $r: \Omega \times \Omega \rightarrow \mathbb{R}$  representing the inner product between images in  $\mathcal{F}$  of vectors:  $r(\mathbf{x}, \mathbf{x}_{\text{opt}}) = \phi(\mathbf{x}) \cdot \phi(\mathbf{x}_{\text{opt}})$ . We use the classical Gaussian kernel,

$$r(\mathbf{x}, \mathbf{x}_{\text{opt}}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_{\text{opt}}\|^2}{2\lambda^2}\right), \quad (8)$$

with scale factor  $\lambda \in \mathbb{R}$  to be adjusted. The LS-SVM parameters  $(\mathbf{w}, b)$  satisfy the primal constrained optimization problem

$$\min_{\mathbf{w}, b, e} \frac{1}{2} \|\mathbf{w}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^n e_i^2, \quad (9)$$

$$\text{s.t. } z_i = \mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, n,$$

With trade-off parameter  $\gamma \in \mathbb{R}^+$  (to be fixed) and relaxation variables  $e_i$  allowing for miss-classification. By writing the Lagrangian of the optimal problem, we introduce the  $\alpha_i \in \mathbb{R}$ , that are the Lagrange multipliers of the constraints. Denoting  $\mathbf{z} = (z_1 \dots z_n)^T$ , the optimality conditions of the Lagrangian are used to derive a linear system for the dual model parameters  $b$  and  $\alpha = (\alpha_1 \dots \alpha_n)^T$ ,

$$\begin{bmatrix} \mathbf{R} + \gamma^{-1} \mathbf{I} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ 0 \end{bmatrix}, \quad (10)$$

where  $\mathbf{R} \in \mathbb{R}^{n \times n}$  is the kernel matrix with  $R_{ij} = r(\mathbf{x}_i, \mathbf{x}_j)$ . This system is solved via a Cholesky factorization (Cawley, 2006), and  $h$  in (7) is expressed in terms of dual model parameters to obtain

$$h(\mathbf{x}) = \sum_{i=1}^n \alpha_i r(\mathbf{x}_i, \mathbf{x}) + b. \quad (11)$$

The LS-SVM classifier depends on two parameters,  $\gamma$  and  $\lambda$ , to be fixed when solving (10). We rely on a Leave-One-Out procedure to determine  $\gamma$  and  $\lambda$  that minimize the Predicted Residual Sum-of-Squares (PRESS) criterion (Allen, 1974). A simple expression for the predicted residuals is proposed in Cawley and Talbot (2003), which is presently minimized for  $\lambda$  and  $\gamma$  with the CMA-ES algorithm (Hansen, 2006).

Finally, the LS-SVM binary classifier is extended to a probabilistic classification, relating  $h$  in (11) to the probability of the class  $C^+$ , denoted  $P(C^+|\mathbf{x})$ . A comparison of several probability models for the LS-SVM classification is provided in Van Calster et al. (2007). We use the sigmoid function (Platt, 1999) and expressing the probability of  $C^+$  as

M. Sacher et al.

Ocean Engineering 157 (2018) 62–72

$$P(C^+|\mathbf{x}) = \frac{1}{1 + \exp(Ah(\mathbf{x}) + B)}. \quad (12)$$

The parameters  $A$  and  $B$  of the sigmoid are determined by minimizing the probability of misclassification, see Platt (1999), Lin et al. (2007). In practice, the probability  $P$  goes to 1 (resp. 0) as the classifier is certain that  $\mathbf{x}$  is feasible and belongs to  $C^+$  (resp. unfeasible and belongs to  $C^-$ ), while a value of  $P = 1/2$  denotes a complete uncertainty in the classification. This can occur because  $\mathbf{x}$  is far from any observations in  $\mathbf{X}_n$  or close to the interface between the two classes.

Following Schonlau (1997) and Basudhar et al. (2012), the probability of feasibility/evaluability is used to derive from the AEI a merit function and select a new point  $\mathbf{x}_{n+1}$  that yields the highest expected improvement while having a high probability of feasibility/evaluability. To this end, Basudhar et al. (2012) uses a modified version of the sigmoid function in (12) that accounts for the distances to the closest classified observations (in  $\mathbf{X}_n$ ). While improving the misclassification rate, this modification results in a discontinuous merit function which prevents the use of gradient-based optimization tools to determine  $\mathbf{x}_{n+1}$ . Further, their strategy requires the addition of  $d + 1$  new observation points per iteration (to maintain isotropy in the observation points distribution) with an associated cost deemed too important in our application. Instead, we use the original definition (12) to extend the unconstrained AEI and we consider a sequential infilling strategy that follows the EGO iterations, by selecting alternatively one of the present definitions for  $\mathbf{x}_{n+1}$ :

$$\mathbf{x}_{n+1} = \underset{\mathbf{x} \in \Omega}{\operatorname{argmax}} \quad AEI(\mathbf{x})P(C^+|\mathbf{x}), \quad (13)$$

$$\mathbf{x}_{n+1} = \underset{\mathbf{x} \in \Omega}{\operatorname{argmax}} \quad AEI(\mathbf{x}) \quad \text{s.t.} \quad P(C^+|\mathbf{x}) \geq \rho, \quad (14)$$

$$\mathbf{x}_{n+1} = \underset{\mathbf{x} \in \Omega}{\operatorname{argmin}} \quad \widehat{\mathcal{P}}_n(\mathbf{x}) \quad \text{s.t.} \quad P(C^+|\mathbf{x}) \geq \rho, \quad (15)$$

$$\mathbf{x}_{n+1} = \underset{\mathbf{x} \in \Omega}{\operatorname{argmax}} \quad [AEI(\mathbf{x})P(C^+|\mathbf{x})(1 - P(C^+|\mathbf{x}))]. \quad (16)$$

The first definition in (13) corresponds to the extension of the AEI favoring points with high chance of feasibility. The second expression (14) maximizes the original AEI definition by enforcing a minimal probability  $\rho$  of feasibility; we use  $\rho = 0.5$ . In (15), the predicted cost is directly minimized and enforced by  $\rho$ . Finally (16) combines the feasibility and infeasibility probabilities to favor areas where the classification is the most uncertain ( $P \sim 0.5$ ), with the objective to improve the exploration of the feasible domain boundaries. In practice, the new points  $\mathbf{x}_{n+1}$  are determined using CMA-ES algorithms without Hansen (2006) or with constraints Arnold and Hansen (2012), depending on the considered definition.

As a final note, we remark that the optimization procedure can generate points that are found unfeasible in the sense that  $\mathcal{P}(\mathbf{x}_{n+1})$  cannot be computed. This is typically the case when  $\mathbf{x}_{n+1}$  corresponds to an impossible geometrical configuration, or to an extreme situation for which the numerical code is not able to converge. In that case, the point is deemed infeasible, setting  $z_{n+1} = -1$ , but no value of the objective function is provided and the construction of the GP model for  $\mathcal{P}$  simply disregard the missing data. As a result, the GP model of  $\mathcal{P}$  and the LS-SVM classification can involve different numbers of observations.

### 3. Hydrofoil optimization setup

This Section introduces the formulation of the constrained optimization problem for the flexible hydrofoil, as defined in collaboration with GTF. Note that some details of the optimization problem are omitted for confidentiality reasons. We also briefly discuss the fluid and structural solvers used in this work.

#### 3.1. Problem formulation

The optimization of the hydrofoil uses control variables  $\mathbf{x}$  which parameterize the geometry and of the hydrofoil at rest, i.e. in the absence of fluid loads, and the elastic properties of the element that gives flexibility to the trailing edge (see below). The optimization concerns the minimization of the hydrofoil drag force over a set of 4 sailing conditions, with an additional set of 4 constraints designed to prevent the cavitation of the flow. Each condition corresponds to a prescribed boat speed and an hydrodynamic lift force developed by the hydrofoil. Note that the 4 drag points and the 4 constraints correspond to different conditions so the evaluation of the objective function  $\mathcal{P}$  and the vector of constraints  $\mathbf{Q}$  at given  $\mathbf{x}$  requires solving 8 fluid-structure interaction problems. The optimization problem is written as

$$\begin{aligned} \min_{\mathbf{x} \in \Omega} \mathcal{P}(\mathbf{x}) &= \sum_{i=1}^4 w_i C_{d_i}(\mathbf{x}), \\ \text{s.t.} \quad \begin{cases} Q_1(\mathbf{x}) = -C_{p_{\min_1}}(\mathbf{x}) \leq \lambda_1 \\ Q_2(\mathbf{x}) = -C_{p_{\min_2}}(\mathbf{x}) \leq \lambda_2 \\ Q_3(\mathbf{x}) = -C_{p_{\min_3}}(\mathbf{x}) \leq \lambda_3 \\ Q_4(\mathbf{x}) = -C_{p_{\min_4}}(\mathbf{x}) \leq \lambda_4 \end{cases} \end{aligned} \quad (17)$$

Here, the  $C_{d_i}$  are the drag force coefficients, the  $w_i > 0$  are prescribed weights, the  $C_{p_{\min_i}}$  are the minimum pressure coefficients, and the  $\lambda_i$  are the cavitation numbers defined by

$$\lambda = \frac{p_0 - p_v}{\frac{1}{2}\rho U_0^2}. \quad (18)$$

We have denoted  $U_0$  and  $p_0$  the reference velocity (that is, the boat-speed) and pressure,  $\rho$  the fluid density, and  $p_v$  the saturated vapor pressure. The reference pressure is taken as  $p_0 = p_{\text{atm}} + \rho gh$ , where  $p_{\text{atm}}$  is the atmospheric pressure,  $g$  is the gravity acceleration, and  $h$  is the immersion depth of the hydrofoil. The minimum of the pressure coefficient is defined by  $C_{p_{\min}} = \frac{p_{\min} - p_0}{\frac{1}{2}\rho U_0^2}$ , where  $p_{\min}$  is the minimal pressure over the hydrofoil surface. The simplest criterion to prevent cavitation is to ensure that the minimal pressure remains higher than the vapor pressure:  $p_{\min} \geq p_v$ . This condition can be expressed in terms of cavitation number  $\lambda$  to obtain the conditions on the minimum pressure coefficient  $C_{p_{\min}}$  to formulate the constraints in (17).

#### 3.2. Hydrofoil parametrization

The geometric and elastic characteristics of the flexible hydrofoil are defined starting from a baseline geometry provided by GTF. The baseline geometry was produced by an optimization procedure using an evolutionary method (Bäck and Schwefel, 1993) and assuming a complete rigidity of the hydrofoil (no elastic effects). The geometry at rest of the flexible hydrofoil is defined by rotating the sections of the baseline geometry (rigid hydrofoil), without changing the sections thickness. We denote  $p(s)$  the law of the sections rotation angle, where  $0 \leq s \leq 1$  is the reduced coordinate along the chord of the baseline geometry. In practice we restrict the rotation to the last 30% of the hydrofoil chord ( $p(s) = 0$ ,  $s \leq 0.7$ ) and use a smooth B-Spline interpolation (Piegl and Tiller, 1997; Wang et al., 2006) of  $p(s)$  between 4 control points  $0.7 < s_1 < \dots < s_4 = 1$ , as illustrated in Fig. 1. We shall refer to the rotation angles  $p_1, \dots, p_4$  at the control points as the geometrical parameters.

Regarding the flexibility of the hydrofoil, the results presented in the following consider a perfectly rigid hydrofoil except for an elastic element located in the backward part of the hydrofoil. As illustrated in Fig. 2, the location of the elastic element is prescribed by the reduced coordinates of its starting and ending points  $0 < s_1 < s_2 < 1$  (resp.  $0 < ds_1 < ds_2 < 1$ ) on the lower (resp. upper) surface of the baseline hydrofoil. The deformable element is made of a homogeneous material having a Young modulus  $E_{\text{LST}}$ ; it is further reinforced by a beam (shown



in black in Fig. 2) on the pressure surface (lower side). The elastic beam has a progressive Young modulus varying linearly between  $s_1$  and  $s_2$ , with mean value  $E_{\text{Beam}}$  and slope  $C_{\text{Beam}}$ . The use of a single, homogeneous, flexible element is supported by preliminary optimization studies briefly presented in the Appendix at the end of the paper. Overall, the parametrization of the deformable part involves 7 parameters (4 geometric coordinates and 3 elastic properties) globally referred to as the structural parameters.

### 3.3. Fluid-structure interaction solver

Given a value of the geometric and structural parameters, the evaluation of the objective function and the constraints of the optimization problem calls for the resolution of 4 FSI problems (one for each condition of the objective function). We assume that these problems have steady solutions for all values of the parameters and we rely on a quasi-steady approach to treat the coupling between the structural and fluid solvers.

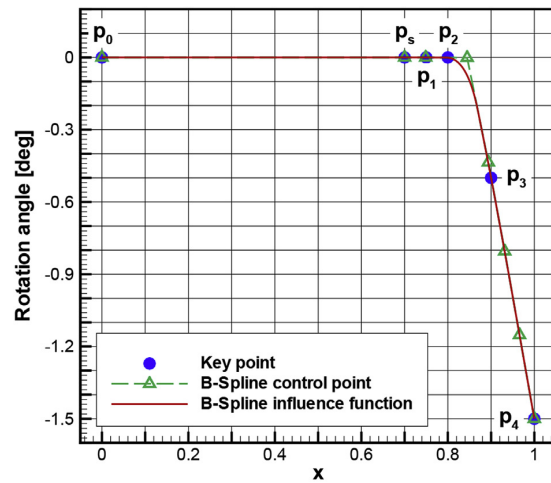
For the flow solver, we rely on XFOIL (Drela, 1989) which has been developed for many years and is widely used for many aerodynamic applications. While based on a two-dimensional airfoil panel code method, XFOIL can provide viscous and inviscid analyses. It incorporates a two-equation integral formulation of the viscous boundary layer and the approximate  $e^N$  envelope method (Van Ingen, 2008); it allows the prediction of transition points and separation bubble, through the resolution of the boundary layer and the transition equations which is based on a Newton method. Further details on XFOIL are given in Drela (1989) and it is compared with other CFD methods and solvers in Morgado et al. (2016). We only mention that the setting of our optimization problem imposes to enforce a prescribed lift force in the flow computation. This is achieved in XFOIL by determining the Angle of Attack (AoA) that yields the requested lift force. Overall, XFOIL is simple to use and has a reasonable computation cost on a classical workstation, in most of the cases, making it a natural candidate to be coupled with a structural solver (see for instance (MacPhee and Beyene, 2013) for a recent FSI study of wind turbine blades using XFOIL).

The ARA software developed by K-Epsilon was used to solve the structural model of the flexible hydrofoil. The solver was originally developed for FSI computations of sails (Durand et al., 2014) and proposes various structural elements (Durand, 2012). In this work, the elastic part of the hydrofoil is modeled by 2D Linear Strain Triangles (LST) (Pedersen, 1973) and Timoshenko beam elements (see Fig. 2). The nonlinear equilibrium solution is computed by a Newton method with Aitken relaxation.

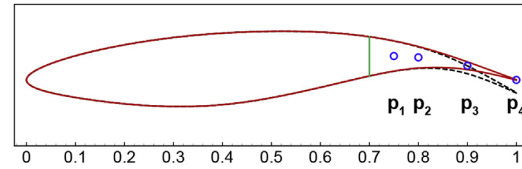
A quasi-monolithic algorithm (Durand, 2012) is finally applied to couple the structural and flow solvers. Briefly, in this algorithm, the resolution of the structural problem is nested inside the iterations of the nonlinear steady flow solver. This approach preserves the convergence rate and stability properties of the monolithic approach. The resulting coupled solver is finally driven by a utility that computes the set of FSI solutions required to estimate the objective function value and its constraints (17). This utility is itself nested into the optimization driver that decides of the sequence of new optimal parameters  $\mathbf{x}_{n+1}$  to be evaluated.

## 4. Results and discussions

We now present the results of the optimization problem. In particular, we contrast the optimal flexible hydrofoil with its rigid counterpart optimized for the same set of conditions and constraints. This comparison concerns the characteristics and performances of the two optimal hydrofoils. We do not provide a comparison in terms of computational times because the two optimizations use different strategies and have significantly different complexities. Indeed, the optimization of the rigid hydrofoil although involving a larger number of optimization variables has an objective function and constraints much faster to compute than in the flexible case (which requires the resolution of FSI problems), enabling the application of evolutionary-based methods (Bäck and Schwefel, 1993).



(a) B-Spline approximation of the rotation angle.



(b) Initial and deformed hydrofoils.

Fig. 1. Illustration of the unloaded geometry definition using a rotation of the baseline geometry. The top plot shows the B-Spline approximation of the rotation rule  $p(s)$  based on the rotation angles  $p_1$  to  $p_4$  at the control points. The bottom plot compares the baseline (red) and unloaded flexible geometries, the green vertical line indicating the chord position  $s = 0.7$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

In the following, the rigid hydrofoil corresponds to the optimized baseline geometry provided by GTF, assuming a perfectly rigid hydrofoil, while the unloaded geometry refers to the rest geometry resulting from our optimization of the flexible hydrofoil, that is, the rigid geometry with the rotations of the trailing edge sections (but without elastic deformations). We stress that the structural and geometrical parameters

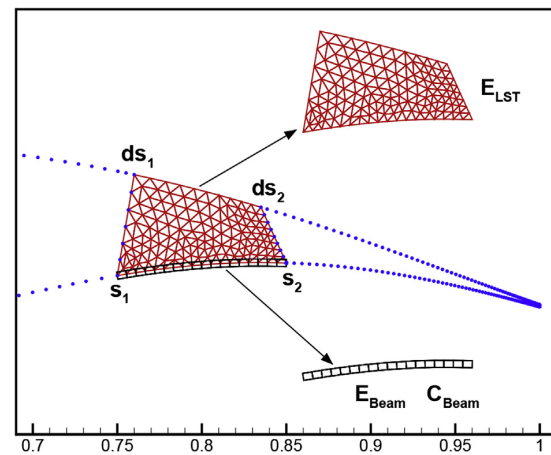


Fig. 2. Illustration of the structural parameters.

characterizing the flexible hydrofoil and are not changing with the boat speed. Thus, the rigid, unloaded, and in fact, all flexible geometries are coinciding if the elastic element is infinitely stiff, because the rigid geometry is optimal in the absence of elastic deformations. For a finite elastic stiffness, one needs, in general, to jointly optimize the geometrical and structural parameters of the flexible hydrofoil to improve the performances compared to the rigid case. Optimizing only for the structural (resp. geometrical) parameters usually leads to a degradation of the performance, compared to the rigid hydrofoil case, unless the geometrical (resp. structural) parameters are fixed close to their joint optimization solution values.

#### 4.1. Selection of the design variables

The computational time for the optimization of the flexible hydrofoil is an important aspect, and several tests were made to determine an appropriate trade-off between the computational and parametrization complexities, before arriving at the set-up described in Section 3.2 with the up to 11 parameters to be optimized (4 geometrical and 7 structural). The number of optimization variables can be further reduced, fixing, for instance, the material properties, the geometry of the deformable element, or changing the number of control points in the parametrization of the rotation rule. These reductions yield different complexities of the optimization problem with different computational times as a result. However, reducing the number of optimization variables also impacts the performance of the optimal design if the fixed parameters are not set close to their optimal values. Fig. 3 shows the evolutions with the computational time of the estimated minimum of the objective function for different numbers of optimization variables considered in the optimization problem. The plot shows that when nine variables are considered the optimization needs  $\approx 16$  times more resources to converge than for only five optimization variables, with close optimal objective values in this case. The case of two optimization variables is even less computationally demanding, but leads to a noticeable drop in the optimal performance because of inappropriately set values for the parameters disregarded in the optimization problem. These effects can be better appreciated from Table 1 which reports the objective function values and associated computational times for the different numbers of design variables to optimize. This underlines the importance of identifying the most influential variables of the optimization problem. In the following, we focus on the case of 9 optimization variables, corresponding to the 4 geometrical and 5 of the 7 structural parameters.  $E_{Beam}$  and  $C_{Beam}$  have been fixed close to their optimization solution values.

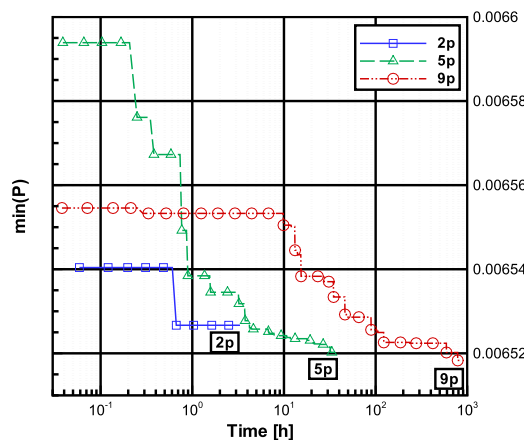


Fig. 3. Convergence of the objective function minimum with the computational time, for different numbers of optimization variables considered.

Table 1

Best objective function values, at convergence, and corresponding computational times for different numbers of optimization variables.

| Dimension | min $\mathcal{P}$ | Time [h] |
|-----------|-------------------|----------|
| 2         | 0.006527          | 3        |
| 5         | 0.006520          | 36       |
| 9         | 0.006518          | 844      |

#### 4.2. Optimal shapes and pressure distributions

Figs. 4 and 5 show for 20 and 40 kts of boat-speed respectively, the shapes of the rigid and flexible hydrofoils. These shapes are depicted at their effective angle of attack giving the prescribed lift force. The pressure coefficient distributions are also reported in the top part of the Figures. At low boat-speed (20 kts), in Fig. 4, the rigid and deformed hydrofoils have noticeable differences in term of angle of attack: the rigid one is higher. Regarding the pressure coefficient distributions, the rigid case presents a minimum at the leading edge of  $\approx -2.25$ , whereas in the flexible case the minimum peaks to  $C_p \approx -1.1$  only, at the same location. This can be explained as a less cambered geometry generally requires a higher angle of attack to achieve the same lift force, while the minimum pressure coefficient usually increases with the angle of attack. Anyway, the results indicate a poorer tolerance to cavitation for the rigid hydrofoil in these conditions.

For 40 kts of boat-speed, shown in Fig. 5, the two angle of attack are negative and the rigid one presents again the largest deviation from zero angle of attack. The deformation of the trailing edge is now pronounced, in the flexible case, as it can be appreciated using dashed black contour which corresponds to the unloaded flexible geometry (which comes from the GTF optimizations with flap laws). The trailing edge deflection is  $\approx 9.5$  deg. The minimum pressure coefficients are now located approximately at mid-chord with  $C_{p_{min}} \approx -0.5$  in the two cases. The distortions in the pressure coefficient distribution observed at  $\approx 0.8$  percent of the chord in the case of the flexible hydrofoil are due to the transitions between the rigid parts and the deformable element constituting the flexible hydrofoil.

The flow around the flexible hydrofoil at 40 kts of boat-speed has also been computed with a URANS transition model in order to validate the pressure coefficient distribution predicted by XFOIL. For that purpose, a structured mesh of the fluid domain around the hydrofoil has been

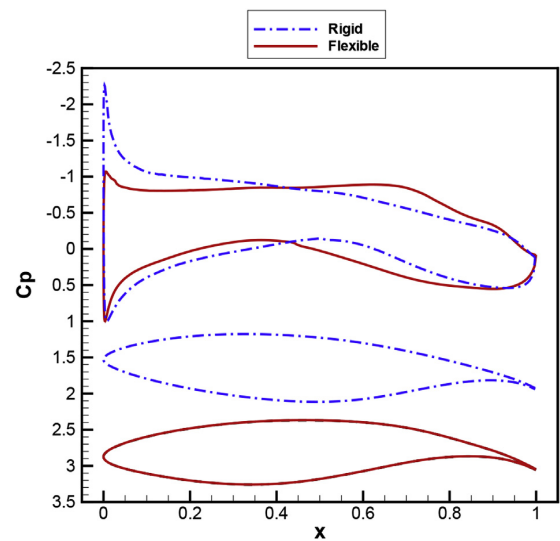


Fig. 4. Shapes and distributions of  $C_p$  at 20 kts of boat-speed for the rigid and flexible hydrofoils.



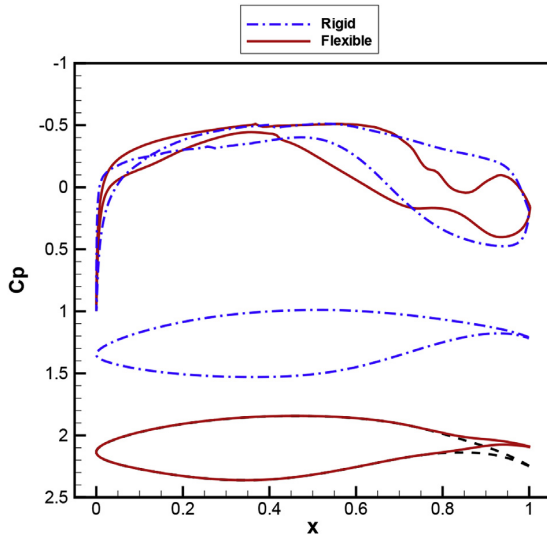


Fig. 5. Shapes and distributions of the pressure coefficient  $C_p$  at 40 kts of boat-speed for the rigid and flexible hydrofoils.

generated using  $\approx 110\,000$  elements; the flow was subsequently solved using FLUENT with the Transition SST turbulence model (Menter et al., 2006). The computed pressure coefficient distribution is reported in Fig. 6, together with the XFOIL prediction, highlighting the excellent agreement between XFOIL and the Transition SST flow model. Small differences in the pressure coefficients are visible only on the pressure side at  $\approx 90\%$  of the chord length, and the magnitude of the differences in the global hydrodynamic loads are 2.6% and 4.2% for the lift and drag coefficients respectively. Thus, the pressure coefficients computed by XFOIL can be considered reliable, even in the area of the deformable element.

Fig. 7 shows the evolutions of the camber along the chord for the rigid, unloaded and the loaded flexible hydrofoils at several boat-speeds. These geometrical quantities are computed based on the distances in the direction normal to the deformed chord, following the classical conventions. Note that because of the changes of the displacement of the trailing edge and the change in chord length with the boat-speed, for the flexible hydrofoil, the camber law in the forward part before the flexible element is also affected. The flexibility is seen to have an important effect

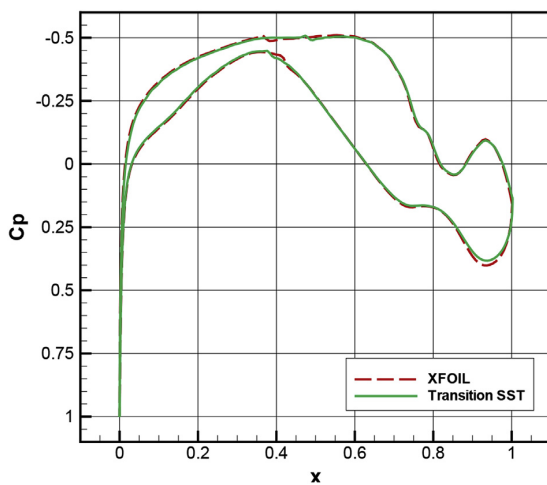


Fig. 6. Distributions of the pressure coefficients  $C_p$  predicted by XFOIL and with the URANS transition models.

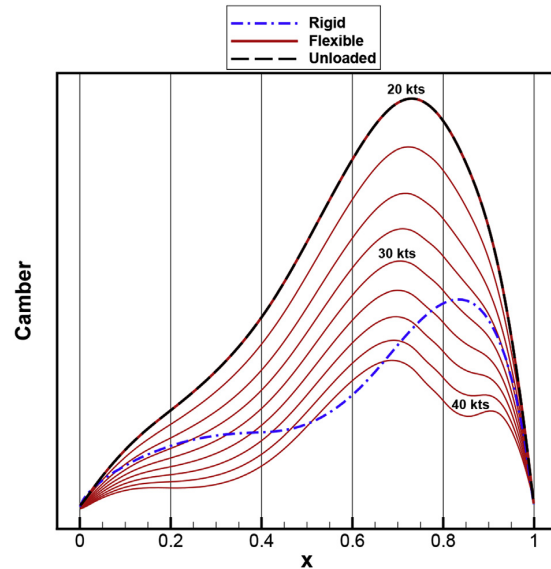


Fig. 7. Flexible hydrofoil camber distribution at several boat-speeds. The rigid hydrofoil case is also shown for comparison.

for a boat-speed larger than 20 kts. In addition, the shapes of the camber distribution of the flexible hydrofoil differ significantly from the rigid one for all the boat-speeds shown. Specifically, the maximum camber of the flexible foil at 20 kts of boat-speed is more than twice that of the rigid one, while at 40 kts of boat-speed it is less by roughly 20%. These evolutions of the apparent camber distribution with the boat-speed (and load in general) is one of the hydrodynamical advantages brought by the flexibility.

#### 4.3. Angles of attack and cavitation criterion

The effective angle of attack of the optimal hydrofoils is reported in Fig. 8 as a function of the boat-speed. The previous observations regarding the angle of attack are confirmed. The angle of attack of the rigid hydrofoil almost linearly decreases from  $\approx 3.9$  deg at 20 kts to  $\approx -0.9$  deg at 30 kts of boat-speed, and subsequently continues to decrease,

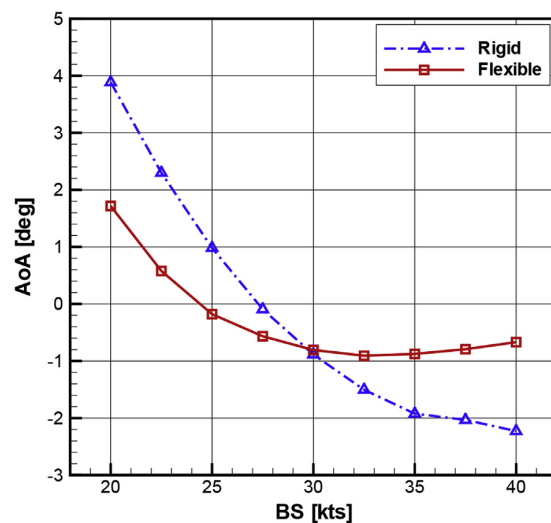


Fig. 8. Effective angle of attack as a function of the boat speed (BS).

but at a lower rate, when the boat-speed increases further. For the flexible case, the maximum of angle of attack is only  $\approx 1.8$  deg at 20 kts and decreases to  $\approx -0.1$  deg at 25 kts of boat-speed. Then, unlike in the rigid case, the flexible hydrofoil keeps an almost constant effective angle of attack between 0 and  $-1$  deg for the rest of boat-speed range, instead of decreasing continuously in the rigid case. Again, the flexibility can explain this behavior, as the deformation of the trailing edge (and reduction of the camber) discharges the foil when the boat-speed increases. This discharge calls for fewer changes in the angle of attack compared to the rigid case. Note that this sort of automatic trimming, through flexibility, may also be beneficial by saving some trimming efforts to the crew members.

Another positive impact of the flexible hydrofoil can be appreciated in Fig. 9 which reports the minimum pressure coefficients for the two optimal hydrofoils, as a function of the boat-speed. The dashed black curve corresponds to the critical value based on the cavitation number (18). We thus expect the optimal hydrofoils to have a minimum pressure coefficient greater than this critical value over the whole range of boat-speeds considered. However, we recall that the cavitation criterion is enforced for a finite set of only 4 distinct values of the boat-speed. For the maximal boat-speed shown, 40 kts, the two optimal foils, in fact, violate by a small margin the criteria based on the cavitation number, which indeed is not explicitly enforced for this boat-speed. As lower boat-speeds are considered, till 25 kts, the two foils satisfy the non-cavitating criterion by a large margin, suggesting that any constraints in this range of boat-speeds would not be active. Finally, when the boat-speed approaches 20 kts the behavior of the minimum pressure coefficient differs between the two hydrofoils: while for the flexible hydrofoil  $C_{p_{\min}}$  remains away from its critical value, it decreases sharply and even becomes significantly lower than its critical value in the rigid case. Comparing the trends of the two hydrofoils, it can be concluded that the flexibility yields a design with minimum pressure coefficient much less dependent on the boat-speed, and therefore having lower chances of violating the non-cavitating conditions for a boat-speed other than in the constraints.

#### 4.4. Drag performances

To complete the comparison between the flexible and rigid hydrofoil, we report in Fig. 10 the drag coefficients as a function of the boat-speed. The actual values are not shown, for confidentiality reasons, but the plot allows to appreciate the improvement brought by the flexible hydrofoil. Except for the lowest boat-speed (20 kts), the flexible hydrofoil has a drag

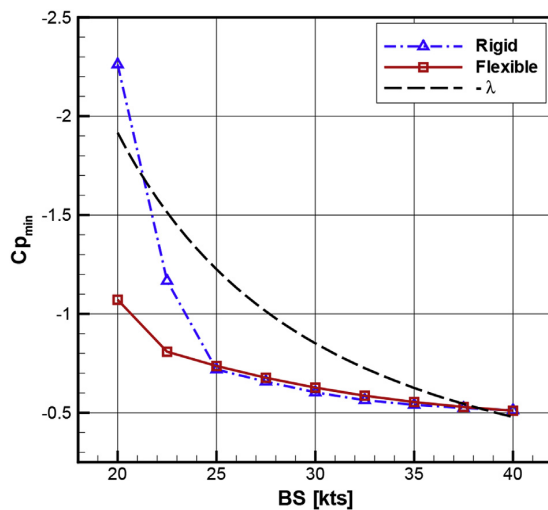


Fig. 9. Minimum pressure coefficients  $C_{p_{\min}}$  as a function of the boat-speed (BS). Also shown is the cavitation number  $\lambda$  given by (18).

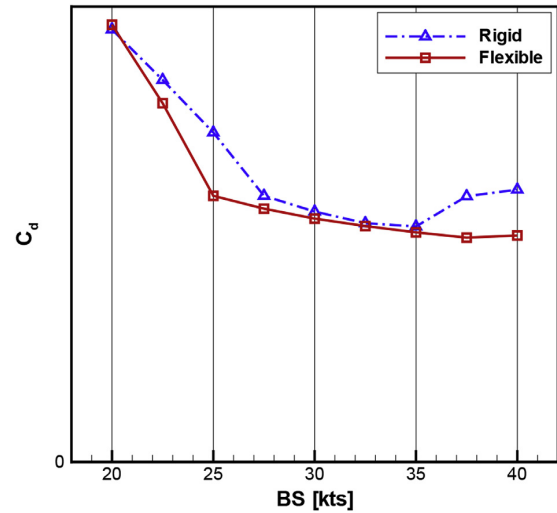


Fig. 10. Drag  $C_d$  of the optimal hydrofoils as a function of the boat-speed (BS).

coefficient  $C_d$  consistently lower than for the rigid one. The improvement is particularly significant at 25 and 40 kts of boat-speed, with roughly 20% drag reduction at the highest boat-speed. Again, these improvements come from the flexibility which allows the hydrofoil to naturally adapt its camber with the boat-speed, with a thinner geometry at rest and lower variations of the angle of attack in the boat-speed range.

For a boat-speed between 20 and 25 kts, the hydrofoil drag is mainly depending on the location of the transition point on the suction side, denoted  $x_{tr}$ . As the boat-speed is lowered, the angle of attack increases to satisfy the required lift force and the transition point  $x_{tr}$  moves toward the leading edge with a sharp increase in the friction drag (see Fig. 10). Note that in XFOIL the location of the transition point is computed whenever it appears to be in the first 40% of the chord. Otherwise, it is set at this maximal location. To highlight the importance of the transition point on the drag of the hydrofoil, Fig. 11 presents the evolutions of the flexible hydrofoil drag and the location of the transition point as functions of the boat-speed. The sharp drop in the hydrofoil drag, around 22.5 kts of boat-speed, is clearly related to the displacement of the transition point away

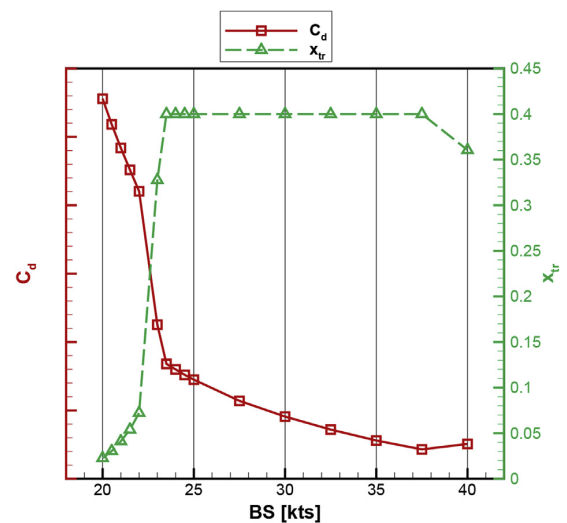


Fig. 11. Drag  $C_d$  and transition point location  $x_{tr}$  for the flexible hydrofoil as functions of the boat-speed (BS).

**Table 2**

Comparison of the hydrodynamical characteristics of the Rigid 30 kts and flexible solutions at 30 kts of boat-speed.

|               | Rigid 30 kts | Flexible | Difference |
|---------------|--------------|----------|------------|
| AoA           | – 0.36°      | – 0.81°  | 0.55       |
| $C_{p_{min}}$ | – 0.605      | – 0.627  | 0.035      |
| $C_d$         | –            | –        | – 0.012    |

from the leading edge. Note that by using different weights and adding more conditions in (17), one could further improve the hydrofoil drag at a low boat-speed, by delaying the displacement of the transition point, but to the detriment of the drag at higher boat-speeds.

#### 4.5. Optimal hydrofoil at 30 kts

To illustrate the importance of the selected conditions used in the definition of the objective function, we optimize another rigid hydrofoil based on the minimization of the drag and non-cavitating constraint in (17) based on the unique condition corresponding to 30 kts of boat-speed. The design variables are reduced to the 4 geometrical angles  $p_{1,...,4}$  and we refer to this optimized hydrofoil as the Rigid 30 kts one.

Table 2 compares the angle of attack (AoA), minimum pressure coefficient  $C_{p_{min}}$  and drag  $C_d$  of the flexible and Rigid 30 kts hydrofoils at 30 kts of boat-speed. The Rigid 30 kts case has a lower incidence, with a drag reduction of roughly 1.2% compared to the flexible case. The two  $C_{p_{min}}$  are relatively close and far from the criterion based on the cavitation number  $\lambda$ . The closeness of the two hydrofoils at this boat-speed of 30 kts can be further appreciated from Fig. 12, where the two distributions of pressure coefficients  $C_p$  and the shapes plotted at the same angle of attack for comparison purposes are shown. The main difference in the pressure coefficient is the smoother character of the distribution in the Rigid 30 kts case, which can be explained by the absence of the localized elastic deformations present in the flexible case. However, the differences between the two distributions are small as one could have expected from the similarity of the shapes and effective angle of attack.

However, as soon as the boat-speed departs from 30 kts, the flexible hydrofoil exhibits a lower drag compared to the Rigid 30 kts one. This can be appreciated in Fig. 13 which reports the differences in the drag coefficients, with positive values in favor of the flexible case. Note that

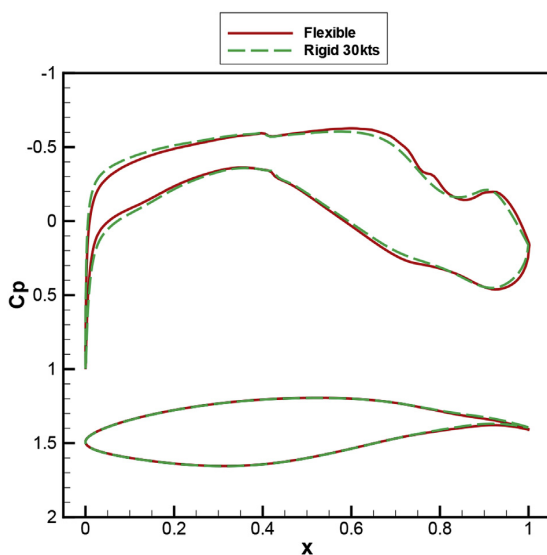


Fig. 12. Shapes and pressure coefficient  $C_p$  distributions for the flexible and Rigid 30 kts solutions at 30 kts of boat-speed.

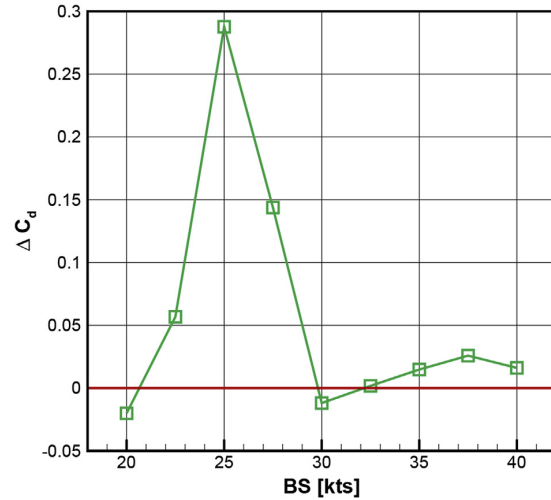


Fig. 13. Difference  $\Delta C_d$  in the drags of the Rigid 30 kts and flexible optima, as a function of the boat-speed (BS).

these consistent improvements were not necessarily expected as the Rigid 30 kts case does not enforce any non-cavitating conditions at boat-speeds other than 30 kts. Therefore, the Rigid 30 kts optimization may have produced a cavitating design with lower drag for other boat-speeds. Note also the higher drag for the flexible hydrofoil, at 20 kts of boat-speed, where the cavitation criterion is actually not satisfied by the rigid solution.

#### 5. Conclusions

An advanced multi-design constrained optimization method has been proposed in this paper for the design of a two-dimensional hydrofoil with a flexible trailing edge. Surrogate-model based optimization, with inequality constraints, have been used to enable the optimization of the hydrofoil at a reasonable computational cost. The surrogates are constructed and updated using a limited number of carefully selected resolutions of the fully nonlinear fluid-structure interaction problem. The nonlinear FSI solver used in this work couples a two-dimensional vortex lattice method code for the flow around the hydrofoil with a static structural solver, based on linear stress triangle and Timoshenko beam elements, for the hydrofoil deformations. The optimization involves 4 parameters describing the rotation of a base geometry, and seven elastic parameters describing the properties of the elastic bound between the rigid leading part and trailing edge of the hydrofoil. These parameters are sought to globally minimize the hydrofoil drag forces in selected conditions (boat speeds and lift forces), while constraints were introduced to ensure noncavitating conditions.

The performances of the resulting optimal flexible hydrofoil have been compared to the rigid case optimized for the same set of conditions. It was found that the flexible hydrofoil performs globally better over the whole set of conditions, compared to the rigid one. In fact, the flexible hydrofoil was found to have drag forces lower than that of the rigid one the whole range of boat speeds, except around 20 knots where the two cases have essentially the same drag force. The improvement in the drag force brought by the flexibility is particularly noticeable for the highest boat speeds ( $\geq 35$  knots) and around 25 knots. A closer inspection of the optimal solutions revealed that not only the flexible hydrofoil performs better for most conditions, but its angle of attack varies with the boat speed in a tighter range to achieve the prescribed lift conditions, compared to the rigid case. In particular, the flexible hydrofoil has a much smaller angle of attack at low boat speed, which immediately translates into a higher minimum of pressure and less

susceptibility to cavitation.

Overall, the optimization clearly evidences the various advantages of considering a hydrofoil with a flexible trailing edge. In particular, it demonstrates that, if correctly designed, the flexible trailing edge allows for an auto-adaption of the hydrofoil camber line and angle of attack with the boat speed, with a global reduction of the drag and lower cavitation risk.

Future developments will concern the optimization of a three-dimensional flexible hydrofoil with the use of more advanced flow models to account for the three-dimensional effects, turbulence, and

more complex cavitation criteria. We are currently exploring multi-fidelity strategies (Park et al., 2016) to tempered the increase in the computational costs in the three-dimensional case. Another aspect requiring further investigation is the dependence of the transition point displacement as a function of the hydrofoil angle of attack at lower boat speed range. Indeed, we believe that the flexible trailing edge could be effective in producing enough lift for the platform to take off at a lower boat speed.

#### Appendix. Preliminary study

We report in this appendix an optimization result to support the definition of the elastic hydrofoil used in the paper. Specifically, we considered a more general definition of the elastic hydrofoil using six consecutive homogeneous elastic elements in the elastic part ranging in the last 30 % of the hydrofoil chord (see Fig. 14). The elastic elements have a fixed size and their respective Young moduli  $E_{i=1,\dots,6}$  are optimized for the same objective function and constraints in (17), together with 3 geometrical parameters (rotations  $p_{1,\dots,3}$  defined in Section 3.2), leading to an optimization problem with 9 design variables. The Young moduli are also constrained to remain in a finite (positive) range of available materials.

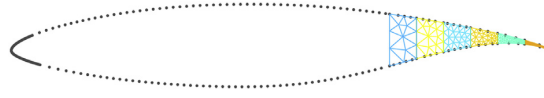


Fig. 14. Flexible hydrofoil composed of 6 elastic elements.

The optimal Young moduli  $E_{1,\dots,6}$  are reported (as a function of the chord length coordinate), in Fig. 15. The values have been divided by a reference value for confidentiality reasons. It is seen that the optimal solution corresponds to a localized flexibility, located at  $\approx 85\%$  of the chord length, and with the first and last elastic elements of the flexible hydrofoil having Young moduli equal to the maximum value allowed. In other words, the optimal design presents very stiff upstream part and trailing edge, separated by a significantly softer region. An idealized version of this solution is the case of a single homogeneous elastic element described in Section 3.2 and depicted in Fig. 2.

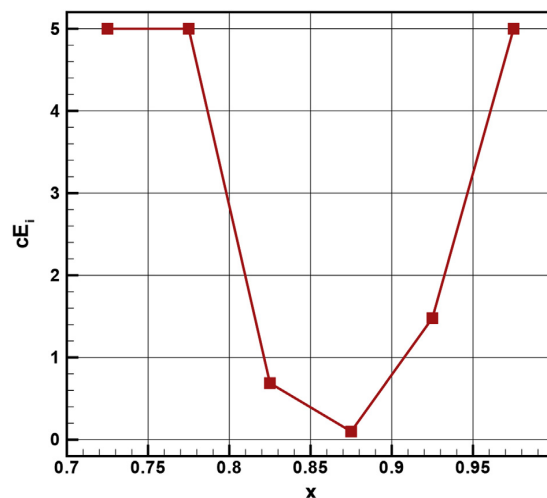


Fig. 15. Optimal distribution of the Young modulus (arbitrary units) with the chord length.

#### References

- Aghajari, N., Schäfer, M., 2015. Efficient shape optimization for fluid–structure interaction problems. *J. Fluids Struct.* 57, 298–313.
- Akabay, D.T., Chae, E.J., Young, Y.L., Ducoin, A., Astolfi, J.A., 2014. Cavity induced vibration of flexible hydrofoils. *J. Fluids Struct.* ISSN: 0889-9746 49, 463–484. <https://doi.org/10.1016/j.jfluidstruct.2014.05.007>. <http://www.sciencedirect.com/science/article/pii/S0889974614001030>.
- Allen, D.M., 1974. The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* 16 (1), 125–127.
- Arnold, D.V., Hansen, N., 2012. A (1+1)-CMA-ES for constrained optimisation. In: Soule, T., Moore, J.H. (Eds.), *GECCO*. ACM Press, Philadelphia, United States, pp. 297–304. <https://doi.org/10.1145/2330163.2330207>. <https://hal.inria.fr/hal-00696268>.
- Bäck, T., Schwefel, H.P., 1993. An overview of evolutionary algorithms for parameter optimization. *Evol. Comput.* 1 (1), 1–23.
- Basudhar, A., Dribusch, C., Lacaze, S., Missoum, S., 2012. Constrained efficient global optimization with support vector machines. *Struct. Multidiscip. Optim.* ISSN: 1615-1488 46 (2), 201–221. <https://doi.org/10.1007/s00158-011-0745-5>. <https://doi.org/10.1007/s00158-011-0745-5>.
- Cawley, G.C., 2006. Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. IEEE, pp. 1661–1668.
- Cawley, G.C., Talbot, N.L., 2003. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recognit.* ISSN: 0031-3203 36 (11), 2585–2592. [https://doi.org/10.1016/S0031-3203\(03\)00136-5](https://doi.org/10.1016/S0031-3203(03)00136-5). <http://www.sciencedirect.com/science/article/pii/S0031320303001365>.
- Coutier-Delgosh, O., Deniset, F., Astolfi, J.A., Leroux, J.-B., 2007. Numerical prediction of cavitating flow on a two-dimensional symmetrical hydrofoil and comparison to experiments. *J. Fluids Eng.* 129 (3), 279–292.
- Drela, M., 1989. *XFOIL: an Analysis and Design System for Low Reynolds Number Airfoils*. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-642-84010-4, pp. 1–12. [https://doi.org/10.1007/978-3-642-84010-4\\_1](https://doi.org/10.1007/978-3-642-84010-4_1).

- Ducoin, A., Deniset, F., Astolfi, J.A., Sigrist, J.-F., 2009. Numerical and experimental investigation of hydrodynamic characteristics of deformable hydrofoils. *J. Ship Res.* 53 (4), 214–226.
- Durand, M., 2012. Interaction fluide-structure souple et légère, application aux voiliers. PhD Thesis. Ecole Centrale de Nantes.
- Durand, M., Leroyer, A., Lothodé, C., Hauville, F., Visonneau, M., Floch, R., Guillaume, L., 2014. FSI investigation on stability of downwind sails with an automatic dynamic trimming. *Ocean. Eng.* 90, 129–139.
- Duvigneau, R., Visonneau, M., 2004. Hydrodynamic design using a derivative-free method. *Struct. Multidiscip. Optim.* ISSN: 1615-1488 28 (2), 195–205. <https://doi.org/10.1007/s00158-004-0414-z>.
- Glaz, B., Friedmann, P.P., Liu, L., 2009. Helicopter vibration reduction throughout the entire flight envelope using surrogate-based optimization. *J. Am. Helicopter Soc.* 54 (1), 12007.
- Hansen, N., 2006. The CMA evolution strategy: a comparing review. In: *Towards a New Evolutionary Computation*. Springer, pp. 75–102.
- Huang, D., Allen, T.T., Notz, W.L., Zheng, N., 2006. Global optimization of stochastic black-box systems via sequential kriging meta-models. *J. Glob. Optim.* 34 (3), 441–466.
- Hueber, F., Caponnetto, G., Poloni, C., 2017. A Passively Morphing Trailing Edge Concept for Sailing Hydrofoil. <https://doi.org/10.20944/preprints201706.0080.v1> preprints.
- Inukai, Y., Horiuchi, K., Kinoshita, T., Kanou, H., Itakura, H., 2001. Development of a single-handed hydrofoil sailing catamaran. *J. Mar. Sci. Technol.* ISSN: 1437-8213 6 (1), 31–41. <https://doi.org/10.1007/s007730170005>.
- Jeong, S., Murayama, M., Yamamoto, K., 2005. Efficient optimization design method using kriging model. *J. Aircr.* 42 (2), 413–420.
- Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. *J. Glob. Optim.* 13 (4), 455–492.
- Kandasamy, M., Peri, D., Ooi, S.K., Carrica, P., Stern, F., Campana, E.F., Osborne, P., Cote, J., Macdonald, N., de Waal, N., 2011. Multi-fidelity optimization of a high-speed foil-assisted semi-planing catamaran for low wake. *J. Mar. Sci. Technol.* ISSN: 1437-8213 16 (2), 143–156. <https://doi.org/10.1007/s00773-011-0119-0>.
- Kleijnen, J.P., 2009. Kriging metamodeling in simulation: a review. *Eur. J. Oper. Res.* ISSN: 0377-2217 192 (3), 707–716. <https://doi.org/10.1016/j.ejor.2007.10.013>. <http://www.sciencedirect.com/science/article/pii/S0377221707010090>.
- Leroux, J.-B., Coutier-Delgosha, O., Astolfi, J.A., 2005. A joint experimental and numerical study of mechanisms associated to instability of partial cavitation on two-dimensional hydrofoil. *Phys. Fluids* 17 (5). <https://doi.org/10.1063/1.1865692>. <http://scitation.aip.org/content/aip/journal/pof2/17/5/10.1063/1.1865692>.
- Li, W., Huysse, L., Padula, S., 2002. Robust airfoil optimization to achieve drag reduction over a range of Mach numbers. *Struct. Multidiscip. Optim.* ISSN: 1615-1488 24 (1), 38–50. <https://doi.org/10.1007/s00158-002-0212-4>.
- Lin, H.-T., Lin, C.-J., Weng, R.C., 2007. A note on Platt's probabilistic outputs for support vector machines. *Mach. Learn.* ISSN: 1573-0565 68 (3), 267–276. <https://doi.org/10.1007/s10994-007-5018-6>.
- MacPhee, D., Beyene, A., 2013. Fluid-structure interaction of a morphing symmetrical wind turbine blade subjected to variable load. *Int. J. Energy Res.* ISSN: 1099-114X 37 (1), 69–79. <https://doi.org/10.1002/er.1925>.
- Menter, F.R., Langtry, R., Likki, S., Suzen, Y., Huang, P., Völker, S., 2006. A correlation-based transition model using local variables—Part I: model formulation. *J. Turbomach.* 128 (3), 413–422.
- Morgado, J., Vizinho, R., Silvestre, M., PÃscoa, J., 2016. {XFOIL} vs {CFD} performance predictions for high lift low Reynolds number airfoils. *Aerosp. Sci. Technol.* ISSN: 1270-9638 52, 207–214. <https://doi.org/10.1016/j.ast.2016.02.031>. <http://www.sciencedirect.com/science/article/pii/S1270963816300839>.
- Papadimitriou, D.I., Papadimitriou, C., 2016. Aerodynamic shape optimization for minimum robust drag and lift reliability constraint. *Aerosp. Sci. Technol.* ISSN: 1270-9638 55, 24–33. <https://doi.org/10.1016/j.ast.2016.05.005>. <http://www.sciencedirect.com/science/article/pii/S1270963816301754>.
- Park, C., Haftka, R.T., Kim, N.H., 2016. Remarks on multi-fidelity surrogates. *Struct. Multidiscip. Optim.* ISSN: 1615-1488 1–22. <https://doi.org/10.1007/s00158-016-1550-y>.
- Pedersen, P., 1973. Some properties of linear strain triangles and optimal finite element models. *Int. J. Numer. Meth. Eng.* ISSN: 1097-0207 7 (4), 415–429. <https://doi.org/10.1002/nme.1620070402>.
- Piegl, L., Tiller, W., 1997. The NURBS Book, 2Nd Ed. Springer-Verlag New York, Inc., New York, NY, USA.
- Platt, J.C., 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv. Large Margin Classif.* 10 (3), 61–74.
- Rasmussen, C.E., Williams, C.K.I., 2006. *Gaussian Processes for Machine Learning*. MIT Press.
- Ribeiro, A., Awruch, A., Gomes, H., 2012. An airfoil optimization technique for wind turbines. *Appl. Math. Model.* ISSN: 0307-904X 36 (10), 4898–4907. <https://doi.org/10.1016/j.apm.2011.12.026>. <http://www.sciencedirect.com/science/article/pii/S0307904X11008146>.
- Richards, P., Johnson, A., Stanton, A., 2001. America's Cup downwind sails-vertical wings or horizontal parachutes? *J. Wind Eng. Ind. Aerod.* ISSN: 0167-6105 89 (14–15), 1565–1577. [https://doi.org/10.1016/S0167-6105\(01\)00135-0](https://doi.org/10.1016/S0167-6105(01)00135-0). <http://www.sciencedirect.com/science/article/pii/S0167610501001350> (bluff Body Aerodynamics and Applications).
- Sacher, M., Hauville, F., Duvigneau, R., Maître, O.L., Aubin, N., Durand, M., 2017. Efficient optimization procedure in non-linear fluid-structure interaction problem: application to mainsail trimming in upwind conditions. *J. Fluids Struct.* ISSN: 0889-9746 69, 209–231. <https://doi.org/10.1016/j.jfluidstructs.2016.12.006>. <http://www.sciencedirect.com/science/article/pii/S0889974616303437>.
- Schonlau, M., 1997. *Computer Experiments and Global Optimization*. Ph.D. thesis. University of Waterloo, Waterloo, Ont., Canada. Canada, aAINQ22234.
- Sedlar, M., Ji, B., Kratky, T., Rebok, T., Huzlik, R., 2016. Numerical and experimental investigation of three-dimensional cavitating flow around the straight {NACA2412} hydrofoil. *Ocean. Eng.* ISSN: 0029-8018 123, 357–382. <https://doi.org/10.1016/j.oceaneng.2016.07.030>. <http://www.sciencedirect.com/science/article/pii/S0029801816302785>.
- Simpson, T., Poplinski, J., Koch, N.P., Allen, J., 2001. Metamodels for computer-based engineering design: survey and recommendations. *Eng. Comput.* ISSN: 1435-5663 17 (2), 129–150. <https://doi.org/10.1007/PL00007198>.
- Srinath, D., Mittal, S., 2010. Optimal aerodynamic design of airfoils in unsteady viscous flows. *Comput. Meth. Appl. Mech. Eng.* ISSN: 0045-7825 199 (29–32), 1976–1991. <https://doi.org/10.1016/j.cma.2010.02.016>. <http://www.sciencedirect.com/science/article/pii/S0045782510000745>.
- Suykens, J., Vandewalle, J., 1999. Least squares support vector machine classifiers. *Neural Process. Lett.* ISSN: 1573-773X 9 (3), 293–300. <https://doi.org/10.1023/A:1018628609742>.
- Van Calster, B., Luts, J., Suykens, J.A.K., Condous, G., Bourne, T., Timmerman, D., Van Huffel, S., 2007. In: *Artificial Neural Networks – ICANN 2007: 17th International Conference, Porto, Portugal, September 9-13, 2007, Proceedings, Part II*, Chap. Comparing Methods for Multi-class Probabilities in Medical Decision Making Using LS-SVMs and Kernel Logistic Regression. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN 978-3-540-74695-9, pp. 139–148. [https://doi.org/10.1007/978-3-540-74695-9\\_15](https://doi.org/10.1007/978-3-540-74695-9_15).
- Van Ingen, J., 2008. The eN method for transition prediction: historical review of work at TU Delft. In: *38th Fluid Dynamics Conference and Exhibit*. American Institute of Aeronautics and Astronautics, Seattle, United States, pp. 1–49.
- Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag New York, Inc., New York, NY, USA, ISBN 0-387-94559-8.
- Wang, W., Pottmann, H., Liu, Y., 2006. Fitting B-spline curves to point clouds by curvature-based squared distance minimization. *ACM Trans. Graph* 25 (2), 214–238.
- WEI, Q., xun CHEN, H., ZHANG, R., 2015. Numerical research on the performances of slot hydrofoil. *J. Hydrodyn. Ser. B* ISSN: 1001-6058 27 (1), 105–111. [https://doi.org/10.1016/S1001-6058\(15\)60462-0](https://doi.org/10.1016/S1001-6058(15)60462-0). <http://www.sciencedirect.com/science/article/pii/S1001605815604620>.

## 9.3 Conclusion

L'optimisation sous contrainte d'un hydrofoil flexible et à paramètres multiples est présentée dans la Section 9.2. Cette optimisation est réalisée à l'aide de la méthode EGO-LS-SVM introduite à la Section 5.3. La performance de l'hydrofoil est évaluée par des calculs IFS utilisant un code fluide potentiel 2D couplé à un code structure permettant la résolution stationnaire de l'équilibre fluide-structure de la partie flexible.

Les performances hydrodynamiques de l'hydrofoil flexible sont comparées à celles d'un second hydrofoil rigide optimisé pour les mêmes conditions. L'apport de la propriété flexible est ainsi illustré en terme d'efforts de traînée sur le range de vitesse rencontré par le voilier. En particulier pour des vitesses supérieures à 35 kts et pour des vitesses proches de 25 kts. En plus de cette réduction de traînée, il est montré que l'hydrofoil flexible nécessite des angles d'attaques inférieurs, en valeur absolue, que ceux de l'hydrofoil rigide. Cela implique directement la tolérance de l'hydrofoil vis-à-vis du régime cavitant aux faibles vitesses du voilier car le pic de pression minimale au niveau du bord d'attaque est alors fortement réduit.

Pour conclure, ces travaux de thèse permettent de valider la méthode EGO-LS-SVM sur un problème physique réel et complexe.





## Optimisation des réglages d'un gréement de voilier par méta-modèles multi-fidélité

### Sommaire

|                                                                                                       |            |
|-------------------------------------------------------------------------------------------------------|------------|
| <b>10.1 Motivations</b>                                                                               | <b>188</b> |
| <b>10.2 Publication soumise dans « <i>Computer Methods in Applied Mechanics and Engineering</i> »</b> | <b>191</b> |
| 10.2.1 Introduction                                                                                   | 191        |
| 10.2.2 Publication                                                                                    | 193        |
| 10.2.3 Conclusion                                                                                     | 215        |
| <b>10.3 Optimisation à six paramètres</b>                                                             | <b>215</b> |
| 10.3.1 Problème d'optimisation                                                                        | 215        |
| 10.3.2 Validation du maillage type « overset »                                                        | 217        |
| 10.3.3 Résultats                                                                                      | 219        |
| <b>10.4 Conclusion</b>                                                                                | <b>226</b> |

Le Chapitre 10 contient des applications d'optimisation par méta-modèles multi-fidélité. Il permet d'illustrer l'utilisation de la méthode EGO formulée en multi-fidélité et précédemment décrite dans la Section 6.2. L'accent est mis sur l'optimisation des réglages d'un gréement de voilier, qui constitue en soi l'un des objectifs de cette thèse. Le plan de ce Chapitre 10 se divise en quatre sections. La Section 10.1 introduit la problématique de l'optimisation des réglages d'un gréement de voilier. Une publication récemment soumise concerne ensuite la Section 10.2. Dans cette publication, la méthode EGO multi-fidélité est validée sur différents problèmes analytiques. Une application originale d'optimisation par méta-modèles multi-fidélité des réglages d'un gréement de voilier, simulé en IFS, fait également partie de cette publication. Deux réglages sont optimisés avec l'utilisation de cinq niveaux de fidélité construits sur les modèles fluide et structures de différentes précisions. Dans la Section 10.3, une autre optimisation des réglages du même gréement de voilier, simulé en IFS, est effectuée.

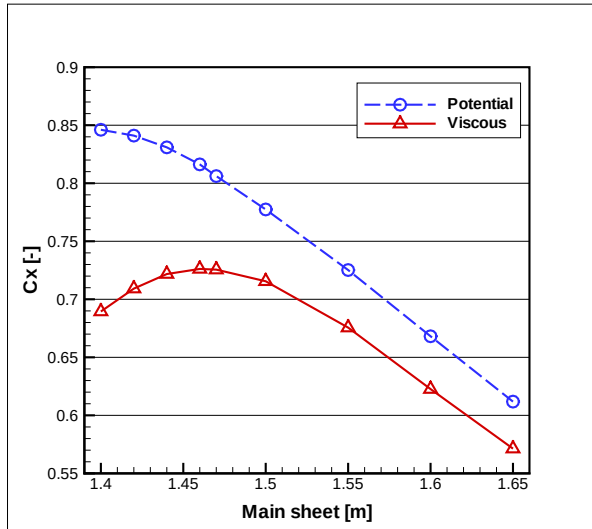


Dans ce dernier cas, le nombre de niveau de fidélité, construit uniquement sur des niveaux de précision du modèle fluide, est réduit à deux et la dimension du problème d'optimisation est portée à six. Finalement, les conclusions du Chapitre 10 sont données dans la dernière Section 10.4.

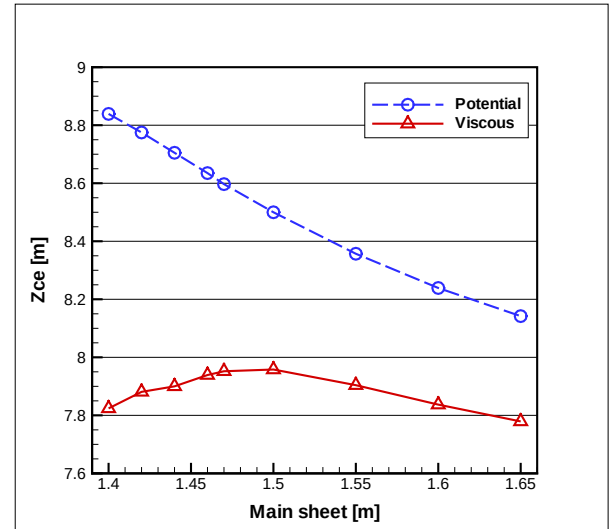
## 10.1 Motivations

L'optimisation de la performance d'un voilier nécessite de déterminer les réglages optimaux des voiles et du gréement. Ainsi, les différents designs des voiles ou de la carène (structure, géométrie, etc.) peuvent être comparés sur leur point de fonctionnement optimal. La modélisation numérique complète du voilier reste très coûteuse et complexe aujourd'hui. En effet, la modélisation aérodynamique des voiles requiert des calculs non linéaires d'interaction fluide-structure. De même, la résistance hydrodynamique de la carène doit être calculée pour différentes attitudes du voilier. Dans le cas où le design de la carène est figé, le coût des simulations hydrodynamiques peut être réduit en construisant une matrice des efforts hydrodynamiques au préalable. La performance du voilier peut alors être déterminée à partir des données hydrodynamiques interpolées, ou encore plus simplement avec une fonction analytique de la performance définie autour du point de fonctionnement du voilier (Sacher et al., 2015).

La résolution du problème d'optimisation des réglages d'un gréement de voilier est également ou plutôt conséquemment compliquée. En effet, en plus de la complexité intrinsèque du problème d'optimisation, sa dimension peut être grande ( $d \approx 30$ ) dans le cas où l'ensemble des réglages du gréement sont considérés. L'utilisation de méta-modèles Gaussiens est alors justifiée afin de réduire le coût de cette optimisation (Sacher et al., 2017c). La modélisation numérique des voiles peut faire intervenir plusieurs niveaux de fidélité construits sur différents niveau de précision des modèles fluide et/ou structure. Par exemple, un même jeu de réglages des voiles peut être simulé à l'aide de deux modèles fluides, parfait et visqueux. En formulation multi-fidélité, dans le processus d'optimisation itératif par méta-modèles des réglages des voiles, les propriétés visqueuses du fluide pourront alors être considérées ou non, en fonction du niveau de fidélité, avec une légère augmentation du coût de calcul. C'est ce qui est illustré sur la Figure 10.1, où des calculs IFS fluide parfait et visqueux ont été réalisés pour différentes longueurs d'écoute de grand-voile ( $L_{\text{sheet}}$ ) d'un Figaro, dans le cas d'une allure au près où seul l'aspect aérodynamique est considéré, l'assiette et la vitesse du bateau étant fixées. Le coefficient aérodynamique propulsif total  $C_x$  est reporté en fonction de  $L_{\text{sheet}}$  sur la Figure 10.1(a) et la position verticale du point d'application de la résultante aérodynamique  $Z_{\text{ce}}$  en fonction de  $L_{\text{sheet}}$  est donnée sur la Figure 10.1(b). On peut ainsi apprécier l'augmentation de  $C_x$ , ici vue comme la performance du voilier, lorsque l'écoute est bordée de 1.65 m à approximativement 1.50 m. On note que les pentes de cette augmentation données par les deux modèles fluides sont très proches et que le modèle fluide parfait estime des valeurs plus élevées de  $C_x$ . Ensuite, lorsque l'écoute de grand-voile continue d'être bordée, les variations de  $C_x$  diffèrent entre les deux modèles. Un extremum de  $C_x$  pour  $\approx 1.46$  m est prédit par le modèle fluide visqueux alors que  $C_x$  continu d'augmenter dans les simulations IFS fluide parfait. On observe les mêmes comportements sur la Figure 10.1(b), avec un léger décalage de l'extremum qui est à  $L_{\text{sheet}} \approx 1.50$  m. Cette chute de la position verticale de la résultante aérodynamique, donnée par le modèle fluide visqueux, provient du décollement fluide situé dans la zone



(a) Coefficient aérodynamique propulsif.



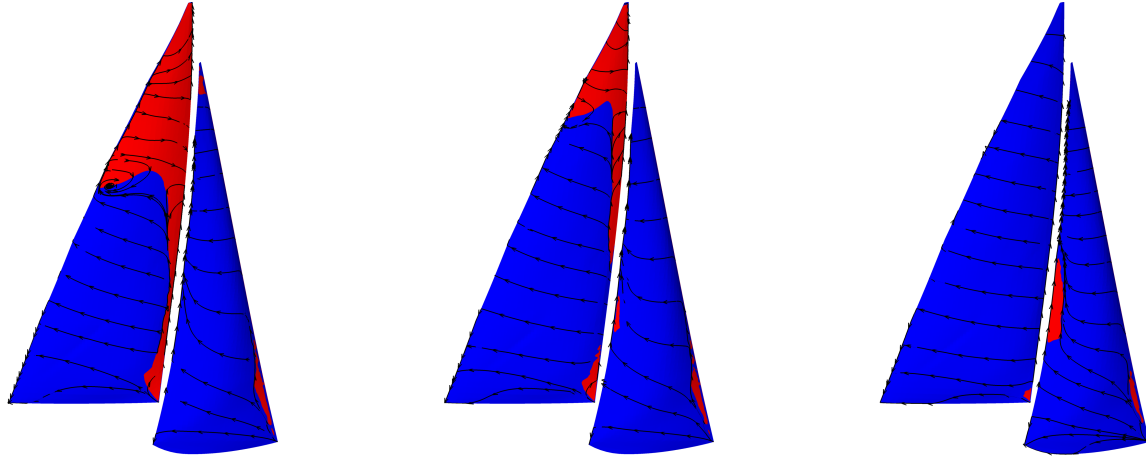
(b) Position verticale du point d'application de la résultante aérodynamique.

**Figure 10.1** – Calculs IFS fluide parfait et visqueux pour différentes longueurs d'écoute de grand-voile. Résultats provenant de (Sacher et al., 2015).

haute de la grand-voile. Ces variations de  $C_x$  et  $Z_{ce}$  s'expliquent assez facilement. En effet, lorsque l'écoute est trop choquée, les incidences locales des profils en tête de grand-voile sont faibles, l'écoulement est donc attaché, et la résultante aérodynamique qui en plus d'être faible provient des zones basses des voiles. Le fait de border la grand-voile implique une augmentation des incidences locales des profils, donc une augmentation des efforts aérodynamiques en tête de grand-voile. Par conséquent,  $C_x$  et  $Z_{ce}$  augmentent, même si un début de décollement apparaît dans la partie supérieure de la grand-voile. L'extremum de  $C_x$  intervient alors lorsque le compromis décollement/propulsion est atteint. Ce phénomène n'est pas modélisé par le fluide parfait car le fluide reste attaché aux voiles, sauf sur les bords où le décollement est imposé pour tous les réglages de la grand-voile. Donc, contrairement au fluide parfait, si l'écoute continue d'être bordée, la zone du décollement prédite par le fluide visqueux s'agrandie,  $Z_{ce}$  diminue et la performance  $C_x$  également.

La Figure 10.2 présente ces variations de la zone du décollement fluide (surfaces rouges sur les voiles) pour trois réglages de la grand-voile. On peut ainsi faire visuellement le lien entre la taille de la zone de décollement fluide, en tête de grand-voile, et la longueur de l'écoute. Sur cet exemple, on sent bien qu'une optimisation en simple fidélité avec un unique modèle fluide parfait ne convergera pas vers un réglage optimal réaliste. Le modèle fluide visqueux est nécessaire et justifie l'utilisation d'une approche d'optimisation multi-fidélité.

Une optimisation simple fidélité à deux paramètres fait également partie de ces résultats préliminaires sur le Figaro. Dans cette application, les deux écoutes (grand-voile et génois) sont optimisées afin de maximiser une fonction performance calculée avec le modèle IFS fluide parfait. Les conditions de navigation sont dans un vent plus soutenu où  $BS = 6.45$  kts,  $TWS = 20$  kts et  $TWA = 45^\circ$ , et la gîte du voilier est fixée à  $25^\circ$ . La fonction performance utilisée ici est définie par (10.1.1), comme le produit de l'effort propulsif  $F_x$  et d'une



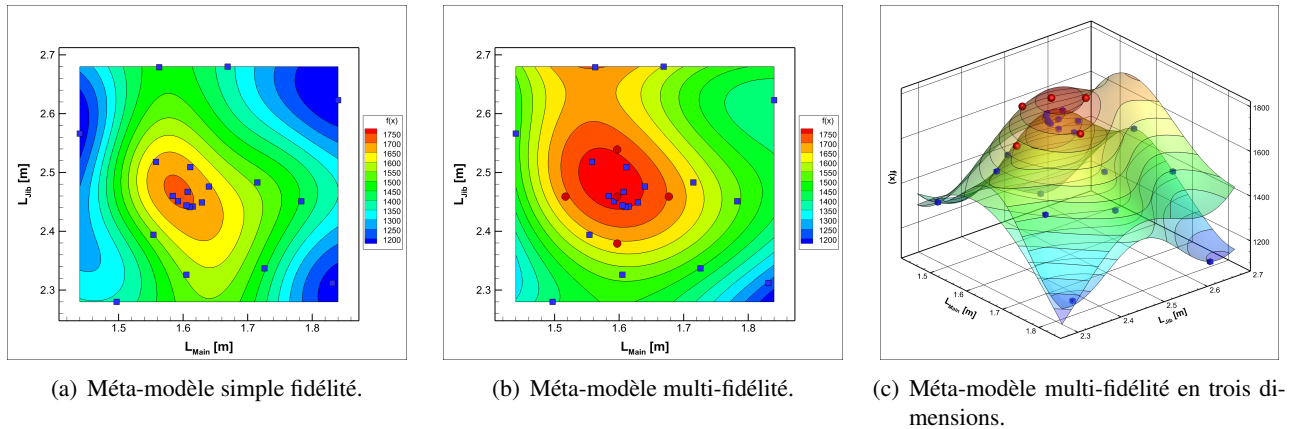
(a) Réglage trop bordé,  $L_{\text{sheet}} = 1.4$  m. (b) Réglage optimal,  $L_{\text{sheet}} = 1.46$  m. (c) Réglage trop choqué,  $L_{\text{sheet}} = 1.6$  m.

**Figure 10.2** – Illustration du décollement fluide pour différents réglages de grand-voile. Résultats provenant de (Sacher et al., 2015).

pénalisation provenant de la gîte du voilier par,

$$\max_{\mathbf{x} \in \Omega} f(\mathbf{x}) = F_x \cos\left(\frac{M_x}{M_{x_{\max}}}\right), \quad (10.1.1)$$

où  $M_x$  est le moment de chavirement qui pénalise  $F_x$  et  $M_{x_{\max}} = 28000 \text{ N} \cdot \text{m}$  est le moment de chavirement maximal du Figaro pris à cette allure. La Figure 10.3(a) présente la reconstruction du méta-modèle simple fidélité de la performance, après avoir atteint la convergence du processus d'optimisation, où les carrés bleus correspondent aux évaluations de  $f(\mathbf{x})$ . Cinq simulations IFS fluide visqueux sont aussi réalisées dans cette



**Figure 10.3** – Résultats de l'optimisation simple fidélité du problème (10.1.1). La construction d'un méta-modèle multi-fidélité utilisant cinq observations d'un niveau supérieur est également présentée.

étude. Ces calculs sont placés autour du réglage optimal obtenu par l'optimisation simple fidélité fluide parfait, comme montré par les cinq cercles rouges sur la Figure 10.3(b). Cette dernière Figure 10.3(b) illustre la

prédiction de la fonction performance donnée par un méta-modèle multi-fidélité, construit sur l'ensemble des observations simple (carrés bleus) et multi-fidélité (cercles rouges). Une illustration en trois dimensions de cette prédiction est également visible sur la Figure 10.3(c). Les deux prédictions simple et multi-fidélité sont très en accord sur la localisation du réglage optimal des voiles et la performance calculée par le modèle fluide parfait sous-estime légèrement celle du modèle fluide visqueux au niveau de l'optimum. Les variations de la fonction performance données par le modèle fluide parfait sont également présentes dans la prédiction multi-fidélité. Cela est particulièrement visible dans les zones lointaines de l'optimum, qui ne sont pourtant pas explorées par le modèle fluide visqueux. Cette dernière application motive l'utilisation d'une approche d'optimisation multi-fidélité pour résoudre l'optimisation de la performance d'un voilier.

## 10.2 Publication soumise dans « *Computer Methods in Applied Mechanics and Engineering* »

### 10.2.1 Introduction

La Section 10.1 précédente introduit les raisons motivant l'utilisation d'une approche d'optimisation par méta-modèles multi-fidélité pour optimiser les réglages d'un gréement de voilier. Cette Section 10.2 contient une publication récemment soumise dans « *Computer Methods in Applied Mechanics and Engineering* » où la méthode proposée dans cette thèse (voir Section 6.2), pour réaliser une optimisation par méta-modèles multi-fidélité est validée. La méthode d'optimisation multi-fidélité, reposant sur la base d'observations non imbriquées, est en particulier éprouvée sur un ensemble de tests analytiques, ainsi que sur un cas concret d'optimisation numérique des réglages d'un gréement de voilier, comportant la résolution de problèmes d'interaction fluide-structure.

La fonction « Hartmann » de dimension six (Picheny et al., 2013) est utilisée pour définir une suite de fonctions contractives, ce qui permet d'inclure la propriété de fidélité dans les différents tests analytiques. Trois niveaux de fidélité sont utilisés, associés à des coûts d'évaluation  $\beta_l$ , afin d'étudier les effets que peuvent avoir les différences de réponses des niveaux, la présence d'un bruit dans la réponse d'un niveau intermédiaire et une évolution non uniforme des coûts  $\beta_l$ , sur l'efficacité de la méthode d'optimisation. La robustesse et l'efficacité de la formulation « non imbriquée » sont comparées à la formulation « imbriquée » ainsi qu'à l'approche simple fidélité où le modèle de plus haute fidélité est uniquement utilisé.

Le problème d'optimisation de la performance du voilier concerne l'optimisation des longueurs des écoutes de la grand-voile et du génois du Figaro, dans un cas de navigation au près. La faible dimension du problème ( $d = 2$ ) facilite ici l'illustration et l'analyse des résultats. Cinq niveaux de fidélité sont utilisés, dans un but plutôt démonstratif que réaliste, où les différences des modèles concernent à la fois les modèles fluides et structures. Le problème d'optimisation est résolu avec la méthode d'optimisation par méta-modèles multi-fidélité, construits sur la base d'observations non imbriquées, ainsi que par une optimisation par méta-modèles simple fidélité avec le niveaux de plus haute fidélité. L'efficacité de l'approche multi-fidélité en terme de coût d'évaluation est illustrée par cette application. Des propriétés intéressantes de la méthode, comme le fait de

pouvoir naturellement peu évaluer la fonction performance donnée par un niveau de fidélité de « faible utilité », sont également présentées dans cette publication.

## 10.2.2 Publication

# A Non-Nested Infilling Strategy for Multi-Fidelity based Efficient Global Optimization

Matthieu Sacher<sup>a,\*</sup>, Olivier Le Maître<sup>b</sup>, Régis Duvigneau<sup>c</sup>, Frédéric Hauville<sup>a</sup>, Mathieu Durand<sup>d</sup>, Corentin Lothodé<sup>e</sup>

<sup>a</sup>Naval Academy Research Institute - IRENAV CC600, 29240 BREST Cedex 9, France

<sup>b</sup>LIMSI - CNRS, Rue John Von Neumann, 91400 Orsay, France

<sup>c</sup>Université Côte d'Azur, Inria, CNRS, LJAD, 2004 Route des Lucioles, 06902 Valbonne, France

<sup>d</sup>Sirli Innovations, 600 Chemin des Aspres, 06410 Biot, France

<sup>e</sup>K-Epsilon, 1300 Route des Crêtes, 06560 Valbonne, France

## Abstract

Meta-model based optimization methods have become very popular in recent years for the global optimization of complex systems with high computational costs. These methods rely on the sequential construction of a statistical surrogate model, using a dataset of computed performance values, which is refined according to a given infilling strategy (Efficient Global Optimization). Multi-fidelity models have been recently proposed to build surrogates that are able to combine observations computed with different fidelity levels. However, the use of multi-fidelity surrogates in the Efficient Global Optimization framework is still a research topic because of the complexity to deal with the mixed fidelities. In this work, we propose a new infilling strategy to perform multi-fidelity surrogate-based optimization with non-nested datasets. The proposed method is able to refine the observation datasets according to a single fidelity level. Consequently, the computational cost is significantly reduced with respect to nested methods and the approach is able to avoid possible useless fidelity levels for the global optimization. The method is applied to several analytical problems in order to demonstrate its efficiency and finally to a complex application of a fully nonlinear fluid-structure interaction system. The latter problem concerns the sail trimming optimization of a Figaro racing yacht, in which the objective function can be computed by five fidelity levels, including differences in term of both structure and fluid models.

**Keywords:** Global optimization, Gaussian process model, multi-fidelity, non-nested datasets

## 1. Introduction

Efficient Global Optimization (EGO) strategies [Jones et al. 1998] have been used in various application domains, such as aerodynamic drag reduction [Jeong et al. 2005], vibration reduction for rotating aircrafts [Glaz et al. 2009], optimization of FSI problems [Aghajari and Schäfer 2015] and sail trimming optimization [Sacher et al. 2017]. The EGO efficiency has been demonstrated for the optimization of complex systems with costly objective function evaluation [Simpson et al. 2001]. Gaussian Processes [Kleijnen 2009] (GP) are usually used within the EGO method to build surrogate models that approximate the costly objective function  $f$  to be optimized. A statistical criterion accounting for the Expected Improvement (EI), usually referred to as *merit function*, is maximized to select a new design point, at which the objective function should be evaluated. This new evaluation is used to improve the accuracy of the surrogate and proceed with the selection of a new design point.

A physical problem can often be approximated by numerical models with different accuracy levels and associated computation costs. As example, this may correspond to a finite element model with a set of different mesh refinements. In the perspective of exploiting this property, multi-fidelity (MF) models [Kennedy and O'Hagan 2000] have been developed in order to build surrogates that are able to use observations obtained from different fidelity models.

\*Corresponding author

Email address: matthieu.sacher@ecole-navale.fr (Matthieu Sacher)

However, the use of MF surrogates in EGO framework is not very widespread. The EGO formulation in the MF framework has been introduced in [Forrester et al. 2007], but the approach accounts for two-fidelity levels only (also called *co-kriging*) and the merit function dependency to the fidelity level is not treated, since the authors propose to use the highest fidelity surrogate prediction to compute the merit function. In the work of Elsayed [2015], a CFD problem which includes seven design variables with two levels of fidelity is investigated. Five infilling strategies are proposed in this paper to select cheap or expensive observations and reduce the uncertainty of MF surrogates. This is then used to perform a genetic-based optimization. However, this is not a formal EGO application since the approach does not account for the fidelity level in the optimization part and it is well known that an infilling strategy based only on model variances is very poor in comparison to *EI*-based criteria [Picheny et al. 2013]. Le Gratiet and Cannamela [2012] have proposed a nested infilling strategy to deal with MF surrogate-based optimizations. The nested formulation means that the observation datasets are nested: at a fidelity level  $l$ , the observation dataset is a subset of the observations of the  $l - 1$  level. In this work, the merit function is based on the maximum of variance in the highest fidelity model. Then, an approximation of integrated variance predictor [Plutowski et al. 1994] is used to compute the potential variance reduction in the highest surrogate due to the computation of the  $l$  first levels (datasets are nested). This is moreover counterbalanced by a computational time ratio between levels  $l$  and  $l - 1$ . To sum up, the method has the advantage of including a fidelity choice in the infilling strategy, but has several drawbacks: need of nested datasets, favor low simulation levels and use of a poor merit function.

In this work, we propose to extend the original and powerful *EI*-based merit function to the case of non-nested and multi-fidelity sources of observations. This is conducted by formulating a new merit function that selects a new design point and includes the fidelity level choice at which the objective function should be evaluated. This infilling method takes into account the computational cost and the potential variance reduction in the highest surrogate region of interest (the one which has to be optimized) due the computation of a level  $l$ . Several advantages arise from this new approach. The main one concerns the reduction of the computational cost regarding to single-fidelity EGO or nested MF methods. The non-nested property gives also rise to positive aspects. Indeed, the MF method naturally selects the best fidelity levels to employ in terms of cost/variance-reduction and can avoid possible useless fidelity levels.

The paper is organized as follows. The optimization methods are presented in Section 2, where the classical single-fidelity EGO is detailed in Section 2.1. The construction of MF surrogates and the current non-nested proposed infilling strategy are then provided Section 2.2. Analytical tests problems are considered in Section 3, that allow efficiency comparisons of the non-nested MF EGO to the nested MF EGO and single-fidelity EGO approaches. The optimization of a fully nonlinear Fluid-Structure Interaction (FSI) problem is subsequently considered in Section 4. It consists in the maximization of an aerodynamic performance function for a Figaro racing yacht. This is performed by optimizing the sail trimming at a fixed point of sail. Two-trimming variables are considered and the performance function is modeled by five fidelity levels that include differences in term of both structure and fluid models. Finally, the conclusion of this work is proposed in Section 5.

## 2. GP-based optimization

The objective is to estimate the solution  $\mathbf{x}_{\text{opt}}$  of the following generic optimization problem

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}), \quad (1)$$

where  $\mathbf{x}$  is the vector of design parameters,  $\Omega \subset \mathbb{R}^d$  is the admissible optimization domain and  $f : \Omega \mapsto \mathbb{R}$  the cost function. The global optimum of  $f$  can be very costly to find, in particular when its evaluation is numerically expensive. A classical approach is then to use surrogate models in place of  $f$  to reduce the computational burden related to the optimization of complex systems [Simpson et al. 2001]. In this work, we consider the use of Gaussian processes (GP) [Kleijnen 2009] which, owing to their statistical nature, provide both a prediction of the cost function and a measure of the uncertainty (variance) in the prediction. These features are appealing in optimization, as they can be used to derive rigorous optimization strategies based on the maximization of the Expected Improvement (EI) criterion. These methods are globally referred to as the Efficient Global Optimization (EGO) [Jones et al. 1998] methods. A summary of the single-fidelity construction of the GP model for  $f$  and the resulting EGO strategy is provided in Section 2.1. The multi-fidelity case is presented in Section 2.2 and the current non-nested proposed EGO approach is detailed.

### 2.1. Single-fidelity EGO

Consider a set of  $n$  training inputs points  $\mathbf{X}_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , each associated to a noisy observation  $y_i$  of the cost function  $f$ . It is assumed that  $y_i = f(\mathbf{x}_i) + \varepsilon_i$ , where the  $\varepsilon_i$  are Gaussian measurement noises, assumed for simplicity independent and identically distributed with variance  $\sigma_\varepsilon^2$ . The GP construction considers that  $f(\mathbf{x})$  is a realization of a zero-mean multivariate Gaussian process with covariance function  $C_f$ . In this work, we consider the multidimensional squared exponential covariance functions defined by

$$C_f(\mathbf{x}, \mathbf{x}'; \Theta) \doteq \theta_1 \prod_{i=1}^d \exp\left(\frac{-(x_i - x'_i)^2}{2l_i^2}\right) + \theta_2, \quad (2)$$

where  $\Theta = \{\theta_1, \theta_2, l_1, l_2, \dots, l_d\}$  is a vector of hyper-parameters to be inferred from the observations. Denoting  $\mathbf{C}(\Theta) \in \mathbb{R}^{n \times n}$  the covariance matrix with entries  $C_{i,j}(\Theta) \doteq C_f(\mathbf{x}_i, \mathbf{x}_j; \Theta)$ ,  $1 \leq i, j \leq n$ , the joint Gaussian distribution of the noisy observation vector,  $\mathbf{Y}_n = (y_1, \dots, y_n)^T$ , and the predicted observation  $y(\mathbf{x})$  is given by

$$\begin{pmatrix} \mathbf{Y}_n \\ y(\mathbf{x}) \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{C} + \sigma_\varepsilon^2 \mathbf{I} & \mathbf{k}(\mathbf{x}) \\ \mathbf{k}^T(\mathbf{x}) & \kappa(\mathbf{x}) + \sigma_\varepsilon^2 \end{bmatrix}\right). \quad (3)$$

In (3) we have denoted  $\kappa(\mathbf{x}) \doteq C_f(\mathbf{x}, \mathbf{x}; \Theta)$ ,  $\mathbf{k}(\mathbf{x}) \doteq (C_f(\mathbf{x}, \mathbf{x}_1; \Theta) \dots C_f(\mathbf{x}, \mathbf{x}_n; \Theta))^T$  and  $\mathbf{I}$  the identity matrix of  $\mathbb{R}^n$ . From the conditional rules of joint Gaussian distributions [Rasmussen and Williams 2006], the *best* prediction  $\hat{f}(\mathbf{x})$  of  $f(\mathbf{x})$ , *i.e.* the mean of  $y$ , and the prediction variance  $\hat{\sigma}_f^2(\mathbf{x})$  are given by

$$\hat{f}(\mathbf{x}) = \mathbf{k}^T(\mathbf{x}) (\mathbf{C}(\Theta) + \sigma_\varepsilon^2 \mathbf{I})^{-1} \mathbf{Y}_n, \quad (4)$$

$$\hat{\sigma}_f^2(\mathbf{x}) = \kappa(\mathbf{x}) + \sigma_\varepsilon^2 - \mathbf{k}^T(\mathbf{x}) (\mathbf{C}(\Theta) + \sigma_\varepsilon^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}). \quad (5)$$

The hyper-parameters  $\Theta$  and noise variance  $\sigma_\varepsilon^2$  can be determined by maximizing the log-marginal likelihood [Rasmussen and Williams 2006] using an evolution strategy algorithm [Hansen 2006]. More details on GP meta-models can be found in Rasmussen and Williams [2006].

Let  $\hat{\mathbf{x}}_n$  be the optimum of  $\hat{f}(\mathbf{x})$ . It is expected that  $\hat{\mathbf{x}}_n \approx \mathbf{x}_{\text{opt}}$  if the approximation error  $\hat{f} - f$  is small enough. The advantage of minimizing  $\hat{f}$  instead of  $f$  is that GP models are usually inexpensive to evaluate compared to the original cost function. To control the error in the approximation, one proceeds sequentially by adding progressively new observation points in the area of interest. A deterministic optimization procedure would choose the next point  $\mathbf{x}_{n+1}$  as the optimal point of  $\hat{f}$ . However, the GP model provides probabilistic information that can be exploited to propose more robust strategies based on merit functions, which combine the prediction and its variance. In this work, we use the Augmented Expected Improvement (*AEI*) merit function [Huang et al. 2006], which estimates the expected progress in the cost, taking into account the noise in the observed values and the prediction variance:

$$AEI(\mathbf{x}) = EI(\mathbf{x}) \left( 1 - \frac{\sigma_\varepsilon}{\sqrt{\hat{\sigma}_f^2(\mathbf{x}) + \sigma_\varepsilon^2}} \right), \quad (6)$$

where the Expected Improvement  $EI(\mathbf{x})$  is defined by

$$EI(\mathbf{x}) = \hat{\sigma}_f(\mathbf{x}) [\Phi(u(\mathbf{x})) + \phi(u(\mathbf{x}))], \quad (7)$$

$$u(\mathbf{x}) = \frac{\hat{f}(\mathbf{x}_{*,n}) - \hat{f}(\mathbf{x})}{\hat{\sigma}_f(\mathbf{x})}, \quad (8)$$

with  $\Phi$  and  $\phi$  the cumulative and density functions of the standard Gaussian distribution, and  $\mathbf{x}_{*,n} \in \mathbf{X}_n$  is the current effective best solution (see Huang et al. [2006]). The optimum  $\mathbf{x}_{n+1}$  of the *AEI* is added to  $\mathbf{X}_n$ , and  $f$  is evaluated at the new point providing  $y_{n+1}$ . Setting  $n \leftarrow n+1$ , a new iteration can start updating the GP model with the new observation. Overall, each iteration requires one computation of the cost and the resolution of two optimization problems: a first one for the hyper-parameters of the GP model, and a second one to find the *AEI* optimum. The iterations of the GP-based optimization problem are continued until a stopping criterion is satisfied or the resources allocated to the optimization have been exhausted.



## 2.2. Multi-fidelity EGO

In this work, we use the recursive formulation of Le Gratiet and Garnier [2014], which reduces the complexity of building a  $s$ -level surrogate by using  $s$  independent models. Indeed, the original fidelity approach from Kennedy and O'Hagan [2000], is usually computationally expensive since the number of observations becomes large, as it requires covariance matrix inversions with dimension equals to the sum of the observations at all fidelity levels. In addition to provide identical multi-fidelity predictive mean and variance, the recursive approach [Le Gratiet and Garnier 2014] has the main advantage of giving the predictive mean and variance at each fidelity level.

We consider a computer output  $y_s(\mathbf{x})$  that we want to model with its lower approximations  $y_l(\mathbf{x})_{l=1,\dots,s-1}$ . The outputs are sorted by order of fidelity from the less accurate  $y_1(\mathbf{x})$  to the most accurate  $y_s(\mathbf{x})$ , and they are modeled by Gaussian Processes  $Z^l(\mathbf{x})_{l=1,\dots,s}$ ,  $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ . The considered recursive formulation of the MF model is  $\forall l = 2, \dots, s$ :

$$\begin{cases} Z^l(\mathbf{x}) = \rho_{l-1} Z^{l-1}(\mathbf{x}) + \delta^l(\mathbf{x}), \\ Z^{l-1}(\mathbf{x}) \perp \delta^l(\mathbf{x}), \end{cases} \quad (9)$$

where  $\rho_{l-1}$  is a kind of regression parameter,  $\perp$  means that  $Z^{l-1}(\mathbf{x})$  is independent of  $\delta^l(\mathbf{x})$  and  $Z^{l-1}(\mathbf{x})$  is a Gaussian process with distribution  $[Z^{l-1}(\mathbf{x})|\mathbf{Y}_{n_{l-1}}^{l-1}]$ .  $Z^1(\mathbf{x})$  and  $\delta^1(\mathbf{x})$  have the same distribution and  $\mathbf{Y}_{n_l}^l = y_l(\mathbf{X}_{n_l}^l)$ , where  $\mathbf{X}_{n_l}^l$  is the experimental design set at level  $l$  that is composed of  $n_l$  observations. The recursive assumption in (9) comes from the Markov property [Kennedy and O'Hagan 2000], where  $\text{cov}(Z^l(\mathbf{x}), Z^{l-1}(\mathbf{x}')|Z^{l-1}(\mathbf{x})) = 0, \forall \mathbf{x} \neq \mathbf{x}'$ , means that we can learn no more about  $Z^l(\mathbf{x})$  from any other  $Z^{l-1}(\mathbf{x}')$  for  $\mathbf{x} \neq \mathbf{x}'$ . The multi-fidelity surrogate at level  $l = 2, \dots, s$  is then given by:

$$[Z^l(\mathbf{x})|\mathbf{Y}_{n_l}^l] \sim \mathcal{N}(\widehat{f}_l(\mathbf{x}), \widehat{\sigma}_{f_l}^2(\mathbf{x})), \quad (10)$$

where

$$\widehat{f}_l(\mathbf{x}) = \rho_{l-1} \widehat{f}_{l-1}(\mathbf{x}) + \mathbf{k}_l^T(\mathbf{x}) (\mathbf{C}_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1} (\mathbf{Y}_{n_l}^l - \rho_{l-1} y_{l-1}(\mathbf{X}_{n_l}^l)), \quad (11)$$

$$\widehat{\sigma}_{f_l}^2(\mathbf{x}) = \rho_{l-1}^2 \widehat{\sigma}_{f_{l-1}}^2(\mathbf{x}) + \kappa_l(\mathbf{x}) + \sigma_{\epsilon_l}^2 - \mathbf{k}_l^T(\mathbf{x}) (\mathbf{C}_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1} \mathbf{k}_l(\mathbf{x}). \quad (12)$$

The mean (11) and variance (12) responses are the model predictions at level  $l$  expressed in function of the  $l-1$  first level predictions; we have thus a recursive multi-fidelity surrogate as introduced in (9). In (11), the formulation assumes that points in datasets are nested:  $\mathbf{X}_{n_s}^s \subseteq \mathbf{X}_{n_{s-1}}^{s-1} \subseteq \dots \subseteq \mathbf{X}_{n_1}^1$ . In the case of non-nested databases, we use the prediction  $\widehat{f}_{l-1}(\mathbf{X}_{n_l}^l)$  instead of  $y_{l-1}(\mathbf{X}_{n_l}^l)$ .  $\mathbf{C}_l(\Theta_l) \in \mathbb{R}^{n_l \times n_l}$  is the covariance matrix with entries  $C_{l_{ij}}(\Theta_l) \doteq C_{F_l}(\mathbf{x}_i, \mathbf{x}_j; \Theta_l)_{\mathbf{x}_i, \mathbf{x}_j \in \mathbf{X}_{n_l}^l}$ , we denote  $\kappa_l(\mathbf{x}) \doteq C_{F_l}(\mathbf{x}, \mathbf{x}; \Theta_l)$ ,  $\mathbf{k}_l(\mathbf{x}) \doteq (C_{F_l}(\mathbf{x}, \mathbf{x}_1; \Theta_l)_{\mathbf{x}_1 \in \mathbf{X}_{n_l}^l} \dots C_{F_l}(\mathbf{x}, \mathbf{x}_{n_l}; \Theta_l)_{\mathbf{x}_{n_l} \in \mathbf{X}_{n_l}^l})^T$  and  $\mathbf{I}_l$  is the identity matrix of  $\mathbb{R}^{n_l}$ . The covariance function  $\text{cov}(Z^l(\mathbf{x}), Z^l(\mathbf{x}'))$  derived from the recursive formulation [Kennedy and O'Hagan 2000] is computed with

$$C_{F_l}(\mathbf{x}, \mathbf{x}'; \Theta_l) = \sum_{j=1}^l \left( \prod_{i=j}^{l-1} \rho_i^2 \right) C_f(\mathbf{x}, \mathbf{x}'; \Theta_l). \quad (13)$$

The hyper-parameters  $\Theta_l$ , noise variances  $\sigma_{\epsilon_l}^2$  and regression parameters  $\rho_{l-1}$  can be determined independently at each fidelity level thanks to the Markov property [Kennedy and O'Hagan 2000]. This is performed by maximizing the log-marginal likelihood [Rasmussen and Williams 2006] in (14) at each level by using an evolution strategy algorithm [Hansen 2006]:

$$\begin{aligned} \mathcal{L}_l(\rho_{l-1}, \Theta_l, \sigma_{\epsilon_l}^2) &= -\frac{n_l}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{C}_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l| - \frac{1}{2} \mathbf{U}_l^T (\mathbf{C}_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1} \mathbf{U}_l, \\ \mathbf{U}_l &= \begin{cases} \mathbf{U}_1 = \mathbf{Y}_{n_1}^1 \\ \mathbf{U}_{l>1} = \mathbf{Y}_{n_l}^l - \rho_{l-1} y_{l-1}(\mathbf{X}_{n_l}^l) \end{cases}, \quad l = 1, \dots, s. \end{aligned} \quad (14)$$

In (14), the prediction  $\widehat{f}_{l-1}(\mathbf{X}_{n_l}^l)$  is used in case of non-nested observations, instead of  $y_{l-1}(\mathbf{X}_{n_l}^l)$ .

A one-variable example (from [Forrester et al. 2007]) of MF surrogate is given in Figure 1. The red curve corresponds to an expensive function ( $f_e$  in (15)) that has been computed four times (red squares). The blue curve is a cheaper estimate ( $f_c$  in (16)) with eleven available observations (blue circles). The green dot-dashed SF curve is a single fidelity model of the four expensive points and the black dashed MF curve is a multi-fidelity model which uses both expensive and cheap observations.

$$f_e(x) = (6x - 2)^2 \sin(12x - 4), \quad (15)$$

$$f_c(x) = \frac{1}{2}f_e(x) + 10\left(x - \frac{1}{2}\right) + 5. \quad (16)$$

The SF model gives a poor prediction of the expensive function while the MF model is very close in both local and global minimum.

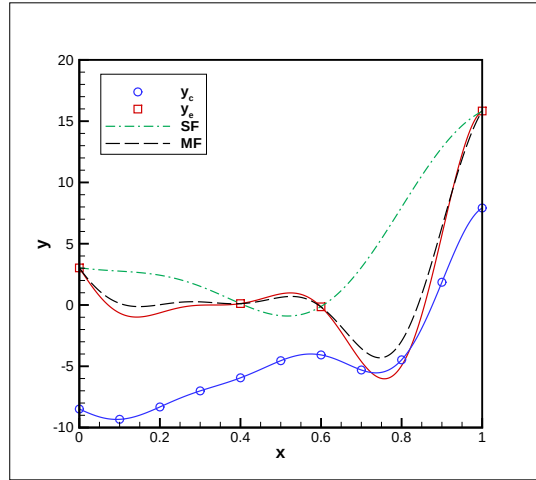


Figure 1: One-variable MF example [Forrester et al. 2007]. The multi-fidelity GP model uses both expensive ( $y_e$ ) and cheap ( $y_c$ ) observations.

We now present our proposition of infilling method for EGO that uses multi-fidelity surrogates. The objective is to use the powerful Augmented Expected Improvement (AEI) [Huang et al. 2006] merit function, and to include the fidelity level choice property for non-nested dataset cases. We firstly define a computational cost ratio  $\frac{\beta_s}{\beta_l}$ , where  $\beta_s$  is the computing cost of the highest level. This cost ratio is used to favor lower fidelity level computations. We then need to counterbalance this computational weight for model evaluations  $l$  that reduce the variance in the multi-fidelity model  $s$ . For this, we compute the variance prediction of the model  $s$  in  $\mathbf{x}$  by considering that  $\mathbf{x}$  has been computed through the model  $q$ :

$$\begin{aligned} \widehat{\sigma}_{f_s}^2|_q(\mathbf{x}) &= \rho_{s-1}^2 \widehat{\sigma}_{f_{s-1}}^2(\mathbf{x}) + \kappa_s(\mathbf{x}) + \sigma_{\epsilon_s}^2 - \mathbf{k}_s^T(\mathbf{x}) \left( \mathbf{C}_s(\Theta_s) + \sigma_{\epsilon_s}^2 \mathbf{I}_s \right)^{-1} \mathbf{k}_s(\mathbf{x}), \\ \widehat{\sigma}_{f_l}^2(\mathbf{x}) &= \begin{cases} \rho_{l-1}^2 \widehat{\sigma}_{f_{l-1}}^2(\mathbf{x}) + \kappa_l(\mathbf{x}) + \sigma_{\epsilon_l}^2 - \mathbf{k}_l^T(\mathbf{x}) \left( \mathbf{C}_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l \right)^{-1} \mathbf{k}_l(\mathbf{x}), & \forall l \neq q, \\ \rho_{q-1}^2 \widehat{\sigma}_{f_{q-1}}^2(\mathbf{x}) + \kappa_q(\mathbf{x}) + \sigma_{\epsilon_q}^2 - \begin{bmatrix} \mathbf{C}_q(\Theta_q) + \sigma_{\epsilon_q}^2 \mathbf{I}_q & \mathbf{k}_q(\mathbf{x}) \\ \mathbf{k}_q^T(\mathbf{x}) & C_{F_q}(\mathbf{x}, \mathbf{x}; \Theta_q) + \sigma_{\epsilon_q}^2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{k}_q(\mathbf{x}) \\ \kappa_q(\mathbf{x}) \end{bmatrix}, & l = q. \end{cases} \end{aligned} \quad (17)$$

More precisely, the variance prediction  $\widehat{\sigma}_{f_s}^2|_q(\mathbf{x})$  in the surrogate  $s$  is commonly computed  $\forall l = 1, \dots, s$ ,  $l \neq q$  with the recursive formulation (12) and at the level  $l = q$ , the last equation in (17) is used. In (17) we assume that hyperparameters have not been modified when adding the new computation of level  $q$  in  $\mathbf{x}$ . The variance prediction (17) requires the computation of a matrix-inversion at each  $\mathbf{x}$  or  $q$  level. This is performed by using partitioned matrix in order not to increase the numerical cost with the matrix-inversion lemma:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}. \quad (18)$$

In (18), we have  $A^{-1} = (C_q(\Theta_q) + \sigma_{\epsilon_q}^2 \mathbf{I}_q)^{-1}$  that has already been computed  $\forall q = 1, \dots, s$  during the initial steps of MF surrogates constructions. Finally, we propose to define the  $AEI$  merit function as

$$AEI_l(\mathbf{x}) = EI_s(\mathbf{x}) \left( 1 - \frac{\sigma_{\epsilon_l}}{\sqrt{\widehat{\sigma}_{f_s}^2(\mathbf{x}) + \sigma_{\epsilon_l}^2}} \right) \frac{\beta_s}{\beta_l} \max \left( 0, \left( 1 - \frac{\widehat{\sigma}_{f_s}^2 |l(\mathbf{x})|}{\widehat{\sigma}_{f_s}^2(\mathbf{x})} \right) \right), \quad (19)$$

where the multi-fidelity Expected Improvement  $EI_s(\mathbf{x})$  is defined by

$$EI_s(\mathbf{x}) = \widehat{\sigma}_{f_s}(\mathbf{x}) [\Phi(u_s(\mathbf{x})) + \phi(u_s(\mathbf{x}))], \quad (20)$$

$$u_s(\mathbf{x}) = \frac{\widehat{f}_s(\mathbf{x}_{*,n_1,\dots,n_s}) - \widehat{f}_s(\mathbf{x})}{\widehat{\sigma}_{f_s}(\mathbf{x})}, \quad (21)$$

with  $\Phi$  and  $\phi$  the cumulative and density functions of the standard Gaussian distribution, and  $\mathbf{x}_{*,n_1,\dots,n_s} \in \mathbf{X}_{n_1,\dots,n_s}^{1,\dots,s}$  is the current effective best solution of all the observations. When using nested level observations, we do not need the prediction  $\widehat{f}_{l-1}(\mathbf{X}_{n_l}^l)$  in (11) and (14), and the multi-fidelity based merit function (19) is modified to

$$AEI_l(\mathbf{x}) = EI_s(\mathbf{x}) \left( 1 - \frac{\sigma_{\epsilon_l}}{\sqrt{\widehat{\sigma}_{f_s}^2(\mathbf{x}) + \sigma_{\epsilon_l}^2}} \right) \frac{\sum_{i=1}^s \beta_i}{\sum_{i=1}^l \beta_i} \max \left( 0, \left( 1 - \frac{\widehat{\sigma}_{f_s}^2 |1,\dots,l(\mathbf{x})|}{\widehat{\sigma}_{f_s}^2(\mathbf{x})} \right) \right). \quad (22)$$

In contrary to the single-fidelity case, the merit function is now dependent on the fidelity level  $l$ . This implies to perform a maximization of the  $AEI_l(\mathbf{x})$  at each level to find the optimum  $\mathbf{x}_{n_l+1} = \arg \max_{l \in [1,s], \mathbf{x} \in \Omega} AEI_l(\mathbf{x})$ . The optimum  $\mathbf{x}_{n_l+1}$  is added to  $\mathbf{X}_{n_l}^l$ , and  $f_l$  is evaluated at the new point providing  $y_{l_{n_l+1}}$ . Setting  $n_l \leftarrow n_l + 1$ , a new iteration can start updating the multi-fidelity GP models with the new observation of level  $l$ . The iterations of the MF GP-based optimization problem are continued until a stopping criterion is satisfied or the resources allocated to the optimization have been exhausted. One iteration step of the non-nested multi-fidelity EGO approach is summarized in the Algorithm 1 below.

---

**Algorithm 1:** Multi-fidelity non-nested EGO

---

**Input:**  $\mathbf{X}_{n_1,\dots,n_s}^{1,\dots,s}$ ,  $\mathbf{Y}_{n_1,\dots,n_s}^{1,\dots,s}$

**Output:**  $\mathbf{x}_{n_l+1}$ ,  $l$

```

1 for $l \leftarrow 1$ to s do
2 Solve $(\rho_{l-1}, \Theta_l, \sigma_{\epsilon_l}^2) = \arg \max$ (14) !Build MF surrogates
3 Compute $(C_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1}$
4 Save ρ_{l-1} , Θ_l , $\sigma_{\epsilon_l}^2$, $(C_l(\Theta_l) + \sigma_{\epsilon_l}^2 \mathbf{I}_l)^{-1}$
5 for $l \leftarrow 1$ to s do
6 Solve $\mathbf{x}_l = \arg \max$ (19) !Find $\mathbf{x}_{n_l+1}$ at each fidelity level
7 Save $\mathbf{x}_l$
8 $(\mathbf{x}_{n_l+1}, l) \leftarrow \arg \max_{\mathbf{x} \in \mathbf{X}_{1,\dots,l}, l \in [1,s]} AEI_l(\mathbf{x})$!Next point to evaluate at level l
```

---

Figure 2 shows a sequence example of non-nested multi-fidelity EGO on the one-variable function  $f_e$ . The MF surrogate prediction  $\widehat{f}_2(x)$  is the black curve and the dot-dashed blue and dashed red curves are the  $AEI_l(x)$  merit functions (right vertical axis). The initial datasets are composed of 5 and 4 observations for respectively the first and second fidelity levels. The computing costs are  $\beta_1 = 1$  and  $\beta_2 = 10$  for this illustrative example. At the first EGO iteration ( $N_f = 1$  in Figure 2(a)), the algorithm favors a cheap computation close to  $x \approx 0.6$ . This improves the prediction  $\widehat{f}_2(x)$  close to global optimum (see Figure 2(b)) and increases the expected improvement of the high fidelity level, which implies the computation of  $f_e$  at  $N_f = 3$  and  $N_f = 4$  in this region. At end of the optimization,  $AEI_l(x)$  are very small and the global optimum of  $f_e$  is well captured by  $\widehat{f}_2(x)$  (see black and red continuous curves in Figure 2(f)).

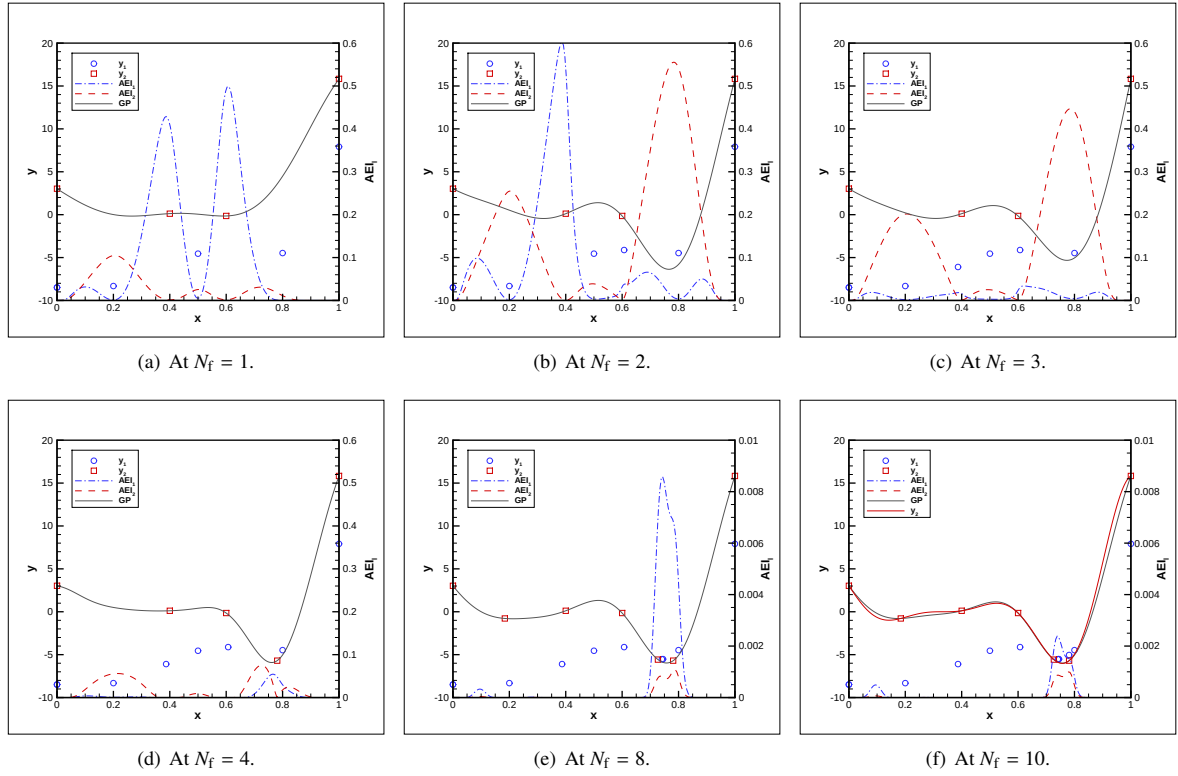


Figure 2: Non-nested EGO sequence example of one-variable function with two fidelity levels.

### 3. Analytical test problems

In this Section, we illustrate the computational efficiency of the multi-fidelity EGO approach by solving some optimization problems defined analytically. In particular, we compare the cases of non-nested and nested datasets on several multi-fidelity problems including level differences in the response function, costs and noise. The analytical problems are defined by the contractive function  $U_k(\mathbf{x})$ ,

$$\begin{cases} U_{k+1}(\mathbf{x}) = \frac{1}{2} \left( \frac{U(\mathbf{x})^2}{U_k(\mathbf{x})} + U_k(\mathbf{x}) \right), \\ U_0 \in \mathbb{R}, \quad \mathbf{x} \in \Omega \subset \mathbb{R}^d. \end{cases} \quad (23)$$

This allows to have several fidelity levels  $k$  in the application examples. In this work,  $U_0 = 10$  and 3 fidelity levels are used with  $k = 1$  for the first level,  $k = 3$  for the second, and the highest level is the true function  $U(\mathbf{x})$ . The test problems to be minimized are constructed with the *Hartmann* [Picheny et al. 2013] function of dimension 6, defined

as

$$\begin{aligned}
 U(\mathbf{x}) &= \sum_{i=1}^4 \alpha_i \exp \left( - \sum_{j=1}^6 A_{ij} (x_j - P_{ij})^2 \right), \\
 \boldsymbol{\alpha} &= (1, 1.2, 3, 3.2)^T, \\
 \mathbf{A} &= \begin{pmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix}, \\
 \mathbf{P} &= 10^{-4} \begin{pmatrix} 1312 & 1696 & 5569 & 124 & 8283 & 5886 \\ 2329 & 4135 & 8307 & 3736 & 1004 & 9991 \\ 2348 & 1451 & 3522 & 2883 & 3047 & 6650 \\ 4047 & 8828 & 8732 & 5743 & 1091 & 381 \end{pmatrix}, \\
 \mathbf{x} &\in [0, 1]^6,
 \end{aligned} \tag{24}$$

where the optimum location is  $\mathbf{x}_{\text{opt}} = [0.20169, 0.150011, 0.476874, 0.275332, 0.311652, 0.6573]$ .

The analytical multi-fidelity problem defined in (23) allows level differences in terms of responses  $U_k$ , but the optimum location  $\mathbf{x}_{\text{opt}}$  is the same at each fidelity level. Therefore, we have also introduced the following shift in the optimization problem:

$$x'_i = x_i + \frac{\delta}{k}, \quad \forall i = 1, \dots, 6, \tag{25}$$

where  $\delta \in \mathbb{R}$  is fixed, depending of the considered problem. The ratio  $\delta/k$  induces differences in the optimum location as function of the fidelity level. This shift is only applied at the first ( $k = 1$ ) and second ( $k = 3$ ) levels.

In this work, the efficiency of the multi-fidelity EGO approach is shown by solving 3 types of problems that are presented in following Sections. For each one, an identical and initial sampling scheme, using Latin Hypercube Sampling (LHS), of 20, 15 and 10 observations is set for the fidelity levels 1, 2 and 3 respectively. A maximum of  $N_f = 400$  evaluations is used for all tests. The computational costs of each problem are reported in Table 1 below. In

|                       | $\beta_1$ | $\beta_2$ | $\beta_3$ |
|-----------------------|-----------|-----------|-----------|
| Problem (3.1) & (3.2) | 1         | 100       | 1000      |
| Problem (3.3)         | 1         | 800       | 1000      |

Table 1: Computational costs of the analytical problems.

the following, Section 3.1 concerns the first problem, where (23) is minimized by using (24) for several  $\delta = 0, 0.5$  and 1. The second application in Section 3.2 is identical to the first one, except that a noisy response has been introduced in the middle-fidelity level 2. The cases  $\delta = 0$  and 1 are solved in this Section 3.2. Finally, Section 3.3 provides the last test which is identical to the first one, except for the cost evolutions (see Table 1). The cases  $\delta = 0$  and 1 are solved in this Section 3.3. Each analytical problem is solved with the non-nested and nested MF approaches, and the highest level is also optimized by considering a SF surrogate. Results are reported as functions of the computational cost.

### 3.1. Impact of the shift

Figure 3 shows the distances to the optimum as functions of the computational cost when solving the problem (3.1) with  $\delta = 0$ . Black circles are the SF model values, red squares come from the nested approach and blue triangles are given by the non-nested method. We note that both nested and non-nested approaches have faster convergence trends until a cost of  $\approx 300000$  in term of optimum location (Figure 3(a)) and prediction (Figure 3(b)) than the SF formulation. After this cost and until the end of the  $N_f = 400$  evaluations, the SF approach gives more precise results of the optimum location and prediction. This is explained by the fact that even if  $\delta = 0$  in the analytical definition of (3.1), the MF predictions at levels 1 and 2 may have slightly different optimum locations. These differences have thus

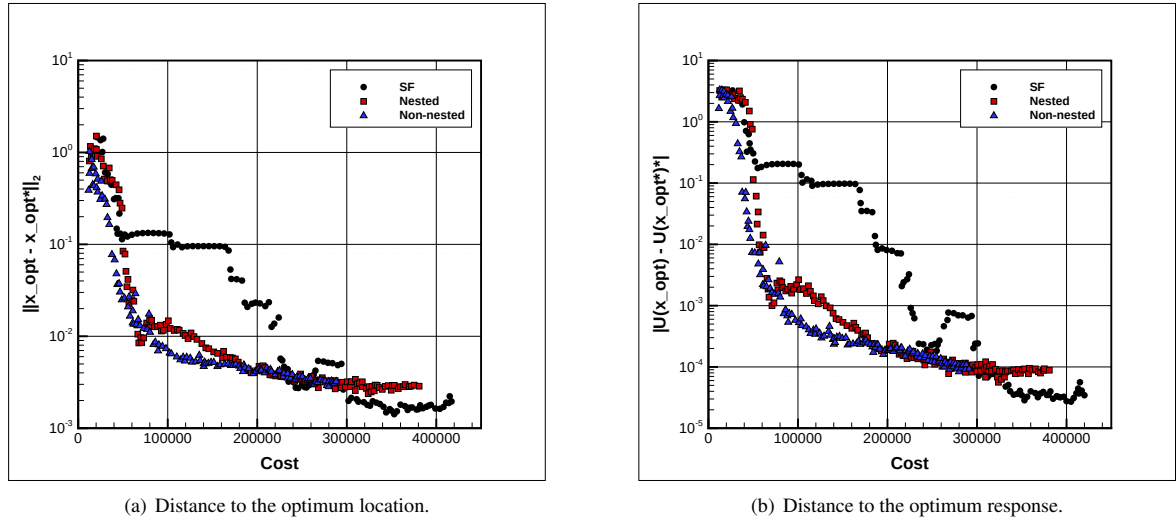


Figure 3: Distances to the optimum as functions of the cost. Case of MF problem (3.1) with  $\delta = 0$ .

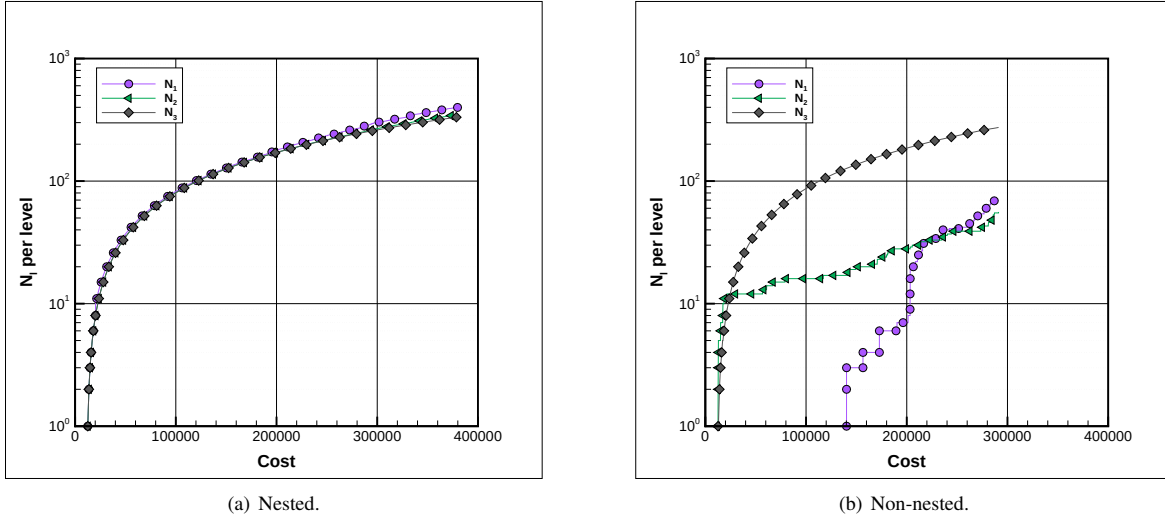
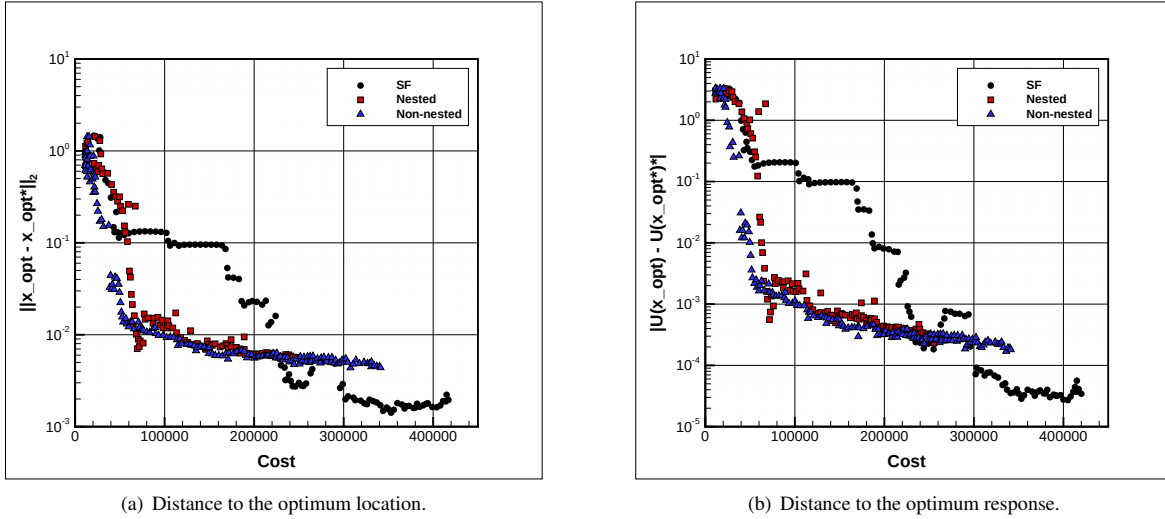
an impact in the multi-fidelity prediction of the highest level 3. By comparing the MF approaches, we can also see that the non-nested method is slightly more efficient in the convergence trend than the nested one.

Figure 4 shows the number of call per fidelity level as function of the cost for both nested and non-nested approaches. The highest fidelity level appears to be the most promising in the nested case since the 3 curves are almost overlaid (except during the last iterations) in Figure 4(a). In Figure 4(b), we can appreciate the non-nested case and the advantages of using such infilling strategy. The final computing cost is reduced because each evaluation of the highest fidelity level does not require the intermediate fidelity computations. At a cost  $\approx 100000$ , both nested and non-nested methods count  $\approx 100$  new observations of the highest fidelity level. The nested formulation has also at least 100 new observations of the levels 1 and 2, while the non-nested approach has 0 and 16 new observations of level 1 and 2 respectively. At the same cost, we see in Figure 3 that the non-nested method gives better predictions than the non-nested one. The non-nested approach is therefore more efficient since the method does not use of poor level evaluations to build the fidelity level 3, except when the intermediate computations help of reducing the highest fidelity level uncertainty.

Figure 5 shows the distances to the optimum as functions of the computational cost when solving the problem (3.1) with  $\delta = 0.5$ . In contrary to the previous application, the optimum location is now varying between the fidelity levels ( $\delta \neq 0$ ). Both MF methods have close convergence trends in the optimum location and prediction. The distances to the optimum are slightly higher in comparison to the case where  $\delta = 0$ . This comes from the surrogate models of the lower levels that induce a less precise global prediction than the highest one alone. Consequently, the SF optimization becomes more powerful if the allowed cost is  $\approx 250000$ , otherwise the MF formulations are much more efficient. Indeed, the multi-fidelity EGO method is very fast to determine a precise prediction of the optimum thanks to the exploration of the variability domain by low fidelity levels observations, that are computed with low costs.

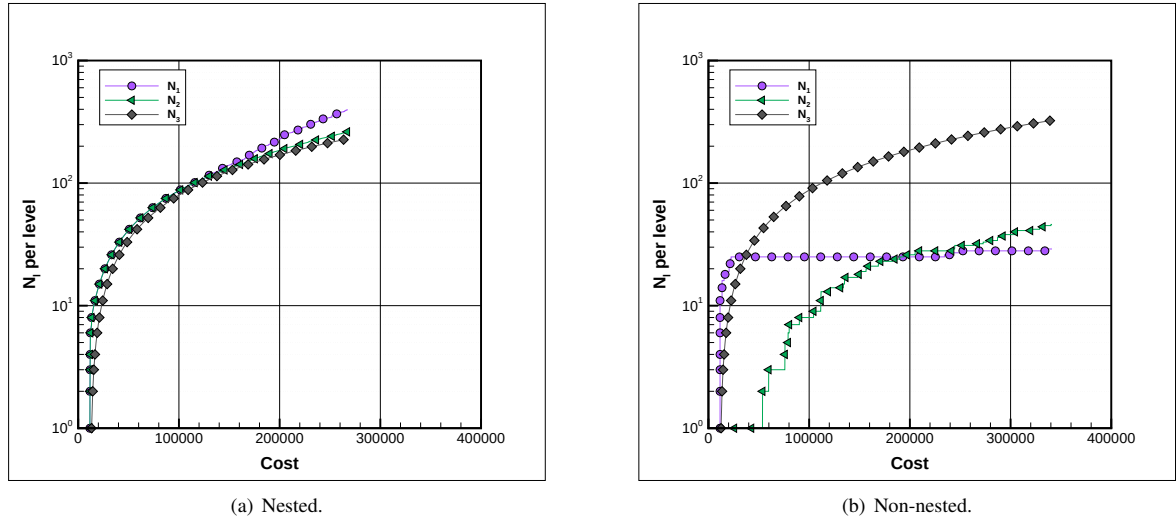
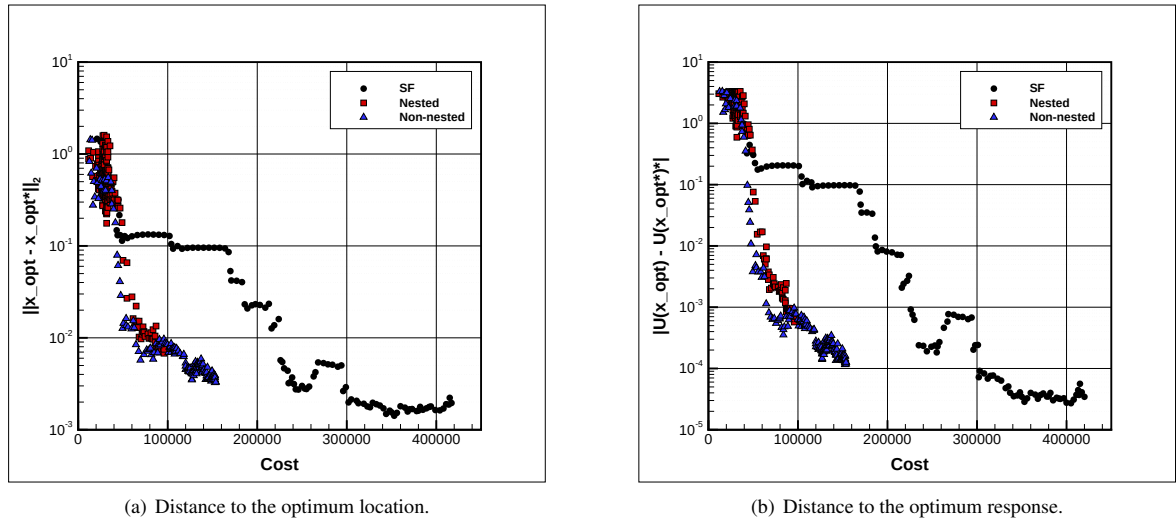
Figure 6 shows the number of call per fidelity level as function of the cost for both nested and non-nested approaches. The nested results are reported in Figure 5(a) and we observe that lower fidelity levels 1 and 2 are now computed from the beginning of the optimization. It is also the case for the non-nested method (see Figure 5(b)) with evaluations of level 1 alone. Indeed, when  $\delta \neq 0$  the multi-fidelity EGO is performed on surrogates that include prediction differences that are functions of  $\mathbf{x}$ , in addition to have fidelity differences in term of global mean responses. The scale parameters  $\rho_{l-1}$  can be build as functions of  $\mathbf{x}$ , as suggested in [Le Gratiet and Garnier 2014], to take into account such local differences; however this will increase the computing time and complexity of the surrogate constructions.

Figure 7 shows the distances to the optimum as functions of the computational cost when solving the problem (3.1) with  $\delta = 1$ . The response differences in the intermediate fidelity levels are even more increased in this case.

Figure 4: Number of call per level as function of the cost. Case of MF problem (3.1) with  $\delta = 0$ .Figure 5: Distances to the optimum as functions of the cost. Case of MF problem (3.1) with  $\delta = 0.5$ .

Consequently, most of evaluations at the end of the two multi-fidelity surrogates based optimizations correspond to level 1 and the total computing cost is much lower than the SF optimization (see Figures 7 and 8).

Figure 8 shows that the MF optimizations need more computations of the lowest level to sufficiently reduce MF variances due to these differences between the fidelity levels predictions. Indeed, when  $\delta = 1$  the low fidelity levels induce larger variances in the highest level 3 than when  $\delta = 0$ . It is thus necessary to further evaluate the lowest level 1 when  $\delta = 1$ , in order to reduce the large variances in the promising improvement areas of the level 3. As already discussed, the two MF methods are more efficient than the SF one at the beginning of the optimization thanks to the low fidelity and cheap observations. The efficiency comparison between the nested and non-nested approaches is not clearly shown in this example, but this analytical application shows that the multi-fidelity infilling strategy is robust against the possible differences in the fidelity level predictions.

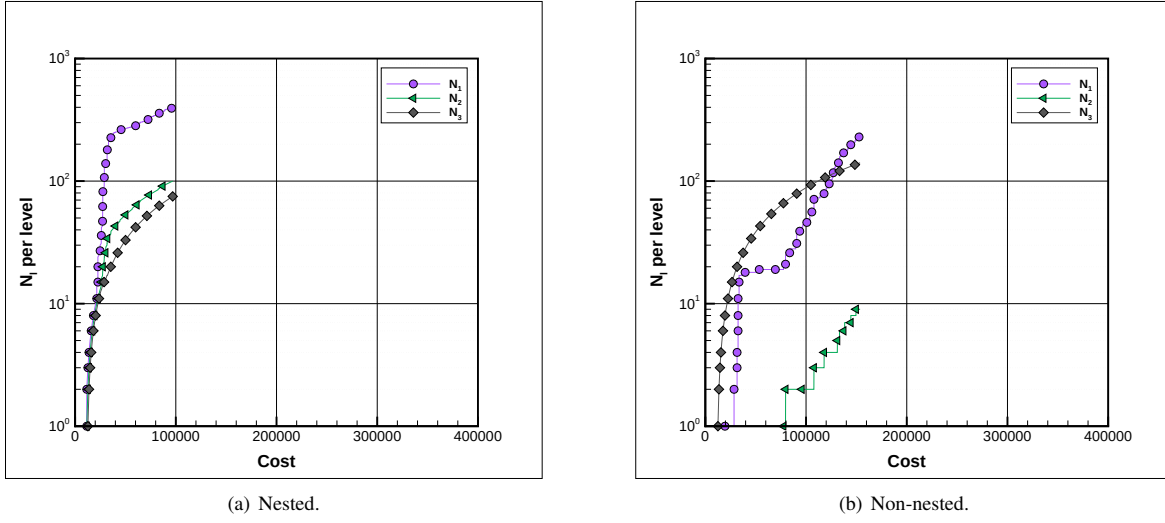
Figure 6: Number of call per level as function of the cost. Case of MF problem (3.1) with  $\delta = 0.5$ .Figure 7: Distances to the optimum as functions of the cost. Case of MF problem (3.1) with  $\delta = 1$ .

### 3.2. Impact of noise

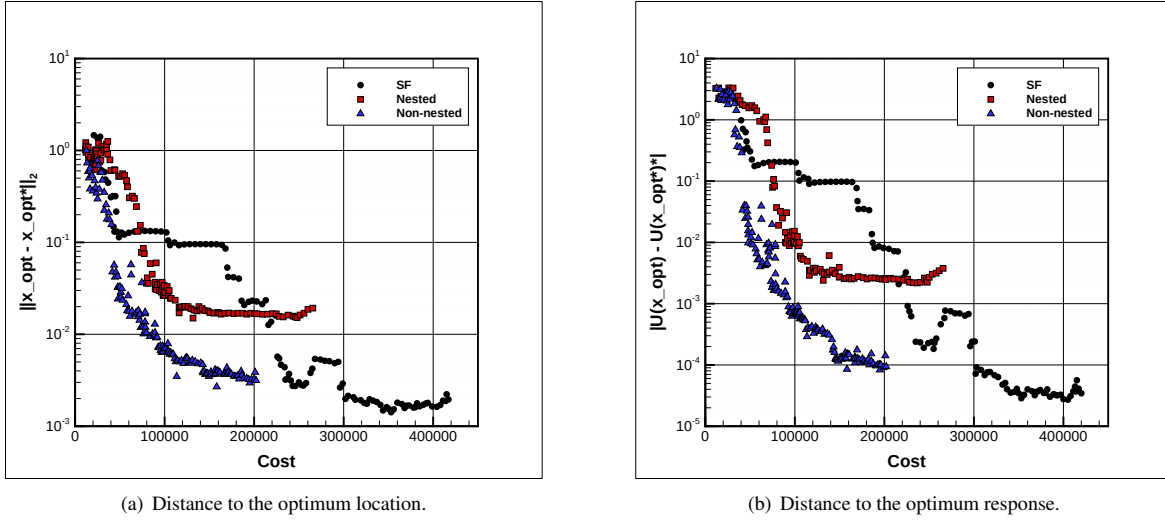
The problem (3.2) is solved in this Section. The computing costs are identical to the problem (3.1) and a noisy response has been introduced in the intermediate level 2 with:  $U'_3(\mathbf{x}) = U_3(\mathbf{x})(1 + 10\mathcal{U}(0,1))$ , where  $\mathcal{U}(0,1)$  is a uniform random number in  $[0,1]$ .

Figure 9 shows the distances to the optimum as functions of the computational cost when solving the problem (3.2) with  $\delta = 0$ . The non-nested approach is clearly the most efficient in this example. Both optimum location and prediction are better captured by this method from the beginning of the optimization. This is in accordance to the number of call per level as function of the cost given in Figure 10. Indeed we see that the noisy level 2 is the most called in the non-nested method (Figure 10(b)) while the level 1 is almost forgotten. The multi-fidelity infilling strategy needs of the level 2 evaluations in order to reduce the large variances in the promising areas of the highest



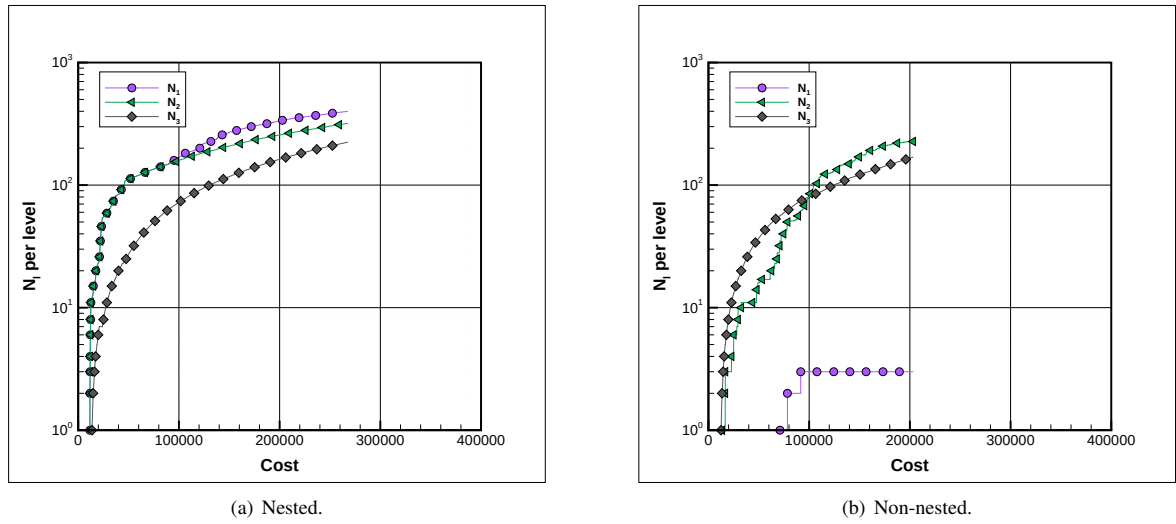
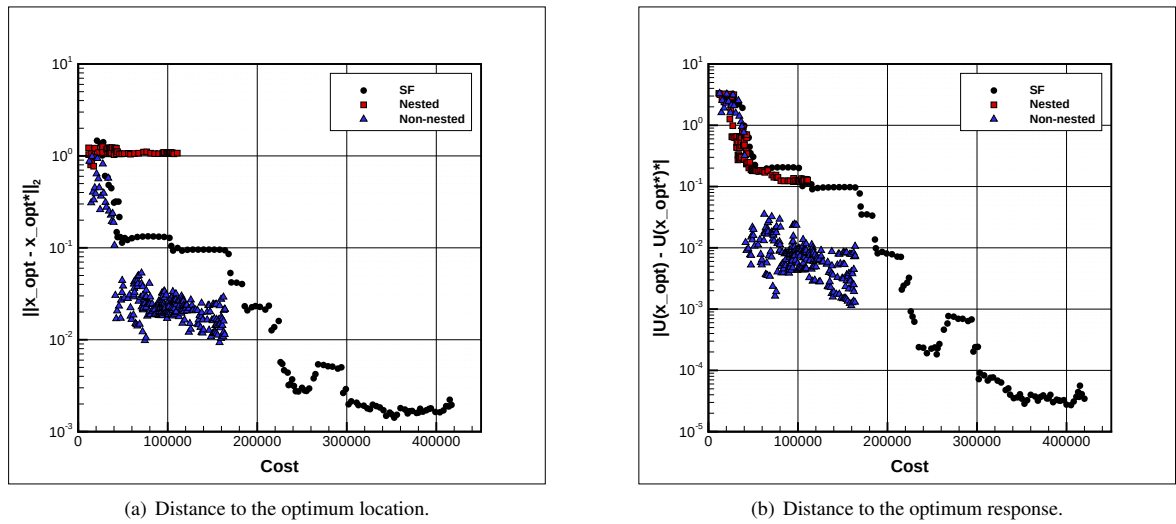
Figure 8: Number of call per level as function of the cost. Case of MF problem (3.1) with  $\delta = 1$ .

level 3. In the nested strategy (Figure 10(a)), the level 2 is also much more called than the level 3. However, this does

Figure 9: Distances to the optimum as functions of the cost. Case of MF problem (3.2) with  $\delta = 0$ .

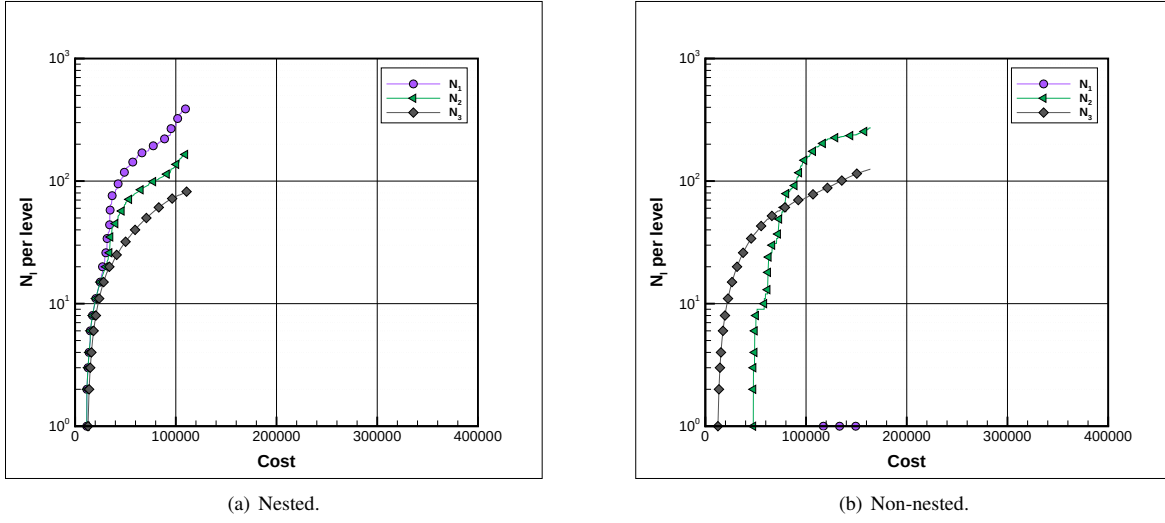
not help of improving the MF prediction of the optimum in the Figures 9(a) and 9(b). This is due to the nested property which force to compute the levels 1 and 2 when the level 3 is chosen by the infilling criterion. As a consequence, noisy observations from the level 2 are added in the nested MF surrogate construction. Whereas in the non-nested strategy, the computations of level 2 are only added if variance reductions of the highest MF surrogate are expected, therefore robustness is automatically brought by this non-nested formulation.

Figure 11 shows the distances to the optimum as functions of the computational cost when solving the problem (3.2) with  $\delta = 1$ . In this complex application example, the nested MF EGO is not able to converge on the optimum location, while the non-nested MF EGO has some difficulties but is able to find good predictions of the optimum in comparison to the SF EGO at the same optimization spent costs. It is important to recall that in this application, the

Figure 10: Number of call per level as function of the cost. Case of MF problem (3.2) with  $\delta = 0$ .Figure 11: Distances to the optimum as functions of the cost. Case of MF problem (3.2) with  $\delta = 1$ .

SF optimization is not penalized by the noise, since it is only included in the fidelity level 2.

The number of call per level as function of the cost in Figure 12 gives the same comments as before. These applications show the robustness and efficiency of the non-nested multi-fidelity strategy against possible noisy intermediate responses.

Figure 12: Number of call per level as function of the cost. Case of MF problem (3.2) with  $\delta = 1$ .

### 3.3. Impact of cost

The problem (3.3) is solved in this Section. The evolution of the computing costs is different from the previous cases, as  $\beta_2 = 800$ . This cost produces a non negligible step between the level 1 and 2 and it may correspond to a real application problem where several solver models are used.

Figure 13 shows the distances to the optimum as functions of the computational cost when solving the problem (3.3) with  $\delta = 0$ . Both MF methods are able to find the optimum within the 400 iterations. The non-nested MF approach is however the fastest to converge.

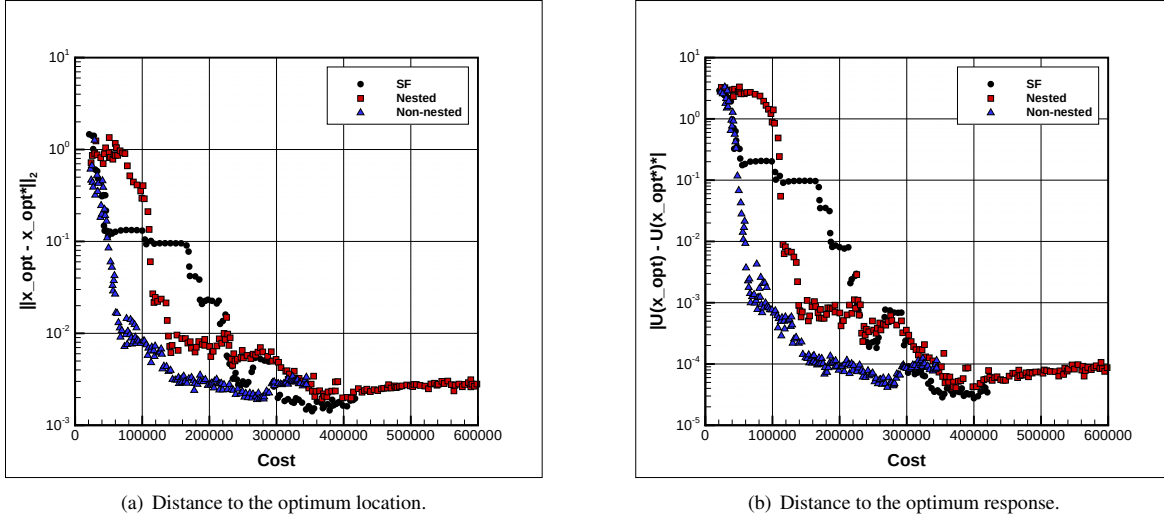
Figure 13: Distances to the optimum as functions of the cost. Case of MF problem (3.3) with  $\delta = 0$ .

Figure 14 shows the number of call per fidelity level as function of the cost for both nested and non-nested approaches. The levels 1 and 3 are favored by the nested strategy, while the non-nested one uses the levels 2 and

3 from the beginning of the optimization until a cost of  $\approx 275000$ . This explains the fast convergence of the non-nested method, while after this cost the level 1 start to be used which induces a slightly increase of the distances to the optimum in Figure 13. As already explained, the MF strategy evaluates the lowest fidelity levels when the

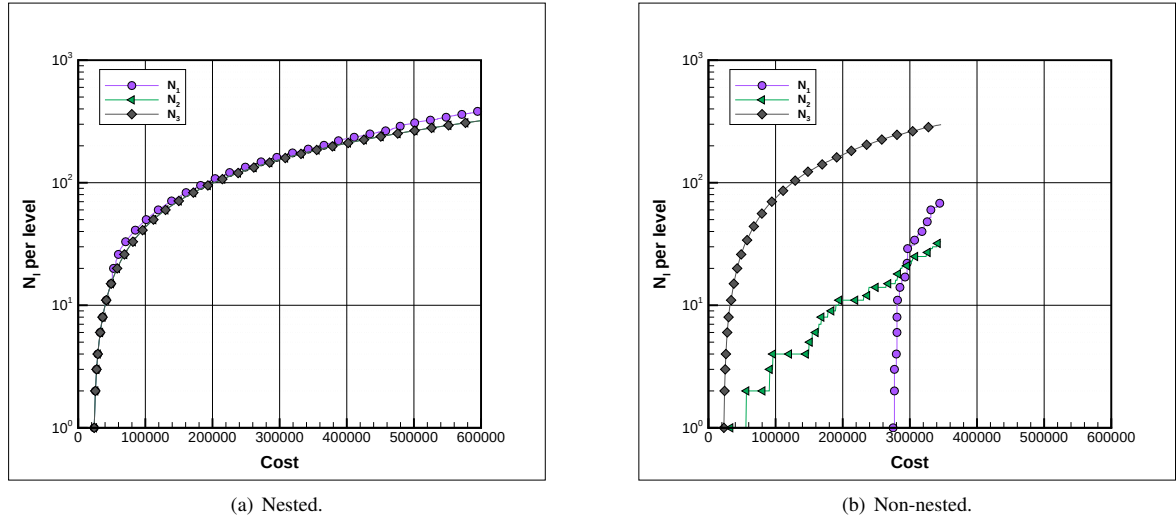
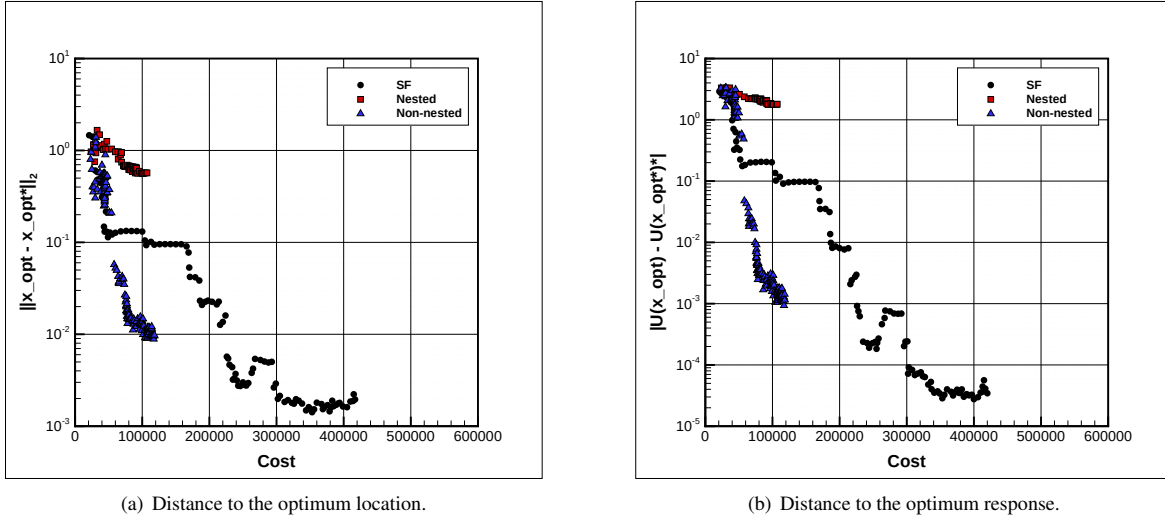
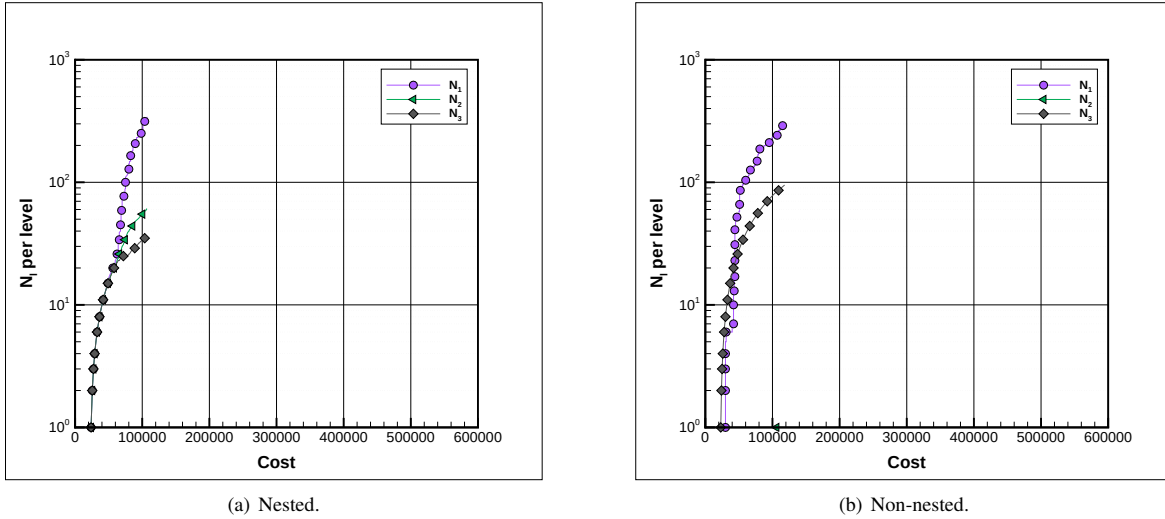


Figure 14: Number of call per level as function of the cost. Case of MF problem (3.3) with  $\delta = 0$ .

compromise cost versus variance-reduction in the expected improvement areas of the highest level 3 is maximal. At the beginning of the optimization, the non-nested approach does not use the level 1, as shown by Figure 14(b), since this level is already known enough. However, at the end of the optimization, this fidelity level 1 becomes a level of choice because the variance in the highest level 3 mainly comes from this level 1. Consequently, new evaluations of level 1 are performed (see Figure 14(b)) and the precisions of the optimum location and prediction are slightly decreased (see Figure 13(b)).

Figure 15 shows the distances to the optimum as functions of the computational cost when solving the problem (3.3) with  $\delta = 1$ . The optimum locations are now varying with the fidelity levels. The nested optimization is decreasing the distances to the optimum but at a very low convergence speed, which is moreover slower than the SF approach. The non-nested EGO is however very fast to decrease the distances to the global optimum and more efficient than the SF optimization.

We easily understand these results when looking at the number of call per level as function of the computational cost in Figure 16. Indeed, the nested formulation has to compute the costly fidelity level 2 for each computation of the highest level. Consequently, the first level is favored even if its response is poor in term of precision. In contrary the non-nested method is able to directly compute the highest level 3 without any calculations of the level 2. As we can see in Figure 16(b), the level 2 is indeed not called during the whole optimization process: the infilling criterion favors the cheap level 1 and costly but the most precise level 3 to complete the databases. This a good property of the proposed non-nested method that is able to automatically skip useless fidelity models.

Figure 15: Distances to the optimum as functions of the cost. Case of MF problem (3.3) with  $\delta = 1$ .Figure 16: Number of call per level as function of the cost. Case of MF problem (3.3) with  $\delta = 1$ .

#### 4. Application to nonlinear FSI problem

We now apply the non-nested EGO on a realistic problem corresponding to the optimization of sail trimmings. The physics of the sail system involves very complex phenomena, such as nonlinear Fluid-Structure Interaction (FSI) effects and aero-elastic instabilities. Moreover, the modeling of real sailing conditions is still an open challenge because of the large uncertainties in the prediction of wind and sea states. To our knowledge, the sails optimization has thus been limited so far to somehow idealized situations. For instance, sail shape optimizations (without accounting for the full FSI problem) were reported in [Rousselon 2008], while the numerical trimming of two-dimensional sails was considered in [Chapin et al. 2008]. Regarding three-dimensional FSI problems, the authors of [Ranzenbach et al. 2013] mention an optimization of the trimming of sails, but within an inviscid flow approximation and few details are provided on the actual optimization procedure used. Recently, three-dimensional FSI sail trimming optimizations,

using single-fidelity EGO, have been performed in both experimental and numerical viscous ways [Sacher et al. 2017].

The present work aims at using of the non-nested EGO method to reduce the numerical cost of optimizing complex sail systems with realistic physical models (e.g. nonlinear FSI and turbulent flows). Indeed, the numerical models of sail systems are composed of several fidelity levels, that may concern both the structural and fluid models. The following optimization problem includes five fidelity levels (in structure and fluid) on a two-parameter objective function in order to allow easy surrogate visualizations. The application is also solved with the highest fidelity level alone to perform SF and MF performance comparisons. The optimization problem and the numerical solvers are presented in Section 4.1. Results are then given in Section 4.2.

#### 4.1. Numerical setting

##### 4.1.1. Problem definition

The considered model is a *Figaro* sailing boat equipped with sails that have been designed by the *Incidence Sails* company. The sailing conditions are a boat speed of 5.6 kts, a true wind speed of 8 kts at 10 m height and a true wind angle of 50 deg. The trim, heel and heading angles are 0, 10 and 0 deg respectively. The two trimming parameters to be optimized are the main sheet and jib sheet lengths (see Figure 17). The performance (objective) function  $f_s(\mathbf{x})$  to

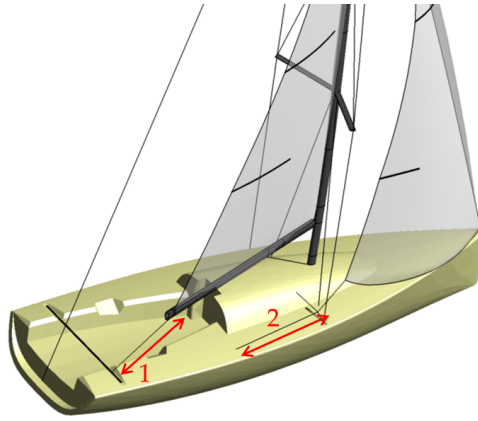


Figure 17: Design variables.

maximize is a linear combination of the thrust force  $F_{X_s}$  and the side force  $F_{Y_s}$  which induces an added hydrodynamic drag and leeway:

$$\begin{aligned} f_s(\mathbf{x}) &= F_{X_s}(\mathbf{x}) + 0.05F_{Y_s}(\mathbf{x}), \\ \mathbf{x} &\in \Omega \subset \mathbb{R}^2. \end{aligned} \quad (26)$$

The variability domain is  $\Omega = [1.5, 1.7] \times [2.45, 2.65]$ , which correspond to 20 cm of variability for each parameters. Five fidelity levels are used in this optimization. This includes differences in term of both structure and fluid models. Table 2 describes the model properties of each level. The structural differences concern the sails and the rig that can be rigid or flexible. Two flow solvers are used, inviscid and viscous, and two types of limit condition, with associated number of fluid cells, on the sails are available for the level 4 and 5. The computing cost are also given in Table 2, where we can see the important step between levels with inviscid and viscous flow models.

##### 4.1.2. Numerical models

We assume that the FSI problem has a steady solution for all values of the trimming parameters in the optimization range, such that a quasi-steady approach can be used to couple the structural and fluid solvers. These solvers are now briefly described.

| Level | Sails    | Rig      | Fluid model | Wall condition & Mesh | Cost $\beta_i/\beta_1$ |
|-------|----------|----------|-------------|-----------------------|------------------------|
| 1     | Rigid    | Rigid    | Inviscid    | –                     | 1.0                    |
| 2     | Flexible | Rigid    | Inviscid    | –                     | 3.3                    |
| 3     | Flexible | Flexible | Inviscid    | –                     | 4.2                    |
| 4     | Flexible | Flexible | Viscous     | Wall function (2.3M)  | 148.0                  |
| 5     | Flexible | Flexible | Viscous     | Solved wall (3.6M)    | 1430.4                 |

Table 2: Model properties at each fidelity level.

**Structural solver.** We rely on the ARA software developed by K-Epsilon for the structural model of the sail. The sail model involves Timoshenko beams (mast, boom, battens), cables (sheet) and Constant Strain Triangles (CST) membrane elements (cloth) of various types in large displacement formulation [Augier 2012]. A wrinkle model is added to capture the local sail deformations with a reasonable number of membrane elements [Katz and Plotkin 2001]. See Durand [2012] or Nakashino and Natori [2005] for further details on CST elements and the wrinkle model. The nonlinear solution is obtained by means of a Newton method with Aitken relaxation. The geometry of the sails is imported from the SAILPACK software developed by BSG Developments, to best fit the real geometry.

**Flow solvers.** The structural solver has been coupled with two different flow solvers. First, it is coupled with the steady and unsteady inviscid flow solver AVANTI [Roux et al. 2008]. The inviscid solver is based on the Vortex Lattice Method (VLM) [Katz and Plotkin 2001] where the sail surface is discretized in rectangular panels supporting doublets distributions. The Kutta condition is enforced at the trailing edge to generate vorticity particles that model the wake and its nonlinear dynamics. This inviscid model (see Figure 18(a)) assumes slip conditions over all the sail surface and therefore cannot account for detached flow situations. Its domain of validity is then limited to low-camber sails trimmings and a moderate angle of attacks. The steady inviscid solution is identified with the limit as  $t \rightarrow \infty$  of the unsteady solutions. The structural solver is also coupled with the parallel (MPI) incompressible viscous flow solver ISIS-CFD (of FINE<sup>TM</sup>/Marine). This software solves the Unsteady Reynolds-Average Navier-Stokes Equations (URANS) with finite volume discretization, accommodating both structured and unstructured meshes. The fluid domain is meshed to fit the sail geometries using HEXPRESS<sup>TM</sup>, a semi-automated mesh generator. Note that the mast is not represented in the fluid model (see Figure 18(b)). Regarding the turbulence modeling, we used the two equations shear stress transport (SST)  $k - \omega$  model [Menter et al. 2003].

**Coupled problem.** The FSI problems are solved coupling the structural and flow solvers through a quasi-monolithic algorithm [Durand 2012], which is an implicit coupling procedure adapted to partitioned solvers. Briefly, the resolution of the structural problem is nested inside the iterations on the nonlinear steady flow solution. This approach preserves the convergence and stability properties of the monolithic approach. As the sail sustains large deformations, the flow solver must account for the change of the boundary geometry. This is easily achieved in the inviscid solver where only the sail surface is discretized. This is less trivial and computationally more demanding in the case of the viscous solver, which requires a volume mesh. In this work we use the mesh deformation propagation (MDP) method [Durand et al. 2014], proceeding from the sail boundary toward the inside of the fluid domain. More details on the solvers and the coupling algorithms can be found in [Durand 2012].

#### 4.2. Optimization results

Results of the MF surrogate-based EGO of the sail trimming problem (26) are now presented. The initial sampling schemes, using LHS, are composed of 9, 7, 5, 3 and 1 observations for the fidelity levels 1, 2, 3, 4 and 5 respectively. Figure 19 shows the convergence process of the computed level as function of  $N_f$ . The value of  $AEI_f(\mathbf{x}_{n_i})$  is used to color the curve and it indicates the expected improvement in the computed levels as function of  $N_f$ . Thus, we observe a decrease of the merit function and the procedure is converged at  $N_f \approx 60$ . We also qualitatively note that level 2 and 3 are the ones that have been computed the most of time during the optimization.

The number of calls per fidelity level as function of  $N_f$  is plotted in Figure 20(a). The level 3 has been computed most of the times, while level 1 has approximately the same number of calls than level 4 and 5. We also note that during the first 20 iterations, the level 4 is the second fidelity level of choice. These observations can be explained with Figure 20(b), where the mean prediction of the performance function  $f_i(\mathbf{x})$  is plotted as function of the two trimming

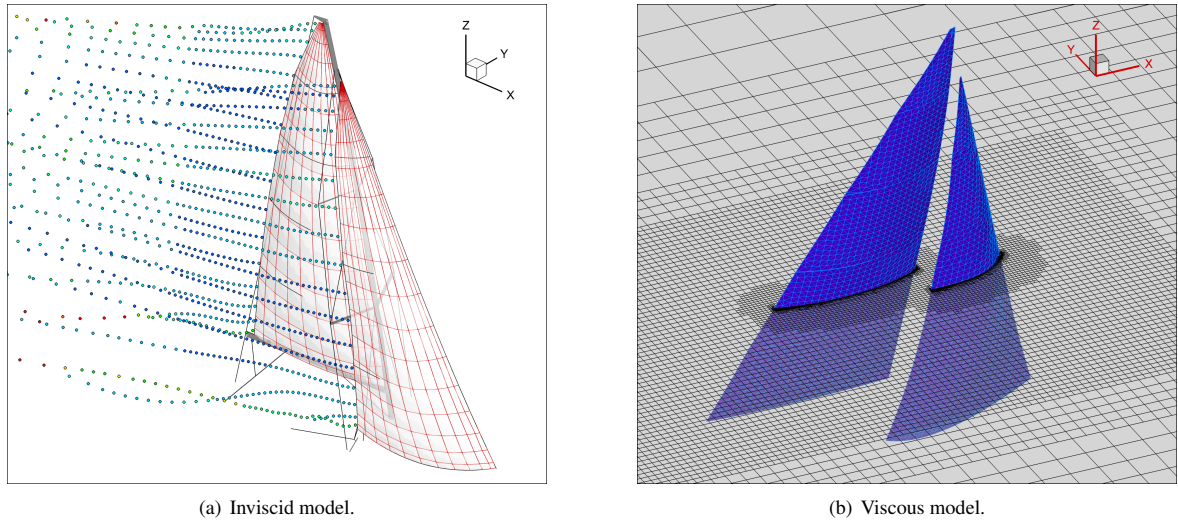


Figure 18: Fluid and structure meshes.

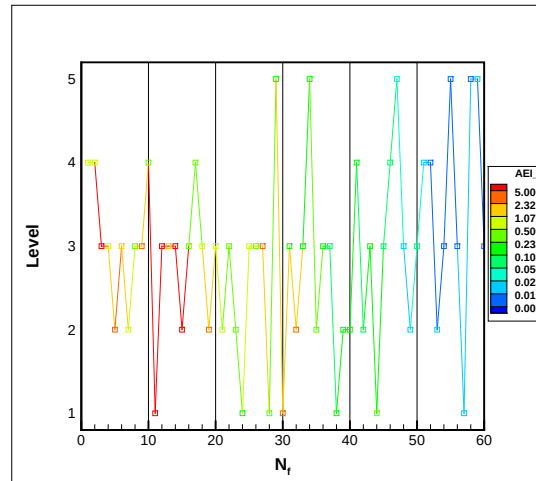


Figure 19: MF EGO convergence when solving (26).

parameters at each fidelity level. The surface colors correspond to the color used in Figure 20(a) to differentiate the levels. We note that level 1 is closer than level 2 and 3 in term of mean response distance to highest fidelity level. However, the performance prediction given by the level 1 has very low variabilities compared to the other levels. Consequently, the non-nested EGO approach does not compute a large number of observations from this level because it does not provide variance reduction in level 5, even if its computing cost is very low. The levels 2 and 3 overestimate the response of the performance function but they contain visible variabilities that are useful for the MF surrogate of level 5. The level 4 is very close to the highest fidelity level in term of mean response and variabilities. This is not surprising, since both models use the same structural and flow solvers, and differences are only on the sails limit condition and mesh refinements. To the end, the non-nested approach is able to explore  $\Omega$  with low cost levels by avoiding useless fidelity levels and to refine promising areas with the two highest fidelity levels.

Figure 21 provides sailor point of views of optimal sail trimmings that are given by each fidelity level. The sails and rigs colors correspond to fidelity levels colors already used in Figure 20(a). Those views allow to understand the



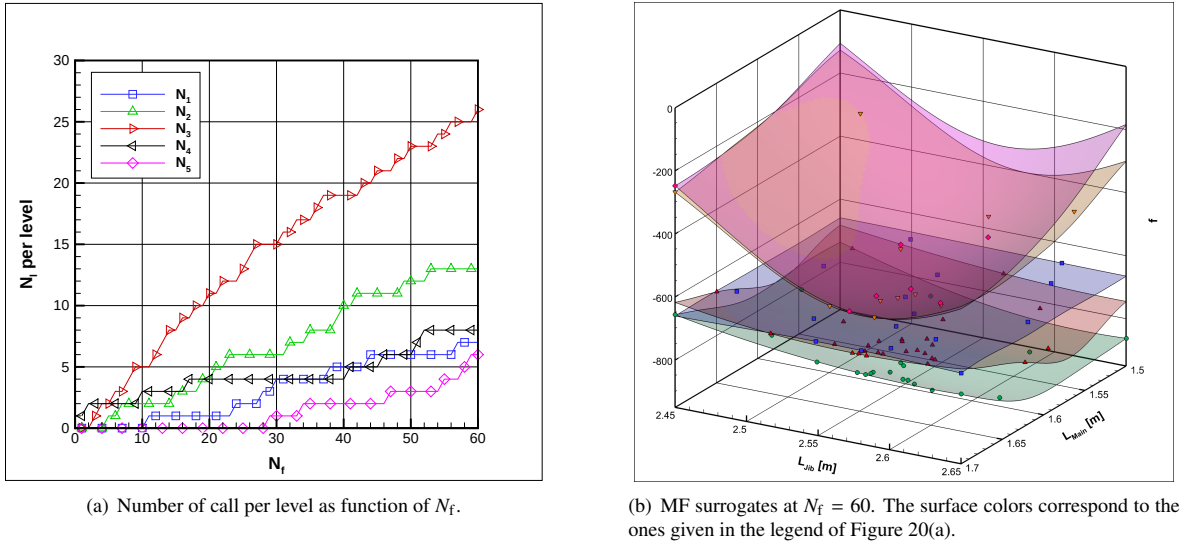


Figure 20: Number of call per level and MF surrogates at convergence when solving (26).

differences in performance responses of each level that are shown in Figure 20(b). The two viscous model 4 and 5 are very close in term of jib and main shapes. Differences are much larger in the fidelity levels with inviscid flow. This is particularly true for the main sail drafts that are larger than the viscous ones. This comes from the inviscid flow property which models an attached flow on sails, leading to overestimations of aerodynamic forces and gives rise to sail shapes with large drafts. It is in accordance with the performance predictions in Figure 20(a), where the most is overestimated the prediction, the most is larger the sail draft.

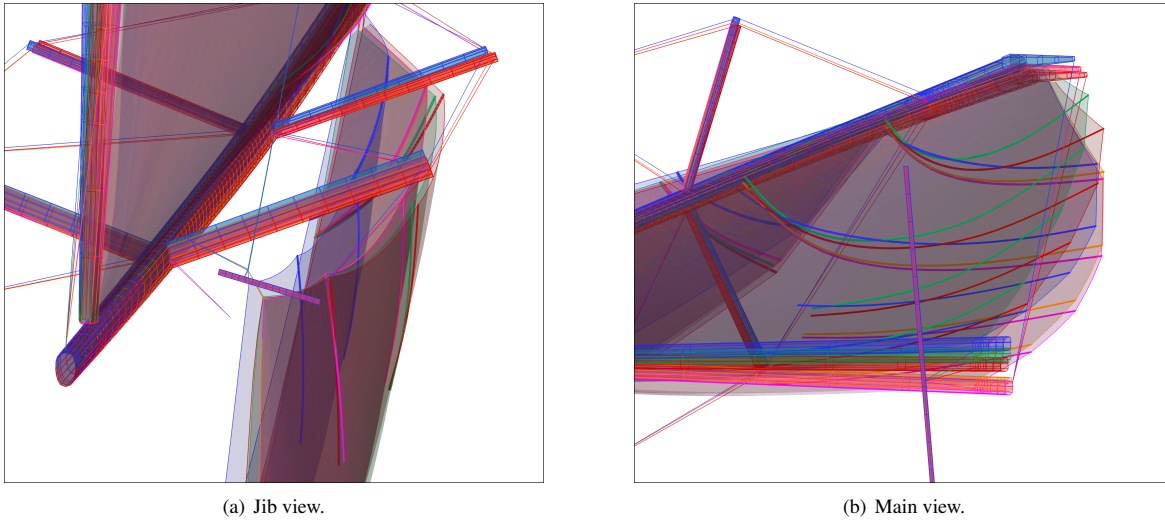


Figure 21: Optimal trimming at each fidelity level when solving (26).

The trimming optimization problem (26) has also been solved with a single-fidelity EGO approach, by using the highest fidelity level 5. An initial LHS sampling of 5 observations is used for the SF optimization. The convergences of both SF and MF optimal trimming as function of the computing cost are reported in Figure 22. The non-nested MF

approach allows to divide the computing time by  $\approx 3$  in comparison to the SF method. To the end, the main and jib lengths parameters are very close with differences that are  $\approx 0.25$  cm for the main and  $\approx 0.55$  cm for the jib.

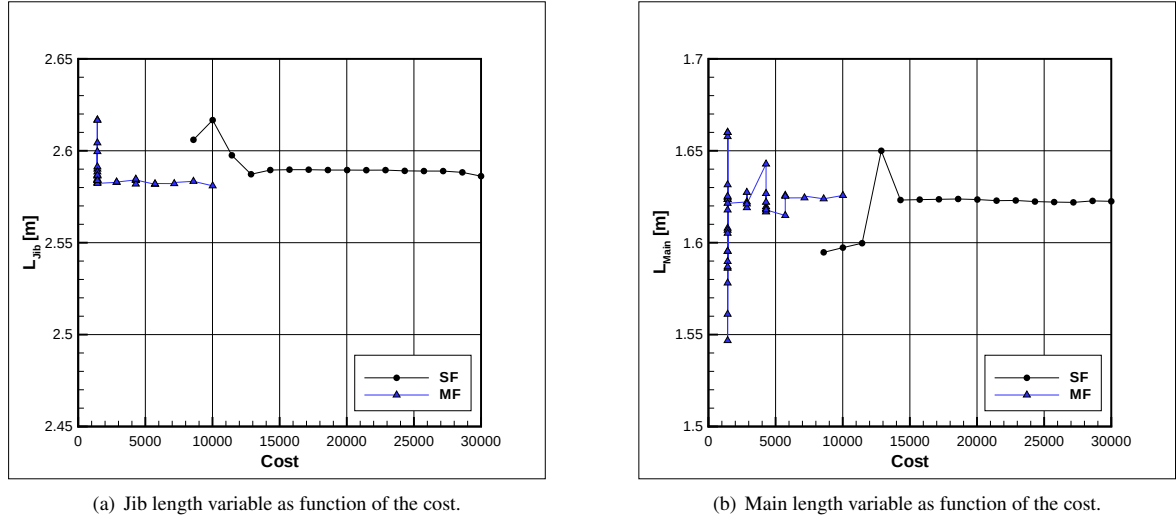


Figure 22: Optimal trimmings of SF and MF surrogates when solving (26).

## 5. Conclusion

We have presented in this paper a new infilling formulation for multi-fidelity surrogate-based optimizations. This allows to use Efficient Global Optimization (EGO) with surrogate models based on observations of different fidelity levels. The approach uses nested or non-nested datasets cases and can be applied with the powerful Expected Improvement (*EI*) based merit functions. Specifically, the sequential approach finds the most promising design variable set at which the objective function should be evaluated by also including its fidelity level choice. This is performed by counterbalancing the computational cost level and the potential uncertainty reduction in the highest fidelity level, that has to be optimized, due to the evaluation of the design variable set at the candidate level.

The proposed non-nested approach has been tested on several analytical problems, that allowed to validate the method and demonstrate its efficiency in comparison to the nested multi-fidelity or single-fidelity EGO approaches. Indeed, it has been shown that, compared to the nested formulation, the non-nested one was more efficient in terms of numerical cost, but also more robust against noisy responses of intermediate fidelity levels and high variations in the cost level evolutions.

Finally, the method was also used in the optimization of a fully nonlinear numerical fluid-structure interaction problem. Here we aimed at maximizing a performance function of a Figaro racing yacht by optimizing the sail trimmings. The problem was two-parameter and composed of five fidelity levels that include differences in term of structure and fluid models. On this example, the non-nested multi-fidelity EGO has been compared to the single-fidelity EGO. This complex application highlighted the interest of using a non-nested multi-fidelity approach to significantly reduce the computational cost.

Future works will concern the formulation of multi-fidelity surrogate-based optimization in the parallel computing framework. Indeed, the computational costs and the multi-fidelity responses can be used to define a formulation of merit function which allows several evaluations at each iteration of the EGO. This would help of reducing even more the global cost of the optimization.

## References

- D. R. Jones, M. Schonlau, W. J. Welch, Efficient Global Optimization of Expensive Black-Box Functions, *Journal of Global optimization* 13 (4) (1998) 455–492.
- S. Jeong, M. Murayama, K. Yamamoto, Efficient Optimization Design Method Using Kriging Model, *Journal of aircraft* 42 (2) (2005) 413–420.
- B. Glaz, P. P. Friedmann, L. Liu, Helicopter Vibration Reduction throughout the Entire Flight Envelope Using Surrogate-Based Optimization, *Journal of the American Helicopter Society* 54 (1) (2009) 12007.
- N. Aghajari, M. Schäfer, Efficient shape optimization for fluid–structure interaction problems, *Journal of Fluids and Structures* 57 (2015) 298 – 313.
- M. Sacher, F. Hauville, R. Duvigneau, O. L. Maître, N. Aubin, M. Durand, Efficient optimization procedure in non-linear fluid-structure interaction problem: Application to mainsail trimming in upwind conditions, *Journal of Fluids and Structures* 69 (2017) 209 – 231, ISSN 0889-9746, doi:<http://dx.doi.org/10.1016/j.jfluidstructs.2016.12.006>, URL <http://www.sciencedirect.com/science/article/pii/S0889974616303437>.
- T. Simpson, J. Poplinski, N. P. Koch, J. Allen, Metamodels for Computer-based Engineering Design: Survey and recommendations, *Engineering with Computers* 17 (2) (2001) 129–150, ISSN 1435-5663, doi:10.1007/PL00007198, URL <http://dx.doi.org/10.1007/PL00007198>.
- J. P. Kleijnen, Kriging metamodeling in simulation: A review, *European Journal of Operational Research* 192 (3) (2009) 707 – 716, ISSN 0377-2217, doi:<http://dx.doi.org/10.1016/j.ejor.2007.10.013>, URL <http://www.sciencedirect.com/science/article/pii/S0377221707010090>.
- M. C. Kennedy, A. O'Hagan, Predicting the Output from a Complex Computer Code When Fast Approximations Are Available, *Biometrika* 87 (1) (2000) 1–13, ISSN 00063444, URL <http://www.jstor.org/stable/2673557>.
- A. I. Forrester, A. Sobester, A. J. Keane, Multi-fidelity optimization via surrogate modelling, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 463 (2007) 3251–3269, ISSN 1364-5021, doi:10.1098/rspa.2007.1900, URL <http://rspa.royalsocietypublishing.org/content/463/2088/3251>.
- K. Elsayed, Optimization of the cyclone separator geometry for minimum pressure drop using Co-Kriging, *Powder Technology* 269 (Supplement C) (2015) 409 – 424, ISSN 0032-5910, doi:<https://doi.org/10.1016/j.powtec.2014.09.038>, URL <http://www.sciencedirect.com/science/article/pii/S0032591014008377>.
- V. Picheny, T. Wagner, D. Ginsbourger, A benchmark of kriging-based infill criteria for noisy optimization, *Structural and Multidisciplinary Optimization* 48 (3) (2013) 607–626, ISSN 1615-1488, doi:10.1007/s00158-013-0919-4, URL <https://doi.org/10.1007/s00158-013-0919-4>.
- L. Le Gratiet, C. Cannamela, Kriging-based sequential design strategies using fast cross-validation techniques with extensions to multi-fidelity computer codes, URL <https://hal.archives-ouvertes.fr/hal-00744432>, working paper or preprint, 2012.
- M. Plutowski, S. Sakata, H. White, Cross-validation estimates IMSE, in: *Advances in neural information processing systems*, 391–398, 1994.
- C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.
- N. Hansen, The CMA Evolution Strategy: A Comparing Review, in: *Towards a new evolutionary computation*, Springer, 75–102, 2006.
- D. Huang, T. T. Allen, W. I. Notz, N. Zheng, Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models, *Journal of global optimization* 34 (3) (2006) 441–466.
- L. Le Gratiet, J. Garnier, Recursive co-kriging model for Design of Computer experiments with multiple levels of fidelity, *International Journal for Uncertainty Quantification* 4 (5) (2014) 365–386, URL <https://hal.archives-ouvertes.fr/hal-01108813>.
- N. Rousselon, Optimization for Sail Design, in: *modeFRONTIER Conference*, 2008.
- V. G. Chapin, R. Neyhousser, G. Dulland, P. Chassaing, Design optimization of interacting sails through viscous CFD, in: *INNOVSail, Innovation in high performance sailing Yacht*, Lorient, 2008.
- R. Ranzenbach, D. Armitage, A. Carrau, Mainsail Planform Optimization for IRC 52 Using Fluid Structure Interaction, in: *The 21st Chesapeake Sailing Yacht Symposium, SNAME*, 2013.
- B. Augier, Etudes expérimentales de l'interaction fluide-structure sur surface souple: application aux voiles de bateaux., PhD Thesis, Université de Bretagne Occidentale, 2012.
- J. Katz, A. Plotkin, *Low-Speed Aerodynamics*, Cambridge University Press, second edn., 2001.
- M. Durand, Interaction fluide-structure souple et légère, application aux voiliers, PhD Thesis, Ecole Centrale de Nantes, 2012.
- K. Nakashino, M. C. Natori, Efficient modification scheme of stress-strain tensor for wrinkled membranes, *AIAA journal* 43 (1) (2005) 206–215.
- Y. Roux, M. Durand, A. Leroyer, P. Queutey, M. Visonneau, J. Raymond, J. M. Finot, F. Hauville, A. Purwanto, Strongly coupled VPP and CFD RANSE code for sailing yacht performance prediction, in: *3rd High Performance Yacht Design Conference*, 2–4 December, Auckland, 215–225, 2008.
- F. R. Menter, M. Kuntz, R. Langtry, Ten Years of Industrial Experience with the SST Turbulence Model, *Turbulence, heat and mass transfer* 4 (2003) 625–632.
- M. Durand, A. Leroyer, C. Lothodé, F. Hauville, M. Visonneau, R. Floch, L. Guillaume, FSI investigation on stability of downwind sails with an automatic dynamic trimming, *Ocean Engineering* 90 (2014) 129–139.

### 10.2.3 Conclusion

L'application de la méthode d'optimisation par méta-modèles à fidélité multiples est illustrée dans la Section 10.2.2. La prise en compte, au cours du processus d'optimisation itératif, de la sélection des niveaux de fidélité des nouvelles observations est présentée. Cette propriété permet, dans le cas d'observations non imbriquées, d'évaluer en priorité et en relation avec les coûts d'évaluation, les modèles apportant une réduction des incertitudes de prédictions dans les zones d'intérêt du modèle de plus haute fidélité.

La méthode est éprouvée sur différents tests analytiques, présentant des variétés sur la fidélité des réponses, sur l'inclusion d'une réponse bruitée dans un niveau intermédiaire et sur une évolution non uniforme des coûts d'évaluation. Les approches « imbriquée » et « non imbriquée » sont confrontées par ces tests analytiques, afin de démontrer l'intérêt de la méthode « non imbriquée ».

Un problème d'optimisation concernant une application concrète où les réglages d'un gréement de voilier sont optimisés est également présenté dans cette contribution. Cinq niveaux de fidélité sont utilisés pour illustrer la robustesse de la méthode, et le gain du coût global de l'optimisation est comparé à la résolution directe du niveau de plus haute fidélité par une approche simple fidélité.

Pour conclure, ces travaux de thèse décrits dans cette contribution permettent de valider la méthode d'optimisation par méta-modèles multi-fidélité sur différents cas analytiques, ainsi que sur une application concrète comportant la résolution de calculs d'interaction fluide-structure.

## 10.3 Optimisation à six paramètres

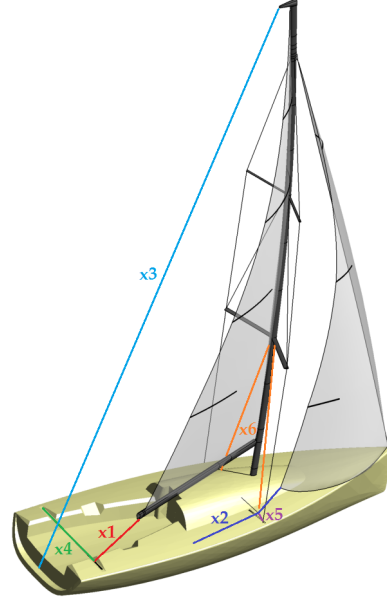
La Section 10.2 précédente met en évidence l'intérêt d'une optimisation EGO en multi-fidélité pour deux paramètres et cinq niveaux de fidélité. L'idée de cette Section 10.3 est d'augmenter la complexité du problème avec six réglages à optimiser. Le problème d'optimisation est défini à la Section 10.3.1 et la validation du maillage volumique fluide du type « overset » utilisé spécifiquement pour résoudre ce problème d'optimisation concerne la Section 10.3.2. Finalement, les résultats de l'optimisation sont donnés dans la Section 10.3.3.

### 10.3.1 Problème d'optimisation

L'objectif de cette optimisation est de déterminer les réglages optimaux  $\mathbf{x}_{\text{opt}}$  d'un Figaro dans un cas de navigation du type « reaching », qui correspond à un angle de vent réel d'environ 60 à 80 degrés (voir Section 1.1). Idéalement, un VPP devrait être utilisé afin de résoudre l'équilibre aéro-hydrodynamique du voilier. Cela permettrait de calculer la vitesse BS du voilier en fonction des réglages du gréement, et donc également l'angle du vent apparent AWA et sa vitesse apparente AWS. Dans cette phase de développement de la méthode d'optimisation, le problème est simplifié en maintenant AWA et AWS constants. L'angle du vent réel est de 65° et la vitesse du vent réel est de 18 kts ; pour le voilier, la gîte est de 20° et la vitesse est de 8.18 kts. Cette allure du type « reaching » a l'avantage de faire intervenir plus de réglages que dans le cas du près serré présenté dans la Section 10.2, et donc d'augmenter la dimension du problème d'optimisation. Les six réglages principaux à optimiser sont énoncés ci-dessous avec leurs domaines de variabilité respectifs. La Figure 10.4 illustre aussi

ces réglages sur le modèle numérique du Figaro.

1. Écoute de grand-voile,  $x_1 \in [1.85, 2.25]$  m ;
2. Écoute de génois,  $x_2 \in [2.75, 3.15]$  m ;
3. Pataras,  $x_3 \in [16.05, 16.50]$  m ;
4. Chariot de grand-voile,  $x_4 \in [1.90, 2.40]$  m ;
5. Barber de génois,  $x_5 \in [0.01, 0.30]$  m ;
6. D1,  $x_6 \in [4.45, 4.49]$  m.



**Figure 10.4** – Réglages du Figaro à optimiser.

Dans cette optimisation par méta-modèles, une formulation multi-fidélité à deux niveaux ( $s = 2$ ) est utilisée. Les différences dans la modélisation concernent uniquement le modèle fluide : pour  $l = 1$ , le modèle fluide est parfait et pour  $l = 2$ , le modèle fluide est visqueux. Le modèle structure est identique pour les deux niveaux de fidélité. Le problème d'optimisation consiste à maximiser une fonction performance  $f_l(\mathbf{x})$ , construite sur le point de fonctionnement du Figaro au « reaching », où l'effort propulsif  $F_x$  est pénalisé par l'effort de dérive  $F_y$ , ainsi que par le moment de chavirement maximal du voilier  $M_{x_{\max}}$ . Il aurait également été possible d'appliquer une contrainte  $M_{x_l} = M_{x_{\max}}$ , car le voilier navigue généralement à son moment de chavirement maximal pour ce genre d'allure. Néanmoins, cela complexifierait le problème d'optimisation en le rendant contraint. Le problème étant déjà assez complexe en soit, on préfère appliquer une pénalisation et travailler avec un problème d'optimisation non contraint. Ce problème d'optimisation s'écrit,

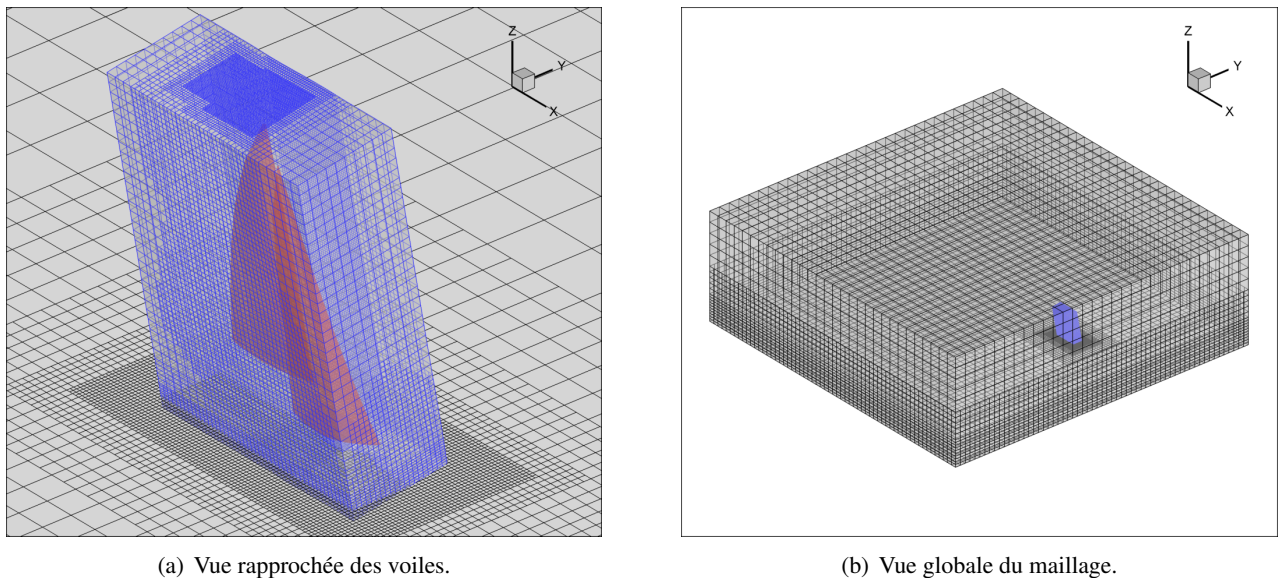
$$\max_{\mathbf{x} \in \Omega} f_l(\mathbf{x}) = F_{x_l} + 0.08F_{y_l} - \max(M_{x_l} - M_{x_{\max}}, 0), \quad (10.3.1)$$

où le moment de chavirement maximal vaut  $M_{x_{\max}} = 28000 \text{ N} \cdot \text{m}$ . La méthode EGO en multi-fidélité, avec la fonction de mérite (6.2.5) prenant en compte les possibles points non évaluables, présentée dans la Section 6.2 est appliquée ici. Les coûts respectifs des deux niveaux de fidélité sont  $\beta_1 = 1$  et  $\beta_2 = 100$ ; les plans d'expériences construits à partir d'hypercubes latins sont composés de 100 observations pour  $l = 1$  et de 10 observations pour  $l = 2$ .

### 10.3.2 Validation du maillage type « overset »

La résolution de (10.3.1) fait intervenir des calculs IFS fluide visqueux des voiles où les déformations du maillage volumique sont importantes, de par la dimension élevée ( $d = 6$ ) du problème d'optimisation. Un maillage volumique « classique » comme utilisé dans l'application de la Section 10.2 n'est malheureusement pas adapté ici. Un maillage du type « overset » (Steger and Benek, 1987; Deng et al., 2015) est utilisé à la place du maillage classique. Le maillage « overset » est composé de sous domaines maillés avec un recouvrement entre les domaines. Les points de ce maillage sont de trois types, « masqués », « interpolés » ou « calculés », suivant leur localisation dans les zones de possibles recouvrements du domaine de calcul. Cette méthode permet de s'affranchir des déformations du maillage dans le cas de simulations où des solides indéformables sont en mouvement. Elle nécessite néanmoins une construction adaptée des sous domaines maillés où les tailles des cellules doivent être du même ordre dans les zones de recouvrement, afin d'assurer la conservation des interpolations entre les sous domaines. Ces méthodes de maillages « overset » sont toujours en développement et elles trouvent des applications variées telles que le secteur naval (Shen et al., 2015) ou l'aviation (Peron, 2014).

Dans les calculs IFS visqueux sur les voiles, la déformation des voiles reste présente avec le maillage « overset ». Cependant, ce type de maillage aide grandement à la déformation des cellules. La Figure 10.5 montre le maillage déformable en « overset » (mailles bleues volumiques), autour des voiles (mailles rouges surfaciques), qui est compris dans le maillage non déformable (mailles grises volumiques). On remarque que



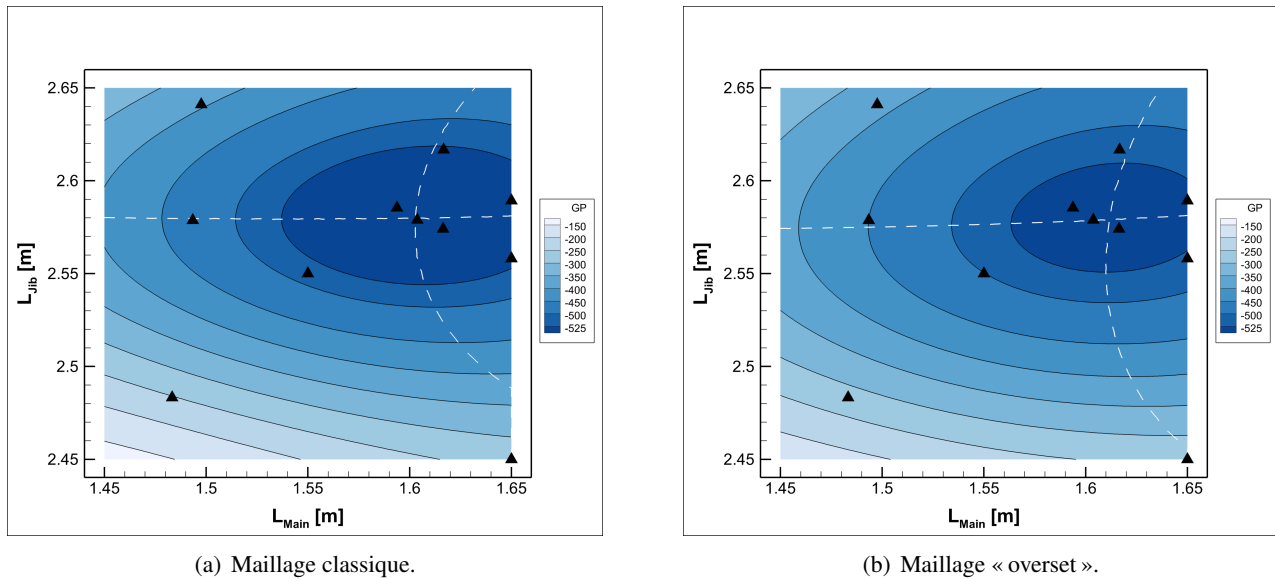
**Figure 10.5** – Maillage fluide visqueux volumique en « overset ».

le sous domaine est maillé verticalement autour des voiles et qu'une rotation autour de l'axe  $X$  est appliquée à l'ensemble afin de respecter la gîte du voilier. Cela permet d'avoir une meilleure qualité des mailles situées dans la couche limite des voiles. La taille du sous domaine (mailles bleues) est de plus assez faible comparée à celle du domaine complet (mailles grises). Cela aide également à la déformation du maillage car les déformations ne



s'appliquent qu'au sous domaine (mailles bleues), en incluant la déformation de ses six frontières externes.

Une validation de la méthode de maillage « overset » a été réalisée pour des calculs IFS visqueux sur les voiles du Figaro. L'objectif de cette validation est de s'assurer que les calculs IFS avec un maillage « overset » donnent des résultats comparables à des calculs IFS avec un maillage classique. L'application de la Section 10.2 où un cas de navigation au près est considéré avec les réglages des écoutes de grand-voile et de génois est reprise pour cette validation numérique. Les observations du niveau  $l = 4$  sont calculées avec un maillage du type « overset » et des comparaisons globales de la performance des voiles sont effectuées à l'aide de méta-modèles à simple fidélité. Cette comparaison est visible sur la Figure 10.6, où la prédiction moyenne du méta-modèle des calculs IFS avec un maillage classique est donnée par la Figure 10.6(a), alors que celle du méta-modèle des calculs IFS avec un maillage « overset » correspond à la Figure 10.6(b). Les lignes blanches discontinues sur chaque Figure sont associées aux deux réglages. Elles nous donnent les réglages optimaux, pour maintenir une performance optimale, lorsque l'un des deux réglages varie sur  $\Omega_i$ . On peut ainsi remarquer que les deux



**Figure 10.6** – Validation IFS numérique globale du maillage overset.

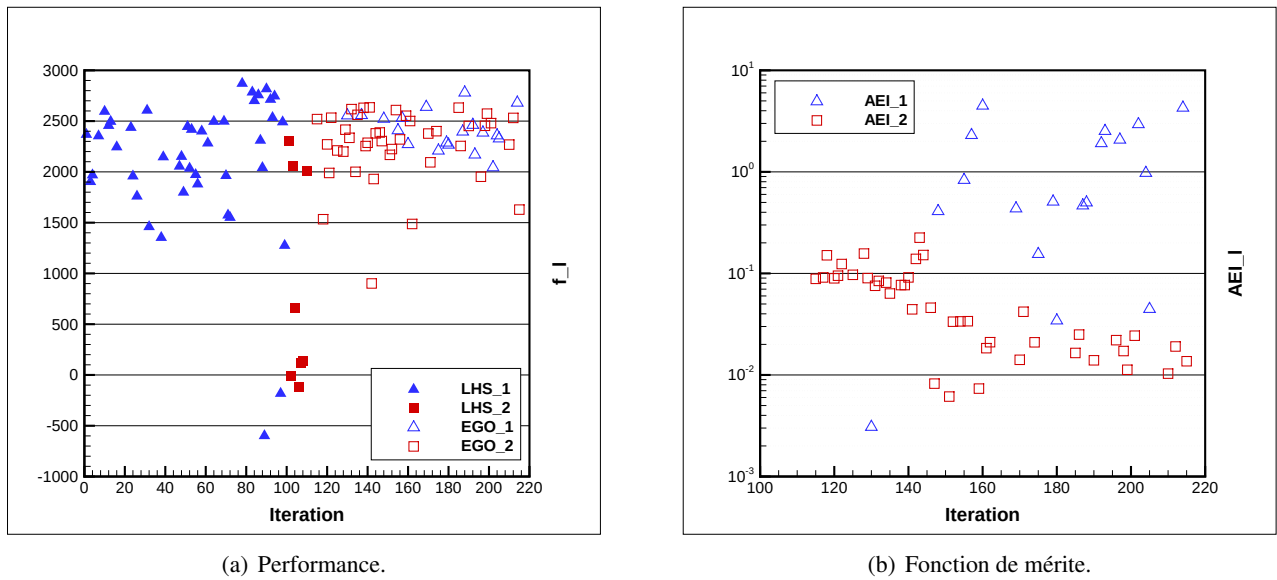
modèles sont très proches sur l'ensemble du domaine de variabilité des deux réglages, mais également au niveau de la position de l'optimum global où la différence maximale observée est de seulement 9 mm. Une comparaison locale au niveau de cet optimum peut également être effectuée. Cela est donné par le Tableau 10.1 qui reporte les efforts aérodynamiques des voiles pour les deux modèles. Les différences restent acceptables avec une sous-estimation de  $\approx 3\%$  par la méthode « overset » sur les efforts d'intérêts pour l'évaluation de la fonction performance. Cette validation numérique de calculs IFS visqueux sur des voiles rend alors possible la résolution du problème d'optimisation (10.3.1) à l'aide d'un maillage du type « overset ».

|             | $F_x$ [N] | $F_y$ [N] | $F_z$ [N] | $M_x$ [N · m] | $M_y$ [N · m] | $M_z$ [N · m] |
|-------------|-----------|-----------|-----------|---------------|---------------|---------------|
| Classique   | 663.9     | −2092.0   | −256.9    | 14956.5       | 5245.7        | −1839.2       |
| Overset     | 642.8     | −2036.5   | −251.8    | 14538.2       | 5088.2        | −1767.2       |
| Différences | 3.2%      | 2.7%      | 2.0%      | 2.8%          | 3.0%          | 3.9%          |

**Tableau 10.1** – Efforts aérodynamiques pour un même réglage optimum (maillages classique et overset).

### 10.3.3 Résultats

Les résultats de l'optimisation du problème (10.3.1) par la méthode EGO multi-fidélité sont présentés ci-après. Les convergences des évaluations de la performance et de la fonction de mérite par niveau de fidélité en fonction des itérations sont tout d'abord données par les Figures 10.7(a) et 10.7(b). Sur ces deux Figures, les points non évaluables ne sont pas tracés, mais l'indice des itérations les prend en compte. Les observations initiales des plans d'expériences sont également reportées sur la Figure 10.7(a). Comme attendu, les évaluations



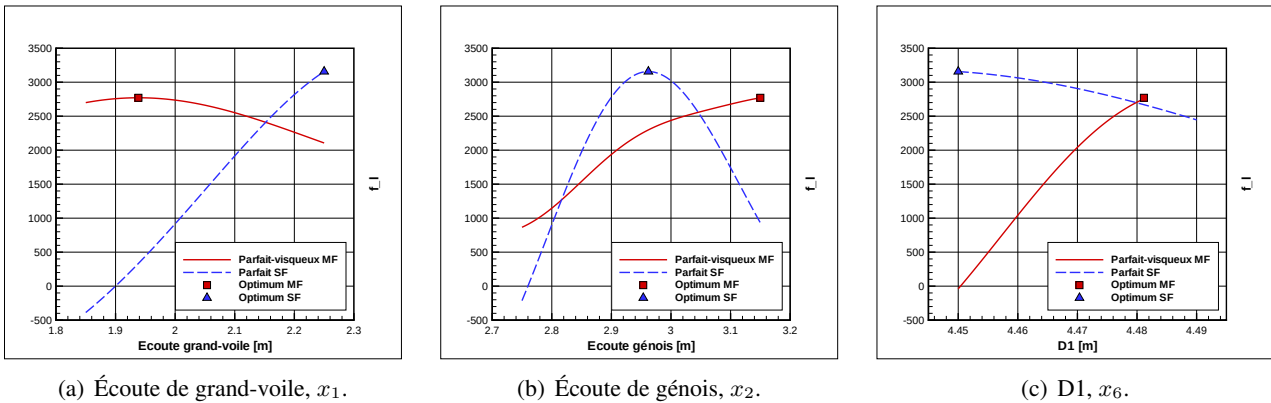
**Figure 10.7** – Convergences de la performance et de la fonction de mérite par niveau de fidélité.

de la performance données par les plans d'expériences sont dispersées (symboles « LHS<sub>1</sub> » et « LHS<sub>2</sub> »). Les premières itérations de l'EGO se concentrent sur le modèle de plus haute fidélité ( $l = 2$ ) et la fonction de mérite  $AEI_2$  est proche de  $10^{-1}$  durant  $\approx 20$  itérations. On note également que de meilleures évaluations de la performance  $f_2(\mathbf{x})$ , comparées au plan LHS, sont observées au cours de ces itérations. A partir de  $\approx 140$  itérations, donc  $\approx 30$  itérations de l'EGO, la fonction de mérite  $AEI_2$  entame une phase de réduction, alors que  $AEI_1$  augmente franchement. Cela signifie que le modèle fluide parfait est responsable de l'incertitude du méta-modèle multi-fidélité dans les zones d'intérêt de la performance, par conséquent la méthode EGO multi-fidélité induit des évaluations du niveau  $l = 1$ . On remarque de plus que ces observations de faible fidélité génèrent de moins bonnes valeurs de  $f_1(\mathbf{x})$  en comparaison de celles obtenues par le plan d'expérience du même niveau.



Ceci est cohérent, puisque ces nouvelles observations sont placées dans des zones d'intérêt du modèle multi-fidélité, et non pas dans celles du niveau  $l = 1$ . Après  $\approx 220$  itérations, on ne peut pas dire qu'une convergence des évaluations de la performance ou de la fonction de mérite soit clairement visible. Cela est parfois typique avec la méthode EGO, où l'incertitude inhérente aux méta-modèles implique une continuité de l'exploration, et rend la définition d'un critère d'arrêt complexe (Chaudhuri and Haftka, 2013). Les nombres de points non évaluables sont respectivement de 58 et 37 pour les niveaux  $l = 1$  et  $l = 2$  à la fin de l'optimisation. On retient de ces courbes de convergences que les meilleures observations de la performance sont déjà déterminées pendant les  $\approx 20$  premières itérations de l'EGO. Le modèle de faible fidélité est quant à lui très utile pour l'exploration des domaines non évaluables.

La Figure 10.8 présente les variations de la performance prédites par le méta-modèle multi-fidélité (légendes « MF ») et par le méta-modèle simple fidélité de niveau  $l = 1$  (légendes « SF ») autour de l'optimum  $\mathbf{x}_{\text{opt}_l} = \arg \max_{\mathbf{x} \in \Omega} f_l(\mathbf{x})$ , lorsque l'on fait varier de manière indépendante les paramètres  $x_1$ ,  $x_2$  ou  $x_6$  de  $\mathbf{x}_{\text{opt}_l}$ . L'objectif est de montrer ici les effets de deux réglages principaux ( $x_1$  et  $x_2$ ) et d'un réglage secondaire ( $x_6$ ) autour des réglages optimaux  $\mathbf{x}_{\text{opt}_1} = [2.25, 2.96, 16.47, 2.40, 0.30, 4.45]^T$  m et  $\mathbf{x}_{\text{opt}_2} = [1.94, 3.15, 16.36, 2.4, 0.13, 4.48]^T$  m. Le réglage optimum qui pourrait être donné par un méta-modèle simple fidélité, construit sur les observations de niveau  $l = 2$  (fluide visqueux), n'est ici pas présenté. En effet, ce résultat n'a pas de signification car le nombre d'évaluations effectuées par le modèle fluide visqueux est trop faible. La valeur de la performance donnée par ce modèle simple fidélité serait localement correcte dans les zones observées, mais ce n'est pas le cas pour l'ensemble du domaine de variabilité des réglages, qui de plus comporte des zones non évaluables, ici principalement détectées dans la formulation multi-fidélité par le modèle  $l = 1$ . On peut tout d'abord remarquer que les prédictions de la performance données par les deux modèles ont des tendances très différentes. Les variations de la performance sont opposées lorsque l'on fait varier indépendamment



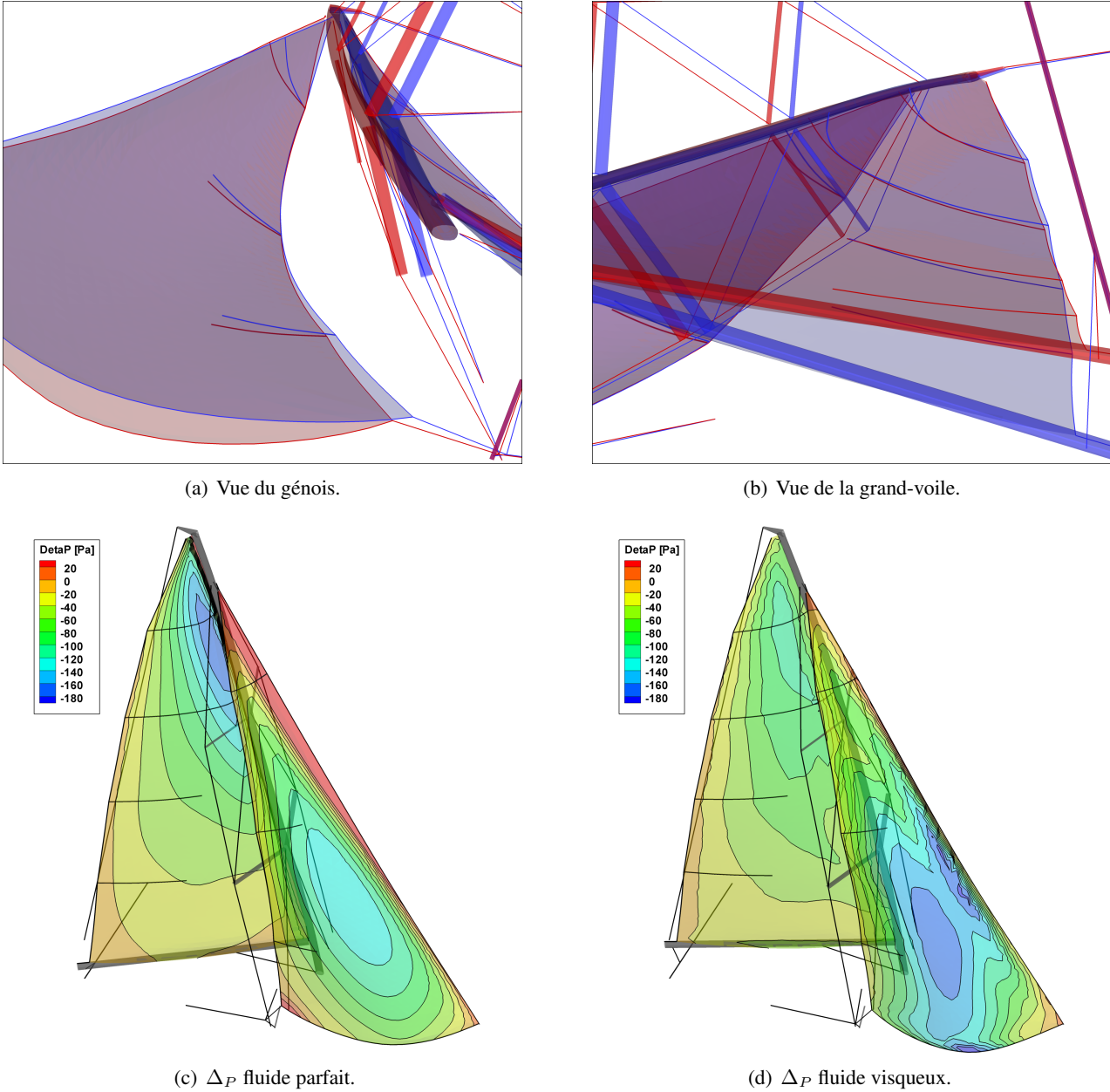
**Figure 10.8** – Effets de l'écoute de grand-voile, de l'écoute de génois et du D1 sur la performance.

$x_1$  ou  $x_6$  (voir Figures 10.8(a) et 10.8(c)), et contrairement au modèle multi-fidélité, la performance présente un extremum pour  $x_2$  dans le cas du modèle simple fidélité fluide parfait (Figure 10.8(b)). Il est important de rappeler que  $\mathbf{x}_{\text{opt}_1}$  et  $\mathbf{x}_{\text{opt}_2}$  sont ici différents. Par conséquent, les variations de la performance lorsqu'un  $x_i$  de  $\mathbf{x}_{\text{opt}_l}$  varie, ne sont pas directement comparables entre les modèles  $l = 1$  et  $l = 2$ .

On peut néanmoins comprendre que les disparités proviennent principalement des différences entre les modèles fluide parfait et visqueux. Plus particulièrement, elles viennent du fait que l'écoulement peut être détaché sur les voiles dans le cas du fluide visqueux. Dans le cas de l'écoute de grand-voile, le fait de choquer ouvre la chute et augmente alors le vrillage de la voile. L'augmentation du vrillage permet de réduire le moment de chavirement, mais cela peut aussi réduire l'effort propulsif si la voile est trop vrillée, car le haut de la grand-voile ne porte presque plus. C'est ce que l'on voit ici sur la Figure 10.8(a) avec le modèle multi-fidélité où  $f_2(\mathbf{x})$  diminue si la voile est trop vrillée, alors que  $f_1(\mathbf{x})$  continue d'augmenter car l'angle d'incidence de l'écoulement sur la grand-voile reste positif jusqu'à la limite maximale imposée à 2.25 m. Si cette limite sur  $x_1$  est augmentée, le comportement de  $f_1(\mathbf{x})$  ressemblerait à celui de  $f_2(\mathbf{x})$ , avec un maximum où  $f_1(\mathbf{x}) > f_2(\mathbf{x})$  et pour une longueur d'écoute de grand-voile plus grande. Pour l'écoute de génois, les effets sont assez similaires à l'écoute de grand-voile, avec un impact plus prononcé sur le creux de la voile, car son point d'écoute n'est pas tenu par une bôme, comme c'est le cas sur la grand-voile. L'extremum de la performance, sur la Figure 10.8(b), donné par le modèle fluide parfait provient des réglages issus du vecteur  $\mathbf{x}_{\text{opt}_1}$ . En effet, dans ce cas la grand-voile est très ouverte et peu vrillée car le pataras est relâché et le chariot de grand-voile sous le vent (visible par la suite sur les Figures 10.9 et 10.11). La propulsion du voilier vient donc principalement de la zone haute de la grand-voile et peu du génois, ce qui n'est normalement pas le cas en navigation réelle. Ainsi, si le génois est bordé, le moment de chavirement augmente et  $f_1(\mathbf{x})$  est pénalisée. À l'inverse, si il est trop choqué, l'effort de propulsion diminue. Dans le cas de l'optimum  $\mathbf{x}_{\text{opt}_2}$ , la grand-voile est plus vrillée et l'effort propulsif provient principalement du génois (réglages plus réalistes). Le fait de trop border le génois entraîne une pénalisation de  $f_2(\mathbf{x})$  par le moment de chavirement, mais également par l'effort propulsif, si la zone de décollement fluide du génois augmente. De même que précédemment pour  $x_1$ , si la limite maximale de  $x_2$  imposée à 3.15 m est augmentée, alors un maximum de  $f_2(\mathbf{x})$  est attendu. Enfin, les D1 participent à la tenue latérale du mât et ils permettent de contrôler son cintre ; ce qui modifie entre autres, l'ouverture de la chute de la grand-voile ainsi que son creux. Plus les D1 sont tendus, plus le creux augmente et se rapproche du bord d'attaque. Les variations de  $f_1(\mathbf{x})$  sont faibles vis-à-vis de  $x_6$  car les réglages de  $\mathbf{x}_{\text{opt}_1}$  conduisent à un mât très peu cintré. Dans le cas du modèle multi-fidélité, les D1 ont un effet très prononcé sur la performance  $f_2(\mathbf{x})$ . Lorsque les D1 sont repris, le cintre du mât diminue et la grand-voile perd son vrillage. Cela impacte directement la performance car le moment de chavirement augmente alors fortement.

La Figure 10.9 présente de manière qualitative les résultats des réglages optimaux donnés par les méta-modèles simple et multi-fidélité, qui sont calculés sur le modèle numérique du Figaro avec leurs modèles fluide-structure respectifs. Des visualisations du génois et de la grand-voile sont illustrées par les Figures 10.9(a) et 10.9(b), où les voiles bleues correspondent au modèle  $l = 1$  et celles en rouges proviennent du modèle  $l = 2$ . Les différences de pression calculées sur les voiles sont également reportées dans le cas du fluide parfait (Figure 10.9(c)) et du fluide visqueux (Figure 10.9(d)). On peut alors confirmer par l'observation des deux visualisations des voiles, que le mât est bien cintré avec le modèle de haute fidélité, ce qui n'est pas le cas avec le modèle de simple fidélité fluide parfait. On peut même remarquer que le mât fait une forme de « S » sur ce modèle, ce qui ne correspond pas à un réglage réaliste du gréement. La partie haute de la grand-voile présente un creux très prononcé et avancé pour  $l = 1$ , en plus de ne pas être vrillée. À l'inverse, celle donnée par  $l = 2$  est

bien plus plate et vrillée. Le génois du modèle  $l = 2$  est bien plus creux que celui du modèle  $l = 1$ , qui de plus « prend à contre » dans la partie haute de son guidant. Ces observations sont confirmées par les Figures

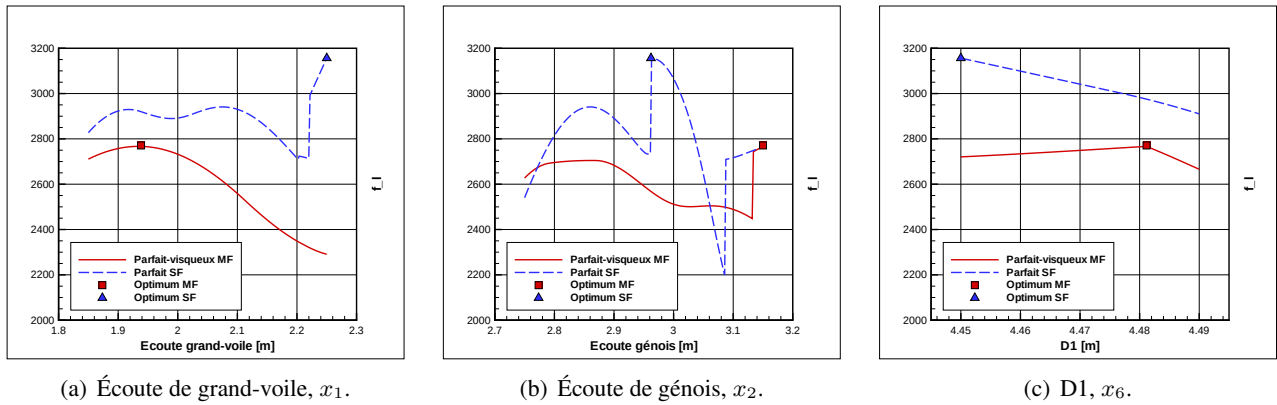


**Figure 10.9** – Réglages optimaux donnés par les méta-modèles  $l = 1$  (voiles bleues) et  $l = 2$  (voiles rouges). Les différences de pression sur les voiles, calculées par les deux modèles fluide, sont également présentées.

10.9(c) et 10.9(d) qui reportent les différences de pression sur les voiles. On voit que la grand-voile du modèle fluide parfait travaille principalement dans sa partie haute (forts  $\Delta_P$  négatifs), alors que le génois apporte peu d'effort propulsif et génère des  $\Delta_P$  positifs sur les zones du guidant qui sont à contre. Le réglage optimum du modèle multi-fidélité est également bien plus réaliste en terme de  $\Delta_P$  (voir Figure 10.9(d)). La grand-

voile est suffisamment vrillée pour ne pas générer un grand moment de chavirement qui est pénalisant pour la performance du voilier. Le génois travaille principalement dans sa partie basse (faible  $M_{x_2}$ ) et apporte un effort propulsif important (forts  $\Delta_P$  négatifs). On peut également remarquer la prise en compte, par le modèle fluide visqueux, des tourbillons créés par sa bordure, où les valeurs de  $\Delta_P$  sont négatives.

Les Figures à suivre ont toujours pour objectif d'illustrer les effets des réglages des voiles sur la performance, dans les cas simple et multi-fidélité, avec cette fois l'ajout de leurs dépendances, dans le cas où l'on souhaite maintenir une performance optimale. Plus concrètement, les réglages  $x_1$ ,  $x_2$  et  $x_6$  varient toujours de manière indépendante sur leurs domaines respectifs, et les cinq autres réglages sont déterminés afin de maintenir une performance optimale pour chaque valeur discrétisée de  $x_i$  par :  $\mathbf{x}_{\text{opt}_{l_i}} = \arg \max_{\mathbf{x} \in \Omega \setminus \Omega_i} f_l([x_i, \mathbf{x}])$ . Une fois résolu, ces problèmes d'optimisation à cinq paramètres permettent d'illustrer l'influence d'un réglage sur une performance maintenue optimale, mais également l'influence des variations conjointes des cinq autres réglages vis-à-vis du réglage d'intérêt  $x_i$ . La Figure 10.10 présente les variations de  $f_l(\mathbf{x})$ , maintenue optimale, pour les deux niveaux de fidélité lorsque  $x_1$ ,  $x_2$  ou  $x_6$  varient indépendamment. On note que pour les trois



**Figure 10.10** – Effets de l'écoute de grand-voile, de l'écoute de génois et du D1 sur la performance maintenue optimale.

réglages,  $f_1(\mathbf{x})$  et  $f_2(\mathbf{x})$  varient dans des intervalles bien plus faibles que ceux présentés par la Figure 10.8, ce qui est assez logique puisque la performance est ici maintenue optimale. Les tendances sont maintenant plus en accord entre les modèles simple fidélité et multi-fidélité. On remarque également que  $f_1(\mathbf{x})$  présente de fortes discontinuités, en particulier pour l'écoute de génois sur la Figure 10.10(b). Ces discontinuités proviennent bien de changement des localisations de l'optimum dans  $\Omega$ , et non pas d'une sous discrétisation des tracés qui contiennent ici une répartition uniforme de 200 points le long de l'axe des abscisses. L'analyse de ces résultats est complexe car pour chaque valeur d'un réglage d'intérêt  $x_i$  les cinq autres réglages varient. On peut cependant confirmer que les D1 ont une faible influence sur la performance, et cela pour les deux modèles simple et multi-fidélité, comme montré par la Figure 10.10(c). L'écoute de grand-voile a peu d'influence sur  $f_2(\mathbf{x})$  localement autour de l'optimum, mais ce n'est plus le cas lorsque l'écoute est choquée. L'écoute de génois apparaît comme le paramètre le plus sensible localement autour de l'optimum pour les deux modèles  $l = 1$  et  $l = 2$ .

La Figure 10.11 donne accès aux variations conjointes des cinq réglages, pour maintenir une performance

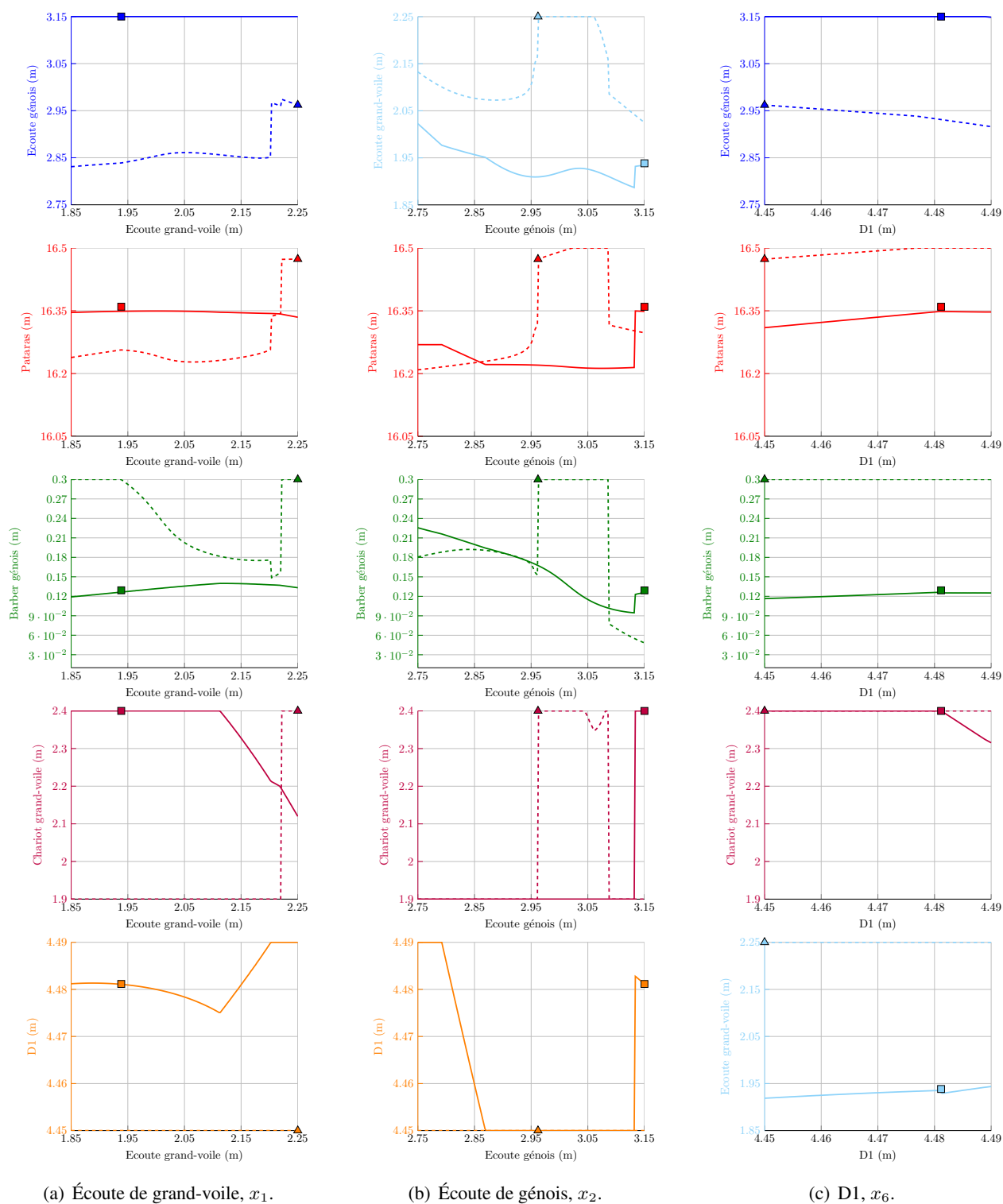
optimale, lorsqu'un réglage,  $x_1$  sur la Figure 10.11(a),  $x_2$  sur la Figure 10.11(b) ou  $x_6$  sur la Figure 10.11(c) varie. Pour chacune de ces Figures, cinq graphiques verticaux présentent les variations des cinq autres réglages (axes des ordonnées en couleur) que celui d'intérêt (axes des abscisses en noir). Les variations sont tracées pour le modèle simple fidélité fluide parfait (courbes discontinues) et pour le modèle multi-fidélité (courbes continues), lorsque la performance est maintenue optimale. Les localisations des réglages optimaux  $\mathbf{x}_{\text{opt}_1}$  (triangles) et  $\mathbf{x}_{\text{opt}_2}$  (carrés) sont également données sur cette Figure 10.11.

On peut alors faire le lien, avec cette Figure 10.11, entre les variations des réglages de  $\mathbf{x}_{\text{opt}_l}$  et celles de  $f_l(\mathbf{x})$  localement autour des optimums, précédemment données par la Figure 10.10. Ainsi, si l'écoute de grand-voile varie entre 1.85 et 2.25 m, on peut observer sur la Figure 10.11(a) que cela impacte principalement les réglages  $x_4$ ,  $x_5$  et  $x_6$  de  $\mathbf{x}_{\text{opt}_2}$ . Ces variations sont faibles localement autour de  $\mathbf{x}_{\text{opt}_2}$ , mais on peut remarquer que lorsque l'écoute de grand-voile est choquée, les D1 sont légèrement repris jusqu'à  $x_1 \approx 2.11$  m. Si  $x_1$  continue d'être choquée, les D1 sont franchement relâchés en même temps que le chariot de grand-voile est progressivement remonté au vent pour faire porter le bas de la grand-voile (voile fortement vrillée). Cet effet est ici nécessaire pour réduire la puissance apportée par la grand-voile en ouvrant sa chute et ainsi limiter le moment de chavirement du voilier. Les variations dans  $\mathbf{x}_{\text{opt}_1}$ , lorsque  $x_1$  varie, sont quant à elles bien plus prononcées. Si l'écoute de grand-voile est légèrement bordée en partant de son réglage optimum (complètement choquée), on observe sur la Figure 10.11(a) que  $x_2$ ,  $x_3$ ,  $x_4$  et  $x_5$  varient alors très fortement. En particulier, le pataras  $x_3$  est franchement repris afin d'ouvrir la grand-voile, toujours dans le but de limiter le moment de chavirement du voilier.

Dans le cas de l'écoute de génois (voir Figure 10.11(b)),  $\mathbf{x}_{\text{opt}_1}$  est situé au niveau de la frontière entre deux points de fonctionnement des voiles pour les réglages  $x_1$  (axe bleu),  $x_3$  (axe rouge),  $x_4$  (axe vert) et  $x_5$  (axe violet). Il n'y a que  $x_6$  (D1, axe orange) qui ne présente pas de variations autour de l'optimum  $\mathbf{x}_{\text{opt}_1}$ . Comme déjà présenté, cet optimum  $\mathbf{x}_{\text{opt}_1}$  correspond à un réglage non réaliste où la grand-voile est très ouverte mais peu vrillée. Le deuxième point de fonctionnement correspond à un optimum local de  $f_1(\mathbf{x})$  où l'écoute de génois vaut  $x_2 \approx 2.86$  m (voir Figure 10.10(b)). Dans ce cas, comme le montre la Figure 10.11(b), le pataras est franchement repris, l'écoute et chariot de grand-voile aussi, et le barber de génois également. Les D1 sont à la limite d'être relâchés. On se rapproche alors d'un réglage réaliste comme  $\mathbf{x}_{\text{opt}_2}$ , donné par le modèle multi-fidélité fluide parfait et visqueux. On peut faire le même type d'analyse avec les réponses données par le modèle  $l = 2$ , qui ici sont cependant moins marquées en terme de variations.

Enfin, la faible variation de la performance maintenue optimale lorsque les D1 varient est confirmée par la Figure 10.11(c). Ceci est vrai pour les deux modèles simple et multi-fidélité, malgré le fait que l'on puisse noter une rupture de pente au niveau du réglage optimum sur tous les réglages, sauf pour  $x_2$ , dans la réponse donnée par le modèle de plus haute fidélité.

D'un point de vue méthode, on peut apprécier le fait que la formulation multi-fidélité a la propriété de ne pas voir la prédiction du méta-modèle de plus haute fidélité dégradée par des prédictions non corrélées de niveaux de fidélités inférieures. On note également l'intérêt d'utiliser une méthode d'optimisation par méta-modèles qui, en plus de ses qualités d'optimisation, permet aussi de faire des études complémentaires sur les variables à optimiser vis-à-vis de la fonction objective, par l'intermédiaire des méta-modèles.



**Figure 10.11** – Effets de l’écoute de grand-voile, de l’écoute de génois et du D1 sur les cinq réglages, pour une performance maintenue optimale. Les courbes discontinues et continues correspondent respectivement aux réglages donnés par les méta-modèles simple et multi-fidélité. Les symboles  $\Delta$  et  $\square$  sont les optimums respectifs de  $x_{opt1}$  et  $x_{opt2}$ .

## 10.4 Conclusion

L'utilisation de la méthode EGO formulée en multi-fidélité est illustrée dans le Chapitre 10. La problématique de l'optimisation des réglages d'un gréement de voilier afin d'en maximiser la performance est également présentée. L'intérêt de l'approche EGO multi-fidélité, dans le but de réduire le coût numérique de l'optimisation, est démontré sur des cas numériques concrets incluant des calculs d'interaction fluide-structure sur des voiles. La méthode est appliquée en faible dimension avec cinq niveaux de fidélité où les différences de modélisation se situent dans les modèles fluide et structure. Elle est également appliquée sur un cas en haute dimension ( $d = 6$ ) avec deux niveaux de fidélité (fluides parfait et visqueux). L'intérêt d'une approche d'optimisation par méta-modèles est présenté par cette dernière étude. En effet, contrairement aux méthodes d'optimisation sans méta-modèles, l'EGO donne accès à des études complémentaires qui permettent, comme par exemple dans cette application, d'analyser les influences des paramètres d'optimisation vis-à-vis de la performance.

Une publication récemment soumise est incluse dans ce Chapitre 10. Des problèmes analytiques présentant des propriétés variées sont utilisés afin de valider la méthode EGO multi-fidélité. Des comparaisons avec l'approche classique où les méta-modèles multi-fidélité sont construits sur la base d'observations imbriquées sont également effectuées. La meilleure efficacité de la méthode EGO multi-fidélité utilisant des observations non imbriquées est illustrée.

Pour conclure, ces travaux de thèse permettent de valider la méthode EGO multi-fidélité sur des problèmes analytiques, mais également sur des cas physiques réels et complexes. Les perspectives d'améliorations des applications sont multiples. En effet, les optimisations des réglages de gréements seraient plus réalistes si l'équilibre aéro-hydrodynamique du voilier était résolu pour chaque jeu de réglages testé. Cela permet de prendre en compte l'influence des réglages sur la vitesse du voilier et donc sur la vitesse et sur l'angle du vent apparent. Les attitudes du voilier pourraient aussi être libérées afin de se rapprocher de la réalité physique. Dans ce cas, les calculs IFS sont plus complexes, mais le problème d'optimisation est simplifié car il n'est plus nécessaire d'utiliser une pénalisation dans la définition de la fonction performance. Par exemple, si l'objectif est de maximiser la vitesse du voilier et que la gîte est libre, alors les réglages conduisant à un important moment de chavirement seront naturellement pénalisés par une gîte excessive du voilier. La méthode de maillage du type « overset » est bien adaptée pour ce genre d'application. En effet, la gîte libre du voilier peut être directement imposée aux mailles « overset », ce qui évite d'avoir des déformations provenant de la rotation en gîte du voilier sur le maillages principal. Les conditions limites de vent réel peuvent aussi être appliquées au niveau des frontières du maillage principal, et en résolvant l'équilibre aéro-hydrodynamique, la vitesse du voilier peut alors être appliquée au maillage « overset » et principal. Cette méthode permet d'éviter la forte augmentation du coût de calcul lorsque les conditions limites, appliquées sur les frontières du maillage principal, utilisent le vent apparent.







# Conclusion générale et perspectives

La propulsion vélique principale ou auxiliaire des navires de transport de fret ou de passagers est aujourd'hui en plein essor. La conception de ces navires et des voiliers de compétitions a de plus en plus recours aux simulations numériques afin de limiter les coûts importants des essais expérimentaux. De nombreux travaux scientifiques ont été effectués numériquement et expérimentalement dans le but de mieux comprendre la physique des voiliers. Néanmoins, l'étude de la performance des voiliers reste aujourd'hui dans un contexte généralement descriptif, par le manque de travaux scientifiques portant sur des méthodes d'optimisation adaptées à l'optimisation de cette performance. Les coûts associés à ces types d'optimisation sont importants car la modélisation de la physique du voilier est complexe. En effet, l'évaluation de la performance d'un voilier fait intervenir un couplage aéro-hydrodynamique et des problèmes d'interaction fluide-structure. De plus, la fonction performance à optimiser est généralement dépendante d'un grand nombre de paramètres. Des méthodes d'optimisation avancées doivent alors être développées afin de pouvoir résoudre ce genre de problèmes d'optimisation.

Cette thèse s'inscrit dans la continuité des travaux de recherche, portant sur le domaine de la voile, menés au sein de l'institut de recherche de l'École-Navale. Le cœur de la thèse concerne le développement et l'utilisation de méthodes d'optimisation avancées, pour résoudre la problématique de l'optimisation de la performance des voiliers. Le montage collaboratif de la thèse entre l'IRENav, le LIMSI-CNRS et l'INRIA permet d'apporter les compétences scientifiques respectives de ces laboratoires, nécessaires aux développements mathématiques des méthodes d'optimisation et des applications portant sur la performance du voilier. Des stratégies et des méthodes d'optimisations avancées sont ainsi développées et validées sur des fonctions analytiques et sur différents cas d'applications réelles, complexes et originales modélisant la performance des voiliers.

Dans la Partie I, le contexte général et spécifique du voilier est introduit. Son principe de fonctionnement, reposant sur l'équilibre aéro-hydrodynamique entre l'air sur les voiles et l'eau sur la carène et les appendices, est présenté. Les notions de la mécanique des fluides et d'interaction fluide-structure, décrivant la physique du voilier, sont également abordées dans cette Partie I. Le concept de la performance du voilier qui peut être estimée avec des programmes de prédiction de vitesse reposant sur des modèles numériques et des résultats d'essais est largement commenté. L'optimisation de cette performance nécessite une bonne compréhension de la physique du voilier et de sa modélisation numérique. La présentation des essais expérimentaux, indispensables à la validation des modèles développés et dont certaines mesures peuvent directement être exploitées

dans les programmes de prédiction de vitesse, clôt cette Partie I.

La Partie II présente les méthodes d'optimisation, reposant sur l'utilisation de méta-modèles, développées et appliquées au cours de cette thèse. L'emploi de méta-modèles construits à partir de processus Gaussiens est détaillé dans le Chapitre 4, en particulier pour les étapes de construction des méta-modèles, mais également pour ce qui est de leurs applications en terme d'analyse de sensibilité, de classification et d'optimisation globale. La méthode itérative d'optimisation par méta-modèles Gaussiens, amplement utilisée dans cette thèse, est ainsi pleinement décrite.

Le Chapitre 5 porte sur la thématique de l'optimisation globale et sous contraintes, effectuée à partir de méta-modèles. Le problème des points non évaluables, qui peuvent apparaître lors de la résolution d'un problème d'optimisation sur une application réelle, est tout d'abord introduit dans ce Chapitre 5. La prise en compte de contraintes d'inégalité approchées par des méta-modèles Gaussiens est ensuite décrite. Enfin, l'introduction d'un modèle de classification binaire, permettant la gestion de tout type de contraintes et donc en particulier des points non évaluables, est proposée dans ce Chapitre 5.

Le dernier Chapitre 6, de la Partie II, concerne l'utilisation de méta-modèles Gaussiens ayant des niveaux de fidélité hiérarchisés. Les méthodes de constructions de ces méta-modèles sont présentées et une nouvelle formulation de l'optimisation par méta-modèles multi-fidélité est proposée. Cette méthode permet d'introduire le choix du niveau de fidélité utilisé dans le processus d'optimisation itératif, à partir d'une estimation de la réduction d'incertitude dans les zones d'intérêts du méta-modèle de plus haute fidélité. Elle permet de plus l'utilisation de méta-modèles multi-fidélité construits sur la base d'observations non imbriquées.

La dernière Partie III regroupe les différentes contributions de cette thèse. Les contributions sont présentées sous la forme d'articles publiés ou soumis dans des revues scientifiques. Le Chapitre 7 reprend une application originale de la méthode d'optimisation par méta-modèles Gaussiens (Sacher et al., 2017c). Elle concerne l'optimisation de la performance d'un modèle réduit de grand-voile d'Imoca pour laquelle deux réglages sont optimisés, afin de maximiser une fonction performance mesurée expérimentalement, dans la soufflerie de l'université d'Auckland. Cette application permet de montrer la nécessité d'insérer le bruit expérimental dans la construction des méta-modèles Gaussiens, pour mener à bien l'optimisation. Cette même optimisation est résolue numériquement. Elle met en œuvre un modèle couplé fluide-structure pour le calcul de la même fonction performance. Des comparaisons expérimentales et numériques sont effectuées, afin de valider les calculs d'interaction fluide-structure. Une analyse de sensibilité est aussi menée sur le modèle numérique, dans le but de quantifier les variations de la performance provenant des incertitudes expérimentales, autour du réglage optimal. Cela permet de conforter les résultats numériques vis à vis de l'expérience, mais également d'étudier la robustesse du réglage optimal de la grand-voile.

Dans le Chapitre 8, la méthode de classification introduite dans le processus d'optimisation par méta-modèles Gaussiens est validée sur des applications comprenant des points non évaluables (Sacher et al., 2018b). Les applications correspondent principalement à des problèmes d'optimisation analytiques présentant des propriétés et niveaux de difficultés variés. L'efficacité de la méthode est comparée à une méthode d'optimisation

reposant sur la stratégie d'évolution. Elle est également comparée à une autre approche originale, reposant sur la classification pour la gestion des contraintes, mais comprenant plusieurs différences avec la méthode proposée dans cette thèse. La méthode proposée ici s'est révélée plus efficace dans toutes ces applications analytiques, excepté sur un problème complexe sur-contraint. Un problème d'optimisation concret fait également partie de ce Chapitre 8. Il correspond à l'optimisation sous contraintes de multiples paramètres d'une section bidimensionnelle d'un hydrofoil flexible. Les résultats portent ici sur l'efficacité de la méthode de classification des contraintes et des points non évaluables, comparée à une approche où les contraintes sont approchées par des méta-modèles Gaussiens et à la méthode de stratégie d'évolution. L'intérêt d'une vision binaire de l'ensemble des contraintes par le modèle de classification est ainsi clairement démontré.

Le Chapitre 9 concerne l'optimisation de l'hydrofoil flexible succinctement introduite dans le Chapitre 8. La présentation du problème d'optimisation, nécessitant la résolution de calculs d'interaction fluide-structure, est détaillée et l'analyse des résultats se concentre ici sur les performances hydrodynamiques de l'hydrofoil (Sacher et al., 2018a). Les variables d'optimisation sont de natures géométriques et structurelles, et les contraintes permettent d'assurer un régime d'écoulement non cavitant sur plusieurs points de fonctionnement de l'hydrofoil. L'apport de la flexibilité est en particulier présenté en comparant les performances de l'hydrofoil flexible optimisé, à un hydrofoil purement rigide et optimisé pour les mêmes conditions. En effet, l'hydrofoil flexible génère des efforts de traînée inférieurs à l'hydrofoil rigide, et cela pour une large gamme de vitesses du voilier. Sa meilleure tolérance vis-à-vis du régime cavitant est également montrée.

Le dernier Chapitre 10, de la Partie III, illustre l'utilisation de la méthode d'optimisation par méta-modèles multi-fidélité. La thématique concerne l'optimisation numérique des réglages d'un voilier dans le but d'en maximiser sa performance. L'intérêt d'une formulation multi-fidélité est tout d'abord introduit à partir de résultats antérieurs à cette thèse. Ils correspondent à des calculs d'interaction fluide-structure sur le gréement d'un voilier, pour lesquels deux modèles fluides (parfait et visqueux) sont comparés. La méthode d'optimisation par méta-modèles multi-fidélité est ensuite validée sur un ensemble de problèmes analytiques ayant des propriétés variées. La formulation multi-fidélité construite sur la base d'observations non imbriquées est comparée à la méthode originale où les observations sont imbriquées. L'approche non imbriquée se révèle plus efficace. Un problème d'optimisation de la performance d'un voilier est ensuite résolu en faible dimension, mais avec de multiples niveaux de fidélité. Deux réglages sont optimisés avec cinq modèles numériques présentant des variétés de fidélité en terme de modélisations fluide et structure. Le problème d'optimisation des deux réglages du voilier est également résolu en considérant une formulation simple fidélité avec le modèle de plus haute fidélité. L'apport de la méthode d'optimisation multi-fidélité pour la réduction du coût de l'optimisation est démontré. Une seconde optimisation des réglages du voilier est réalisée dans ce Chapitre 10. Elle est de plus haute dimension, avec six réglages à optimiser, et utilise deux niveaux de fidélité avec des différences de modélisation sur le modèle fluide uniquement : parfait et visqueux. La nécessité d'utiliser un modèle fluide visqueux, permettant la prise en compte des effets de décollements fluide, est illustrée par cette application. En effet, le réglage optimal du gréement et des voiles donné par le méta-modèle simple fidélité, avec le fluide parfait, ne correspond pas à un réglage réaliste. Cela est heureusement corrigé par l'apport de la formulation multi-fidélité qui utilise les deux modèles fluides. Les variations de la performance autour des réglages optimaux (simple et multi-fidélité)

sont également illustrées, de même que les variations conjointes des réglages, lorsque la performance est maintenue optimale.

En conclusion, cette thèse se caractérise par le développement de méthodes d'optimisation avancées permettant la résolution de problèmes physiques complexes, tels que l'optimisation de la performance des voiliers. L'optimisation par méta-modèles Gaussiens est pleinement exploitée dans cette thèse et ses multiples qualités pour optimiser la performance des voiliers ont été exposées. Les propriétés statistiques de ces méta-modèles donnent de l'efficacité au processus d'optimisation itératif, en bénéficiant d'un compromis naturel entre l'exploitation et l'exploration, des domaines de variabilité des paramètres d'optimisation. Les méta-modèles Gaussiens ont également la capacité d'utiliser des observations possiblement bruitées. C'est une faculté essentielle lorsque l'on souhaite optimiser des problèmes physiques réels, tels que la performance des voiliers. La prédiction de la variance donnée par les méta-modèles Gaussiens est aussi une propriété importante, car elle fournit une mesure directe des incertitudes de la prédiction moyenne des méta-modèles. Ces deux prédictions (moyenne et variance) des méta-modèles Gaussiens sont également très utiles à la réalisation d'études complémentaires à l'optimisation, comme par exemple pour faire de l'analyse de sensibilité ou bien pour quantifier les effets conjoints de paramètres d'une optimisation. Enfin, il est important de retenir que ces méthodes d'optimisation par méta-modèles Gaussiens ont de l'intérêt si les coûts de construction des méta-modèles sont inférieurs aux coûts d'évaluation des fonctions de performance à optimiser.

Les perspectives sont quant à elles nombreuses, tant sur les méthodes d'optimisation que sur leurs applications. Pour le développement des méthodes, la thématique de la multi-fidélité présente plusieurs axes de recherche. L'optimisation par méta-modèles multi-fidélité pourrait par exemple utiliser des fonctions de mérite générant plusieurs observations à chaque itération globale du processus itératif. La notion de budget rentrerait alors possiblement en compte dans l'optimisation, afin de définir les niveaux de fidélité et le nombre d'observations à évaluer. Concrètement, à une itération donnée, on peut se demander s'il est plus intéressant, pour un même coût d'évaluation, d'évaluer plusieurs observations de faible fidélité ou bien une seule observation de haute fidélité. Le nombre d'observations dans ce genre d'approche multi-fidélité pourra néanmoins augmenter rapidement, ce qui aura l'effet néfaste d'augmenter les coûts numériques de construction des méta-modèles. Il serait alors intéressant de développer des méthodes de « sparsification » afin de contrôler cette augmentation du coût. Une approche qui serait basée sur la même idée que celle proposée dans cette thèse, pour l'optimisation par méta-modèles multi-fidélité, serait d'estimer les observations apportant peu de réduction de variance dans les méta-modèles, pour ensuite les oublier. Le processus d'optimisation itératif serait alors mené avec un même nombre fini d'observations, idéalement sélectionnées pour leurs apports dans la construction des méta-modèles. La gestion des points non évaluables dans l'optimisation multi-fidélité pourrait également être améliorée. Il serait intéressant de réussir à formuler mathématiquement un véritable modèle de classification multi-fidélité. Les problèmes d'optimisation multi-objectif sont aujourd'hui très présents dans l'industrie. Pareillement, la formulation multi-fidélité adaptée à ces applications est un important axe de recherche. De même, les problèmes d'optimisation de haute dimension sont toujours d'actualité et ils nécessitent aussi des développements.

D'un point de vue des applications, les modèles numériques sont aujourd'hui suffisamment validés pour les cas d'écoulements « généraux ». Les essais expérimentaux ont cependant encore un grand intérêt dès lors que la physique se complexifie, comme par exemple pour les écoulements à fort décrochement, les régimes de transition laminaire-turbulent, ou encore en présence de phénomènes de cavitation. L'emploi de la méthode d'optimisation par méta-modèles multi-fidélité serait alors justifié pour ces problèmes physiques complexes, où les modèles numériques composeraient les niveaux de faible fidélité, alors que l'expérience donnerait le niveau de plus haute fidélité.



# Références bibliographiques

- I. H. Abbott and A. E. Von Doenhoff. *Theory of wing sections, including a summary of airfoil data*. Courier Corporation, 1959. (page 31)
- M. M. Adankon and M. Cheriet. Model selection for the LS-SVM. Application to handwriting recognition. *Pattern Recognition*, 42(12) :3264 – 3270, 2009. New Frontiers in Handwriting Recognition. (page 95)
- M. M. Adankon, M. Cheriet, and A. Biem. Semisupervised Learning Using Bayesian Interpretation : Application to LS-SVM. *IEEE Transactions on Neural Networks*, 22(4) :513–524, April 2011. (page 95)
- D. M. Allen. The relationship between variable selection and data agumentation and a method for prediction. *Technometrics*, 16(1) :125–127, 1974. (page 94)
- A. M. Alvarez Meza, G. Daza Santacoloma, C. D. Acosta Medina, and G. Castellanos Dominguez. Parameter selection in least squares-support vector machines regression oriented, using generalized cross-validation. *Dyna*, 79(171) :23–30, 2012. (page 95)
- D. V. Arnold and N. Hansen. A (1+1)-CMA-ES for Constrained Optimisation. In T. Soule and J. H. Moore, editors, *GECCO*, pages 297–304, Philadelphia, United States, July 2012. ACM, ACM Press. (page 101)
- N. Aubin. *Fluid-structure interaction on yacht sails*. PhD thesis, Université de Bretagne occidentale - Brest, 2017. (pages vii, 2, 11, 14, 41, 119, et 120)
- N. Aubin, B. Augier, J. Deparday, and M. Sacher. To curl or not to curl : Wind tunnel investigations of spinnaker performance. In *INNOVSAIL International Conference on Innovation in High Performance Sailing Yachts*, pages 139–150, 2017. (page 18)
- C. Audet, A. J. Booker, J. Dennis Jr, P. D. Frank, and D. W. Moore. A surrogate-model-based method for constrained optimization. *AIAA paper*, 4891, 2000. (page 84)
- B. Augier. *Etudes expérimentales de l'interaction fluide-structure sur surface souple : application aux voiles de bateaux*. PhD thesis, Université de Bretagne occidentale - Brest, 2012. (pages 2 et 42)
- B. Augier, P. Bot, F. Hauville, and M. Durand. Experimental validation of unsteady models for fluid structure interaction : Application to yacht sails and rigs. *Journal of Wind Engineering and Industrial Aerodynamics*, 101 :53 – 66, 2012. (page 35)



- B. Augier, P. Bot, F. Hauville, and M. Durand. Dynamic behaviour of a flexible yacht sail plan. *Ocean Engineering*, 66 :32 – 43, 2013. (page 35)
- F. Aurenhammer. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, 23(3) :345–405, 1991. (page 147)
- N. Ayat, M. Cheriet, and C. Suen. Automatic model selection for the optimization of SVM kernels. *Pattern Recognition*, 38(10) :1733 – 1745, 2005. (page 95)
- S. Barnett. *Matrix methods for engineers and scientists*. McGraw-Hill, 1979. (page 61)
- A. Basudhar, C. Dribusch, S. Lacaze, and S. Missoum. Constrained efficient global optimization with support vector machines. *Structural and Multidisciplinary Optimization*, 46(2) :201–221, 2012. (pages 81, 85, 98, et 170)
- G. Baudat and F. Anouar. Feature vector selection and projection using kernels. *Neurocomputing*, 55(1–2) :21 – 38, 2003. Support Vector Machines. (page 94)
- L. Bonfiglio, P. Perdikaris, S. Brizzolara, and G. Karniadakis. Multi-fidelity optimization of super-cavitating hydrofoils. *Computer Methods in Applied Mechanics and Engineering*, 332 :63 – 85, 2018. (pages 19 et 110)
- N. Bose. *Hydrofoils : design of a wind propelled flying trimaran*. PhD thesis, University of Glasgow, 1982. (page 171)
- P. Bot, I. M. Viola, R. G. Flay, and J.-S. Brett. Wind-tunnel pressure measurements on model-scale rigid downwind sails. *Ocean Engineering*, 90 :84–92, 2014. (page 1)
- D. Büche, N. N. Schraudolph, and P. Koumoutsakos. Accelerating evolutionary algorithms with gaussian process fitness function models. *IEEE Tran. on Systems, Man, and Cybernetics - Part C : Applications and Reviews*, 35(2) :183–194, 2005. (page 3)
- G. C. Cawley. Leave-one-out cross-validation based model selection criteria for weighted LS-SVMs. In *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pages 1661–1668. IEEE, 2006. (pages 94 et 95)
- G. C. Cawley and N. L. Talbot. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recognition*, 36(11) :2585 – 2592, 2003. (pages 95 et 96)
- G. C. Cawley and N. L. Talbot. Fast exact leave-one-out cross-validation of sparse least-squares support vector machines. *Neural Networks*, 17(10) :1467 – 1475, 2004. (page 94)
- G. C. Cawley and N. L. Talbot. Preventing over-fitting during model selection via bayesian regularisation of the hyper-parameters. *Journal of Machine Learning Research*, 8(Apr) :841–861, 2007. (page 96)

- G. C. Cawley and N. L. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11(Jul) :2079–2107, 2010. (page 96)
- V. Chapin, N. D. Carlan, and P. Heppel. A multidisciplinary computational framework for sailing yacht rig and sails design through viscous FSI. In *The 20th Chesapeake Sailing Yacht Symposium*, pages 1–17, Annapolis, USA, 2011. (page 18)
- T. Charvet, F. Hauville, and S. Huberson. Numerical simulation of the flow over sails in real sailing conditions. *Journal of Wind Engineering and Industrial Aerodynamics*, 63(1-3) :111–129, 1996. (pages 17 et 35)
- A. Chaudhuri and R. T. Haftka. A stopping criterion for surrogate based optimization using EGO. In *10th world congress on structural and multidisciplinary optimization*, pages 20–24, 2013. (page 220)
- C. P. Chen. A rapid supervised learning neural network for function interpolation and approximation. *IEEE Transactions on Neural Networks*, 7(5) :1220–1230, 1996. (page 56)
- A. Cimarelli, A. Leonforte, and D. Angeli. Direct numerical simulation of the flow around a rectangular cylinder at a moderately high reynolds number. *Journal of Wind Engineering and Industrial Aerodynamics*, 174 :39 – 49, 2018. (page 34)
- D. D. Cox and S. John. A statistical method for global optimization. In *[Proceedings] 1992 IEEE International Conference on Systems, Man, and Cybernetics*, pages 1241–1246 vol.2, Oct 1992. (page 76)
- N. Cressie. The origins of kriging. *Mathematical Geology*, 22(3) :239–252, Apr 1990. (page 60)
- T. Crestaux, O. Le Maître, and J.-M. Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7) :1161 – 1172, 2009. Special Issue on Sensitivity Analysis. (page 69)
- C. Darwin. *The Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life*. John Murray, 1873. (page 52)
- A. Day. Sail optimisation for maximal speed. *Journal of wind engineering and industrial aerodynamics*, 63 (1-3) :131–154, 1996. (page 17)
- J. de Baar, S. Roberts, R. Dwight, and B. Mallol. Uncertainty quantification for a sailing yacht hull, using multi-fidelity kriging. *Computers & Fluids*, 123 :185–201, 2015. (page 18)
- C. A. De Moura and C. S. Kubrusly. The Courant-Friedrichs-Lewy (CFL) Condition. *Commun. Pure Appl. Math*, 10(2) :363–371, 2013. (page 34)
- F. J. DeBord, J. Reichel, B. Rosen, and C. Fassardi. Design optimization for the international America’s Cup class. In *Transactions of the 2002 SNAME Annual Meeting*, 2002. (page 17)
- G. Deng, A. Leroyer, E. Guilmineau, P. Queutey, M. Visonneau, J. Wackers, and A. del Toro Llorens. Verification and validation of resistance and propulsion computation. In *Proceedings of Tokyo 2015 Workshop on CFD in Ship Hydrodynamics*, 2015. (page 217)

- J. Deparday. *Etudes expérimentales de l'Interaction fluide-structure sur les voiles de bateaux au portant*. PhD thesis, Université de Bretagne occidentale - Brest, 2016. (pages [vii](#), [2](#), [41](#), et [42](#))
- J. Deparday, P. Bot, F. Hauville, B. Augier, and M. Rabaud. Full-scale flying shape measurement of offwind yacht sails with photogrammetry. *Ocean Engineering*, 127 :135 – 143, 2016. (page [42](#))
- R. Dony et al. Karhunen-loeve transform. *The transform and data compression handbook*, 1 :1–34, 2001. (page [59](#))
- T. Doyle, M. Gerritsen, and G. Iaccarino. Towards sail-shape optimization of a modern clipper ship. *Center for Turbulence Research, Annual Research Briefs*, 2002 :215–224, 2002. (page [18](#))
- O. Dubrule. Cross validation of kriging in a unique neighborhood. *Journal of the International Association for Mathematical Geology*, 15(6) :687–699, Dec 1983. (page [64](#))
- S. A. Dudani. The Distance-Weighted k-Nearest-Neighbor Rule. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(4) :325–327, April 1976. (page [71](#))
- M. Durand. *Interaction fluide-structure souple et légère, application aux voiliers*. PhD thesis, Ecole Centrale de Nantes (ECN), 2012. (pages [2](#), [30](#), [35](#), et [36](#))
- M. Durand, A. Leroyer, C. Lothodé, F. Hauville, M. Visonneau, R. Floch, and L. Guillaume. FSI investigation on stability of downwind sails with an automatic dynamic trimming. *Ocean Engineering*, 90 :129–139, 2014. (pages [2](#), [18](#), et [35](#))
- R. Duvigneau. Numerical Optimization : Basic Concepts and Algorithms. INRIA, 2015. (pages [vii](#) et [49](#))
- R. Duvigneau, J. Labroquère, and E. Guilmineau. Comparison of turbulence closures for optimized active control. *Computers & Fluids*, 124 :67–77, January 2016. (page [3](#))
- K. Elsayed. Optimization of the cyclone separator geometry for minimum pressure drop using Co-Kriging. *Powder Technology*, 269(Supplement C) :409 – 424, 2015. (page [110](#))
- V. Falkner. *The solution of lifting-plane problems by vortex-lattice theory*. Ministry of Supply, Aeronautical Research Council, 1947. (page [31](#))
- J. B. Fallow. America's Cup sail design. *Journal of wind engineering and industrial aerodynamics*, 63(1-3) : 183–192, 1996. (page [17](#))
- B. Farmanesh and A. Pourhabib. Sparse Pseudo-input Local Kriging for Large Non-stationary Spatial Datasets with Exogenous Variables. *arXiv preprint arXiv :1508.01248*, 2015. (page [65](#))
- R. G. Flay. A twisted flow wind tunnel for testing yacht sails. *Journal of Wind Engineering and Industrial Aerodynamics*, 63(1-3) :171–182, 1996. (pages [1](#) et [40](#))

- R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, 1987. (page 49)
- A. I. Forrester, A. Sobester, and A. J. Keane. Multi-fidelity optimization via surrogate modelling. *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, 463(2088) :3251–3269, 2007. (pages viii, 106, 107, 109, et 110)
- A. I. J. Forrester, A. Sobester, and A. J. Keane. *Constraints*, pages 117–139. John Wiley & Sons, Ltd, 2008. (page 85)
- F. Fossati. *Aero-Hydrodynamics and the Performance of Sailing Yachts : The Science Behind Sailboats and Their Design*. McGraw-Hill Education, 2010. (pages vii, 13, et 29)
- F. Fossati, S. Muggiasca, I. M. Viola, and A. Zasso. Wind tunnel techniques for investigation and optimization of sailing yachts aerodynamics. In *Proc. 2nd High Performance Yacht Design Conference (HPYD2), Auckland, New Zealand, Feb*, pages 14–16, 2006. (page 40)
- F. Fossati, S. Muggiasca, I. Bayati, and C. Bertorello. Lecco innovation hub sailing yacht lab project. In *Proceedings of the 3rd International Conference Innovation in High Performance Sailing Yachts, Lorient France*, volume 1, 2013. (page 42)
- F. Fossati, I. Bayati, F. Orlandini, S. Muggiasca, A. Vandone, G. Mainetti, R. Sala, C. Bertorello, and E. Begovic. A novel full scale laboratory for yacht engineering research. *Ocean Engineering*, 104 :219 – 237, 2015. (pages 2 et 42)
- F. Fossati, I. Bayati, S. Muggiasca, A. Vandone, G. Campanardi, T. Burch, M. Malandra, et al. Pressure Measurements on Yacht Sails : Development of a New System for Wind Tunnel and Full Scale Testing. In *The 22nd Chesapeake Sailing Yacht Symposium*, pages 84–97, 2016. (pages 2 et 42)
- S. Gallay. *Algorithmes de couplage RANS et écoulement potentiel*. PhD thesis, Ecole Polytechnique, Montreal (Canada), 2016. (page 31)
- J. Gardner, M. Kusner, K. Q. Weinberger, J. Cunningham, and Z. Xu. Bayesian Optimization with Inequality Constraints. In T. Jebara and E. P. Xing, editors, *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 937–945. JMLR Workshop and Conference Proceedings, 2014. (pages 81 et 84)
- R. Garrett. *The Symmetry of Sailing : The Physics of Sailing for Yachtsmen*. Sheridan House, 1996. (pages vii, 16, 28, et 29)
- S. Ghosal and P. Moin. The Basic Equations for the Large Eddy Simulation of Turbulent Flows in Complex Geometry. *Journal of Computational Physics*, 118(1) :24 – 37, 1995. (page 33)
- M. N. Gibbs. *Bayesian Gaussian processes for regression and classification*. PhD thesis, Citeseer, 1998. (page 64)

- D. E. Goldberg and J. H. Holland. Genetic algorithms and machine learning. *Machine learning*, 3(2) :95–99, 1988. (page 49)
- K. Graf and O. Müller. Photogrammetric investigation of the flying shape of spinnakers in a twisted flow wind tunnel. In *Proceedings 19th Chesapeake Sailing Yacht Symposium*, 2009. (page 40)
- L. L. Gratiet, S. Marelli, and B. Sudret. *Metamodel-Based Sensitivity Analysis : Polynomial Chaos Expansions and Gaussian Processes*, pages 1–37. Springer International Publishing, Cham, 2016. (page 71)
- E. Guilmineau. *Contribution a la prediction du décrochage sur des ailes en incidence au moyen des equations de navier-stokes-reynolds*. PhD thesis, École doctorale Sciences pour l’ingénieur (Nantes), 1995. (page 33)
- I. Guyon, A. Saffari, G. Dror, and G. Cawley. Model selection : Beyond the bayesian/frequentist divide. *Journal of Machine Learning Research*, 11(Jan) :61–87, 2010. (page 96)
- J. Hadamard. Sur les problemes aux derive espartielles et leur signification physique. *Bulletin of Princeton University*, 13 :1–20, 1902. (page 48)
- H. Hansen. *Enhanced wind tunnel techniques and aerodynamic force models for yacht sails*. PhD thesis, The University of Auckland, 2006a. (pages vii et 14)
- N. Hansen. *The CMA Evolution Strategy : A Comparing Review*, pages 75–102. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006b. (pages 53, 54, 65, 74, 148, et 170)
- N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, 9(2) :159–195, 2001. (pages 49 et 51)
- N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary computation*, 11(1) :1–18, 2003. (page 54)
- N. Hansen, D. V. Arnold, and A. Auger. Evolution strategies. In *Springer handbook of computational intelligence*, pages 871–898. Springer, 2015. (pages 53 et 54)
- H. Hattori and Y. Nagano. Direct numerical simulation of turbulent heat transfer in plane impinging jet. *International Journal of Heat and Fluid Flow*, 25(5) :749 – 758, 2004. Selected papers from the 4th International Symposium on Turbulence Heat and Mass Transfer. (page 34)
- P. Hémon. *Vibrations des structures couplées avec le vent*. Editions Ecole Polytechnique, 2006. (page 13)
- G. Hepner. Artificial neural network classification using a minimal training set. Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing*, 56(4) :469–473, 4 1990. (page 71)

- R. Hooke and T. A. Jeeves. "Direct Search" Solution of Numerical and Statistical Problems. *Journal of the ACM (JACM)*, 8(2) :212–229, 1961. (page 49)
- D. Huang, T. T. Allen, W. I. Notz, and N. Zeng. Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models. *Journal of global optimization*, 34(3) :441–466, 2006. (pages 77 et 111)
- K. Huang, D. Zheng, J. Sun, Y. Hotta, K. Fujimoto, and S. Naoi. Sparse learning for support vector classification. *Pattern Recognition Letters*, 31(13) :1944 – 1951, 2010. Meta-heuristic Intelligence Based Image Processing. (page 94)
- A. Jameson. Aerodynamic Design via Control Theory. *Journal of Scientific Computing*, 3(97-0101) :233–260, 1988. (page 51)
- A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur. Asymptotic normality and efficiency of two sobol index estimators. *ESAIM : Probability and Statistics*, 18 :342–364, 2014. (page 70)
- L. Jiao, L. Bo, and L. Wang. Fast Sparse Approximation for Least Squares Support Vector Machine. *IEEE Transactions on Neural Networks*, 18(3) :685–697, May 2007. (page 94)
- M. Johnson, L. Moore, and D. Ylvisaker. Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, 26(2) :131 – 148, 1990. (page 67)
- D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization*, 21(4) :345–383, 2001. (page 76)
- D. R. Jones, M. Schonlau, and W. J. Welch. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4) :455–492, 1998. (pages 76 et 77)
- W. Jones and B. Launder. The prediction of laminarization with a two-equation model of turbulence. *International journal of heat and mass transfer*, 15(2) :301–314, 1972. (page 33)
- A. Keane and P. Nair. *Computational Approaches for Aerospace Design : The Pursuit of Excellence*. John-Wiley and Sons, 2005. (page 51)
- J. M. Keller, M. R. Gray, and J. A. Givens. A fuzzy k-nearest neighbor algorithm. *IEEE transactions on systems, man, and cybernetics*, 15(4) :580–585, 1985. (page 147)
- J. Kennedy and R. Eberhart. Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948 vol.4, Nov 1995. (page 49)
- M. C. Kennedy and A. O'Hagan. Predicting the Output from a Complex Computer Code When Fast Approximations Are Available. *Biometrika*, 87(1) :1–13, 2000. (pages 105, 106, 107, et 110)
- H. Kerhascoet. *De la mesure du vent au pilotage automatique d'un voilier : modélisation, optimisation & application du traitement du signal prédictif*. PhD thesis, Université Bretagne Loire, 2017. (pages 14 et 41)



- J. Keuning. *Approximation of the calm water resistance on a sailing yacht based on the Delft systematic yacht hull series*. Delft University of Technology, Faculty of Mechanical Engineering and Marine Technology, Ship Hydromechanics Laboratory, 1999. (page 38)
- J. Keuning and M. Katgert. The influence of heel on the bare hull, resistance of sailing yacht. In *The 2nd International Conference on Innovation in High Performance Sailing Yachts, Lorient, France*, 2010. (page 38)
- J. Keuning, R. Onnink, A. Versluis, and A. van Gulik. The bare hull resistance of the delft systematic yacht hull series. In *14th International HISWA Symposium on Yacht Design and Yacht Construction, Amsterdam, The Netherlands*, 1996. (page 38)
- T. G. Kolda, R. M. Lewis, and V. Torczon. Optimization by direct search : New perspectives on some classical and modern methods. *SIAM review*, 45(3) :385–482, 2003. (page 51)
- R. Korpus. Performance prediction without empiricism : A rans-based vpp and design optimization capability. In *Proc. 18 th Chesapeake Sailing Yacht Symposium*, 2007. (page 16)
- B. R. Kowalski and C. F. Bender. K-Nearest Neighbor Classification Rule (pattern recognition) applied to nuclear magnetic resonance spectral interpretation. *Analytical Chemistry*, 44(8) :1405–1411, 1972. (page 71)
- D. G. Krige. *Lognormal-de Wijsian geostatistics for ore evaluation*. Johannesburg : South African Institute of mining and metallurgy, 1981. (page 60)
- M. Kuss and C. E. Rasmussen. Assessing approximate inference for binary Gaussian process classification. *Journal of machine learning research*, 6(Oct) :1679–1704, 2005. (page 73)
- J. Laurenceau and P. Sagaut. Efficient response surfaces of aerodynamic functions with Kriging and Cokriging. *AIAA Journal*, 46(2) :498–507, 2008. (page 3)
- L. Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments*. PhD thesis, Université Paris-Diderot - Paris VII, Oct. 2013. (page 105)
- L. Le Gratiet and C. Cannamela. Kriging-based sequential design strategies using fast cross-validation techniques with extensions to multi-fidelity computer codes. working paper or preprint, Oct. 2012. (pages 107 et 110)
- L. Le Gratiet and J. Garnier. Recursive co-kriging model for Design of Computer experiments with multiple levels of fidelity. *International Journal for Uncertainty Quantification*, 4(5) :365–386, 2014. (pages 106, 107, et 108)
- L. Le Gratiet, C. Cannamela, and B. Iooss. A bayesian approach for global sensitivity analysis of (multifidelity) computer codes. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1) :336–363, 2014. (pages 70 et 71)

- D. Le Pelley, D. Morris, and P. Richards. Aerodynamic force deduction on yacht sails using pressure and shape measurements in real time. In *4th High Performance Yacht Design Conference*, pages 28–37, 2012. (page 42)
- H. Lee, Y. Jo, D.-J. Lee, and S. Choi. Surrogate model based design optimization of multiple wing sails considering flow interaction effect. *Ocean Engineering*, 121 :422–436, 2016. (page 18)
- R. L. Levin and L. Larsson. Sailing yacht performance prediction based on coupled CFD and rigid body dynamics in 6 degrees of freedom. *Ocean Engineering*, 144 :362 – 373, 2017. (page 16)
- H.-T. Lin, C.-J. Lin, and R. C. Weng. A note on Platt’s probabilistic outputs for support vector machines. *Machine Learning*, 68(3) :267–276, 2007. (page 97)
- M. Lombardi, N. Parolini, A. Quarteroni, and G. Rozza. Numerical simulation of sailing boats : Dynamics, FSI, and shape optimization. In *Variational Analysis and Aerospace Engineering : Mathematical Challenges for Aerospace Design*, pages 339–377. Springer, 2012. (page 18)
- J. López Lázaro, K. D. Brabanter, J. R. Dorronsoro, and J. A. Suykens. Sparse LS-SVMs with L0-norm minimization. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.*, pages 189–194, 2011. (page 94)
- D. J. MacKay. Bayesian interpolation. *Neural computation*, 4(3) :415–447, 1992. (page 64)
- D. J. MacKay. *Information theory, inference and learning algorithms*. Cambridge university press, 2003. (page 71)
- A. Marrel, B. Iooss, B. Laurent, and O. Roustant. Calculations of Sobol indices for the Gaussian process metamodel. *Reliability Engineering & System Safety*, 94(3) :742 – 751, 2009. (page 70)
- Y. Masuyama and T. Fukasawa. Full-scale measurement of sail force and the validation of numerical calculation method. In *13th Chesapeake sailing yacht symposium, SNAME, Annapolis, 1997*, 1997. (page 41)
- Y. Masuyama, Y. Tahara, T. Fukasawa, and N. Maeda. Database of sail shapes versus sail performance and validation of numerical calculations for the upwind condition. *Journal of marine science and technology*, 14 (2) :137–160, 2009. (page 42)
- G. Matheron. Principles of geostatistics. *Economic geology*, 58(8) :1246–1266, 1963. (page 60)
- M. D. McKay, R. J. Beckman, and W. J. Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2) :239–245, 1979. (page 66)
- F. R. Menter, M. Kuntz, and R. Langtry. Ten years of industrial experience with the SST turbulence model. *Turbulence, heat and mass transfer*, 4(1) :625–632, 2003. (page 33)



- J. Mercer. Functions of Positive and Negative Type, and their Connection with the Theory of Integral Equations. *Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, 209(441-458) :415–446, 1909. (page 91)
- Z. Michalewicz and M. Schoenauer. Evolutionary algorithms for constrained parameter optimization problems. *Evolutionary computation*, 4(1) :1–32, 1996. (pages 49 et 51)
- T. P. Minka. *A family of algorithms for approximate Bayesian inference*. PhD thesis, Massachusetts Institute of Technology, 2001. (pages 73 et 113)
- D. Motta, R. Flay, P. Richards, D. Le Pelley, J. Deparday, and P. Bot. Experimental investigation of asymmetric spinnaker aerodynamics using pressure and sail shape measurements. *Ocean Engineering*, 90 :104–118, 2014. (page 18)
- M. M. Munk. The aerodynamic forces on airship hulls. Technical report, NASA, 1924. (page 31)
- J. A. Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7(4) :308–313, 1965. (pages 17, 49, et 51)
- J. E. Oakley and A. O’Hagan. Probabilistic sensitivity analysis of complex models : a bayesian approach. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 66(3) :751–769, 2004. (page 70)
- J.-S. Park. Optimal latin-hypercube designs for computer experiments. *Journal of Statistical Planning and Inference*, 39(1) :95 – 111, 1994. (page 67)
- S. V. Patankar and D. B. Spalding. A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows. In *Numerical Prediction of Flow, Heat Transfer, Turbulence and Combustion*, pages 54–73. Elsevier, 1983. (page 34)
- D. Paulet and D. Presles. *Architecture navale : connaissance et pratique*. Savoir-faire de l’architecture. Editions de La Villette, 1998. (pages vii et 10)
- S. Peron. *Méthode d’assemblage de maillages recouvrants autour de géométries complexes pour des simulations en aérodynamique compressible*. PhD thesis, Ecole nationale supérieure d’arts et métiers - ENSAM, Oct. 2014. (page 217)
- A. Persson, L. Larsson, M. Brown, and C. Finnsgard. Performance evaluation and ranking of 7 rudders for the finn dinghy. In *INNOVSAIL International Conference on Innovation in High Performance Sailing Yachts*, pages 59–68, 2017. (page 18)
- A. Philpott and A. Mason. Advances in optimization in yacht performance analysis. In *The 1st High Performance Yacht Design Conference*, pages 229–236, 2002. (page 17)
- V. Picheny, T. Wagner, and D. Ginsbourger. A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization*, 48(3) :607–626, Sep 2013. (pages 75, 77, 110, et 191)

- J. C. Platt. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. In *Advances in kernel methods - Support Vector*, pages 185–208. MIT Press, 1999a. (page 90)
- J. C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3) :61–74, 1999b. (page 97)
- L. Prandtl. Theory of lifting surfaces. Technical report, NASA, 1920. (page 31)
- P. Z. Qian and C. J. Wu. Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments. *Technometrics*, 50(2) :192–204, 2008. (page 107)
- P. Queutey and M. Visonneau. An interface capturing method for free-surface hydrodynamic flows. *Computers & Fluids*, 36(9) :1481 – 1510, 2007. (pages 15 et 35)
- J. Quiñonero-Candela and C. E. Rasmussen. A unifying view of sparse approximate Gaussian process regression. *Journal of Machine Learning Research*, 6(Dec) :1939–1959, 2005. (page 73)
- R. Ranzenbach, D. Armitage, and A. Carrau. Mainsail planform optimization for irc 52 using fluid structure interaction. In *The 21st Chesapeake Sailing Yacht Symposium, SNAME*, 2013. (page 18)
- C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006. (pages 56, 58, 73, et 107)
- I. Rechenberg. Evolutionsstrategien. In B. Schneider and U. Ranft, editors, *Simulationsmethoden in der Medizin und Biologie*, pages 83–114, Berlin, Heidelberg, 1978. Springer Berlin Heidelberg. (page 52)
- C. Rehbach. Calcul numérique d’écoulements tridimensionnels instationnaires avec nappes tourbillonnaires. *La recherche aérospatiale*, 5 :289–298, 1977. (page 31)
- C. Rehbach. Numerical calculation of three-dimensional unsteady flows with vortex sheets. In *16th Aerospace Sciences Meeting*, page 111, 1978. (page 35)
- S. Rippa. An algorithm for selecting a good value for the parameter  $c$  in radial basis function interpolation. *Advances in Computational Mathematics*, 11(2) :193–210, 1999. (page 56)
- B. Rosen, J. Laiosa, and W. Davis, Jr. CFD design studies for America’s Cup 2000. In *18th Applied Aerodynamics Conference*, page 4339, 2000. (page 17)
- F. Rosenblatt. The Perceptron : A Probabilistic Model for Information Storage and Organization in The Brain. *Psychological Review*, 65(6) :65–386, 1958. (page 85)
- N. Rousselon. Optimization for sail design. In *ModeFrontier Conference*, 2008. (page 18)
- Y. Roux, S. Huberson, F. Hauville, J.-P. Boin, M. Guilbaud, and M. Ba. Yacht performance prediction : Towards a numerical VPP. In *High Performance Yacht Design Conference*, 2002. (page 35)

- M. Sacher, F. Hauville, P. Bot, and M. Durand. Sail trimming FSI simulation - comparison of viscous and inviscid flow models to optimise upwind sails trim. In *Proceedings of 5th High Performance Yacht Design Conference. Auckland, New-Zealand*, 2015. (pages viii, 188, 189, et 190)
- M. Sacher, F. Hauville, R. Duvigneau, O. Le Maître, N. Aubin, and M. Durand. Experimental and numerical optimizations of an upwind mainsail trimming. In *THE 22nd CHESAPEAKE SAILING YACHT SYMPOSIUM*, 2016. (page 121)
- M. Sacher, M. Durand, E. Berrini, F. Hauville, R. Duvigneau, O. Le Maître, and J. A. Astolfi. Flexible hydrofoil optimization for the 35th America’s cup with constrained EGO Method. In P. Bot, editor, *INNOVSAIL International Conference on Innovation in High Performance Sailing Yachts*, INNOVSAIL 2017, 4th Edition, pages 193–206, Lorient, France, June 2017a. CVET and Ecole Navale. (page 173)
- M. Sacher, R. Duvigneau, O. Le Maître, M. Durand, E. Berrini, F. Hauville, and J. A. Astolfi. Surrogates and Classification Approaches for Efficient Global Optimization (EGO) with Inequality Constraints. In *SIAM Optimization*, Vancouver, Canada, May 2017b. (page 148)
- M. Sacher, F. Hauville, R. Duvigneau, O. Le Maître, N. Aubin, and M. Durand. Efficient optimization procedure in non-linear fluid-structure interaction problem : Application to mainsail trimming in upwind conditions. *Journal of Fluids and Structures*, 69 :209 – 231, 2017c. (pages 3, 35, 121, 188, et 230)
- M. Sacher, M. Durand, Élisabeth Berrini, F. Hauville, R. Duvigneau, O. L. Maître, and J.-A. Astolfi. Flexible hydrofoil optimization for the 35th America’s Cup with constrained EGO method. *Ocean Engineering*, 157 : 62 – 72, 2018a. (pages 4, 172, et 231)
- M. Sacher, R. Duvigneau, O. Le Maître, M. Durand, É. Berrini, F. Hauville, and J.-A. Astolfi. A classification approach to efficient global optimization in presence of non-computable domains. *Structural and Multidisciplinary Optimization*, Apr 2018b. (pages 4, 148, et 230)
- M. Sasena, P. Papalambros, and P. Goovaerts. Global optimization of problems with disconnected feasible regions via surrogate modeling. In *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*. American Institute of Aeronautics and Astronautics Inc., 2002a. (page 84)
- M. J. Sasena, P. Papalambros, and P. Goovaerts. Exploration of metamodeling sampling criteria for constrained global optimization. *Engineering optimization*, 34(3) :263–278, 2002b. (page 84)
- M. Schonlau. *Computer Experiments and Global Optimization*. PhD thesis, University of Waterloo, Waterloo, Ont., Canada, Canada, 1997. AAINQ22234. (pages 76, 81, et 83)
- M. Seeger, C. K. I. Williams, and N. D. Lawrence. Fast Forward Selection to Speed Up Sparse Gaussian Process Regression. In *IN WORKSHOP ON AI AND STATISTICS 9*, 2003. (page 65)
- Z. Shen, D. Wan, and P. M. Carrica. Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering. *Ocean Engineering*, 108 :287–306, 2015. (page 217)

- M. Shur, P. Spalart, M. Strelets, and A. Travin. Detached-eddy simulation of an airfoil at high angle of attack. In W. Rodi and D. Laurence, editors, *Engineering Turbulence Modelling and Experiments 4*, pages 669 – 678. Elsevier Science Ltd, Oxford, 1999. (page 34)
- A. J. Smola and P. L. Bartlett. Sparse greedy Gaussian process regression. In *Advances in neural information processing systems*, pages 619–625, 2001. (page 65)
- A. Sóbester, S. J. Leary, and A. J. Keane. On the Design of Optimization Strategies Based on Global Response Surface Approximation Models. *Journal of Global Optimization*, 33(1) :31–59, Sep 2005. (page 77)
- I. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1) :271 – 280, 2001. The Second IMACS Seminar on Monte Carlo Methods. (pages 68, 69, et 70)
- I. M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4) :407–414, 1993. (page 70)
- J. L. Steger and J. A. Benek. On the use of composite grid schemes in computational aerodynamics. *Computer Methods in Applied Mechanics and Engineering*, 64(1-3) :301–320, 1987. (page 217)
- M. L. Stein. *Interpolation of Spatial Data - Some Theory for Kriging*. Springer Science & Business Media, 2012. (page 63)
- M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2) :111–147, 1974. (page 94)
- J. Suykens and J. Vandewalle. Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*, 9 (3) :293–300, 1999. (pages 71, 85, et 92)
- J. Suykens, J. D. Brabanter, L. Lukas, and J. Vandewalle. Weighted least squares support vector machines : robustness and sparse approximation. *Neurocomputing*, 48(1–4) :85 – 105, 2002. (page 94)
- J. A. K. Suykens, L. Lukas, and J. Vandewalle. Sparse approximation using least squares support vector machines. In *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*, volume 2, pages 757–760, 2000. (page 94)
- F. Tagliaferri, A. Philpott, I. Viola, and R. Flay. On risk attitude and optimal yacht racing tactics. *Ocean Engineering*, 90 :149 – 154, 2014. Innovation in High Performance Sailing Yachts - INNOVSAIL. (page 17)
- V. Torczon. On the convergence of the multidirectional search algorithm. *SIAM journal on Optimization*, 1(1) : 123–145, 1991. (page 51)
- O. Şugar Gabor, A. Koreanschi, and R. M. Botez. A new non-linear vortex lattice method : Applications to wing aerodynamic optimizations. *Chinese Journal of Aeronautics*, 29(5) :1178 – 1195, 2016. (page 31)

- B. Van Calster, J. Luts, J. A. K. Suykens, G. Condous, T. Bourne, D. Timmerman, and S. Van Huffel. *Comparing Methods for Multi-class Probabilities in Medical Decision Making Using LS-SVMs and Kernel Logistic Regression*, pages 139–148. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. (page 97)
- T. Van Gestel, J. Suykens, B. De Moor, and J. Vandewalle. Bayesian Inference for LS-SVM on Large Data Sets using the Nystrom Method. In *World Congress on Computational Intelligence - International Joint Conference on Neural Networks, Honolulu, USA*, pages 2779–2784, May 2002. (page 95)
- T. van Gestel, J. A. Suykens, B. Baesens, S. Viaene, J. Vanthienen, G. Dedene, B. de Moor, and J. Vandewalle. Benchmarking Least Squares Support Vector Machine Classifiers. *Machine Learning*, 54(1) :5–32, 2004. (page 93)
- P. Van Oossanen. Predicting the speed of sailing yacht. *Transactions - Society of Naval Architects and Marine Engineers*, 101 :337–397, 1993. (page 15)
- V. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995. (pages 85, 86, et 92)
- V. Vapnik. *The Support Vector Method of Function Estimation*, pages 55–85. Springer US, Boston, MA, 1998a. (page 86)
- V. Vapnik. *Statistical learning theory*, volume 1. Wiley New York, 1998b. (pages 86 et 90)
- G. N. Venderplaats. *Numerical optimization techniques for engineering design*. McGraw-Hill, 1984. (page 49)
- I. Viola and A. Arredondo-Galeana. The leading-edge vortex of yacht sails. In *INNOVSAIL International Conference on Innovation in High Performance Sailing Yachts*, pages 115–126, 2017. (page 2)
- I. M. Viola and R. G. Flay. Full-scale pressure measurements on a sparkman and stephens 24-foot sailing yacht. *Journal of Wind Engineering and Industrial Aerodynamics*, 98(12) :800–807, 2010. (page 1)
- I. M. Viola and R. G. Flay. Sail aerodynamics : Understanding pressure distributions on upwind sails. *Experimental Thermal and Fluid Science*, 35(8) :1497–1504, 2011a. (page 1)
- I. M. Viola and R. G. Flay. Sail pressures from full-scale, wind-tunnel and numerical investigations. *Ocean Engineering*, 38(16) :1733–1743, 2011b. (page 1)
- I. M. Viola, S. Bartesaghi, T. Van-Renterghem, and R. Ponzini. Detached eddy simulation of a sailing yacht. *Ocean Engineering*, 90 :93–103, 2014. (pages 2 et 18)
- I. M. Viola, M. Sacher, J. Xu, and F. Wang. A numerical method for the design of ships with wind-assisted propulsion. *Ocean Engineering*, 105 :33–42, 2015. (page 2)
- R. Von Mises. *Mathematical Theory of Probability and Statistics*. Academic press, 1964. (page 61)

- J. Wackers, G. Deng, A. Leroyer, P. Queutey, and M. Visonneau. Adaptive grid refinement for hydrodynamic flows. *Computers & Fluids*, 55 :85 – 100, 2012. (page 15)
- D. C. Wilcox. Reassessment of the scale-determining equation for advanced turbulence models. *AIAA journal*, 26(11) :1299–1310, 1988. (page 33)
- C. K. Williams and D. Barber. Bayesian classification with Gaussian processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12) :1342–1351, 1998. (page 73)
- P. Wolfe. Convergence conditions for ascent methods. *SIAM review*, 11(2) :226–235, 1969. (page 50)
- H.-Y. Xu, C.-L. Qiao, H.-Q. Yang, and Z.-Y. Ye. Delayed detached eddy simulation of the wind turbine airfoil S809 for angles of attack up to 90 degrees. *Energy*, 118 :1090 – 1109, 2017. (page 34)
- Z. Yan, Y. Yang, and Y. Ding. An experimental study of the hyper-parameters distribution region and its optimization method for support vector machine with gaussian kernel. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 6(5) :437–446, 2013. (page 96)
- J. Zhu and T. Hastie. Kernel logistic regression and the import vector machine. *Journal of Computational and Graphical Statistics*, 14(1) :185–205, 2012. (page 71)







# MÉTHODES AVANCÉES D'OPTIMISATION PAR MÉTA-MODÈLES – APPLICATION À LA PERFORMANCE DES VOILIERS DE COMPÉTITION

**RÉSUMÉ** L'optimisation de la performance des voiliers est un problème difficile en raison de la complexité du système mécanique (couplage aéro-élastique et hydrodynamique) et du nombre important de paramètres à optimiser (voiles, gréement, etc.). Malgré le fait que l'optimisation des voiliers est empirique dans la plupart des cas aujourd'hui, les approches numériques peuvent maintenant devenir envisageables grâce aux dernières améliorations des modèles physiques et des puissances de calcul. Les calculs aéro-hydrodynamiques restent cependant très coûteux car chaque évaluation demande généralement la résolution d'un problème non linéaire d'interaction fluide-structure. Ainsi, l'objectif central de cette thèse est de proposer et développer des méthodes originales dans le but de minimiser le coût numérique de l'optimisation de la performance des voiliers. L'optimisation globale par méta-modèles Gaussiens est utilisée pour résoudre différents problèmes d'optimisation. La méthode d'optimisation par méta-modèles est étendue aux cas d'optimisations sous contraintes, incluant de possibles points non évaluables, par une approche de type classification. L'utilisation de méta-modèles à fidélités multiples est également adaptée à la méthode d'optimisation globale. Les applications concernent des problèmes d'optimisation originaux où la performance est modélisée expérimentalement et/ou numériquement. Ces différentes applications permettent de valider les développements des méthodes d'optimisation sur des cas concrets et complexes, incluant des phénomènes d'interaction fluide-structure.

**Mots-clés :** Optimisation globale ; méta-modèles ; processus Gaussien ; classification ; multi-fidélité ; voilier ; mécanique des fluides ; interaction fluide-structure ; études numériques-expérimentales.

## ADVANCED SURROGATE-BASED OPTIMIZATION METHODS – APPLICATION TO RACING YACHT PERFORMANCE

**ABSTRACT** Sailing yacht performance optimization is a difficult problem due to the high complexity of the mechanical system (aero-elastic and hydrodynamic coupling) and the large number of parameters to optimize (sails, rigs, etc.). Despite the fact that sailboats optimization is empirical in most cases today, the numerical optimization approach is now considered as possible because of the latest advances in physical models and computing power. However, these numerical optimizations remain very expensive as each simulation usually requires solving a non-linear fluid-structure interaction problem. Thus, the central objective of this thesis is to propose and to develop original methods aiming at minimizing the numerical cost of sailing yacht performance optimization. The Efficient Global Optimization (EGO) is therefore applied to solve various optimization problems. The original EGO method is extended to cases of optimization under constraints, including possible non computable points, using a classification-based approach. The use of multi-fidelity surrogates is also adapted to the EGO method. The applications treated in this thesis concern the original optimization problems in which the performance is modeled experimentally and/or numerically. These various applications allow for the validation of the developments in optimization methods on real and complex problems, including fluid-structure interaction phenomena.

**Keywords:** Global optimization ; surrogate models ; Gaussian process ; classification ; multi-fidelity ; sailing yacht ; fluid mechanics ; fluid-structure interaction ; numerical-experimental studies.