



**HAL**  
open science

# Une description morphologique précise des grains pour l'étude numérique des propriétés mécaniques du ballast ferroviaire

Noura Ouhbi

► **To cite this version:**

Noura Ouhbi. Une description morphologique précise des grains pour l'étude numérique des propriétés mécaniques du ballast ferroviaire. Infrastructures de transport. Université Paris-Est, 2017. Français. NNT : 2017PESC1164 . tel-01730240

**HAL Id: tel-01730240**

**<https://pastel.hal.science/tel-01730240>**

Submitted on 13 Mar 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.



UNIVERSITÉ  
— PARIS-EST



N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

## THÈSE

pour obtenir le grade de

**DOCTEUR de l'Université Paris-Est**

Spécialité : **Génie civil**

préparée au laboratoire **Laboratoire Navier**

dans le cadre de l'École Doctorale **Science, Ingénierie Environnement**

présentée et soutenue publiquement

par

**Noura OUHBI**

le 26 Octobre 2017

Titre:

**Description précise de la morphologie des grains pour l'étude numérique du comportement mécanique du ballast ferroviaire**

Directeur de thèse: **Jean-Noël ROUX**

Co-directeur de thèse: **Charles VOIVRET**

### Jury

M. Emilien AZEMA,	Rapporteur
M. Gaël COMBE,	Rapporteur
M. Patrick RICHARD,	Examinateur
M. Karam SAB,	Examinateur
M. Jean-François FERELLEC,	Invité
M. Guillaume PERRIN,	Invité
M. Jean-Noël ROUX,	Directeur de thèse
M. Charles VOIVRET,	Co-directeur de thèse

---

# Résumé

Le ballast est la couche granulaire, formée de grains rocheux irréguliers de taille centimétrique, qui supporte la voie ferrée par l'intermédiaire des traverses. Son rôle est crucial dans la transmission des efforts statiques et dynamiques imposés par la circulation des trains vers le sol supportant la voie. Pour comprendre le phénomène de tassement du ballast, qui croît fortement avec la vitesse des trains, et qui entraîne des opérations coûteuses de maintenance, il est important de bien connaître l'influence de la morphologie des grains sur le comportement mécanique du ballast. C'est l'objectif de cette thèse, dans laquelle est développée une technique innovante de caractérisation de la morphologie fondée sur la décomposition orthogonale aux valeurs propres (« POD »). Cette technique, appliquée à des données expérimentales détaillées (scans 3D) permet d'identifier un choix optimal de fonctions de forme. Il est alors possible de générer numériquement des ensembles granulaires représentatifs du ballast avec différents degrés de raffinement selon le nombre de fonctions de base retenu.

Pour passer aux applications mécaniques, les grains, ainsi modélisés par un nombre limité de coefficients, doivent ensuite par une opération de maillage être assimilés à des polyèdres convexes dont le comportement mécanique est simulé par la méthode des éléments discrets (dynamique des contacts, avec le logiciel LMGC90). On détermine alors des caractéristiques de base comme l'angle de frottement interne et la compacité dans l'état critique de cisaillement quasi-statique maintenu. Ces propriétés sont stabilisées à partir d'un nombre assez faible de fonctions de forme (de l'ordre de 20).

Dans le domaine de la description morphologique, l'approche fondée sur la POD ouvre d'intéressantes perspectives pour améliorer les caractérisations usuelles. Elle permet de différencier les populations de grains de ballast selon leur degré d'usure ou la carrière dont ils proviennent. Elle conduit également à une méthode d'identification des faces et des courbures des grains proposée dans ce travail.

# Abstract

The ballast is the granular layer formed of irregular centimeter-sized rock grains, which supports the railway track through the sleepers. Its role is crucial in the transmission of static and dynamic forces imposed by the movement of trains towards the ground supporting the track. To understand the phenomenon of ballast settlement, which increases significantly with the speed of the trains, and which requires costly maintenance operations, it is important to understand the influence of the grain shape on the mechanical behaviour of ballast. This is the objective of this thesis. First, an innovative method for particle shape characterization is developed, based on the Proper Orthogonal Decomposition (POD). When applied to highly accurate experimental data (3D scan of ballast grains), this technique enables the identification of an optimal set of shape functions, such that the shape information is condensed in a minimum number of coefficients. It becomes then possible to numerically generate granular assemblies that are representative of the morphology of ballast grains with different degrees of refinement according to the number of retained shape functions.

In order to investigate the mechanical properties of ballast, the simulated grains modeled by a limited number of coefficients have to be subjected to a meshing procedure and turned into convex polyhedra, suitable for numerical simulations of mechanical behaviours via the discrete element method (contact dynamics, using the LMGC90 software). We determine some fundamental characteristics such as the internal friction angle and the solid fraction in the critical state of quasistatic steady shear flow. These properties are stabilized after including a rather small number of shape functions in to the model (on the order of 20).

The method also opens promising avenues for morphological characterization applications. It proves able to distinguish ballast grains with varying degrees of wear, or produced in different quarries. It also leads to a new procedure for identifying faces and curvatures.

# Introduction

En tant que mainteneur et propriétaire de l'infrastructure ferroviaire française, SNCF doit en réaliser une gestion industriellement performante. Au-delà des aspects économiques primordiaux, le succès de cette gestion repose grandement sur une infrastructure à grande capacité, à faible coût de maintenance et à longue durée de vie. La solution empirique et continuellement améliorée au cours de l'histoire ferroviaire est la voie ballastée. Notamment, depuis la mise en service des premiers tronçons à grandes vitesses (1981) la technologie ballastée a considérablement progressé pour maintenant supporter des vitesses commerciales supérieures à 300 km/h avec des efforts de maintenance acceptables.

Pour atteindre ce résultat, des composants de la voie comme les traverses, les attaches ou les rails ont fait et font toujours, l'objet d'améliorations. Malheureusement, ce n'est pas le cas du ballast, dont les spécifications restent largement d'origine empirique et très probablement sous-optimales.

La qualité du ballast est liée aux fonctions qui lui sont assignées : la transmission et la répartition des charges exercées par les véhicules sur la plate-forme, l'amortissement des vibrations grâce à ses propriétés rhéologiques (dissipation de l'énergie vibratoire par frottement, collision ou attrition des grains de ballast), l'ancrage des traverses dans le sens longitudinal et transversal, et le drainage rapide des eaux zénithales.

Il est donc clair que les attentes relatives au ballast sont principalement liées à ses performances mécaniques en tant que matériau granulaire, avec des propriétés de frottement interne et d'amortissement. Son caractère de matériau grossier intervient directement pour sa perméabilité, mais aussi par les effets possibles du rapport de la taille des grains à celle des objets solides avec lesquels le matériau entre en contact. A long terme et sous l'effet des sollicitations dynamiques dues aux circulations, le ballast se tasse de manière inhomogène (tassement différentiel) et se dégrade (rupture des grains, abrasion de leur surface) créant ainsi des défauts de géométrie. Ces défauts sont ensuite traités par des opérations de maintenance de type bourrage qui rétablissent un nivellement acceptable en remaniant localement le ballast sous les traverses. Cependant, au cours de sa vie, la voie perd progressivement sa capacité à être maintenue efficacement. On comprend donc l'importance de limiter le tassement différentiel et d'augmenter la maintenabilité des voies afin de maximiser leur durée d'exploitation.

Dans ce sens, quelques années après la mise en service des premières lignes à grande vitesse, des travaux de recherches ont abouti à une augmentation des exigences sur la qualité de la roche utilisée pour produire le ballast ferroviaire. Les propriétés mécaniques des grains résultent de celles de la roche mère dont ils sont extraits et ne présentent donc pas de moyens d'amélioration au-delà d'une sélection

drastique des carrières. Au contraire, la géométrie des grains (taille et forme) est clairement un levier d'optimisation des performances du ballast.

Afin de relever le défi industriel fort que constitue, pour SNCF, l'amélioration des performances instantanées et de long terme de la voie ballastée, on doit donc mener à bien des recherches assez fondamentales sur le rôle de la géométrie des grains dans les propriétés mécaniques du matériau granulaire, en tant que matériau désordonné, à structure discrète.

Si la simulation numérique, grain par grain, des propriétés mécaniques des assemblages granulaires, par les méthodes dites d'éléments discrets, connaît depuis plusieurs décennies un essor considérable, l'immense majorité des études concerne encore des grains de forme simplifiée, et le plus souvent sphérique. Ce n'est que relativement récemment que l'on a pu simuler des objets de formes variées. Des formes assez complexes ont pu être modélisées par assemblage de grains plus simples. Le développement d'une méthode de simulation de grains tridimensionnels de forme polyédrique quelconque, mise en œuvre dans le code de calcul LMGC90, représente une avancée significative. Ce code est disponible et son usage est bien maîtrisé dans le département Lignes, Voie et Environnement de Ingénierie et Projets à SNCF Réseau. C'est l'outil que nous avons choisi d'utiliser dans cette thèse.

Les travaux de modélisation numérique de grains de formes diverses, bien qu'encore relativement peu nombreux, ont déjà donné quelques résultats significatifs et encourageants. On a pu ainsi mettre en évidence le niveau de frottement interne, bien supérieur à celui des grains sphériques, des assemblages de polyèdres. Diverses études se sont attachées à quantifier les effets des variations d'un certain paramètre de forme : grains plus ou moins allongés, plus ou moins facettés, plus ou moins convexes, etc.

Cependant, la modélisation numérique de granulats avec des formes complexes, tels que le ballast, se trouve rapidement confrontée, en plus de la difficulté numérique propre à la simulation elle-même, à des problèmes d'identification et de représentation. En effet, on peut raffiner presque indéfiniment la description géométrique d'objets irréguliers. Ainsi la question de la représentation numérique des grains de ballast est au cœur de cette thèse. En d'autres termes, comment reproduire les bonnes propriétés morphologiques, et par conséquent mécaniques, sans pour autant introduire un nombre excessif de paramètres ?

C'est afin d'apporter une réponse à cette question que ce travail de thèse propose une nouvelle méthode d'identification et de modélisation de la forme des grains. A partir de la description, obtenue expérimentalement par numérisation tridimensionnelle (scanner 3D), de la surface des grains comme nuage de points, on propose de modéliser un grain comme une combinaison linéaire de fonctions de formes de richesse croissante, les fonctions de base étant déterminées pour leur pertinence statistique. Cette modélisation de la géométrie des grains, sans paramétrage a priori de leur forme, permettra d'engendrer numériquement, à la demande, des équivalents statistiques virtuels de grains de ballast, tout en hiérarchisant l'information morphologique, ou encore de comparer des populations de grains. Les variations morphologiques permises par la modélisation proposée sont accompagnées de simulations des propriétés mécaniques collectives de l'assemblage de grains, ce qui permettra de corrélérer la forme aux performances recherchées. Comme dans nombre d'études précédentes des effets de forme, on s'attachera aux propriétés mécaniques les plus

robustes : compacité maximale, frottement interne à l'état critique du matériau dans un cisaillement quasi-statique stationnaire, connectivité.

Dans le Chapitre 1, nous présentons succinctement la structure de la voie ferrée ballastée ainsi que le rôle de chacun de ses constituants. Nous nous focalisons alors sur le ballast en tant que couche granulaire dans la voie, puis en tant que matériau : son rôle, et les critères fixés pour son utilisation. L'un de ces critères est au centre de cette étude : la forme des grains. Nous présentons ensuite les différents outils utilisés pour la caractérisation de la forme des grains de ballast. Nous terminons par une discussion sur l'effet de la forme sur les propriétés mécaniques, mis en évidence dans plusieurs études expérimentales et numériques, puis une formulation de la problématique.

Le Chapitre 2 est consacré à la caractérisation de la forme des grains de ballast. Nous proposons une approche innovante permettant une caractérisation optimale de la morphologie des grains, dont nous détaillons les étapes, les résultats, les avantages et les limitations. Nous établissons une comparaison de cette approche avec les méthodes existantes.

Le Chapitre 3 relève un défi important sur la génération de formes réelles de grains de ballast. A la suite d'une brève présentation des méthodes existantes de génération de grains, nous présentons l'approche proposée, et nous discutons les résultats de son application à notre ensemble de grains de ballast. Une attention particulière est prêtée à la caractérisation des faces/courbures des grains comme moyen de validation de l'approche de génération.

Le Chapitre 4 traite de deux étapes essentielles à notre étude : le maillage et la simulation. Après une description des méthodes utilisées pour mailler l'ensemble des grains, nous en détaillons les résultats, ainsi que l'effet de cette opération sur la géométrie des grains en fonction du raffinement choisi sur la forme. En deuxième lieu, nous introduisons les concepts généraux de la méthode de Dynamique des Contacts comme méthode choisie à la simulation numérique de nos grains de ballast représentés par des polyèdres irréguliers.

Le Chapitre 5 est consacré à l'étude du comportement quasi-statique du ballast et à l'effet de forme.

Enfin, le Chapitre 6 s'intéresse à une application de la méthode de caractérisation développée au Chapitre 2, pour deux types de données : des grains à différents degrés d'usure, et des grains provenant de différentes carrières. Nous explorons la performance de la méthode, comparée à d'autres approches possibles, à discerner des grains de forme différente. Nous en détaillons alors les résultats, pour enfin conclure.





# Table des matières

Résumé . . . . .	iii
Abstract . . . . .	iv
Introduction . . . . .	v
Table des matières . . . . .	ix
<b>I Ballast : Le matériau aux mille pouvoirs</b>	<b>1</b>
<b>1 Contexte industriel</b>	<b>3</b>
1 La voie ferrée ballastée . . . . .	4
2 Le ballast . . . . .	7
2.1 Fonctions et spécifications . . . . .	7
2.2 Maintenance et enjeux . . . . .	9
2.3 Mesure et caractérisation géométrique du ballast . . . . .	12
2.4 Effet connu de la forme sur le comportement mécanique du ballast . . . . .	19
3 Synthèse . . . . .	23
3.1 Discussion . . . . .	23
3.2 Le ballast dans tous ses états . . . . .	24
<b>II Ballast : Les nuages aux mille points</b>	<b>27</b>
<b>2 Caractérisation de la forme des grains de ballast</b>	<b>29</b>
1 Lumière sur le vocabulaire de la « forme » . . . . .	30
2 Méthodes de caractérisation de la forme de particules . . . . .	31
2.1 Caractérisation scalaire . . . . .	31
2.2 Caractérisation par méthodes spectrales . . . . .	37
3 Caractérisation par POD . . . . .	40
3.1 Principes de la Décomposition Orthogonale aux Valeurs Propres (POD) . . . . .	40
3.2 Mesure et acquisition des données : du réel au virtuel . . . . .	47
3.3 Prétraitement . . . . .	49
3.4 Décomposition de la forme : résultats . . . . .	52
3.5 La caractérisation par POD vs la caractérisation scalaire . . . . .	59
3.6 La POD vs les harmoniques sphériques . . . . .	62
4 Conclusion . . . . .	64
	ix

<b>3</b>	<b>Génération de grains virtuels</b>	<b>67</b>
1	Méthodes de génération de grains virtuels . . . . .	68
1.1	Génération par indicateurs de Fourier . . . . .	68
1.2	Génération par harmoniques sphériques . . . . .	70
2	Caractérisation des faces/courbures des particules 3D . . . . .	71
2.1	Normales et variation de surface aux sommets . . . . .	71
2.2	Procédure de clustering & détection de faces et arêtes . . . . .	74
3	Méthodes proposées pour la génération des grains . . . . .	75
3.1	Méthode paramétrique : copule gaussienne . . . . .	77
3.2	Application au ballast . . . . .	78
3.3	Méthode non paramétrique : Multivariate Gaussian Kernel Density Estimation (Multivariate G-KDE) . . . . .	79
3.4	Application au ballast . . . . .	82
4	Validation . . . . .	84
4.1	Validation par indicateurs classiques . . . . .	84
4.2	Validation par les indicateurs de faces/courbures . . . . .	87
5	Conclusion . . . . .	89
<b>III</b>	<b>Ballast : Les polyèdres aux mille facettes</b>	<b>91</b>
<b>4</b>	<b>Maillage &amp; simulation</b>	<b>93</b>
1	Maillage : du nuage de points au polyèdre . . . . .	94
1.1	Revue bibliographique . . . . .	94
1.2	Méthode proposée et résultats . . . . .	98
1.3	Effet du maillage sur la forme . . . . .	103
2	Introduction aux simulations numériques discrètes . . . . .	105
2.1	Méthodes de résolution . . . . .	105
2.2	La dynamique des contacts (CD) . . . . .	108
3	Conclusion . . . . .	112
<b>5</b>	<b>Etude de l'effet de la forme sur le comportement mécanique</b>	<b>115</b>
1	Etat d'équilibre statique . . . . .	116
1.1	Paramètres numériques . . . . .	116
1.2	Vérifications de l'équilibre mécanique . . . . .	118
1.3	Etude préliminaire des effets de forme et du frottement . . . . .	120
1.4	Synthèse . . . . .	121
2	Cisaillement simple maintenu . . . . .	121
2.1	Paramètres numériques . . . . .	121
2.2	Vérifications de la qualité des calculs numériques . . . . .	122
2.3	Effets de la forme et du frottement . . . . .	123
2.4	Comparaison aux grains historiques de ballast . . . . .	127
3	Conclusion . . . . .	130

---

<b>IV</b>	<b>Ballast : Les applications</b>	<b>135</b>
<b>6</b>	<b>Applications</b>	<b>137</b>
1	La classification . . . . .	139
2	Tri de populations de grains de ballast selon le degré d'usure . . . . .	140
2.1	Méthodes de caractérisation . . . . .	143
2.2	Comparaison des performances . . . . .	149
3	Tri de populations de grains de ballast provenant de différentes carrières	151
3.1	Méthodes de caractérisation . . . . .	151
3.2	Comparaison des performances . . . . .	155
4	Conclusion . . . . .	157
	<b>Conclusion</b>	<b>159</b>
	<b>Bibliographie</b>	<b>163</b>



## Première partie

# Ballast : Le matériau aux mille pouvoirs



# Chapitre 1

## Contexte industriel



Dans ce premier chapitre, une synthèse bibliographique est présentée dans le but d'expliquer le cadre, les motivations et le cheminement de ce travail de thèse.

Tout d'abord, nous commencerons par une brève description de la structure de la voie ferrée ballastée, des fonctions de ses composants, pour nous focaliser ensuite sur le ballast. Nous décrirons ses fonctions et ses spécifications actuelles mais également les enjeux auxquels SNCF est confrontée avec ce composant fondamental. Ensuite, nous présenterons les différents moyens de mesure de la forme concernant le ballast, puis les études expérimentales et numériques sur l'effet de la forme sur son comportement mécanique. Enfin, nous détaillerons la démarche scientifique proposée dans ce manuscrit afin de répondre aux questions que nous nous poserons au fur et à mesure de l'évolution de ce chapitre.

## 1 La voie ferrée ballastée

La voie ferrée ballastée est constituée de différents composants capables de supporter divers types de sollicitations longitudinales, transversales et verticales imposées par la circulation des trains(Fig. 1.1).

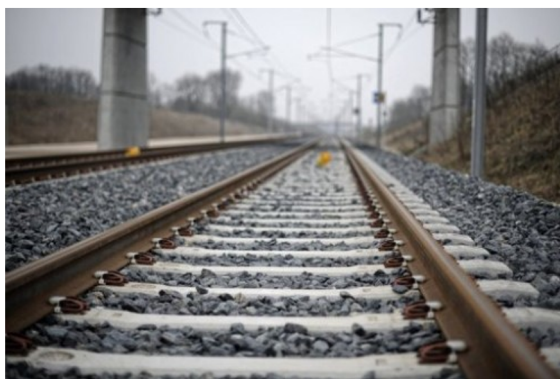


FIGURE 1.1 – La voie ballastée

Les lignes à grande vitesse (LGV) sont généralement composées de deux voies juxtaposées, et permettent une vitesse commerciale de 300 Km/h. La circulation des trains entraîne des sollicitations verticales et transversales qui engendrent des déformations irréversibles : des tassements verticaux et des étalements horizontaux de la voie, ainsi que des efforts longitudinaux (freinage, accélération). Afin d'assurer la durabilité et la résistance de la voie à ces sollicitations, ses caractéristiques sont normées en termes de géométrie et de caractéristiques mécaniques.

La voie à grande vitesse a un écartement standard de 1.435 m entre les deux files de rails. La plate-forme est la structure d'assise sur laquelle la voie repose. Elle fait 14 m de large avec une emprise de 50 m environ. Sa qualité dépend de la nature géotechnique du sol et des conditions géologiques et hydrologiques locales.

La structure d'une (LGV) ballastée est présentée sur la Fig. 1.2. Ses constituants peuvent être classés en deux groupes :

**L'armement** : il assure entre autres la transmission des efforts à la structure d'assise. Il est composé des rails, des attaches rail-traverse, des semelles sous rails et des traverses.

**La structure d'assise** : composée du ballast et des sous-couches sur lesquelles il repose. Sur LGV, l'épaisseur de la couche de ballast est en général de 30 cm sous les traverses.

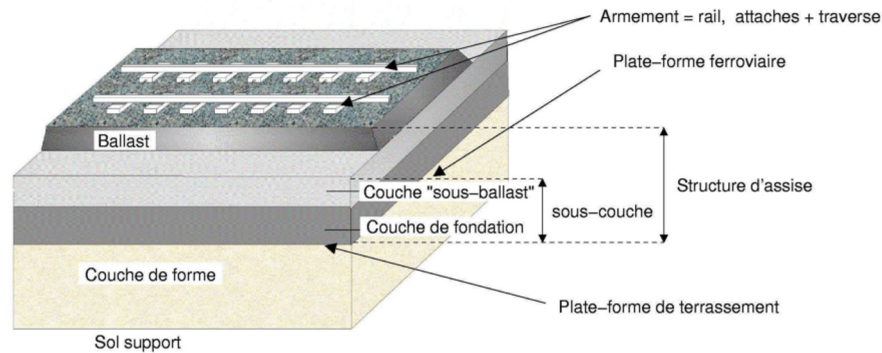


FIGURE 1.2 – Structure d'une LGV

### a. Les rails

En plus de guider les roues du train dans le sens transversal, les rails assurent la transmission des charges aux traverses. Ce sont deux fils parallèles de barres d'acier profilées mis bout à bout pour former la voie ferrée, et sont en contact avec les roues des véhicules. Les rails sont soumis à différents types de sollicitations statiques (charge des trains à l'arrêt) et dynamiques (circulations sur la voie), et sont dimensionnés pour garantir une bonne répartition des charges sur les appuis discrets que constituent les traverses. L'assemblage de plusieurs barres élémentaires de longueur fixe produit des Longs Rails Soudés (L.R.S.) qui sont soudés en voie et fixés sur les traverses par l'intermédiaire des attaches.

### b. Les attaches rail-traverse

Les attaches assurent la fixation du rail sur la traverse. Une semelle en élastomère dite semelle sous-rails est posée entre le rail et la traverse et permet une absorption importante des vibrations générées par le passage des trains. Elle permet de filtrer les efforts dynamiques produits par les charges exercées sur les traverses, notamment pour les hautes fréquences. Une deuxième fonction de la semelle sous-rails est qu'elle empêche l'abrasion de la traverse son usure sous le patin de rail. La Fig. 1.3 montre un exemple des attaches utilisées actuellement et de la semelle sous-rails.

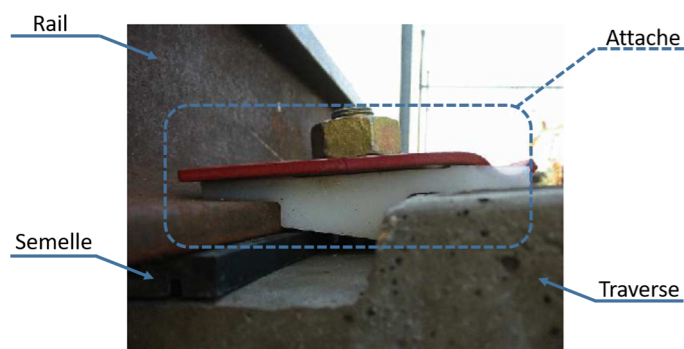


FIGURE 1.3 – Attache de type nabla

### c. Les traverses

Les traverses assurent la transmission des efforts entre le rail et la structure d'assise. De plus, elles permettent le maintien de l'écartement des deux files de rail et d'une certaine inclinaison du rail. Comme montré dans la Fig. 1.4, elles sont constituées d'une partie en béton (traverses mono-bloc) ou de deux parties (traverses bi-bloc). Les traverses utilisées classiquement mesurent 2.415 m de long et pèsent 245 kg. La distance (entraxe) entre deux traverses consécutives est généralement de 0.6 m.

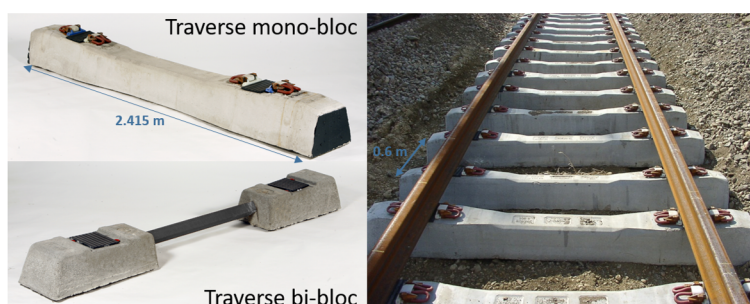


FIGURE 1.4 – Traverses mono-bloc et bi-bloc (gauche); traverses mono-blocs mises en place sur la voie (droite).

### d. Structure d'assise

Ce système multi-couche est dimensionné pour garantir le niveau suffisant de rigidité, de drainage et de tenue mécanique de la voie. De manière générale, la structure d'assise est constituée de la couche de ballast, de la sous-couche, et de la couche de forme. Selon le cas d'une voie classique ou d'une LGV, la composition de la sous-couche varie (Fig. 1.5). Une étude approfondie de cette structure peut être trouvée dans [1].

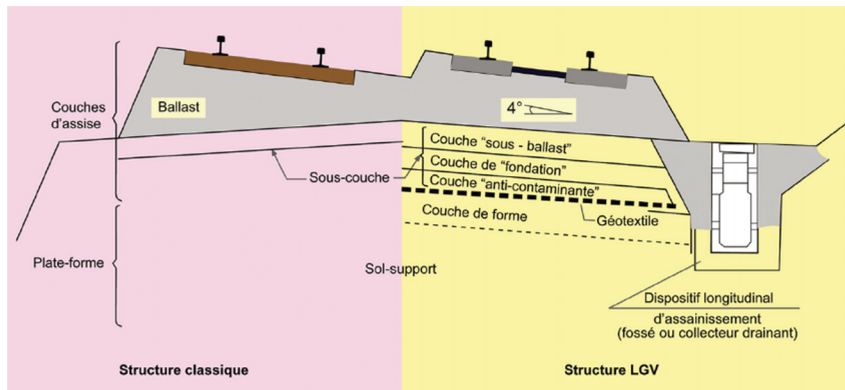


FIGURE 1.5 – Structure d’assise : cas d’une voie classique (gauche) et d’une LGV (droite) [2]

Chaque couche de la structure d’assise est dimensionnée pour remplir une fonction bien précise. Dans le cas d’une LGV, hormis la couche de ballast à laquelle une attention particulière est prêtée dans ce travail de thèse, la sous-couche protège la plate-forme contre l’érosion et assure une bonne répartition des charges au niveau de la plate-forme. Elle est constituée de 3 couches. D’abord, la couche de sous-ballast, dimensionnée pour améliorer la portance du sol, la répartition des charges et la protection de la plate-forme contre l’érosion et le gel. Ensuite, la couche de fondation protège la plate-forme contre l’action des engins lors de la construction de la voie. Enfin, la couche anti-contaminante qui protège les sous-couches supérieures de la remontée de particules fines de la plate-forme, et qui n’est pas toujours indispensable. La sous-couche repose sur la plate-forme qui est constituée d’une couche de forme à haute capacité de portance et à élasticité élevée.

De manière générale, la grande vitesse nécessite la maîtrise de la stabilité du matériel roulant. Une très grande qualité des infrastructures est alors exigée. Un élément crucial de la structure d’assise soutenant l’armement est au centre de ce travail : la couche de ballast. Nous y consacrons alors la suite du chapitre.

## 2 Le ballast

### 2.1 Fonctions et spécifications

La couche de ballast est placée dans la partie supérieure de la structure d’assise, sous les traverses. Elle est composée d’un matériau granulaire : le ballast (Fig. 1.6). Les grains de ballast proviennent du concassage de roches extraites dans des carrières de pierres dures (granit, rhyolite, diorite, grès).

Le ballast constitue un élément crucial pour la voie. Il remplit plusieurs fonctions :

1. La transmission et la distribution des efforts dus aux sollicitations statiques et dynamiques des circulations à la plate-forme.
2. L’absorption des vibrations acoustiques et mécaniques induites par les circulations.



FIGURE 1.6 – Le ballast : le matériau (gauche), et la structure couche dans la voie(droite)

3. L’ancrage longitudinal et latéral de la voie : la résistance au cisaillement permet de contenir les chargements longitudinaux et latéraux.
4. Le drainage immédiat des eaux pluviales dans la voie.
5. Une mise en place et une maintenance aisée de la voie.

Pour répondre de façon optimale aux fonctions qui lui sont assignées, le ballast doit satisfaire des critères de qualité définis dans la norme européenne NF EN 13450 (Août 2003) [3]. Ces critères portent sur 4 aspects :

- Granulométrie : le diamètre des grains de ballast (dimension maximale des grains lors du tamisage) varie entre 25 et 61.5 mm. Sur les lignes ferroviaires françaises actuelles, le fuseau de refus du ballast employé est représenté sur la Fig. 1.7.

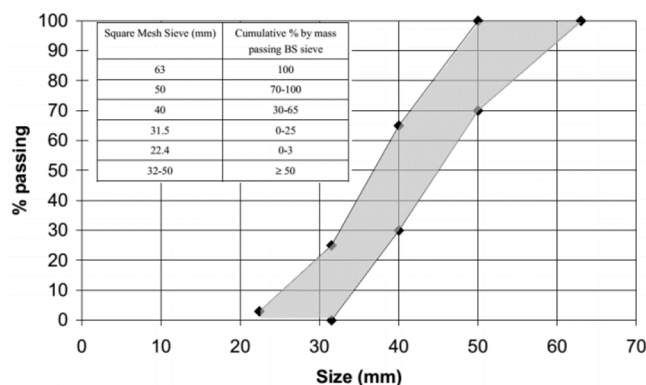


FIGURE 1.7 – Fuseau granulaire du ballast utilisé par SNCF [3, 4].

- Dureté : le niveau de dureté du matériau ballast est évalué grâce au coefficient de dureté globale (DRG). Afin de calculer ce coefficient, deux tests sont utilisés : l’essai au Micro-Deval [5], effectué à l’état humide et caractérisant la résistance à l’attrition, et l’essai Los-Angeles [6] qui caractérise la fragilité aux chocs. La valeur du DRG exigée varie en fonction du trafic et de la nature des traverses.

- Forme : les grains de ballast sont de forme polyédrique et possèdent des arêtes vives. Comme nous le détaillerons plus tard, les arêtes vives sont une caractéristique d'une angularité élevée et permettent d'assurer une bonne résistance au cisaillement. L'angularité, n'étant pour l'instant définie qu'intuitivement par l'apparence. D'une autre part, les grains allongés nuisent à la stabilité de la couche compactée, car ils sont susceptibles de se rompre et donc de perturber la stabilité du ballast : les aiguilles et les éléments plats ne doivent pas dépasser 92 mm. Le pourcentage en masse d'aiguilles ne doit pas dépasser 7%. Les éléments plats sont caractérisés par un coefficient d'aplatissement global qui représente le pourcentage total en masse de passant à l'ensemble des grilles et ne doit pas dépasser 12% [3].

Sur la voie, l'homogénéité du ballast est définie par la proportion d'éléments friables (se dégradant facilement en poudre) ou altérés qui ne doit pas dépasser 3%. Enfin, pour un fonctionnement optimal, les grains de ballast doivent être expurgés du sable, de la terre, des poussières, des débris et des fines pour ne pas influencer la rugosité et le frottement entre grains qui sont à l'origine de la résistance macroscopique. La masse relative des éléments passant au tamis de 1.6 mm doit être inférieure à 0.5%.

## 2.2 Maintenance et enjeux

La circulation des trains et les sollicitations qui en résultent entraînent un tassement progressif et hétérogène de la couche de ballast (tassement différentiel à cause du réarrangement des grains) dont la vitesse est impossible à prévoir. Ce tassement peut engendrer des défauts de nivellement qui nuisent au confort et à la sécurité des voyageurs s'ils sont trop importants. Pour identifier ces défauts de géométrie, des rames spéciales circulent sur les voies et permettent d'obtenir des mesures pour déterminer l'état de la voie, et préconiser une réduction de vitesse ou planifier les opérations de maintenance nécessaires.

L'opération de maintenance la plus courante est le bourrage. Cette opération permet de rendre à la voie son profil initial en corrigeant le nivellement longitudinal et transversal par vibration et serrage des grains de ballast. Elle est réalisée par le biais d'engins appelés bourreuses (Fig.1.8), et peut être appliquée à une partie ou à l'intégralité de la voie. Le bourrage est fait par cycles dont chacun est composé des étapes suivantes :

1. La traverse est soulevée jusqu'à une hauteur nécessaire pour rétablir une bonne géométrie.
2. Les bourroirs plongent dans le ballast et effectuent une vibration. Le ballast est ramené et compacté sous la traverse suite au serrage des bourroirs.
3. Les bourroirs sont remontés et la traverse est relâchée.

Certains paramètres (fréquence de vibration, force de serrage, etc.) permettent de contrôler l'intensité du bourrage.

Malgré une grande maîtrise des infrastructures ferroviaires et une bonne connaissance de leur fonctionnement, il reste difficile de cerner totalement les mécanismes de leur dégradation avec le temps. En particulier, la durée de vie d'une voie ferrée selon les seuils de sécurité fixés est généralement surestimée jusqu'à 3 fois sa durée

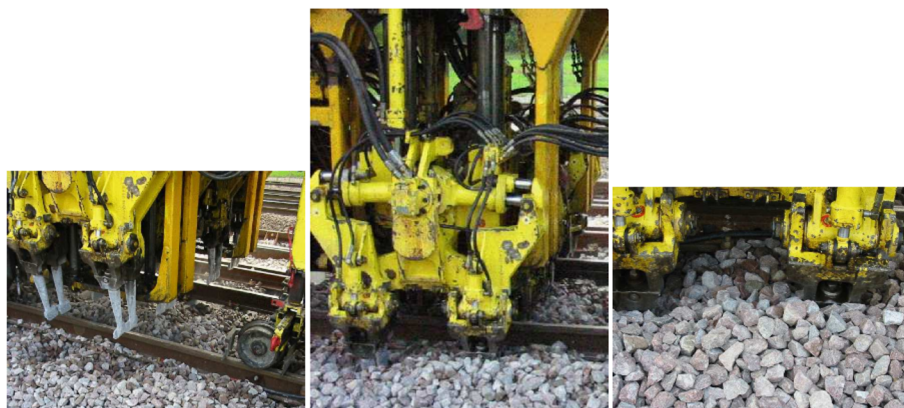


FIGURE 1.8 – Bourrage

réelle, et une voie se voit renouvelée au bout de 15 ans au lieu de 40 ans. Ceci engendre des opérations de maintenance plus fréquentes, et donc plus coûteuses, notamment avec l'augmentation de la vitesse de circulation sur LGV. Ces vitesses augmentent également les sollicitations imposées aux grains de ballast. Ces sollicitations provoquent avec le temps l'usure des grains, ce qui nécessite de plus en plus de bourrage. Or, le bourrage n'est pas totalement sans effet sur les grains. En effet, cette opération endommage elle-même les grains et provoque leur rupture (Fig. 1.9). La rupture des grains produit des particules plus petites et plus fines qui altèrent la granulométrie du milieu. Le bourrage permet, là encore, par la vibration de chasser les particules fines en émoussant les arêtes des grains. L'impact du bourrage et l'effet de chaque étape de son cycle n'est toujours pas complètement déterminé, mais la performance du bourrage diminue indubitablement avec le temps [7] alors que sa fréquence augmente (Fig. 1.10), sans oublier que certaines fonctions du ballast (comme le drainage) sont dégradées. Au bout de plusieurs opérations de maintenance, une opération de renouvellement du ballast (RB) ou de la voie et du ballast (RVB) doit être réalisée.



FIGURE 1.9 – Quelques grains de ballast suite au bourrage

La dégradation géométrique de la voie étant essentiellement due au tassement des couches de la structure d'assise, si la voie est fondée sur un bon sol support, le tassement est principalement causé par la couche de ballast [8]. Il est alors indispensable de bien comprendre le comportement non linéaire du ballast, et les facteurs influant ses performances afin de les optimiser, de limiter le tassement différentiel et d'augmenter la maintenabilité des voies pour maximiser leur durée d'exploitation.

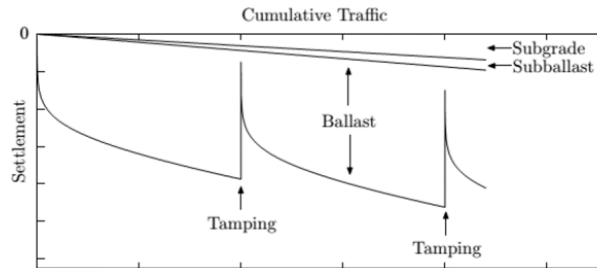


FIGURE 1.10 – Contribution relative des éléments de la structure d'assise au tassement de la voie : avec le temps, les opérations de maintenance sont plus fréquentes et moins performantes [7]

Les fonctions du ballast précédemment citées montrent que les attentes relatives à ce matériau sont principalement liées à ses performances mécaniques en tant que matériau granulaire, avec des propriétés de frottement interne (points 1 et 3) et d'amortissement (point 2). Son caractère de matériau grossier intervient directement pour sa perméabilité (point 4), mais aussi par les effets possibles du rapport de la taille des grains à celle des objets solides avec lesquels le matériau entre en contact (effets d'échelle).

Le comportement mécanique du ballast est défini par les aspects suivants :

- Sa minéralogie.
- La taille et la forme des grains.

Quelques années après la mise en service des premières lignes à grande vitesse, des travaux de recherche ont abouti à une augmentation des exigences sur la qualité de la roche utilisée pour produire le ballast ferroviaire. Les propriétés mécaniques des grains résultent de celles de la roche mère dont il sont extraits et ne présentent donc pas de moyens d'amélioration au-delà d'une sélection drastique des carrières. Au contraire, la géométrie des grains (taille et forme) est clairement un levier d'optimisation des performances du ballast.

Afin de relever le défi industriel fort que constitue, pour SNCF, l'amélioration des performances instantanées et de long terme de la voie ballastée, on doit donc mener à bien des recherches assez fondamentales sur le rôle de la géométrie des grains dans les propriétés mécaniques du matériau granulaire, en tant que matériau désordonné, à structure discrète.

Dans ce travail de thèse, nous nous intéressons donc à la question de la forme des grains de ballast, et de l'influence de celle-ci sur son comportement mécanique. La réponse à cette question nécessite d'abord une caractérisation de la forme des grains, puis ensuite une modélisation de la forme de manière à pouvoir étudier son influence sur le comportement mécanique à l'aide d'un outil déterminé par la suite.



La caractérisation de la forme débute par une mesure. Intéressons-nous aux différents moyens de mesure employés jusqu'à présent pour caractériser la forme des grains de ballast.

## 2.3 Mesure et caractérisation géométrique du ballast

Les spécifications imposées sur la forme sont pour la plupart qualitatives ou basées sur l'observation humaine. Le degré de précision de ces spécifications est directement lié aux moyens de mesure utilisés pour caractériser la forme. Nous aborderons alors les mesures et les caractérisations auxquelles ces mesures donnent accès selon leur potentiel croissant de précision.

### a. Tamisage

L'analyse granulométrique est une méthode de référence permettant de déterminer la grosseur des grains d'un matériau et le pourcentage des grains de chaque grosseur. L'essai consiste à fractionner au moyen d'une série de tamis (Fig. 1.11) un matériau en plusieurs classes granulaires de tailles décroissantes. On place ainsi une masse connue de grains au sommet d'une colonne de tamis classés par ouverture de maille décroissante de haut en bas puis on fait vibrer la colonne à l'aide d'une tamiseuse. Les dimensions des mailles et le nombre de tamis sont choisis en fonction de la nature de l'échantillon et de la précision attendue.

Les masses des différents refus ou celles des différents tamisats sont rapportées à la masse initiale du matériau, et les pourcentages ainsi obtenus sont exploités, soit sous leur forme numérique, soit sous une forme graphique : la courbe granulométrique (Fig. 1.12).



FIGURE 1.11 – Colonne de tamis

La granulométrie du ballast est représentée par  $d/D$  où  $d \geq 25$  mm est la dimension minimale des grains et  $D \leq 61.5$  mm est la dimension maximale des grains. Chaque matériau employé possède une tolérance granulométrique, puisque la granulométrie est basée sur un échantillonnage. Pour qu'un matériau soit conforme, il

faut que sa courbe granulométrique soit située à l'intérieur d'un fuseau de tolérance appelé fuseau de spécification du matériau. Pour le ballast (Fig. 1.13), le fuseau rouge correspond à la spécification préconisée par la norme. Le fuseau noir, établi suite à des contrôles de conformité à la norme [9, 10], correspond aux limites au-delà desquelles le ballast est refusé [3].

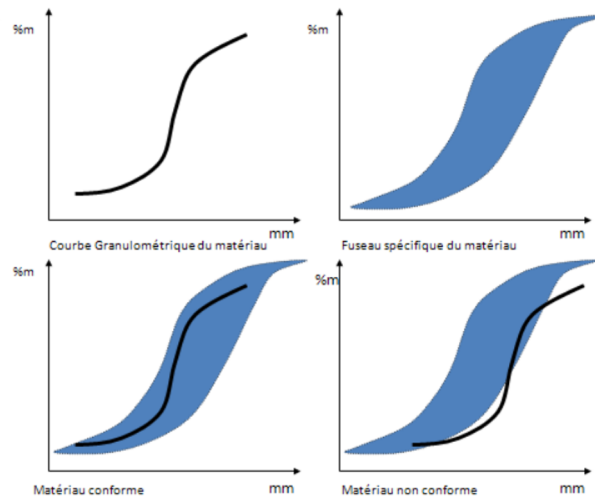


FIGURE 1.12 – Interprétation d'une courbe granulométrique

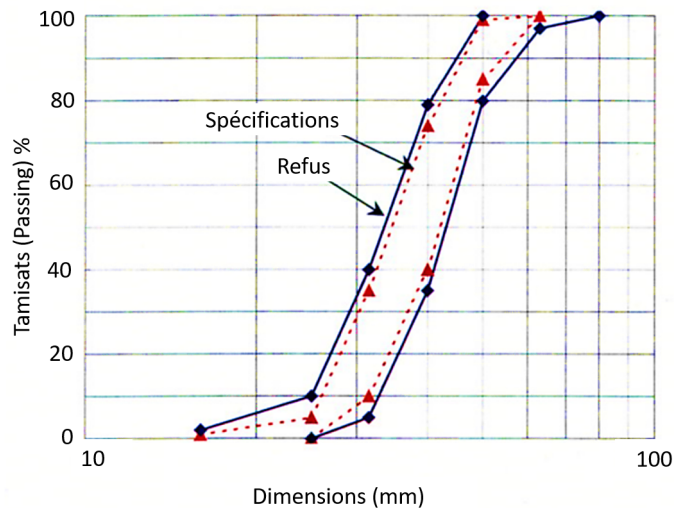


FIGURE 1.13 – Fuseau de spécification et limites de refus pour le ballast

Le tamisage fait partie des différents procédés de mesure de la taille des grains. Malgré toutes les précautions d'usage possibles et la précision du dispositif de mesure, une estimation "correcte" de la granularité suppose que l'échantillon analysé soit représentatif du milieu réel dont il a été extrait. En effet, l'échantillon ne représente qu'une infime partie d'un matériau granulaire dont l'étude totale est impos-

sible. Il convient alors de multiplier les échantillons en s'assurant que la méthode de prélèvement assure une reproductibilité suffisante.

Le tamisage ne fournit que l'estimation d'un diamètre équivalent des grains au sens où les bornes des classes granulométriques sont assimilées à l'ouverture des tamis. Pour des particules très rondes, l'estimation est fiable puisqu'elle correspond à un diamètre de sphère à laquelle on assimile les grains. Toutefois, pour des particules de forme complexe, la seule information d'un diamètre équivalent est réductive. Pour des grains allongés, anguleux et irréguliers comme le ballast, cette mesure reste limitée, mais est la plus exploitée pour qualifier la taille du ballast.

Pour le ballast, l'analyse granulométrique permet de calculer deux indices :

- L'indice de forme (SI) :  $SI = \frac{G}{E}$ , où  $G$  correspond à la dimension de la maille carrée minimale du tamis qui laisse passer le grain, et  $E$  est la distance minimale de deux plans parallèles tangents au grain.
- Le coefficient d'aplatissement : c'est le pourcentage en masse des grains vérifiant  $SI > 1.58$ . Ce pourcentage doit être supérieur à 35%.

### b. Pied à coulisse (Caliper)

C'est un instrument de mesure composé essentiellement de deux parties coulissantes l'une par rapport à l'autre (Fig. 1.14). Il permet de mesurer les cotes extérieures d'un objet.

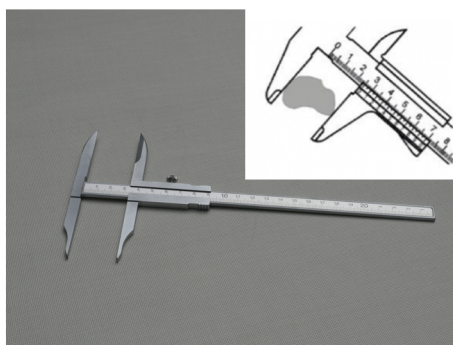


FIGURE 1.14 – Caliper

A partir de cet instrument manuel, une notion équivalente au diamètre classique a été définie [11] : le diamètre de Feret. Il est défini comme la distance normale entre deux plans tangents à la surface de la particule (en  $3D$ ) ou entre deux droites tangentes au contour de la particule (en  $2D$ ). Sa valeur dépend alors de l'orientation de la particule, et a besoin d'être calculée dans plusieurs directions (aléatoires ou déterminées par l'utilisateur) afin d'avoir un sens. Il est également possible de définir les diamètres maximal et minimal de Feret à partir du plus grand et plus petit rectangle entourant le contour de la particule (Fig. 1.15).

Dans le cas où cette mesure est manuelle, le choix de la plus grande et de la plus petite dimension est basé uniquement sur le jugement visuel de l'opérateur. Quoique plus précise que la granulométrie en terme de mesure de la géométrie et de caractérisation (discernement de deux grains de géométrie différente), cette méthode

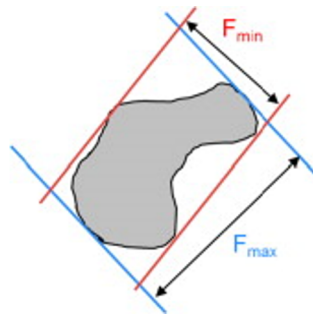


FIGURE 1.15 – Diamètres maximal et minimal de Feret

de mesure manuelle grain par grain nécessite beaucoup de temps, et ne permet qu'un contrôle limité des grandeurs à calculer.

Une méthode de mesure plus rapide et plus performante permet l'accès à la caractérisation donnée par la granulométrie et la mesure manuelle, ainsi qu'à un degré de précision plus élevé de mesure : l'analyse d'images.

### c. Analyse d'images

Ces techniques exploitent en entrée des images (Fig. 1.16). En utilisant un traitement d'image sur ordinateur, il est possible d'obtenir des contours 2D correspondant aux périmètres des particules (Fig. 1.17). Avec l'évolution des techniques d'imagerie et des logiciels de traitement de l'image, le recours à la mesure de la forme par analyse d'images est de plus en plus courant. Quelques exemples sont donnés dans [12, 13].

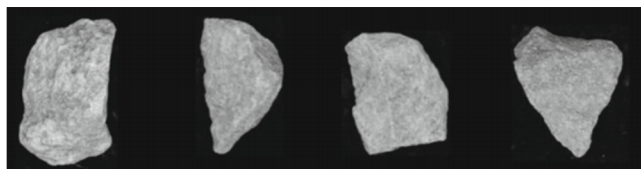


FIGURE 1.16 – Images de particules

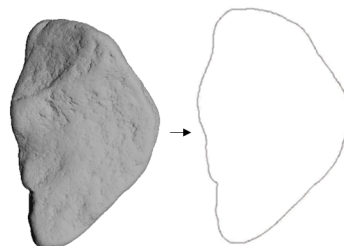


FIGURE 1.17 – Contours 2D récupérés à partir d'images

Les images obtenues permettent classiquement de mesurer plusieurs grandeurs, comme les diamètres de Feret [12], mais également les deux longueurs dites longueurs principales  $L$ ,  $S$  considérées comme la plus grande et la plus petite dimension, détaillées dans le chapitre suivant (Fig. 1.18). Il est également possible de calculer la surface, la rondeur, et la circularité à partir du contour 2D [12].

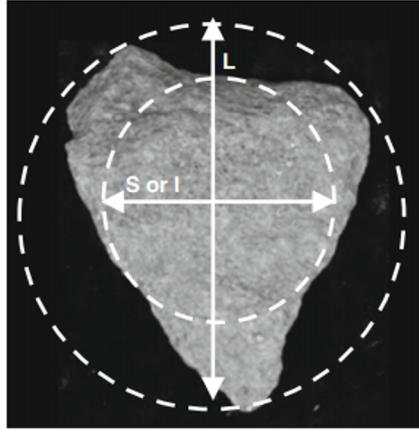


FIGURE 1.18 – Longueurs principales calculées pour un grain de ballast

Plus récemment, [14] proposent, en se basant sur les études faites par [15, 16] pour les particules, de calculer un indice d'angularité et un autre de texture de surface pour les contours 2D des grains de ballast en se basant sur des formes régulières connues.

L'analyse d'images des grains de ballast permet l'accès à plus de paramètres par rapport à l'analyse granulométrique et à la mesure manuelle qui traitent uniquement de la taille des grains. Elle a également un avantage considérable en terme de rapidité et d'automatisation. Toutefois, elle présente une limitation par rapport au cas réel 3D, car elle traite exclusivement un plan considéré représentatif de la forme tridimensionnelle. Un cas particulier de l'application de l'analyse d'images, notamment pour le ballast, est le *VDG40* [17, 18].

En effet, les auteurs proposent de définir et de calculer un paramètre d'angularité des grains de ballast en exploitant des images. Ce paramètre est obtenu à partir de contours 2D de grains à l'aide d'un appareil appelé *VDG40* (Fig. 1.19).

En utilisant cet appareil, les auteurs récupèrent des silhouettes 2D de grains (Fig. 1.20) dont ils extraient les contours décrits par des coordonnées cartésiennes. Ensuite, en utilisant le découpage des contours en segments, proposé par la transformation de Hough [19], les auteurs proposent deux paramètres moyens sans dimension, un de l'angularité (ANG) et l'autre de la rugosité (RUG) :

$$ANG = \frac{1}{n} \sum_{i=1}^n [ang(\alpha_i)] = \frac{1}{n} \sum_{i=1}^n \left[ 1 - \frac{\alpha_i}{180} \right] \quad (1.1)$$

où :  $\alpha_i$  l'angle entre deux segments adjacents et  $n$  est le nombre d'angles pris en compte ( $n \leq 8$ ).

$$RUG = mean \left( \frac{L_{perim} - L_{seg}}{L_{perim}} \right) \quad (1.2)$$

où  $L_{seg}$  est la longueur des segments en pixels, et  $L_{perim}$  la longueur du contour correspondant en pixels (Fig. 1.20).

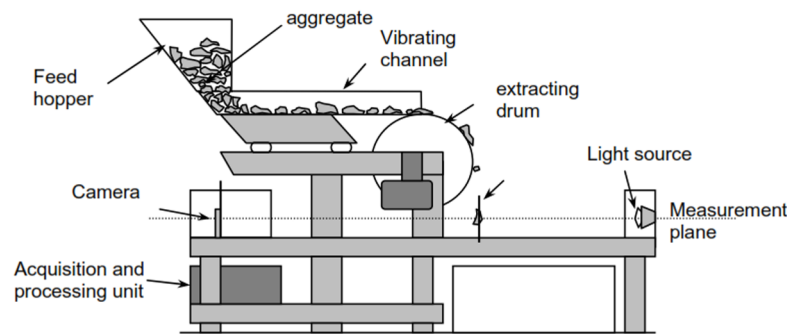


FIGURE 1.19 – Appareil VDG40 [18]

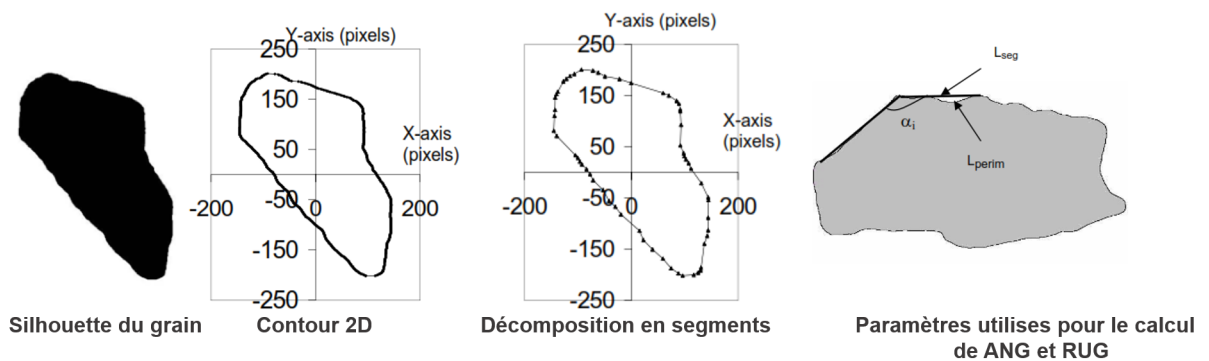


FIGURE 1.20 – Etapes de traitement pour le calcul des paramètres d’angularité et de rugosité selon [18]

Nous abordons maintenant une autre méthode de mesure : le scan 3D.

#### d. Scan 3D

Pour ce qui est de la quantification de la géométrie complète des grains de ballast, il est clair que le tamisage n’est qu’une approximation grossière de la taille des grains et ne permet qu’un accès à une caractérisation limitée de la forme. Les méthodes basées sur l’analyse d’images semblent fournir davantage d’informations [20–22] et éliminer la subjectivité associée à la plupart des outils manuels de mesure.

Néanmoins, plusieurs études [13, 22–24] basées sur l’exploitation d’images ou sur du scan pseudo 3D ont montré la limitation de ces méthodes à fournir une information complète et pertinente sur la forme des grains. Des différences notables entre les mesures 2D et 3D de la forme ont été observées par [25]. En effet, comme

montré dans [26] pour le cas du ballast, l'analyse 2D peut sous-estimer (entre autres) la véritable sphéricité des grains.

Plus récemment, des études proposent le recours au scan 3D afin d'obtenir une information exhaustive sur la forme des grains de ballast et de permettre une mesure plus pertinente de leur morphologie complexe [26–30].

Selon le contexte et le besoin, il existe plusieurs techniques de scan 3D présentées et détaillées dans le chapitre suivant sur la caractérisation de la forme. Quoiqu'il en soit, la sortie du scan 3D correspond à la surface des grains échantillonnés par des points dans l'espace 3D : on parle alors de nuages de points représentés simplement par un ensemble de coordonnées cartésiennes. La résolution et le niveau de précision dépendent alors des méthodes de scan mises en œuvre pour ce faire. Selon le logiciel de traitement utilisé, la sortie peut être un maillage 3D ou un nuage de points (Fig. 1.21).

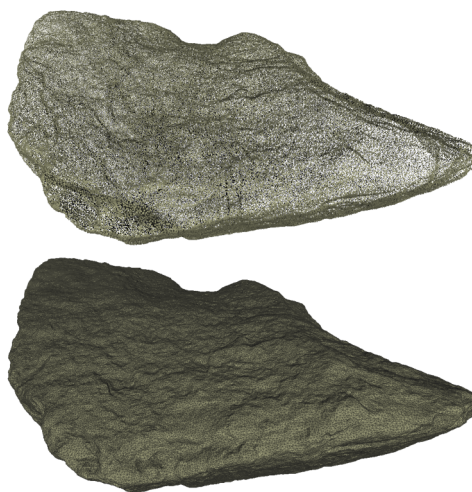


FIGURE 1.21 – Résultats d'un scan 3D d'un grain : nuage de points 3D(haut) ou Maillage 3D (bas).

Pour le ballast, [26] proposent d'utiliser le scan 3D par laser (Fig. 1.22) afin de calculer les trois longueurs principales  $L$ ,  $I$  et  $S$  mesurées en calculant les côtés de la boîte entourant le grain (boîte englobante) comme déjà proposé par [31]. En fonction de ces longueurs, ils ont calculé d'autres paramètres de forme comme le rapport d'élongation ( $I/L$ ), d'aplatissement ( $S/I$ ) comme défini par [32,33] ainsi que la sphéricité [34], et ont défini un indicateur de proximité du grain à un ellipsoïde.

Les résultats des études évoquées précédemment montrent qu'il est possible d'obtenir des informations plus précises et plus objectives comparées à la granulométrie, aux méthodes manuelles de mesure, et à l'analyse d'images sur la forme lorsqu'il s'agit du ballast. Pour d'autres matériaux comme le sable, les techniques de scan 3D sont de plus en plus utilisées pour des fins de caractérisation de la morphologie. Néanmoins pour le ballast, ces méthodes restent encore très peu exploitées pour caractériser la forme et se limitent à des mesures d'indicateurs scalaire simples.

Le peu d'exploitation des méthodes avancées comme le scan 3D est indirectement lié à la simplicité des spécifications retrouvées dans les normes. Dans le but de motiver ou non l'évolution des normes concernant le ballast ainsi que les référentiels

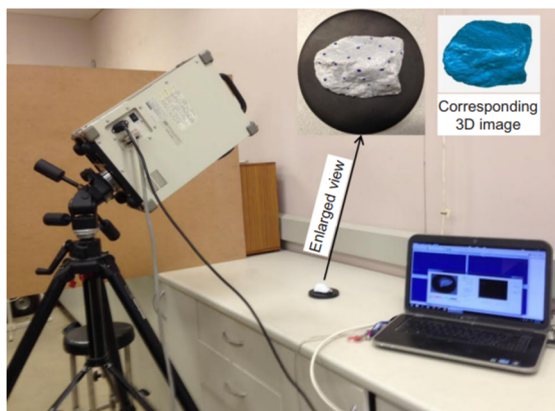


FIGURE 1.22 – Exemple d'un dispositif de scan 3D d'un grain de ballast

internes SNCF décrivant l'application de ces normes, nous choisissons dans ce travail de thèse, d'utiliser ce moyen de mesure associé à une exploitation minutieuse pour caractériser la morphologie des grains.

## 2.4 Effet connu de la forme sur le comportement mécanique du ballast

Une autre motivation de ce travail vient de l'influence observée de la forme sur le comportement mécanique des matériaux granulaires de manière générale, et sur le ballast en particulier. Cet impact est mis en évidence dans plusieurs études expérimentales et numériques sur l'effet de forme.

### a. Etudes expérimentales

Parmi les études expérimentales sur l'effet de forme, nous ne notons pas d'études concernant le ballast. Néanmoins, nous pouvons donner l'exemple des études menées par [35] ou encore par [36] sur les assemblages granulaires de manière générale.

D'abord, [35] se sont intéressés à la mesure de la réponse en contrainte d'assemblages de grains de 14 formes différentes : des sphères, des formes convexes (tétraèdres, bipyramides triangulaires) ainsi que des géométries plus complexes non convexes (hexapodes). Ensuite, à l'aide d'une impression 3D, 5500 particules de chaque forme sont imprimées. En mesurant la relation contrainte-déformation sous compression quasi-statique, les auteurs se sont focalisés sur la réponse macroscopique globale. Pour chaque forme différente, des essais triaxiaux ont été réalisés pour différentes pressions de confinement. En calculant plusieurs paramètres comme le module d'Young etc., les auteurs tirent les conclusions suivantes quant à l'effet de la forme :

1. La forme peut changer le module effectif d'Young et la contrainte d'élasticité du grain d'environ un ordre de grandeur.
2. Les pressions de confinement élevées réduisent la dépendance du module d'Young à la forme.



3. Les interactions entre polyèdres facettés se font à travers plusieurs types de contacts dont les surfaces varient sous compression. D'une autre part, pour d'autres types de particules comme les hexapodes et les sphères, les interactions se font uniquement selon des contacts ponctuels.
4. Enfin, les polyèdres très facettés semblent fournir une résistance plus élevée que d'autres formes, moins sensibles au confinement.

Ces études permettent de souligner l'effet de la forme sur le comportement mécanique macroscopique des grains.

D'un autre côté, [36] se sont focalisés sur l'écoulement de grains sphériques puis allongés par une trémie. L'essai a été modélisé expérimentalement (le silo a été rempli en fermant l'orifice, puis l'orifice a été ouvert pour libérer l'écoulement des grains). Ensuite, à l'aide d'une caméra, la configuration de la trémie bouchée a été explorée. Le système est perturbé par la suite pour générer l'écoulement de nouveau, et ainsi de suite pendant 30 fois pour chaque forme de particules. Les auteurs se sont intéressés à l'orientation, à la distribution de densité et au réarrangement des grains, ainsi que du nombre de grains formant la dernière couche bloquant l'écoulement. Les conclusions des auteurs quant à la forme peuvent être résumées en deux points :

1. On remarque une orientation privilégiée des particules, et donc un meilleur arrangement dans la zone d'écoulement du silo.
2. Le nombre de grains formant la couche bloquante du silo est supérieur dans le cas des grains allongés par rapport aux grains sphériques de même volume.

## **b. Etudes numériques**

Les études numériques sur l'effet de la forme sur le comportement du ballast sont faites par le biais de simulations numériques discrètes (DEM) [37,38] afin d'effectuer différents tests mécaniques. Ces simulations permettent de calculer des paramètres mécaniques qui servent d'éléments de comparaison entre plusieurs configurations de forme. Au-delà de l'outil, une composante cruciale est le modèle 3D représentant les grains en questions, et utilisé comme donnée d'entrée en simulation. En effet, pour explorer l'effet de la forme nous distinguons deux approches : une variation de la primitive géométrique, ou une variation des paramètres de cette primitive. La première approche signifie un choix radical du modèle (sphères, assemblages de sphères, pinacoïdes, polyèdres, etc.), tandis que le deuxième consiste à une variation des paramètres d'un même modèle (angularité, nombre de faces du maillage, etc.). Ces deux approches sont différentes en terme de degré de finesse et de variation : un changement de primitive est grossier/discret en comparaison du changement du paramètre de la primitive, considéré comme fin/continu. En terme de comportement mécanique, nous nous attendons alors à un effet plus significatif de la forme en changeant la primitive qu'en faisant varier un paramètre sur le même modèle. Le choix de l'approche dépend de l'aspect que l'on cherche à mettre en évidence.

[39] soulignent l'importance d'un modèle représentatif de la forme pour les simulations DEM, en modélisant les grains de ballast par des sphères ou par des assemblages de sphères. En proposant une méthode de génération de grains par assemblages de sphères, ils ont montré que les assemblages de sphères mènent à une plus grande homogénéité des contraintes d'autant plus que l'angularité est élevée.

Un test de chargement sur traverse après compaction effectué sur les sphères et les assemblages de sphères a montré que le réarrangement des assemblages de sphères produit une réponse en contrainte-déformation plus réaliste que celle donnée par les sphères.

[40, 41] se sont intéressés aux assemblages de grains polyédriques (modélisant des grains de ballast) sous cisaillement quasi-statique, ainsi que sous chargement triaxial. L'effet de forme est analysé en comparant des assemblages granulaires sous même conditions et même caractéristiques à part le degré de sphéricité des particules. Les conclusions de cette études peuvent être résumées comme suit :

1. Les assemblages de polyèdres présentent une résistance au cisaillement et une dilatance nettement supérieures à celles des assemblages de sphères. Ceci se traduit par un angle de frottement interne plus élevé ( $34^\circ$ ), mais une relation contrainte-dilatance similaire.
2. Les assemblages de polyèdres présentent une connectivité plus importante que celle des sphères.
3. Dans le cas des polyèdres, le réarrangement des grains est plus lent, et la durée de vie des contacts est plus grande que celle des sphères.
4. L'origine de la différence dans la résistance au cisaillement réside dans l'anisotropie des forces. Pour les polyèdres, la présence des contacts face-face entraîne une anisotropie importante des forces, représentant la contribution la plus importante au déviateur des contraintes.

Dans une étude 2D faite par [42], l'effet de l'élongation des particules sur le comportement quasi-statique d'un milieu granulaire cisailé par simulations DEM a été exploré. En utilisant des rectangles aux côtés arrondis (appelés têtes) caractérisés par 8 valeurs différentes valeurs d'un paramètre d'élongation (paramètre  $\eta$  variant de 0 à 1 décrivant une évolution d'un cercle à une ligne fine), les auteurs ont analysé quelques propriétés microscopiques et macroscopiques d'une compression bi-axiale appliquée à des assemblages de 13000 particules. Cette étude a permis les conclusions suivantes quant à l'effet de la forme :

1. La résistance au cisaillement augmente linéairement en fonction du paramètre d'élongation choisi.
2. La mobilisation des forces de frottement et de leur anisotropie augmentent avec le paramètre d'élongation.
3. La transmission des forces se fait essentiellement par des contacts tête côté, ce qui représente la plus grande partie des contacts pour la plupart des particules allongées.
4. La compacité augmente avec l'élongation, puis décroît d'autant plus que les particules sont allongées.
5. Le nombre de coordination ne suit pas cette tendance : Les particules allongées sont plus lâches mais plus fortement connectées.

Nous notons à travers cette études l'effet significatif de la forme même dans un cas bidimensionnel fortement simplifié, autant à l'échelle macroscopique qu'à l'échelle microscopique.

D'autres études 2D [43, 44] et plus tard en 3D [45, 46] par le biais de simulations DEM ont été effectuées, en utilisant des disques ou des polygones et se sont

intéressées à l'effet d'un paramètre d'angularité des particules sur la résistance au roulement, la résistance au cisaillement, la compacité, les anisotropies des forces et les distributions des forces de contact. Les principales conclusions tirées peuvent être résumées comme suit :

1. En comparant des disques à résistance de roulement croissante avec des polygones à nombre de côtés croissant, il est possible d'associer la résistance au roulement au nombre de côtés du polygone.
2. La résistance au cisaillement augmente avec l'angularité, alors que la compacité diminue vers une valeur constante, de telle manière que les assemblages composés de particules angulaires sont plus lâches mais possèdent une plus grande résistance au cisaillement.
3. La stabilité des assemblages dépend fortement de la présence des contacts côté-côté et de la mobilisation des forces de frottement, toutes les deux favorisées dans les cas des particules plus anguleuses.
4. Dans le cas 3D avec des polyèdres (grains de ballast), en faisant varier le nombre de faces de 8 à 596, les auteurs remarquent une augmentation de la résistance au cisaillement avec l'angularité (ici définie comme l'angle moyen externe entre deux faces adjacentes d'un grain maillé (Fig. 1.23)) jusqu'à une valeur de saturation. Quant à la compacité, on note une augmentation vers un pic puis une diminution pour les grains les plus anguleux. On remarque également une anisotropie plus importante des forces, ce qui entraîne l'augmentation de la résistance au cisaillement.

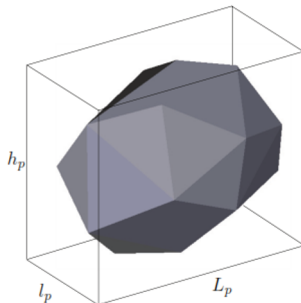


FIGURE 1.23 – Exemple de grains utilisés en simulation

Ainsi, dans le cas 2D comme dans le cas 3D, la stabilité du matériau, le réarrangement des grains, la mobilisation du frottement ainsi que la connectivité sont conditionnés en partie par l'angularité.

Une exploration de l'effet de la non-convexité à l'aide des simulations DEM est proposée par [47]. En utilisant des particules 3D constituées d'assemblages de 4 sphères qui se recouvrent, les auteurs proposent de définir la non-convexité par la position relative des sphères. Les auteurs remarques d'abord une augmentation de l'angle de frottement interne en fonction de la non-convexité, et comme pour l'angularité, une augmentation de la compacité jusqu'à une valeur de pic suivie d'une diminution. Un autre point de ressemblance avec l'angularité est que les assemblages

non convexes sont plus lâches mais présentent une plus grande résistance au cisaillement. A l'échelle microscopique, la non-convexité entraîne davantage de contacts entre les grains et de blocage (enchevêtrement), ce qui favorise l'anisotropie des forces et la mobilisation du frottement.

Dans une étude plus récente [48] proposent d'exploiter le scan 3D de grains de ballast provenant de 5 matériaux différents (grains fracturé, recyclés, etc.) afin d'explorer l'effet de la rondeur (définie comme le rapport des surfaces respectives d'un ellipsoïde de même volume et de la surface de la particule) sur la résistance au cisaillement, avec comme référence de ballast très arrondi des cailloux de rivière (A noter que l'auteur remarque que 90% des grains fracturés ont une rondeur supérieure à 0.75 tandis que 90% des cailloux de rivière ont une rondeur supérieure à 0.6). Par le biais de tests triaxiaux et d'un modèle de rupture de Mohr-Coulomb développé dans ce papier, les résultats principaux consistent à l'augmentation de l'angle de frottement interne avec la rondeur.

## 3 Synthèse

### 3.1 Discussion

Les études expérimentales et numériques présentées montrent que la forme des particules affecte significativement le comportement macroscopique et microscopique des matériaux granulaires. Malgré le faible nombre d'études dédiées au ballast, elles attirent notre attention sur la nécessité de caractériser finement la forme pour espérer évaluer quantitativement son effet sur le comportement mécanique. De plus, dans le cas d'étude numérique cela est intimement couplé au choix d'une discrétisation des grains à travers un type de primitive géométrique et aux variations paramétriques qu'elle permet.

C'est précisément dans ce cadre de la simulation numérique du ballast par éléments discrets que ce travail s'inscrit en continuité de ceux entrepris à SNCF depuis 10 ans [10, 38, 41]. L'objectif est d'aborder le sujet complexe de l'effet de la forme en proposant une approche pertinente. La recherche de pertinence se confronte principalement à deux grandes problématiques que nous aborderons : la modélisation des grains d'une part, et la discrétisation couplée à la mécanique de l'autre.

En ce qui concerne la modélisation des grains, il est capital de disposer d'une information précise et représentative de la variabilité de la forme des grains de ballast. Comme illustré dans ce chapitre, cela est possible en utilisant un scan 3D et en représentant chaque grain comme un nuage de points. Cependant, nos ambitions ne pourront être atteintes que si nous disposons d'un grand nombre de scans de grains uniques et que nous exploitons la totalité de l'information contenue dans cette collection de scans. Disposer d'un grand nombre de scans maximise les chances de travailler sur un échantillon non biaisé et la création d'une telle base de données ne repose que sur l'engagement des moyens nécessaires. Par contre, son exploitation fine, principalement à travers une méthode de caractérisation sensible aux variabilités de forme, est un enjeu central de ce travail et au sein de la communauté scientifique. Un enjeu associé est naturellement l'extension à l'infini de l'échantillon de formes par la proposition d'une méthode de génération de grains virtuels représentatifs. Les réponses apportées à ces enjeux permettront notamment la bonne maîtrise du

paramètre d'entrée qu'est la forme dans les simulations.

Les simulations en éléments discrets nécessitent de changer de discrétisation pour passer d'un nuage de points à un maillage. Au-delà de l'automatisation et du paramétrage du maillage de chaque grain, un des enjeux forts de ce travail est d'explorer l'effet de la dualité forme/maillage sur le comportement mécanique macroscopique et microscopique. En effet, un nuage aussi détaillé soit-il, est représenté ici par un maillage dont le nombre de sommets est fini et limité. La limite est principalement imposée par les temps de calcul admissibles qui augmentent approximativement rapidement avec le nombre de sommets par grain et le nombre de grains. Si le choix d'un nombre de grain dans une simulation peut être raisonnablement fait par rapport à la littérature, le choix de la résolution du maillage (nombre de sommet et taille des éléments) reste une question complexe et peu abordée. En effet, une réponse claire nécessite de séparer, dans la réponse mécanique d'un système, ce qui provient de l'approximation de la forme par un maillage de ce qui provient de la forme elle-même. En cela il est nécessaire d'évaluer l'erreur purement morphologique que produit le maillage. La difficulté principale découle du couplage potentiel de cet effet de maillage à celui du frottement microscopique. En effet, malgré une forme de grains peu fidèle, [10, 41] ont obtenu des résultats mécaniques intéressants en jouant sur le coefficient de frottement microscopique. D'une manière générale, nous formulons ce problème comme « l'équivalence forme-frottement » et nous l'aborderons comme un sous-problème intrinsèque du traitement de l'effet de la forme des grains par simulation numérique discrète.

### 3.2 Le ballast dans tous ses états

Nous choisissons pour ce travail de thèse de suivre la démarche suivante, exploitant plusieurs "états" différents des grains (grains réels, nuages de points, polyèdres), imposés par les moyens de mesures et les différents outils mis en œuvre. Le logigramme de la Fig. 1.24 résume ces étapes. En partant de grains réels, nous effectuons un scan  $3D$  qui fournit en sortie des nuages de points, à partir desquels nous proposons une méthode de caractérisation de la forme, ainsi qu'une méthode de génération de grains virtuels. Suite au maillage, nous passons de nuages de points à des polyèdres facettisés, utilisés dans les simulations numériques dans le but d'évaluer l'effet de la forme sur le comportement mécanique. Nous terminons cette étude par quelques applications des méthodes proposées au long des chapitres. Les chapitres suivants discutent ces étapes en détail, et permettent chacun d'apporter les lumières nécessaires pour éclairer notre chemin vers la compréhension de ce matériau complexe.

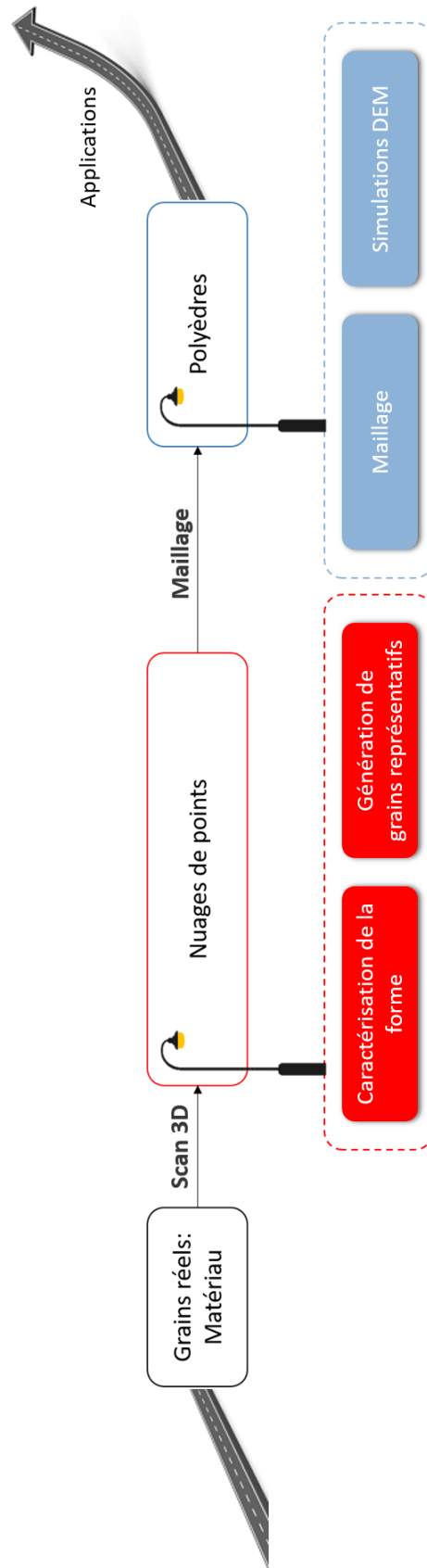


FIGURE 1.24 – Logigramme : plan de l'étude



## Deuxième partie

Ballast : Les nuages aux mille points





## Chapitre 2

### Caractérisation de la forme des grains de ballast

En science comme en industrie, la compréhension du comportement des matériaux granulaires passe inévitablement par la caractérisation de leurs propriétés. La caractérisation est généralement définie comme un processus analytique aboutissant à une définition, une description puis une classification des éléments d'un ensemble. Elle consiste à marquer le caractère de ces éléments, et à mettre en relief leurs traits dominants. Dans ce chapitre, nous nous intéressons particulièrement à la caractérisation de la forme des grains de ballast. Pour répondre aux objectifs que nous nous sommes fixés, la caractérisation de la forme (Fig. 2.1) doit répondre à certains enjeux :

1. La méthode de mesure permettant d'obtenir les données d'entrée doit être facilement réalisable, et doit assurer une bonne représentativité des données du cas réel en terme de forme.
2. La méthode de description de la forme doit être exhaustive, optimale en temps et en quantité d'information (représenter le plus d'information possible avec le moins de fonctions possibles).

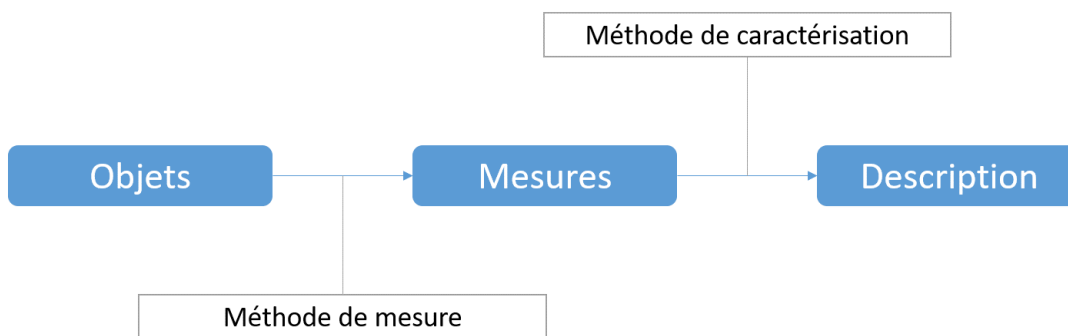


FIGURE 2.1 – Etapes d'un processus de caractérisation

Comme précisé dans le chapitre 1, notre choix de la méthode de mesure de la forme du ballast se porte sur le scan 3D. Après une analyse des différentes méthodes possibles de caractérisation de la forme, nous proposerons une nouvelle méthode de caractérisation. Les étapes de l'application de cette méthode à notre cas, ainsi que les résultats et la validation de sa performance sont détaillés.

## 1 Lumière sur le vocabulaire de la « forme »

Avant de présenter les différentes méthodes de caractérisation, arrêtons-nous d'abord brièvement sur le sens de la forme, et les différents termes utilisés dans la littérature pour la désigner. La forme peut également être désignée par le terme générique "morphologie" pour décrire la représentation géométrique externe d'une particule dans son ensemble. La plupart des auteurs [33, 49] définissent cette représentation sur trois ordres différents selon 3 échelles décroissantes par rapport à la taille du grain :

- 1<sup>er</sup> ordre : dimensions de la particule, longueurs principales, sphéricité, etc.
- 2<sup>me</sup> ordre : angularité de la particules, sa rondeur, etc.

— 3<sup>me</sup> ordre : rugosité/texture de la surface de la particule.

Nous utilisons le terme forme pour rassembler tous ces aspects. Dans la section suivante, nous aborderons les indicateurs relatifs à chaque ordre en allant de l'échelle la plus grossière à la moins grossière.

## 2 Méthodes de caractérisation de la forme de particules

Dans toute cette partie, il est important de distinguer entre la méthode de mesure ( Scan 3D, tomographie, etc.) qui fournit en sortie la mesure, et la méthode de caractérisation qui donne un sens à cette mesure et permet de la qualifier. La Fig. 2.2 permet de faire cette distinction et montre les méthodes étudiées par la suite. Dans ce travail, quelle que soit la méthode de mesure, nous choisissons de nous placer dans le cas 3D, et nous considérons un grain  $G$  représenté par un nuage  $\zeta^N = \{\zeta_i, 1 \leq i \leq N\}$  de  $N$  points référencés par leurs coordonnées cartésiennes  $x_i, y_i, z_i$ .

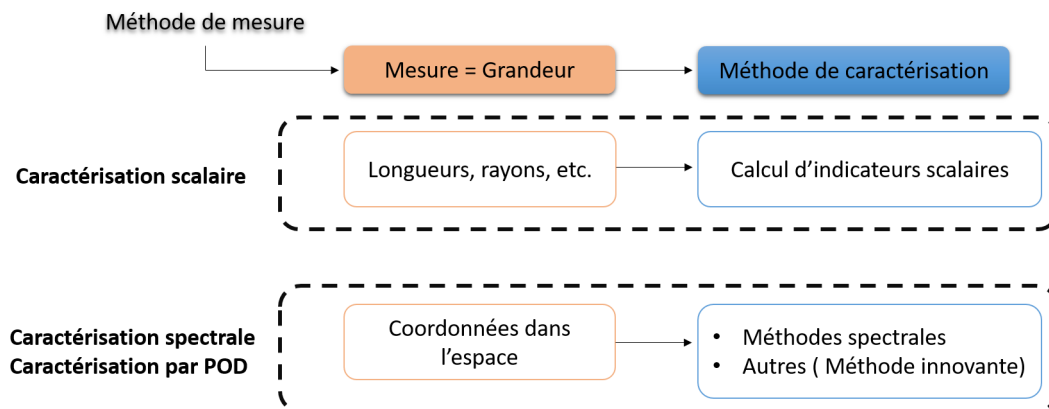


FIGURE 2.2 – Schéma des approches de caractérisation étudiées

### 2.1 Caractérisation scalaire

La caractérisation de la forme des matériaux granulaires est classiquement faite à l'aide d'indicateurs scalaires synthétisant la forme de la particule dans son ensemble. Parmi ces indicateurs, ceux qui sont également des mesures (surfaces, volumes, rayons moyens et longueurs principales) que nous appellerons indicateurs simples, et ceux qui sont calculés à partir de ces mesures (aplatissement, élongation, rapport d'aspect, sphéricité, angularité, rondeur, texture de surface, diagrammes, corrélations) que nous appellerons indicateurs avancés. Chacun de ces indicateurs est alors représenté par un scalaire unique.

## a. Indicateurs simples

**a.1. Surface, volume, rayon moyen** En partant d'une description en coordonnées cartésiennes, une conversion en coordonnées sphériques  $\theta_i$  et  $\phi_i$ , où  $\theta_i \in [0, 2\pi]$  et  $\phi_i \in [0, \pi]$  représentent respectivement la longitude et la colatitude (complémentaire de la latitude) (Fig. 2.3) permet d'exprimer le volume  $V$  et la surface extérieure  $S$  comme suit :

$$V = \int_{\theta=0}^{\theta=2\pi} \int_{\varphi=0}^{\varphi=\pi} \sin(\varphi) \int_{r=0}^{R(\theta,\varphi)} r^2 dr d\theta d\varphi = \frac{1}{3} \int_{\theta=0}^{2\pi} \int_{\varphi=0}^{\pi} \sin(\varphi) R^3(\theta, \varphi) d\theta d\varphi \quad (2.1)$$

$$S = \int_S R^2(\theta, \varphi) \sin(\theta, \varphi) d\theta d\varphi \quad (2.2)$$

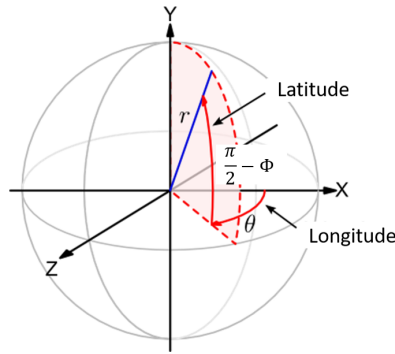


FIGURE 2.3 – Représentation en coordonnées sphériques

Par discrétisation en éléments finis en utilisant une base à pas angulaires constants ( $\Delta\theta_i = 2\pi/N = cte$  et  $\Delta\phi_i = \pi/N = cte$ ), le volume  $V$  et la surface  $S$  et s'expriment sous la forme suivante :

$$V \approx \frac{\Delta\theta\Delta\varphi}{3} \sum_{j=1}^d \sin(\varphi_j) R^3(\theta_j, \varphi_j) \quad (2.3)$$

$$S \approx \Delta\theta\Delta\varphi \sum_{j=1}^d \sin(\varphi_j) R^2(\theta_j, \varphi_j) \quad (2.4)$$

On définit également le rayon moyen du grain  $G$  :

$$R_m = \frac{1}{S} \int_S R^3(\theta, \varphi) \sin(\theta, \varphi) d\theta d\varphi \approx \frac{\Delta\theta\Delta\varphi}{S} \sum_{j=1}^d \sin(\varphi_j) R^3(\theta_j, \varphi_j) \quad (2.5)$$

**a.2. Longueurs principales** Pour qualifier la forme des particules, on calcule classiquement leur « longueur », leur « largeur » et leur « épaisseur ». Plusieurs notations de ces trois grandeurs ont été proposées par [50–52]. En 2D, ces longueurs correspondent à la longueur et largeur de la boîte d'encombrement de la particule. En 3D, ces longueurs sont classiquement définies selon les trois axes d'inertie : elles

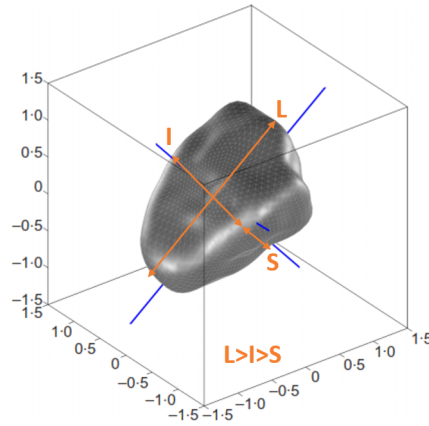


FIGURE 2.4 – Les 3 longueurs principales mesurées sur une particule

correspondent respectivement à la plus grande dimension  $L$ , à la dimension intermédiaire  $I$ , et à la plus petite dimension  $S$  selon ces trois axes (Fig. 2.4).

Le calcul de ces longueurs nécessite donc le calcul du tenseur d'inertie pour le grain en question. Dans le cas d'un grain décrit par un grand nombre de points, ce calcul s'avère fastidieux. Afin de faciliter ce calcul, il existe une méthode qui fournit une approximation des axes d'inertie d'un nuage de points, et des 3 longueurs associées. Le détail en sera donné dans la section suivante.

## b. Indicateurs avancés

Les indicateurs avancés de la forme d'une particule 2D ou 3D sont calculables à partir de combinaisons des mesures présentés ci-dessus.

### b.1. Aplatissement

$$F = \frac{S}{I} \quad (2.6)$$

où  $S$  et  $I$  correspondent respectivement à la plus petite longueur et à la longueur intermédiaire de la particule.

### b.2. Elongation

$$E = \frac{I}{L} \quad (2.7)$$

où  $I$  et  $L$  correspondent respectivement à la longueur intermédiaire et à la plus grande longueur de la particule.

### b.3. Rapport d'aspect

$$a = EF = \frac{S}{L} \quad (2.8)$$

où  $S$  et  $L$  correspondent respectivement à la plus petite et à la plus grande longueur de la particule.

Les valeurs données par le calcul de ces grandeurs sont comprises entre 0 et 1. Elles décrivent un rapport entre les différentes dimensions, qui est fortement corrélé avec la nature de la forme donnée par leurs valeurs. Une particule est d'autant plus aplatie que  $F$  est élevé, et d'autant plus allongée que  $E$  est élevé.

**b.4. Sphéricité** La sphéricité [53] est donnée par :

$$SP = \frac{\sqrt[3]{36\pi V^2}}{S} \quad (2.9)$$

où  $V$  et  $S$  sont respectivement le volume et la surface d'une particule. La sphéricité décrit la proximité entre une forme tridimensionnelle donnée et celle d'une sphère.

Etant donnée la formulation de la sphéricité, nous pouvons remarquer qu'à volume égal, deux particules peuvent avoir une surface égale sans pour autant avoir la même morphologie. En effet, la déviation par rapport à la sphère peut être dans plusieurs directions différentes. Pour cette raison, [54, 55] soulignent la nécessité de coupler la sphéricité avec le "facteur de forme" défini comme le rapport de l'aplatissement et de l'élongation afin de décrire proprement la forme de la particule. Selon [54], la performance de l'élongation et de l'aplatissement à estimer la forme est exactement la même que celle obtenue en utilisant la sphéricité et le facteur de forme, sans pour autant discréditer la sphéricité qui maintient son importance lorsque la comparaison des surfaces est nécessaire.

**b.5. Diagrammes** Les diagrammes font partie des moyens de caractérisation scalaire (en 2D ou en 3D) puisqu'ils consistent en une combinaison des différents indicateurs scalaires précédents. [52, 56, 57] proposent une classification de la forme à partir de diagrammes construits selon les rapports des longueurs principales. Les avantages est les limitations de ces diagrammes sont discutés dans [56, 58–61].

**b.6. Corrélations** Quelques auteurs ont également exploré la relation de corrélation entre différents indicateurs de forme. [62] ont établi la matrice de corrélation entre plusieurs indicateurs scalaires unidimensionnels (dont l'élongation et l'aplatissement) et un indicateur scalaire tridimensionnel (sphéricité). Ils ont montré que tous les indicateurs scalaires choisis sont fortement corrélés avec l'élongation et l'aplatissement, et donc qu'ils peuvent être remplacés par ceux-ci afin de réduire le nombre d'indicateurs utilisés pour la description de la forme. En outre, en augmentant le degré de complexité de l'indicateur choisi (1D, 2D ou 3D), ils ont remarqué une diminution de la corrélation, de telle manière qu'entre les indicateurs 1D et 3D, la corrélation est quasi-nulle.

Dans une représentation tridimensionnelle, qu'il s'agisse du ballast ou d'un autre matériau granulaire, la sphéricité est un indicateur de premier ordre sur la forme, mais il a son importance. Le résultat d'absence de la corrélation entre cet indicateur et d'autres montre que la description de la forme ne peut se faire sans avoir recours à plusieurs indicateurs scalaires, dont le nombre et la combinaison sont d'autant plus complexes que la particule est irrégulière.

Nous abordons à présent des indicateurs scalaires relatifs au deuxième ordre de la forme.

**b.7. Rondeur et angularité** En 2D, la rondeur  $Rnd$  est classiquement mesurée à travers la comparaison du rayon de courbure moyen du contour de la particule, avec celui du plus grand cercle inscrit dans la particule. Plusieurs indices de rondeur proposés par différents auteurs sont cités dans [52]. En 2D, le plus courant est celui proposé par [34,63] :

$$Rnd = \frac{\sum \frac{r}{R}}{N} \tag{2.10}$$

où  $r$  est le rayon de courbure du coin et  $R$  le rayon du plus grand cercle inscrit.  $N$  est le nombre de coins donné par le contour 2D. En 3D, une définition généralisée est proposée par [64] :

$$Rnd = \frac{\sum R_c}{n_c R_{ins}} \tag{2.11}$$

où  $R_c$  les rayons des  $n_c$  sphères remplissant le volume de la particule et  $R_{ins}$  est le rayon de la plus grande sphère inscrite dans la particule.

Ces définitions de la rondeur signifient donc que l'on néglige les éléments de surface donc le rayon de courbure est supérieur à celui du cercle (resp. sphère) inscrit, et que par conséquent des particules de forme différente peuvent avoir la même rondeur (2.5).



FIGURE 2.5 – 2 particules possédant la même rondeur mais de forme différente (Les rayons de courbure supérieurs à celui du plus grand cercle inscrit sont négligés)

En outre, comme expliqué dans [65], le choix de la moyenne des rayons de courbure pose un problème de distribution puisque cette valeur nécessite également la spécification de l'écart-type. Une telle mesure est également incapable de distinguer entre des particules anguleuses possédant des angles différents : le rayon de courbure associé à un cercle inscrit dans un coin anguleux n'est pas associé à une taille unique de l'angle (Fig. 2.6).

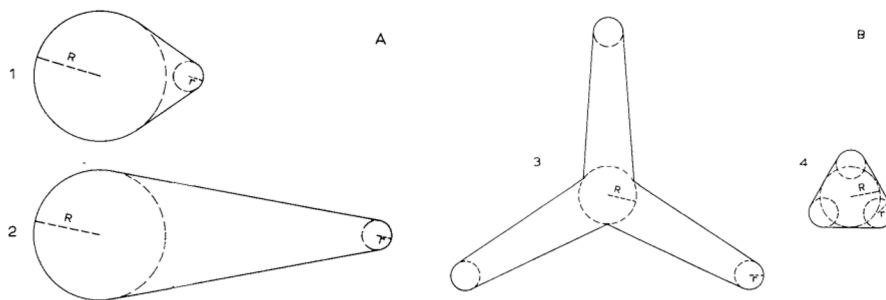


FIGURE 2.6 – Exemples de coins à rondeur identique et angularité différente (A) et exemples de particules à rondeur identique et angularité différente (B) [54]

La conclusion est que l'angularité n'est pas une absence de rondeur, comme évoqué dans [52,54]. C'est un concept différent : la rondeur peut concerner la forme arrondie



des coins (après usure par exemple), tandis que l'angularité prend en compte la relation entre les plans entourant les coins anguleux.

L'angularité est donc une caractéristique de la particule à plus petite échelle. Comme son nom l'indique, que ce soit en  $2D$  ou  $3D$ , elle fait référence aux angles formés par les faces et les arêtes qui définissent une particule. Ces faces font, entre elles, des angles plus ou moins aigus. L'arête peut être plus ou moins vive ou émoussée (Fig. 2.7).

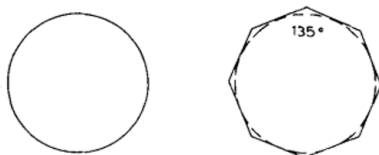


FIGURE 2.7 – Particules avec une degré différent d'angularité

Il est souvent plus évident de décrire l'angularité de manière qualitative. Toutefois, [54] a proposé une mesure de l'angularité, dite degré d'angularité  $A_i$  calculé pour chaque coin d'un contour  $2D$  fermé, en calculant les angles entre les plans entourant tous les coins du contour, ainsi que le rayon du cercle inscrit. Cette définition est extensible au cas  $3D$  en utilisant la projection  $2D$  de la particule sur les trois plans orthogonaux. Une discussion autour des avantages de cette mesure est présentée dans [52]. D'autres propositions ont été faites par [66, 67].

Dans le cas  $3D$ , il s'avère complexe de définir l'angularité. Pour ce qui est du ballast, comme évoqué dans le chapitre précédent, [18] proposent une définition basée sur le calcul de l'angle moyen formé entre les côtés détectés du contours  $2D$ .

La difficulté, là encore, apparaît sur deux aspects :

1. Cette mesure est une moyenne, ce qui pose le même problème de distribution évoqué précédemment, et encore plus présent dans le cas  $3D$  (le ballast en particulier).
2. La question de la convexité se pose. En effet, cette mesure est plus simple à faire dans le cas convexe, puisque la détection des contours convexes est plus évidente.

Une autre manière de calculer l'angularité est présentée dans [45]. Elle consiste à mailler les grains, puis à définir l'angularité comme l'angle externe moyen entre deux faces adjacentes. Cette mesure permet de qualifier l'angularité du grain dans son ensemble, mais elle dépend fortement du nombre de faces, et donc du maillage. Nous traitons cette question de l'effet du maillage sur la forme au chapitre 4.

Enfin, nous abordons un indicateur relatif au troisième ordre : la rugosité.

**b.8. Texture de surface** [15] proposent une définition  $2D$  de la texture de surface (TS) en utilisant les particules fines, et la technique érosion-dilatation (technique de traitement d'images). L'érosion est définie comme une opération morphologique causant la diminution du contour de la particule d'un pixel. La dilatation, quant à elle, est associée au phénomène inverse, c'est-à-dire au gain d'un pixel sur la surface

de la particule. Selon [15], la (ST) s'écrit alors comme suit :

$$ST = \frac{[(A_1 - A_2) * 100]}{A_1} \quad (2.12)$$

où  $A_1$  et  $A_2$  les surfaces respectives avant et après érosion-dilatation. En 3D, cette définition est applicable en moyennant sa valeur pour chaque projection 2D sur les 3 plans principaux.

En 2D, il est à noter que la performance de cet indicateur intéressant dépend fortement du niveau de discrétisation du contour en question, qui fait référence à une méthode connue : l'analyse fractale [68].

### c. Conclusion

Pour un matériau de forme complexe comme le ballast, nous jugeons ces indicateurs comme étant faibles par rapport aux ambitions fixées. Nous nous proposons alors d'étudier d'autres méthodes de caractérisation offrant une meilleure prise en compte des différents aspects de la forme.

## 2.2 Caractérisation par méthodes spectrales

De manière générale, l'analyse spectrale recouvre les techniques de décomposition et de description des signaux, fonctions de temps ou de variables d'espace. L'analyse harmonique correspond à l'une de ces techniques. Un exemple simple et connu en est la décomposition en séries de Fourier en (2D).

### a. Mesures

En 2D comme en 3D, les méthodes spectrales s'appuient sur une mesure spatiale des positions des sommets des grains en question, et ceci quel que soit le moyen de mesure : des images obtenues par photographie en 2D, ou de la photogrammétrie, du scan, ou de la tomographie en 3D.

### b. Caractérisation

**b.1. Cas 2D** Les indicateurs de Fourier ont été introduits pour la première fois pour la caractérisation 2D de la forme par [69, 70], et ont été ensuite utilisés pour la caractérisation des sables [30, 49, 71, 72].

En 2D, le contour fermé d'une particule décrit par  $(R, \theta)$  peut selon la théorie de Fourier être représenté par la somme des séries suivantes :

$$R(\theta) = r_0 + \sum_{n=1}^N [A_n \cos(n\theta) + B_n \sin(n\theta)] \quad (2.13)$$

Où  $R(\theta)$  le rayon dans la direction  $\theta$ ,  $N$  le nombre total d'harmoniques, et  $A_n$  et  $B_n$  les coefficients donnant l'amplitude de chaque harmonique [72].  $r_0 = \frac{1}{N} \sum_{n=1}^N R(\theta_n)$  est le rayon moyen de la particule. Le nombre  $N$  d'harmoniques est égal au nombre

de points utilisés pour discrétiser le contour de la particule. Selon [49,72], l'amplitude normalisée du spectre obtenu par transformée de Fourier permet une caractérisation pertinente d'une population donnée de particules.

Pour chaque harmonique  $n$ , cette amplitude, qu'on appelle aussi mode  $n$  ou indicateur de Fourier d'ordre  $n$  s'écrit de la manière suivante :

$$D_n = \frac{\sqrt{A_n^2 + B_n^2}}{r_0}, 0 \leq n \leq N/2 \quad (2.14)$$

En ce qui concerne les grains de sables, on associe souvent  $D_2$  à l'élongation et les modes  $D_3$  à  $D_8$  aux principales irrégularités du contour de la particule. Ensuite, en montrant que les modes  $D_n$  pour  $n > 8$  diminuent en fonction de l'ordre de l'indicateur, [49] a conclu que ces modes-là constituent de bons indicateurs de rugosité.

Une extension de cette approche au cas 3D est proposée par [73, 74], mais davantage pour la reconstruction et la génération de grains de sable que pour la caractérisation de leur forme. Nous reviendrons plus en détail sur les avantages et les limitations de cette méthode au chapitre consacré à la génération des grains virtuels.

**b.2. Cas 3D** Dans le cas 3D, bien plus complexe, les équivalents des indicateurs de Fourier sont les harmoniques sphériques (SH).

Au cours des dix dernières années, l'utilisation des harmoniques sphériques est de plus en plus courante et touche à plusieurs domaines dont la géologie et la géotechnique pour la reconstruction des objets convexes, sans pouvoir la généraliser au cas-non convexe [30, 75]. Le cas des surfaces non convexes se présente toutefois souvent et doit être pris en compte dans la caractérisation et la reconstruction de la forme des particules. Cette difficulté a été soulevée par plusieurs auteurs dont [76].

Pour remédier à cela, [77] ont proposé une représentation de surfaces non convexes par trois décompositions en harmoniques sphériques selon les trois coordonnées cartésiennes, ainsi que son applicabilité dans le cas des grains de sables.

Intéressons-nous maintenant à la définition mathématique de la méthode, à ses avantages et limitations, et aux exemples traités dans la littérature quant à la caractérisation de la formes des particules.

Partant d'une population de grains décrits par leur coordonnées cartésiennes  $V(x, y, z)$ , la première étape est une conversion en coordonnées sphériques  $\theta$  et  $\phi$ , où  $\theta \in [0, 2\pi]$  et  $\phi \in [0, \pi]$  représentent respectivement la longitude et la colatitude (Fig. 2.3).

On a donc :

$$V(\theta, \phi) = (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))^T \quad (2.15)$$

A partir de cette expression, le développement en harmoniques sphériques s'écrit de la manière suivante :

$$V(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_n^m Y_n^m(\theta, \phi) \quad (2.16)$$

Où :  $Y_n^m(\theta, \phi)$  est la fonction SH :

$$Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos(\phi)) e^{imp\theta} \quad (2.17)$$

où  $P_n^m$  est le polynôme de Legendre d'ordre  $n$  :

$$P_n^m(x) = (1 - x^2)^{m/2} \frac{d^m}{dx^m} p_n(x) \quad (2.18)$$

et :

$$p_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (2.19)$$

L'écriture matricielle de cette décomposition est donc la suivante :

$$A = \begin{pmatrix} y_1^1 & y_1^2 & \dots & y_1^{(n+1)^2} \\ y_2^1 & y_2^2 & \dots & y_2^{(n+1)^2} \\ \dots & \dots & \dots & \dots \\ y_i^1 & y_i^2 & \dots & y_i^{(n+1)^2} \end{pmatrix} \begin{pmatrix} c_x^1 & c_y^1 & c_z^1 \\ c_x^2 & c_y^2 & c_z^2 \\ \dots & \dots & \dots \\ c_x^{(n+1)^2} & c_y^{(n+1)^2} & c_z^{(n+1)^2} \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \dots & \dots & \dots \\ x_i & y_i & z_i \end{pmatrix} \quad (2.20)$$

Pour déterminer  $c_n^m$ , le nombre d'inconnues à trouver par la résolution de  $i$  équations linéaires est  $(n+1)^2$ . Le système a donc des solutions si  $i \geq (n+1)^2$ . Selon [30],  $n = 12$  suffit pour caractériser et reconstruire la forme de la majorité des particules. La résolution du système permet d'obtenir la matrice de coefficients  $c_n^m$ . L'ensemble (ou une partie) de ces coefficients peut être utilisé comme indicateur caractéristique de la forme des particules. Cette décomposition signifie que chaque grain est décrit par une combinaison linéaire de coefficients  $c_n^m$  et de vecteurs de séries harmoniques  $Y_n^m$  qui représentent donc une base de projection du grain.

Qui dit décomposition, dit reconstruction. Pour reconstruire un grain qu'on a décomposé en plusieurs vecteurs  $c_n^m$  et  $Y_n^m$  à un ordre donné, il suffit de choisir le nombre de fonctions à garder, et de faire la combinaison linéaire de celles-ci. Le nombre de coefficients et de vecteurs utilisé définit le degré de précision souhaité sur la forme : plus l'ordre est grand, plus le grain reconstruit se rapproche du grain initial. Le fait de couper la combinaison linéaire à un certain ordre  $k$  s'appelle la troncature au mode  $k$ . [29] proposent une application de cette méthode à deux populations de sables de niveaux différents d'irrégularité. Les comparaisons des surfaces et des volumes calculés pour ces deux populations de grains avant et après leur reconstruction pour plusieurs degrés de troncature montrent que les volumes des particules convergent rapidement à partir de  $n > 10$  pour les deux types de sables. Une convergence moins rapide est observée pour la surface dans le cas des grains à faible irrégularité (à partir de  $n \geq 12$ ), et une erreur de 11% est notable pour les grains plus irréguliers même au-delà de  $n = 20$ , ce qui revient à 450 coefficients environ. En effet, la surface étant un indicateur qui se base davantage sur les détails de texture et de rugosité, est moins évident à approcher avec cette méthode.

Comme pour le cas des indicateurs de Fourier, une interprétation possible de la décomposition en harmoniques sphériques consiste à classer les amplitudes des coefficients selon leur rapport respectif à la forme générale, à la texture de surface, ou aux détails de rugosité : les 3 premiers modes, les modes de  $n = 4$  à  $n = 8$ , puis le reste des modes.

Cette méthode démontre un grand intérêt de par le fait qu'elle prend en compte (jusqu'à un certain ordre) les dépendances complexes entre les différentes propriétés géométriques des grains. Elle permet également la reconstruction des grains à plusieurs degrés quantifiables de précision, bien qu'elle ne convienne pas aussi bien à

des grains de sable à degré élevé d'irrégularité qu'à des grains de faible degré d'irrégularité. Il paraît alors tout à fait légitime, d'appliquer cette méthode à nos grains de ballast, en gardant en tête leur complexité par rapport aux grains de sable, en plus de leur non-convexité.

La méthode des harmoniques sphériques (et les indicateurs de Fourier en 2D) permet théoriquement la génération de nouvelles particules à formes équivalentes de l'échantillon initial. Les avantages et les limitations de cette méthode seront donc davantage discutés au chapitre sur la génération des grains virtuels.

Dans la section suivante, nous présentons la méthode que nous proposons d'appliquer aux grains de ballast pour caractériser leur forme. Cette méthode est basée sur le même principe de décomposition de la forme sur une base de vecteurs, mais permet d'obtenir un résultat plus optimal.

### 3 Caractérisation par POD

La Décomposition Orthogonale aux Valeurs Propres, ou Proper Orthogonal Decomposition (POD) est une méthode d'analyse de données, qui consiste à transformer des variables « corrélées » en nouvelles variables décorrélées les unes des autres [78]. Cette méthode permet d'approximer un système de dimension élevée par un nouveau système de dimension plus faible [79]. Le principe est de déterminer une base de vecteurs orthogonaux qu'on appelle modes propres sur la base des réalisations les plus probables. Dans les différentes études menées sur la méthode, il a été montré que ces modes propres constituent une base optimale au sens de la quantité d'information (qu'on appelle aussi énergie) qu'ils représentent quel que soit le type des données d'entrée.

#### 3.1 Principes de la Décomposition Orthogonale aux Valeurs Propres (POD)

Originellement, la POD a été introduite dans le but d'identifier les structures cohérentes d'un écoulement (Adrian et al. 1996). Dans ce cas-là, la POD peut être une alternative à la décomposition en séries de Fourier (Section méthodes spectrales) en l'absence de l'hypothèse d'homogénéité ou de périodicité des directions de l'écoulement. Pas toujours sous le même nom, la POD touche à beaucoup d'autres domaines scientifiques. En processus stochastiques, on reconnaît la POD sous le nom de Décomposition de Karhunen-Loève (KLD) [80], ou encore analyse d'Hotelling [81]. En analyse de données, elle est plus communément appelée Analyse en Composantes Principales (PCA) [82], et est étroitement liée à la Décomposition aux Valeurs Singulières (SVD) [83]. Ces différentes appellations font référence à des formulations différentes mais équivalentes de la POD. Une réponse à la question de l'équivalence entre ces formulations est apportée dans [84], précisant que le choix se porte sur la PCA dans le cas particulier de variables discrètes et d'un espace de données de dimension finie. Dans la suite, nous justifions notre choix de détailler uniquement ce type de formulation.

Quel que soit le domaine concerné (compression de données, détection des visages, océanographie, analyse de signal, etc.), l'objectif est commun : l'extraction

des propriétés dominantes d'un ensemble de réalisations, tout en optimisant l'espace permettant leur description, ce qui facilite l'analyse de leur structure.

Enfin, la POD fait partie des méthodes du Machine Learning, ou de l'apprentissage automatique [85]. En effet, l'un des objectifs de l'apprentissage automatique est de construire des modèles qui permettent à la machine de mieux comprendre nos données et nous aider à les interpréter en les représentant sous une autre forme.

### a. Notations et prérequis

Soit  $X = (X_1, \dots, X_d)$  un vecteur aléatoire défini sur l'espace probabilisé  $(\Omega, \Gamma, P)$  à valeurs dans  $R^d$ . Pour chaque  $1 \leq i \leq d$ ,  $X_i$  est une variable aléatoire à valeurs dans  $R$ . On considère que  $X$  est de moyenne  $\mu = (\mu_1, \dots, \mu_d)$  telle que :

$$\mu_i = E[X_i] \quad i = 1, \dots, d \quad (2.21)$$

où  $E[X_i]$  est l'espérance mathématique de la variable aléatoire  $X_i$ . On note :  $\mu = E[X]$ .

De la même manière, on définit  $[R_{\mathbf{X}}]$  la matrice de covariance de  $X$  :

$$[R_{\mathbf{X}}] = (v_{ij}) = (cov(X_i, X_j)) = E[(X_i - \mu_i)(X_j - \mu_j)] \quad (2.22)$$

où  $cov(X_i, X_j)$  est la covariance de  $X_i$  et  $X_j$ . Les éléments diagonaux de  $[R_{\mathbf{X}}]$  correspondent aux variances respectives  $var(X_i)$  de  $X_i$  :  $v_{ii} = cov(X_i, X_i) = var(X_i)$ ,  $i = 1, \dots, d$ .

$[R_{\mathbf{X}}]$  est une matrice symétrique par construction, puisque :  $(cov(X_i, X_j)) = (cov(X_j, X_i))$ . Cette matrice décrit le rapport de dépendance linéaire entre les  $X_i$  et  $X_j$ . Ainsi, la covariance de deux variables indépendantes est nulle, bien que la réciproque ne soit pas toujours vraie, puisque la dépendance linéaire n'est qu'un des divers aspects de dépendance.

On définit également la matrice de corrélation  $corr(X)$  :

$$corr(X) = (corr(X_i, X_j)) = \frac{cov(X_i, X_j)}{\sigma_i \sigma_j} \quad (2.23)$$

où  $\sigma_i$  est l'écart-type de  $X_i$ .

Le calcul du coefficient de corrélation est toujours possible. Néanmoins, celui-ci n'arrive pas toujours à rendre compte du caractère de dépendance entre les variables étudiées. En effet, comme la covariance, il n'est pas adapté pour juger de corrélations non linéaires et non linéarisables. Son intérêt devient également limité dans le cas de données très hétérogènes, puisqu'il représente une relation moyenne et que l'on sait que la moyenne n'a pas toujours un sens, notamment si la distribution des données est multimodale.

De la même manière que la covariance, si les deux variables sont totalement indépendantes, alors leur corrélation est égale à 0. La réciproque est cependant fautive, car le coefficient de corrélation indique uniquement une dépendance linéaire. D'autres phénomènes, par exemple, peuvent être corrélés de manière exponentielle, sous forme de puissance, etc.

L'exemple le plus courant dans la littérature est le suivant : supposons que la variable aléatoire  $X_i$  soit uniformément distribuée sur l'intervalle  $[-1; 1]$ , et que

$X_j = X_i^2$ ; alors  $X_j$  est complètement déterminée par  $X_i$ , de sorte que  $X_i$  et  $X_j$  ne sont pas indépendantes, mais leur corrélation vaut bien 0. D'autres exemples sont présentés dans la Fig. 2.8.

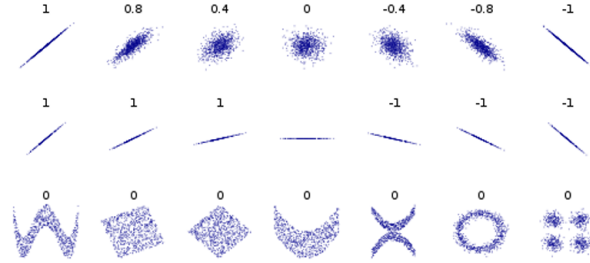


FIGURE 2.8 – Exemples de variables aléatoires dépendantes mais non corrélées : les nuages de points montrent l'évolution d'une variable aléatoire  $X_j$  en fonction d'une autre variable aléatoire  $X_i$ . Les valeurs au-dessus des nuages de points correspondent à la corrélation linéaire entre  $X_i$  et  $X_j$ .

Dans la suite, on suppose que nous disposons d'un ensemble fini de  $n$  réalisations  $X(w_1), \dots, X(w_n)$  du vecteur aléatoire  $X$ . On modélise l'ensemble de ces réalisations par une matrice  $A$  de dimension  $n \times d$  de rang  $r$  où  $r \leq \min(n, d)$ . Les lignes correspondent aux réalisations tandis que les colonnes correspondent aux variables. Dans ce cas, soit  $\hat{\mu}$  et  $[\hat{R}_{\mathbf{X}}]$  les estimateurs (non biaisés) respectifs de  $\mu$  et  $[R_{\mathbf{X}}]$ . Sans aucune perte de généralité, nous supposons que le vecteur aléatoire  $X$  est centré :  $\hat{\mu} = 0$ . Dans ce cas, la matrice de covariance  $[\hat{R}_{\mathbf{X}}]$  s'écrit :

$$[\hat{R}_{\mathbf{X}}] = \frac{1}{n-1} A^T A \quad (2.24)$$

Enfin, on définit les deux normes suivantes :

1. La norme  $L_2$  sur  $R^n$  :

$$\|u\|_2^2 = \sum_{k=1}^n |u_k|^2 \quad u = (u_1, \dots, u_n) \quad (2.25)$$

A partir du produit scalaire :

$$\langle u, v \rangle = u^T \cdot v = \sum_{k=1}^n u_k v_k \quad u, v \in R^n \quad (2.26)$$

2. La norme de Frobenius sur l'espace vectoriel  $M_{n,d}(R)$  de matrices de dimension  $n \times d$  à valeurs dans  $R$  :

$$\|A\|_F^2 = \sum_{i=1}^n \sum_{j=1}^d a_{ij}^2 = \text{tr}(AA^T) \quad (2.27)$$

où  $\text{tr}(S) = \sum_{k=1}^d s_{kk}$  est l'opérateur trace d'une matrice carrée  $S$  de dimension  $d \times d$ .

## b. Analyse en Composantes Principales (PCA)

On considère donc la matrice  $A \in M_{n,d}(R)$  de rang  $r$  représentant le vecteur aléatoire centré  $X$ .

De plus, si les variables sont de natures ou d'unités différentes, on cherchera à les normaliser de façon à traiter les variables de façon symétrique et à ne pas biaiser l'interprétation à cause de l'hétérogénéité des données. Afin de normaliser les variables, la méthode la plus courante est de diviser les valeurs des variables par leur écart-type. On travaillera alors sur des variables centrées et réduites représentées par la matrice  $\hat{A}$ .

Les données de la matrice  $\hat{A}$  sont représentées dans une base qu'on appellera la base initiale. La question qui motive la PCA est la suivante :

Existe-t-il une base orthonormale  $B = \{u_i \in R^d, i = 1, \dots, d\}$  de  $R^d$  (où  $\langle u_i, u_j \rangle = \delta_{i,j}$ ), combinaison linéaire de la base initiale, qui permette d'exprimer nos données sous une forme optimale facilement interprétable ?

En ceci, deux propriétés sont à remarquer :

- Optimalité : trouver une base optimale signifie que l'information sera représentée de manière à permettre une réduction du modèle avec une moindre perte d'information.
- Linéarité : la PCA fait une hypothèse sur la linéarité des données : ceci signifie que le problème se réduit à un changement de base et se réduit alors au choix de la meilleure base. Cette hypothèse est importante, et nous servira également lors du chapitre sur la génération de grains.

De ces deux propriétés, nous sommes amenés à nous poser les questions suivantes :

Quels sont les critères permettant de juger de la qualité de la représentation des données, et quelle base choisir pour représenter les données ?

1- D'abord, sous l'hypothèse de la linéarité, une bonne représentation des données est déterminée par la minimisation de deux éléments : le bruit, et la redondance. Comme expliqué dans [86], ces deux éléments sont directement liés à la minimisation de la covariance entre les variables deux à deux : ceci justifie l'objectif de la PCA de transformer les variables corrélées (dépendance linéaire), en des variables décorrélées.

2- Soit  $U = [u_1, \dots, u_d]$  la matrice dont les colonnes correspondent aux vecteurs de  $B$ , et soit  $V$  le sous-espace vectoriel engendré par la base orthonormale  $B$ . Alors, on définit la matrice suivante :

$$C = \hat{A}U \tag{2.28}$$

$C$  est la matrice de données projetée dans la nouvelle base, on l'appellera dorénavant la matrice des coefficients. Naturellement, le choix de  $U$  n'est pas unique et les polynômes de Legendre ou de Chebychev pourraient tout à fait convenir. Un autre choix pourrait être les harmoniques sphériques, qui formeraient donc une base imposée quelle que soit la nature des données de  $\hat{A}$ . Cependant, gardons en tête que la matrice  $C$  dépend du choix de la base. Puisque la PCA propose un choix optimal, celui-ci doit être donné par la minimisation d'une erreur.

La projection (qu'on appellera également reconstruction) de la matrice  $\hat{A}$  dans



le sous-espace  $V$  s'écrit :

$$A^* = CU^T = \hat{A}UU^T \quad (2.29)$$

Le choix de la base proposée par la PCA minimise la différence au sens des moindres carrés entre  $\hat{A}$  et sa reconstruction avec chaque ordre de troncature de la base  $B$  (base réduite). L'existence et l'optimalité de la solution sont expliquées dans [87, 88].

Pour les raisons évoquées plus haut, la solution consiste à la diagonalisation de la matrice de covariance.

Soit  $cov(A) = [\hat{R}_X]$  la matrice de covariance de dimension  $d \times d$  associée à la matrice  $\hat{A}$  :

$$cov(A) = \frac{1}{n-1} \hat{A}^T \hat{A} \quad (2.30)$$

$cov(A)$  est une matrice symétrique à coefficients réels. Selon le théorème spectral en dimension finie, elle est donc diagonalisable à l'aide d'une matrice de passage orthogonale :

$$cov(A) = UDU^T \quad (2.31)$$

où :

-  $D = (d_{ii}) = (\lambda_i)$  est une matrice diagonale de dimension  $d \times d$  avec  $d_{ii} = 0$  pour  $i > r$  et  $\lambda_1 > \lambda_2 > \dots > \lambda_r > 0$ .

-  $U$  est une matrice orthogonale ( $U^{-1} = U^T$ ) de dimension  $d \times d$ .

Les colonnes  $U_j$  de  $U$  sont les vecteurs propres de  $cov(\hat{A})$ , et sont appelées composantes principales, ou encore modes propres (en référence à la POD). Elles constituent une base orthonormale optimale de projection de  $\hat{A}$ .

Les valeurs propres  $\lambda_j$  constituent les variances associées aux composantes principales  $U_j$ . L'amplitude des variances mesure l'importance de l'information. En effet, les variances sont classées selon l'ordre décroissant, et on dit qu'une composante principale est d'autant plus importante que la variance associée est grande.

### c. Reconstruction et approximation

L'optimalité de la base offre la possibilité de réduire le modèle. La décomposition obtenue plus haut signifie que le vecteur aléatoire  $X$  est représenté par une combinaison linéaire de composantes principales et de coefficients. La réduction du modèle se réduit donc à la troncature de cette combinaison à un ordre donné  $K < d$  :

$$A_{(K)} = C_{(K)}U^T = \hat{A}U_{(K)}U_{(K)}^T \quad (2.32)$$

où :

-  $C_{(K)} = [c_1, \dots, c_K]$  est une matrice de dimension  $d \times K$  dont les colonnes sont les coefficients.

-  $U_{(K)} = [u_1, \dots, u_K]$  est la matrice de dimension  $d \times K$  dont les colonnes sont les composantes principales.

Une troncature est équivalente à une perte d'information à quantifier. Les amplitudes des variances permettent de mesurer cette erreur.

Soit  $K \leq r$  où  $r = rang(A)$  l'ordre de troncature et  $\epsilon(K)$  l'erreur qui y est associée. Cette erreur, minimisée par la PCA au sens de la norme de Frobenius, se

calculé par le rapport de la variance associée à un ordre de troncature, à la variance totale. Elle s'écrit :

$$\epsilon(K) = \frac{\|\hat{A} - A_{(K)}\|_F^2}{\|\hat{A}\|_F^2} = \frac{\sum_{i=K+1}^r \lambda_i}{\sum_{i=1}^r \lambda_i} \quad (2.33)$$

où  $\lambda_i, 1 \leq i \leq r$  sont les valeurs propres issues de la décomposition aux valeurs propres de la matrice de covariance (Eq. 2.31). A partir de l'erreur, on peut définir la quantité d'information dite énergie :

$$\zeta(K) = 1 - \epsilon(K) \quad (2.34)$$

Il est à noter que le choix de la norme 2 est tout à fait possible. Dans ce cas, l'erreur s'écrit :

Sachant que :

$$\|\hat{A}\|_2 = \max_{\|x\|_2=1} \|\hat{A}x\|_2 = \sqrt{\lambda_1} \quad (2.35)$$

On a :

$$\epsilon_2(K) = \frac{\|\hat{A} - A_{(K)}\|_2^2}{\|\hat{A}\|_2^2} = \frac{\lambda_{K+1}}{\lambda_1} \quad (2.36)$$

Au sens de cette erreur, les premières composantes principales contribuent le plus à la variance totale, et contiennent donc l'information la plus importante. La combinaison linéaire constituée par tous les modes correspond à une erreur nulle, et la base est d'autant plus optimale que l'erreur décroît rapidement.

Pour résumer l'ensemble des étapes de la PCA, voici un algorithme récapitulatif :

1. Donnée d'entrée : matrice  $A = (a_{ij})$  de rang  $r$  et de dimension  $n \times d$  où  $n$  est le nombre d'observations et  $d$  est le nombre de variables, représentant le vecteur aléatoire  $X$ .
2. Centrer la matrice  $A$ .
3. Calculer la matrice de covariance  $cov(A) = [\hat{R}_X]$ .
4. Résoudre le problème aux valeurs propres de  $cov(A)$  : vecteurs propres  $u_1, \dots, u_d$  et valeurs propres correspondantes  $\lambda_1, \dots, \lambda_d$ , telles que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \geq 0$ .
5. Pour chaque réalisation, l'approximation à l'ordre  $K \leq r$  s'écrit :  $A_{(K)} = \hat{A}U_{(K)}U_{(K)}^T$ .
6. Calculer l'erreur  $\epsilon(K)$  correspondante à la troncature à l'ordre  $K$  :

$$\epsilon(K) = \frac{\|\hat{A} - A_{(K)}\|_F^2}{\|\hat{A}\|_F^2} = \frac{\sum_{i=K+1}^r \lambda_i}{\sum_{i=1}^r \lambda_i}$$

#### d. Limitations

Malgré le grand potentiel de cette méthode pour la réduction de modèle pour l'extraction des propriétés dominantes d'une structure de données, ainsi que son

énorme champ d'application et ses résultats pertinents dans plusieurs domaines [89, 90], elle possède évidemment des limites. Ceux-ci n'enlève rien de son mérite, mais il est néanmoins important de les prendre en considération quant aux conditions d'applicabilité de la méthode et aux attentes de l'utilisateur.

Ces limitations peuvent être en lien avec la construction de la méthode, ou avec les hypothèses qu'elle fait.

- La linéarité : cette hypothèse fait que l'amélioration de la représentation des données se réduit à un changement de bases. Il en vient que la meilleure représentation n'est pas toujours assurée par la PCA dans le cas de variables ayant des dépendances non linéaires. Des études montrent que l'introduction d'une non-linéarité mène vers une extension de la méthode appelée Kernel PCA [91].
- L'analyse en variance : le fait de relier l'information directement à la variance suppose que les axes (composantes principales) ayant les plus grandes variances portent l'information la plus importante. Ce n'est pas toujours le cas selon les données [86]. Ceci pose la question de la signification physique de l'erreur PCA. Néanmoins, il appartient à l'utilisateur durant l'étape de l'interprétation des données de juger de la qualité de cette erreur et d'effectuer les vérifications nécessaires.
- Analyse non paramétrique : ce point est à la fois une force et une faiblesse de la méthode. En effet, un des avantages considérables de cette approche est la facilité d'application. Il suffit de faire les hypothèses évoquées plus haut et de calculer les résultats correspondants. Aucun paramètre supplémentaire n'a besoin d'être calibré ou ajusté par l'utilisateur en se basant sur l'expérience. Le résultat est unique et indépendant de l'utilisateur. Ce même avantage peut être vu comme un inconvénient : il serait logique de pouvoir incorporer des hypothèses supplémentaires dans le cas où l'utilisateur possède une connaissance a priori de quelques propriétés du système.
- Le fléau de la dimension (curse of dimensionality) [92] : ce problème est un obstacle majeur dans l'apprentissage automatique, qui revient souvent à tirer des inférences d'un nombre réduit d'expériences dans un espace de possibilités de dimension élevée. Dans le cas de la PCA, cette condition fait que la dimension la plus faible doit être relative aux variables qui représentent les propriétés qu'on cherche à explorer, et la dimension la plus élevée doit être relative aux réalisations qui représentent notre échantillon. Un cas contraire présenterait une matrice dite "pathologique". On détaillera ce point lors de l'application de la méthode à notre base de données.

### e. Approches associées

Comme évoqué précédemment, les deux approches associées à la PCA sont la transformation de Karhunen-Loève (KLT) et la Décomposition en Valeurs Singulières (SVD). La première est adaptée aux champs aléatoires génériques et se réduit à la PCA lorsque ceux-ci sont échantillonnés. La deuxième constitue une méthode de résolution de la PCA.

En effet, la résolution par POD peut se faire soit par PCA en résolvant un problème aux valeurs propres qui revient à la diagonalisation de la matrice de cova-

riance, ou par SVD. Pour éviter les calculs fastidieux de multiplications matricielles engendrés par les grandes quantités de données, ainsi qu'à l'instabilité numérique qui peuvent en résulter, les implémentations de POD ont souvent recours à la SVD  $\hat{A} = W[\sigma]L^T$  où  $[\sigma]$  est la matrice diagonale contenant les valeurs singulières de  $\hat{A}$  qui correspondent exactement aux racines carrées des  $\lambda_i$ , et où  $L = U$ . Dans notre cas, l'ensemble de ces opérations est résolu à l'aide de codes implémentés sur le langage Python, ou en utilisant la librairie Scikit-Learn [93] pour l'apprentissage automatique.

## f. Conclusion

Quelle que soit la formulation utilisée et les hypothèses faites, la POD offre plusieurs avantages sur la réduction de modèle, sur la caractérisation de données complexes. Une importante force de la méthode est sa flexibilité selon la nature des données. Pour la caractérisation de la forme des grains de ballast, nous appliquons dans la suite cette approche, et la comparons avec la décomposition en harmoniques sphériques (HS) afin de mettre en lumière son optimalité.

## 3.2 Mesure et acquisition des données : du réel au virtuel

Comme expliqué dans le chapitre précédent, la mesure 3D de la forme constitue un outil performant par rapport aux mesures 2D ne permettant pas de cerner la complexité de la morphologie des matériaux granulaires, et particulièrement celle du ballast. De manière générale, la mesure 3D de la forme trouve de plus en plus ses applications : en industrie aéronautique, en automobile, en nanotechnologie, en médecine, en cosmétique, en puériculture, mais aussi en archéologie et en bijouterie.

Quel que soit l'objet concerné, on peut représenter sa forme 3D par un vecteur de points/sommets dont on connaît les coordonnées cartésiennes  $(x, y, z)$ . Une représentation théorique peut ensuite être formulée afin de décrire la forme de la surface à partir de l'ensemble de ces points. Par conséquent, l'objectif ultime de toutes les méthodes de mesure de la forme 3D (méthodes optiques ou mécaniques) est de déterminer avec une précision "suffisante" en résolution, les coordonnées cartésiennes des points constituant l'objet.

La mesure 3D de la forme permet donc d'obtenir le positionnement d'un échantillonnage de ces points qu'on appelle nuages de points de la surface d'un objet pour ensuite en extrapoler la forme à partir de leur répartition : ce procédé est appelé une reconstruction 3D, et c'est le principe d'un scan 3D.

Pour ce faire, plusieurs méthodes existent et ne cessent d'être améliorées avec l'évolution incessante des techniques d'imagerie. Pour ce qui est des méthodes optiques, deux catégories se distinguent : les méthodes avec contact, et les méthodes sans contact.

- Le scan 3D avec contact : qui sonde la forme par contact physique continu, et qui présente donc l'avantage de la précision, mais peut endommager les objets fragiles, sans oublier le problème de sa lenteur.
- Le scan 3D sans contact : cette même catégorie se décline en deux autres :
  - Le scan 3D sans contact passif : ce type de scanner n'émet aucun rayonnement. Il détecte tout simplement les sources de rayonnement dans l'en-

vironnement ambiant. Pour plus de détails, voir [94].

- Le scan 3D sans contact actif : ce type de scanner émet un rayonnement (X, ultrason, ou lumière) pour ensuite détecter sa réflexion et reconstruire la forme. On retrouve dans cette catégorie les scanners basés sur un décalage de phase (parmi lesquels on retrouve la technique de l'interférométrie), et ceux basés sur une triangulation. Ces méthodes diffèrent en terme de rapidité, de résolution, et d'instrumentation. Pour une description détaillée de toutes ces méthodes, voir [94].

Le compromis entre tous ces aspects nous a conduit au choix du scan 3D sans contact actif par laser à l'aide d'un scanner manuel. Les scanners laser manuels créent des nuages de points 3D à partir du principe suivant : un point ou une ligne laser est projetée sur un objet et un capteur mesure la distance par rapport à la surface.

En pratique, les positions des sommets sont stockées par rapport à un système de coordonnées internes, et la position du scanner (étant lui-même en mouvement) est calculée (différentes méthodes possibles : soit en utilisant des points de repères caractéristiques sur la surface à scanner, ou en utilisant des caméras incorporées permettant les 6 degrés de liberté du scanner). L'information est ensuite recueillie par un ordinateur sous forme de coordonnées 3D des points détectés. On peut choisir de préserver le nuage de points ou de le mailler par triangulation pour construire une surface.

La durée du scan dépend de la complexité et de la dimension de l'objet, mais également de la performance du scanner. Le scanner utilisé dans notre cas est présenté dans la Fig. 2.9.



FIGURE 2.9 – Scanner 3D utilisé

Le nombre de points scannés est de 150000 points/s, à raison de 4000 points par ligne laser. La fréquence de mesure est de 100 Hz, et l'écart moyen entre les points est de  $13 \mu\text{m}$ . On procède au scan de  $n = 975$  grains de ballast neuf, conforme aux normes de la voie ferrée. Le résultat pour chaque grain est sous forme de listes de coordonnées cartésiennes  $(x, y, z)$  formant un nuage de points 3D d'environ 400000 points. En pratique, le scan nécessite environ 15 min /grain (intervention possible de l'opérateur). Un exemple de grains scannés est présentée sur la Fig. 2.10.

Nous avons choisi de travailler sur des nuages de points sans maillage préalable afin de ne pas altérer la forme. L'étape qui suit est l'exploitation de ces nuages de points qui passe d'abord par une préparation de l'échantillon à la POD.

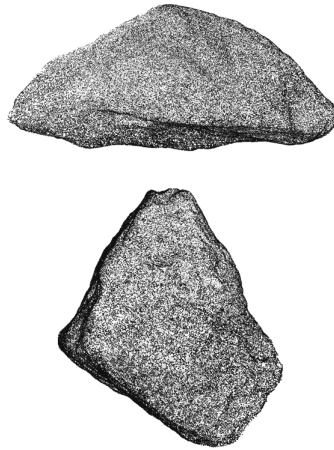


FIGURE 2.10 – Quelques grains scannés (fournis par SNCF)

### 3.3 Prétraitement

Notre objectif, -et celui de la POD également - est de faire ressortir les propriétés communes et dominantes de notre échantillon de grains. En l'occurrence, les propriétés de la forme. Seulement, les nuages de points obtenus peuvent contenir du bruit (perturbations dues à des erreurs de mesure), et n'ont aucune raison d'être centrés autour d'une même origine, et encore moins orientés de la même manière. L'objectif de cette première étape de prétraitement est donc de maximiser la ressemblance entre les grains pour ne pas biaiser l'interprétation des résultats et de maximiser le pouvoir de réduction de la POD : comme le traitement est effectué uniquement sur des positions dans l'espace, toutes les propriétés communes ne pourront être extraites que si les grains sont disposés de la même façon.

Nous effectuons le prétraitement en plusieurs étapes :

#### a. Nettoyage des nuages de points

Nous construisons un filtre basé sur la distance moyenne de chaque points aux points de son voisinage et nous supprimons les points qui dépassent ce périmètre. La Fig. 2.11 montre un exemple de suppression du bruit.

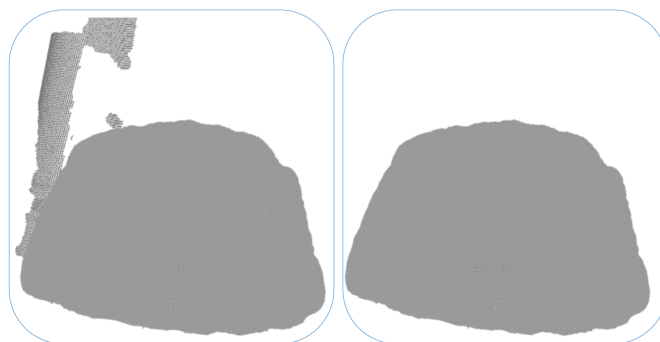


FIGURE 2.11 – Suppression du bruit pour un grain scanné : grain brut (gauche), grain nettoyé (droite)

### b. Réduction du nuage de points

Nous choisissons de réduire le nombre de points pour deux raisons. D’abord, la manipulation d’un grand nombre de points nécessite des temps de calculs importants, autant pour la caractérisation de la forme que pour les calculs mécaniques. Ensuite, comme expliqué précédemment la POD nécessite par construction que le nombre de réalisations (grains) soit nettement supérieur au nombre de variables (directions dans l’espace, et donc sommets). Autrement, la méthode ne sera pas en mesure de nous apporter des résultats solides (comme pour n’importe quel indicateur statistique, la valeur est d’autant plus significative que l’échantillon est grand). Nous faisons un choix aléatoire de points parmi les 400000, pour n’en garder que 10000.

### c. Choix d’une base commune de projection

Afin de construire notre matrice d’entrée à la POD, nous souhaitons que chaque ligne contienne l’information sur un grain, et que chaque colonne corresponde à une direction dans l’espace. Il est alors nécessaire que l’ensemble de ces directions soit le même pour tous les grains. Mais quelle base choisir ?

Notre but étant de conserver le mieux possible la forme initiale des grains réels, la réponse est une base qui préserve la forme. Nous avons deux choix possibles en coordonnées sphériques :

1. Une base  $(\theta, \varphi)$  avec  $\Delta\theta_i = cte$  et  $\Delta\varphi_i = cte$ . C’est une base uniforme. La Fig. 2.12 montre un exemple du résultat de cette projection avec  $d = 800$  points, où  $0 \leq \theta \leq 2\pi$ ,  $-\pi \leq \varphi \leq \pi$ ,  $\Delta\theta_i = \frac{2\pi}{d}$  et  $\Delta\varphi_i = \frac{\pi}{d}$ . L’implémentation et l’exploitation de cette base est simple, mais elle présente l’inconvénient de la concentration aux pôles. Ce choix est donc peu adapté.

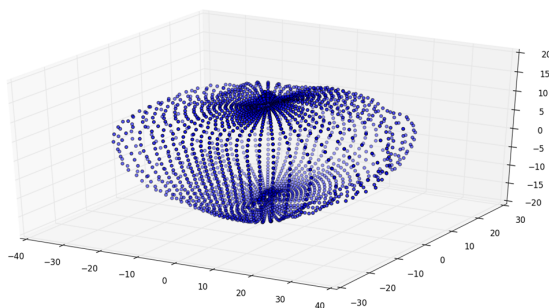


FIGURE 2.12 – Interpolation sur une base à pas constants pour un grain de ballast

2. Une base  $(\theta, \varphi)$  à pas variable. Le choix d’une telle base doit être minutieux, afin de garder une dispersion équi-distante et uniforme des points. On choisit d’appliquer l’algorithme appelé Répulsion des Points. Le principe de cet algorithme détaillé en [95] est de repositionner les points de façon à avoir une répartition uniforme en dépit de la texture de la surface. Pour appliquer cet algorithme, on choisit d’abord le nombre de points cible (ici on a fait le choix

de  $d = 800$  inférieur au nombre de grains  $n=975$  pour les raisons expliquées ci-dessus), pour ensuite les placer d'une manière aléatoire sur la surface.

Une fois tous les points positionnés sur la surface, la seconde étape est une procédure de "relaxation" qui repousse chaque point de ses voisins : pour chaque point  $\mathbf{p}$  dont on définit le voisinage de points  $\mathbf{p}_i, i \in N_p$  avec le rayon  $R$  (toujours par le même algorithme KD-Tree), ces points sont projetés sur le plan tangent à la surface du grain en ce point, puis une force de répulsion est calculée comme suit :

$$F_i(\mathbf{p}) = k(R - |\mathbf{p} - \mathbf{p}_i|) \cdot (\mathbf{p} - \mathbf{p}_i) \quad (2.37)$$

où  $F_i(\mathbf{p})$  est la force appliquée par le point  $\mathbf{p}_i$  sur le point  $\mathbf{p}$ ,  $k$  est une constante et  $R$  est le rayon de répulsion. La force totale appliquée au point  $\mathbf{p}$  est donc :

$$F(\mathbf{p}) = \sum_{i \in N_p} F_i(\mathbf{p}) \quad (2.38)$$

Nous utilisons ensuite cette force de répulsion pour déplacer le point, et on applique cette procédure à tous les points. Fig. 2.13 montre le résultat sur un grain. On voit bien que le problème de concentration des points n'est plus présent.

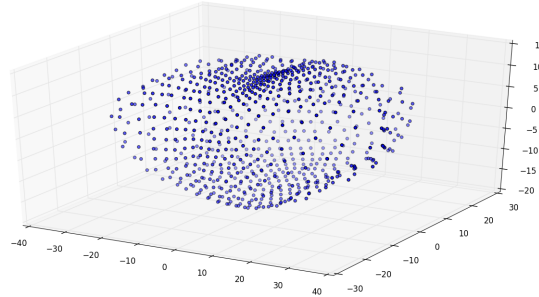


FIGURE 2.13 – Interpolation sur une base obtenue par répulsion de points

L'information sur les grains est donc maintenant sous forme de  $r_{ij} = r_i(\theta_j, \varphi_j), i = 1, \dots, n, j = 1, \dots, d$  où  $\theta_i = (\theta_j)$  et  $\varphi_i = (\varphi_j) \forall i \in 1, \dots, n$ .

#### d. Centrage des grains autour de l'origine

Le centrage est effectué en positionnant le centre des gravité des grains à l'origine.

#### e. Alignement des grains

Nous réorientons tous les grains dans leurs repères principaux. Pour effectuer cette opération, nous avons besoin de déterminer les axes principaux de chaque grain. La méthode classique est de calculer le tenseur d'inertie  $[J]$ , de le diagonaliser, et d'en tirer les 3 axes d'inertie que représentent les vecteurs propres issus de cette diagonalisation, tandis que les valeurs propres, elles, représentent les 2 longueurs principales.



Sur une base à pas constants (comme évoqué dans la section 2.1), ce calcul s'effectue comme suit :

$$[J] = \int_V \rho \{ |\mathbf{GM}|^2 [I] - \mathbf{GM} \otimes \mathbf{GM} \} dV \quad (2.39)$$

Si on centre l'origine du repère sphérique au centre de gravité,  $\mathbf{GM} = r(\theta, \varphi) \mathbf{e}_r(\theta, \varphi)$ , alors on a :

$$\begin{aligned} [J] &= \int_V \rho r^4(\theta, \varphi) \{ [I] - \mathbf{e}_r(\theta, \varphi) \otimes \mathbf{e}_r(\theta, \varphi) \} dr d\theta d\varphi \\ &\approx \rho \frac{\Delta\theta \Delta\varphi}{5} \sum_{i=1}^{N_\theta} \sum_{j=1}^{N_\varphi} \sin(\varphi_j) R^5(\theta_i, \varphi_j) \{ [I] - \mathbf{e}_r(\theta_{i,j}) \otimes \mathbf{e}_r(\theta_{i,j}) \} \end{aligned} \quad (2.40)$$

Enfin, la matrice d'inertie est définie positive et peut donc s'écrire sous la forme :

$$[J] = [D] [\lambda] [D]^T \quad (2.41)$$

avec  $[\lambda]$  une matrice diagonale à éléments strictement positifs, ordonnés dans l'ordre décroissant, et  $[D]$  une matrice orthogonale.

En effectuant la rotation  $\mathbf{GM} \rightarrow [D]^T \mathbf{GM}$  pour tout point M, le grain est alors positionné dans son repère d'inertie propre.

Cette opération nécessite la réinterpolation de chaque grain sur une base à pas angulaires constants, et un nombre important de produits matriciels. Afin de faciliter cette procédure, nous avons recours à une méthode donnant une approximation des axes principaux d'inertie en se basant sur les 3 directions privilégiées de la dispersion des points (méthode basée sur une POD appliquée à chaque grain tout seul).

La matrice à considérer pour chaque grain G est une matrice  $800 \times 3$  où chaque ligne correspond à un point et chaque colonne correspond à une direction parmi les 3 directions  $x, y, z$ . On résout le problème aux valeurs propres : Les valeurs propres sont alors une approximation des longueurs principales, et les vecteurs propres constituent les axes principaux. La rotation se fait tout simplement par multiplication par la matrice contenant ces vecteurs.

Un exemple de grain après l'ensemble de ces transformations est présenté sur la Fig. 2.14.

Une fois notre matrice de grains construite, nous procédons à l'application de la méthode.

### 3.4 Décomposition de la forme : résultats

Nous possédons donc la matrice A de dimension  $n \times d$  (où  $n = 975$  et  $d = 800$ ) et de rang  $r$  :

$$\mathbf{A} = \begin{pmatrix} r_1(\theta_1, \varphi_1) & r_1(\theta_2, \varphi_2) & \dots & r_1(\theta_{d-1}, \varphi_{d-1}) & r_1(\theta_d, \varphi_d) \\ r_2(\theta_1, \varphi_1) & r_2(\theta_2, \varphi_2) & \dots & r_2(\theta_{d-1}, \varphi_{d-1}) & r_2(\theta_d, \varphi_d) \\ \dots & \dots & \dots & \dots & \dots \\ r_n(\theta_1, \varphi_1) & r_n(\theta_2, \varphi_2) & \dots & r_n(\theta_{d-1}, \varphi_{d-1}) & r_n(\theta_d, \varphi_d) \end{pmatrix} \quad (2.42)$$

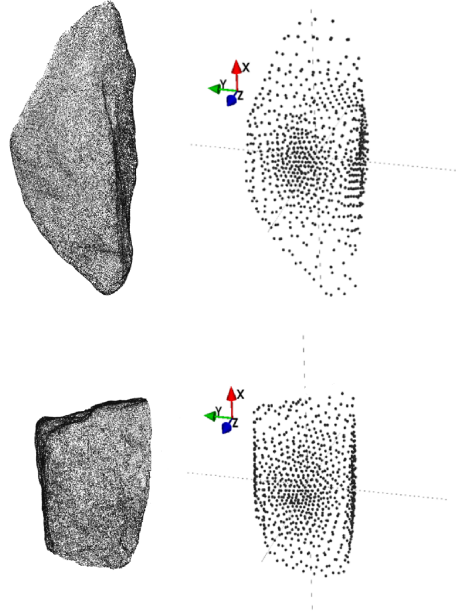


FIGURE 2.14 – Grains à la fin du pré-traitement

En gardant la notation  $r_{ij}$  pour représenter le rayon du grain  $i$  selon la direction  $j$ , on définit le grain moyen par ses rayons  $\bar{r}_j$  :

$$\bar{r}_j = \frac{1}{n} \sum_{k=1}^n r_{kj} = \frac{1}{n} \sum_{k=1}^n r_k(\theta_j, \varphi_j) \quad (1 \leq j \leq d). \quad (2.43)$$

Le grain moyen (Fig. 2.15) est proche d'un ellipsoïde de demi-longueurs principales  $(\bar{L}), (\bar{I})$  et  $(\bar{S})$  telles que  $\bar{S} \leq \bar{I} \leq \bar{L}$  et :

$$\bar{L} \simeq 29.6 \text{ mm}; \bar{I} \simeq 25.7 \text{ mm}; \bar{S} \simeq 21.6 \text{ mm}. \quad (2.44)$$

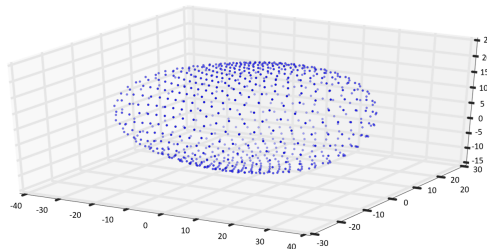


FIGURE 2.15 – Grain moyen

Comme expliqué précédemment, on travaille avec une matrice centrée pour chaque grain. Morphologiquement parlant, on évalue la différence au grain moyen, c'est-à-dire que l'on garde l'écart sur la forme de chaque grain au grain moyen :

$$\Delta r_{ij} = r_{ij} - \bar{r}_j \quad (2.45)$$

Les écarts au grain moyen définis dans l'équation 2.45 définissent la matrice  $\mathbf{A}$ . En considérant  $\|\bar{r}\|$  la norme euclidienne du vecteur  $d$ -dimensionnel  $\bar{r}$ , on a par définition :

$$\|\mathcal{A}\|^2 = n\|\bar{r}\|^2 + \|\mathbf{A}\|^2 \quad (2.46)$$

Le choix de travailler sur la matrice centrée, au-delà de son importance pour le cœur de la POD, interroge sur la quantité d'information que représente cette partie par rapport au grain moyen. La contribution relative  $\delta_1$  de cette variation de forme d'un grain à l'autre, au grain moyen, est exprimée par le rapport suivant :

$$\delta_1 = \frac{\|\mathbf{A}\|}{n^{1/2}\|\bar{r}\|} \simeq 20\% \quad (2.47)$$

Ce faible apport signifie que notre analyse se focalise sur une caractérisation de nature autre que celle du grain moyen, mais ne signifie en aucun cas un jugement sur l'importance de chacune des parties (grain moyen et variation grain-par-grain) dans la forme globale. L'analyse similaire de la contribution de la variation grain-par-grain à la forme globale donne un résultat d'environ 30%. Si on suppose que le grain moyen contient une partie ou toute l'information sur le premier ordre de la forme, nous travaillons alors essentiellement sur les deux autres ordres : l'angularité et la texture de surface.

L'étape suivante est d'appliquer la POD à notre matrice. En résolvant le problème aux valeurs propres de la matrice  ${}^T\mathbf{A} \cdot \mathbf{A}$ , les sorties de la POD consistent aux matrices suivantes :

- La matrice  $U$  : de dimension  $d \times d$  contenant les modes propres. La matrice tronquée à l'ordre  $K$  est  $U_{(K)}$ .
- La matrice  $E$  : de dimension  $d \times d$  contenant les valeurs propres  $\lambda_i$ . La matrice tronquée à l'ordre  $K$  est  $E_{(K)}$ .
- La matrice  $C = \mathbf{A}U$  : de dimension  $nd$  contenant les coefficients. La matrice tronquée à l'ordre  $K$  est  $C_{(K)}$ .

L'erreur relative, comme notée précédemment, s'écrit :

$$\epsilon(K) = \frac{\|\mathbf{A} - \mathbf{A}_{(K)}\|_F^2}{\|\mathbf{A}\|_F^2} = \frac{\sum_{l=K+1}^r \lambda_l}{\sum_{l=1}^r \lambda_l}. \quad (2.48)$$

De la même manière,  $\zeta(K) = 1 - \epsilon(K)$  définit la quantité d'information conservée sur la forme, en ne gardant que les  $K$  premiers modes comme suit :

$$\mathbf{A}^{(K)} = C^{(K)}U^{(K)} \quad (2.49)$$

et la matrice totale  $\mathcal{A}^{(K)}$  est formée comme suit :

$$\mathcal{A}^{(K)} = C^{(K)}U^{(K)} + \bar{r} \quad (2.50)$$

La forme représentée par la matrice  $\mathbf{A}$  est donc décomposée en un ensemble de coefficients  $C$  et d'un ensemble de modes  $U$  qu'on appelle également les fonctions

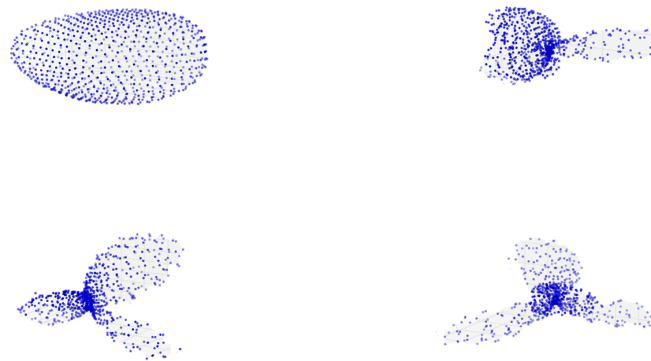


FIGURE 2.16 – Premiers modes propres issus de la POD

de forme. La représentation géométrique des premiers modes est présentée sur la Fig. 2.16.

Plusieurs informations peuvent être tirées de la représentation des modes.

1. Vu que l'on s'intéresse à la caractérisation de la forme de particules, un élément de comparaison important dans ce travail est la décomposition en harmoniques sphériques. La Fig. 2.17 montre les premières harmoniques sphériques. Nous notons une certaine similarité entre les modes POD et les modes HS. Il est à noter que les modes HS sont indépendants des données d'entrée à décomposer, contrairement aux modes POD qui sont propres à chaque matrice de grains. Une comparaison plus détaillée des deux décompositions sera présentée par la suite.

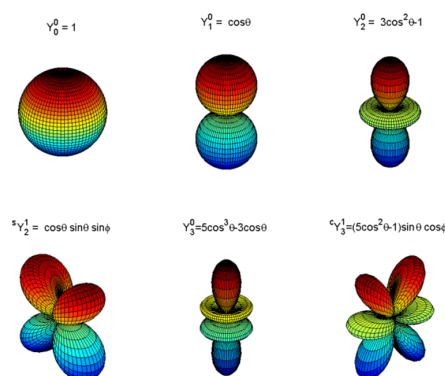


FIGURE 2.17 – Premiers modes propres issus des harmoniques sphériques

2. Il est important de noter la ressemblance du mode 1 avec le grain moyen, mais également les différences : le mode 1 semble contenir une information liée en partie à la taille du grain mais essentiellement à son élongation, puisqu'il a un rapport d'aspect plus faible que le grain moyen, et une élongation plus grande que celui-ci. On pourrait alors se demander à quoi ressemble le plus petit et le

plus gros grain qui s'obtiennent par la reconstruction suivante :

$$Gg = C_{1max}u_1 + \bar{r} \quad (2.51)$$

et :

$$Pg = C_{1min}u_1 + \bar{r} \quad (2.52)$$

où  $C_{1min}$  et  $C_{1max}$  correspondent respectivement à la valeur minimale et maximale du premier coefficient POD. La Fig. 2.18 montre ces grains.

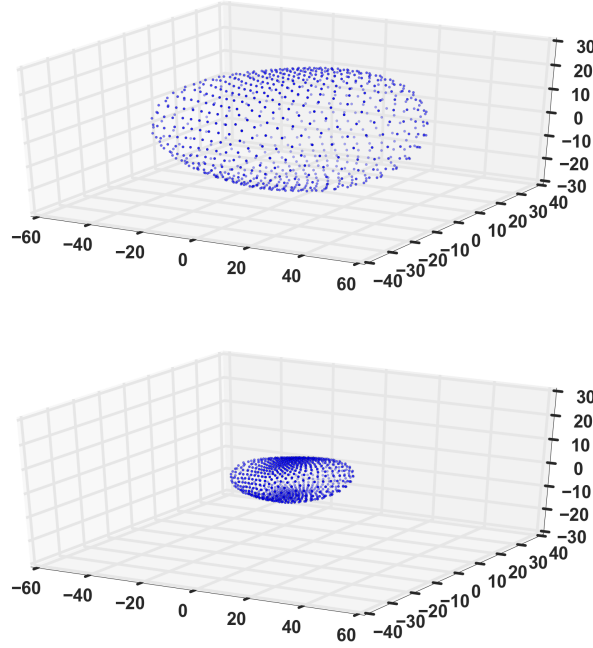


FIGURE 2.18 – Plus gros grain (haut) et plus petit grain (bas)

Quoique le second mode puisse changer l'indice du plus gros/petit grain, à partir du grain moyen, l'ajout du premier mode/coefficient montre que l'on peut obtenir deux tailles significativement différentes des grains. Le premier mode, et donc le premier coefficient contient bien une information sur la taille. Le résultat donné par Eq. 2.47 semble prendre davantage de sens : le taux de contribution de la variation grain-par-grain de la forme au grain moyen relève de l'élément commun qu'ils partagent : la taille du grain, contenue en partie dans le premier mode reconstruit et en partie dans le grain moyen. Comme évoqué plus haut, il s'avère que nous travaillons sur les trois ordres de la forme, même en écartant le grain moyen : une partie du premier ordre, ainsi que les deux ordres inférieurs.

Pour ce qui est des coefficients, les distributions statistiques (PDF) des six premiers coefficients sont présentées sur la Fig. 2.19. Les coefficients possèdent des distributions non gaussiennes toute centrées autour de 0. Les valeurs du premier coefficient se situent entre  $c_{1min} \approx -291$  et  $c_{1max} \approx 238$ . Cet intervalle se rétrécit significativement dès le deuxième coefficient, dont les valeurs se situent entre  $c_{2min} \approx -143$  et  $c_{2max} \approx 99$ . Les valeurs des autres coefficients se resserrent ensuite davantage autour de 0. La décroissance rapide de  $C_1$  à  $C_i$ ,  $i > 1$  confirme le résultat obtenu sur le premier mode par rapport au reste des modes : le premier coefficient

contient une information supplémentaire sur la taille des grains, leur élongation et leur aplatissement. Le reste des modes relèvent alors essentiellement des deux ordres inférieurs : l'angularité et la texture de la surface. Enfin, l'aspect non gaussien des distributions n'est pas à négliger : ceci montre la complexité de la forme des grains de ballast.

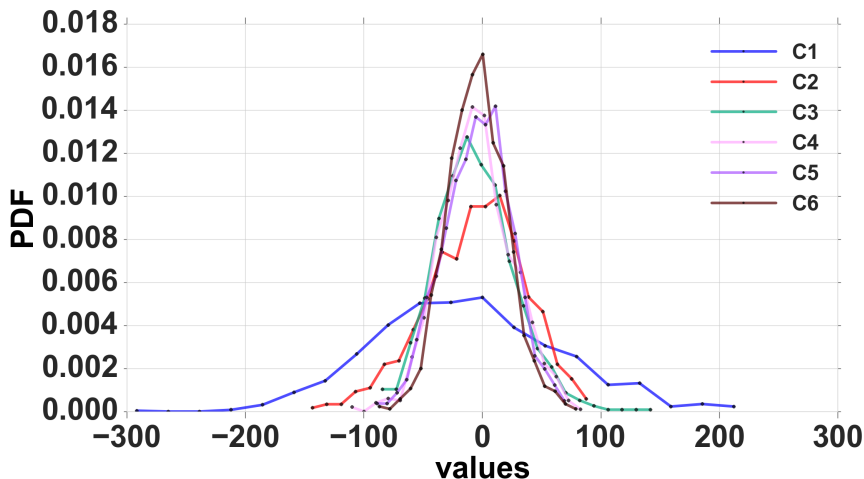


FIGURE 2.19 – Les distributions statistiques (PDFs) des premiers coefficients POD

La décomposition en modes et en coefficients est spécifique à la population de grains constituant nos données d'entrée, et englobe les propriétés de forme communes entre ces grains, des plus dominantes aux moins dominantes (au sens de leur contribution à la quantité d'information  $\zeta$ ). La variation de l'ordre de troncature correspond au nombre de modes retenus et contrôle le degré de détail sur la forme, et permet de reconstruire notre population de grain avec différentes précisions, quantifiées par l'erreur relative  $\epsilon$ . La Fig. 2.20 montre l'évolution de cette erreur en fonction du nombre de modes conservés. Par construction, l'erreur  $\epsilon(K)$  diminue en augmentant le nombre de modes et est égale à 0 pour  $K = p$  (ici  $p = d$  puisque  $d < n$ ).

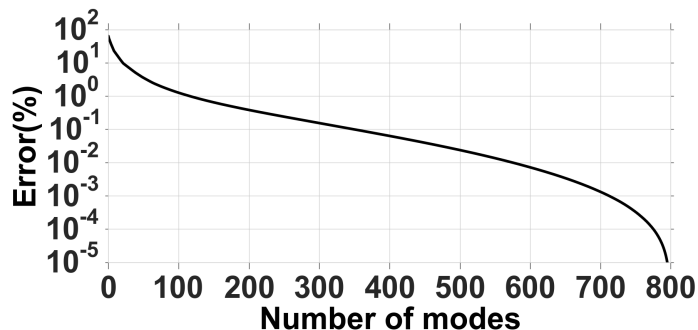


FIGURE 2.20 – Erreur relative sur la forme en fonction du nombre de modes

En observant l'évolution de l'erreur, nous remarquons bien que la méthode a un excellent potentiel de réduction. En effet, en gardant uniquement 20 modes au

lieu de 800 (2.5% du nombre total des modes), l'erreur demeure en dessous de 10%. Ceci s'explique aussi par la forme du 1<sup>er</sup> mode, dont la forme est déjà proche d'un ellipsoïde, tandis que le reste des modes, présentant des lobes positifs et négatifs, est associé aux zones à courbure élevée. La pertinence de l'erreur définie à quantifier la variation de la forme est d'autant plus grande que l'échantillon de grains est grand.

Ensuite, nous choisissons d'évaluer un autre type de performance de cette décomposition : son potentiel prédictif, c'est-à-dire sa faculté à représenter des formes de grains extérieurs à l'échantillon à partir duquel elle a été établie. Nous utilisons pour cela une méthode d'estimation de fiabilité basée sur une technique d'échantillonnage : la validation croisée, ou "cross-validation" [96–98]. Le but est de tester la faculté de la base construite à partir d'un sous-échantillon appelé d'apprentissage (training dataset), de représenter un autre sous-échantillon appelé échantillon de test ou de validation (testing/validation dataset).

Le choix le plus classique consiste à diviser notre échantillon en deux sous-échantillons : 70% pour l'apprentissage, et 30% pour le test : c'est la validation conventionnelle. Toutefois, la réduction de l'échantillon d'apprentissage à partir duquel la base POD est extraite est pour nous contraignante par construction de la méthode, et est préjudiciable à la décomposition optimale de la forme. La validation croisée, et la méthode du "leave-one-out" (LOO) [99], en particulier, présente l'avantage de ne pas réduire l'échantillon d'apprentissage.

Le leave-one-out (LOO) est un cas particulier du "Leave-p-out" pour  $p = 1$ . Le principe est simple :  $p$  réalisations sont utilisées comme échantillon de test, et le reste comme échantillon d'apprentissage, puis l'opération est répétée pour tous les tirages de  $p$  éléments dans  $n$ . Le nombre  $N_t$  de tests du modèle s'élève alors à  $C_n^p$ , où  $C_n^p$  est le coefficient binomial [100]. Pour de grandes valeurs de  $n$ , le calcul  $N_t$  fois peut devenir très fastidieux. Par exemple, pour  $n = 100$  et  $p = 30$ ,  $C_{100}^{30} \approx 3 \times 10^{25}$ .

Dans notre cas, avec  $p = 1$ , nous sélectionnons à chaque fois un grain  $g_{init} = (r_l(\theta_j, \varphi_j))$  où  $l \in 1, \dots, n$  dans notre échantillon qui n'a pas servi à la construction de la base POD, et nous le reconstruisons avec la base obtenue ( $g_{recon} = g_{init}UU^T + \bar{r}$ ). L'erreur quadratique est ensuite calculée entre le grain initial  $g_{init}$  et le grain reconstruit  $g_{recon}$  avec tous les modes, et ceci est répété  $C_n^1 = n$  fois, c'est-à-dire pour tous les grains. La moyenne des erreurs quadratiques obtenues constitue l'erreur de prédiction. Dans notre cas, cette erreur vaut  $3.25 \cdot 10^{-5}$ , ce qui prouve que notre base POD représente de manière satisfaisante la forme des grains.

Notre objectif étant d'utiliser les populations reconstruites à différents modes pour des fins de simulations numériques, nous avons choisi de reconstruire nos grains avec les seuils d'erreur présentés sur le Tab. 2.1.

TABLE 2.1 – Nombre de modes et erreur correspondante sur la forme. L'erreur diminue à 0 pour 800 modes.

Error	20%	15 %	10%	5%	1%	0.1%
Mode number	11	16	22	40	117	350

Ces seuils d'erreur sont traduits morphologiquement par des grains à formes différentes. Un exemple de grain reconstruit à ces mêmes seuils d'erreurs (degrés de précision) est présenté sur la Fig. 2.21. On voit que la forme se rapproche de la forme réelle en augmentant le nombre de fonctions de forme. De la même manière

que pour l'évolution de l'erreur, nous remarquons que 117 modes, correspondant à 99% de l'information se traduit également par une bonne approximation visuelle du grain initial. Néanmoins, il est à noter que la traduction morphologique de 11 modes, représentant 80% de l'information pose la question sur l'efficacité de l'erreur à caractériser les aspects de l'angularité et de texture de surface.

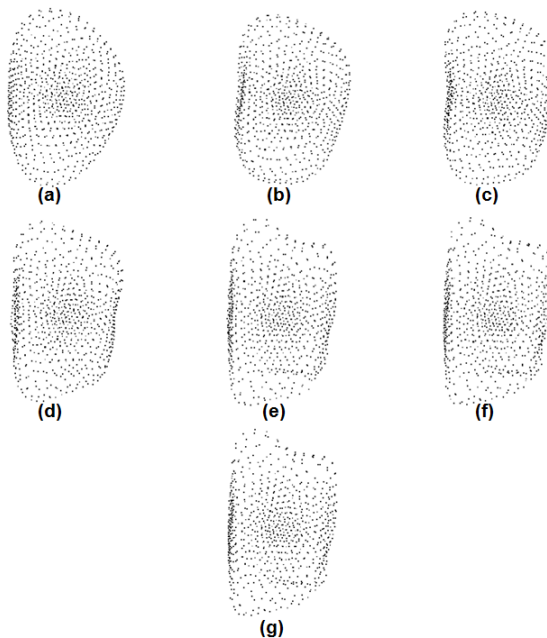


FIGURE 2.21 – Evolution de la forme en fonction du nombre de modes pour suivant le Tab. 2.1

Pour évaluer l'optimalité de la POD appliquée à nos grains, il est intéressant d'établir le lien de cette méthode de caractérisation de la forme avec les indicateurs scalaires et spectraux. Pour ce faire, nous comparons les résultats de cette approche avec ceux des méthodes évoquées précédemment.

### 3.5 La caractérisation par POD vs la caractérisation scalaire

Nous mesurons donc des indicateurs comme la surface  $\mathcal{S}_i$ , le volume  $\Omega_i$ , et le rayon moyen  $\bar{R}_i$ ,  $1 \leq i \leq n$  ainsi que l'élongation  $E$ , l'aplatissement  $F$  et le rapport d'aspect  $a$ , pour tous les grains, reconstruits aux seuils d'erreurs présentés dans le Tab. 2.1.

Nous comparons donc les distributions statistiques d'échantillons de même dimension  $n = 975$ .

Les Fig. 2.22 et 2.23 présentent les fonctions de densité de probabilité (PDF) de  $\mathcal{S}$ ,  $\Omega$ ,  $\bar{R}$ ,  $E$ ,  $F$  et  $a$  pour les grains reconstruits et les grains scannés.

Le degré d'approximation de ces distributions à celle des grains scannés peut être quantifié à l'aide des valeurs de Kolmogorov et Smirnov (KS). C'est un test statistique de vraisemblance, qui quantifie la probabilité que deux échantillons suivent la même distribution. Ces valeurs de vraisemblance, calculées pour les 6 indicateurs,



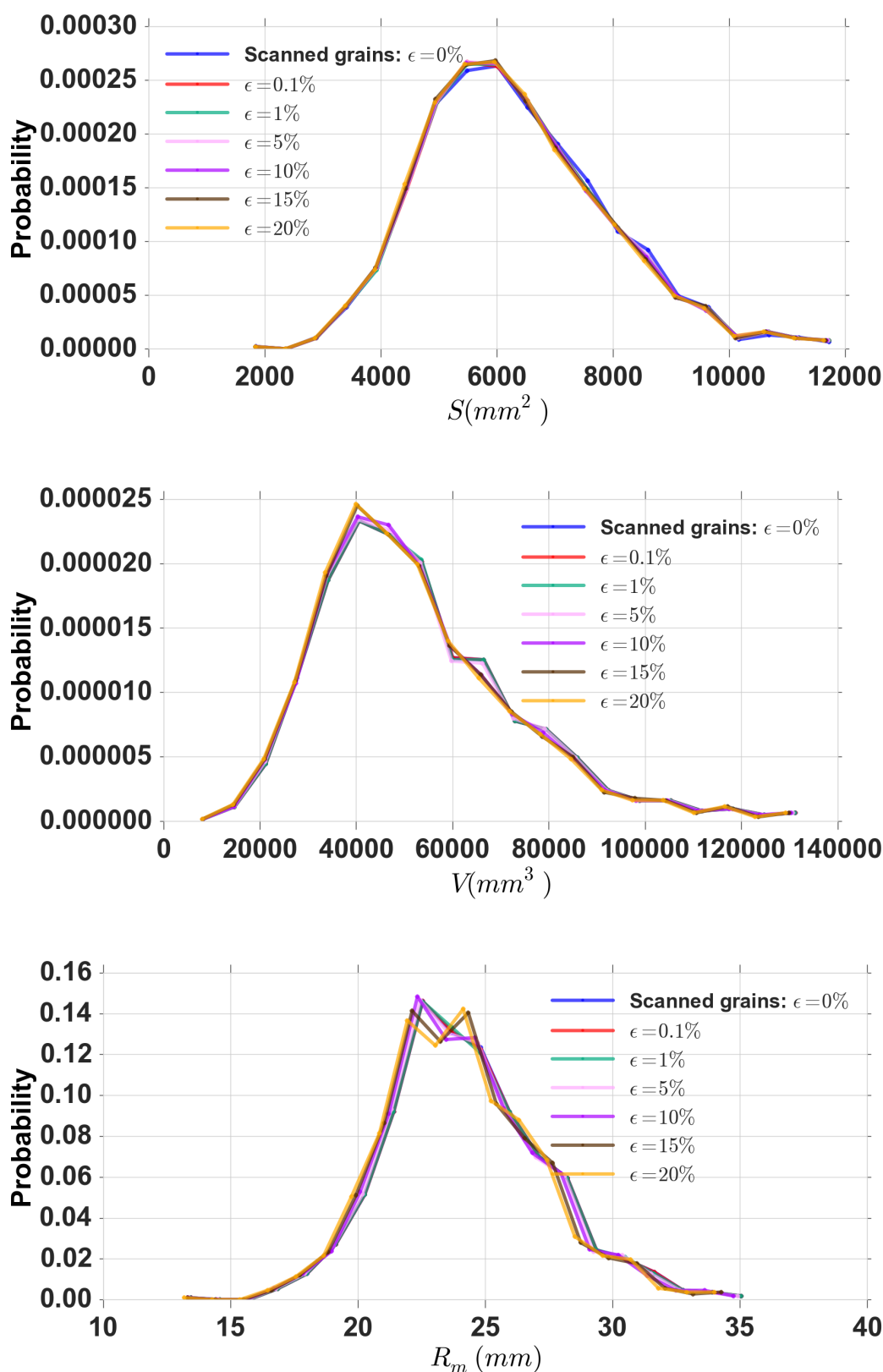


FIGURE 2.22 – PDFs de  $S$ ,  $V$ ,  $\bar{R}$  pour différents seuils d'erreur, comparées avec celle des grains scannés

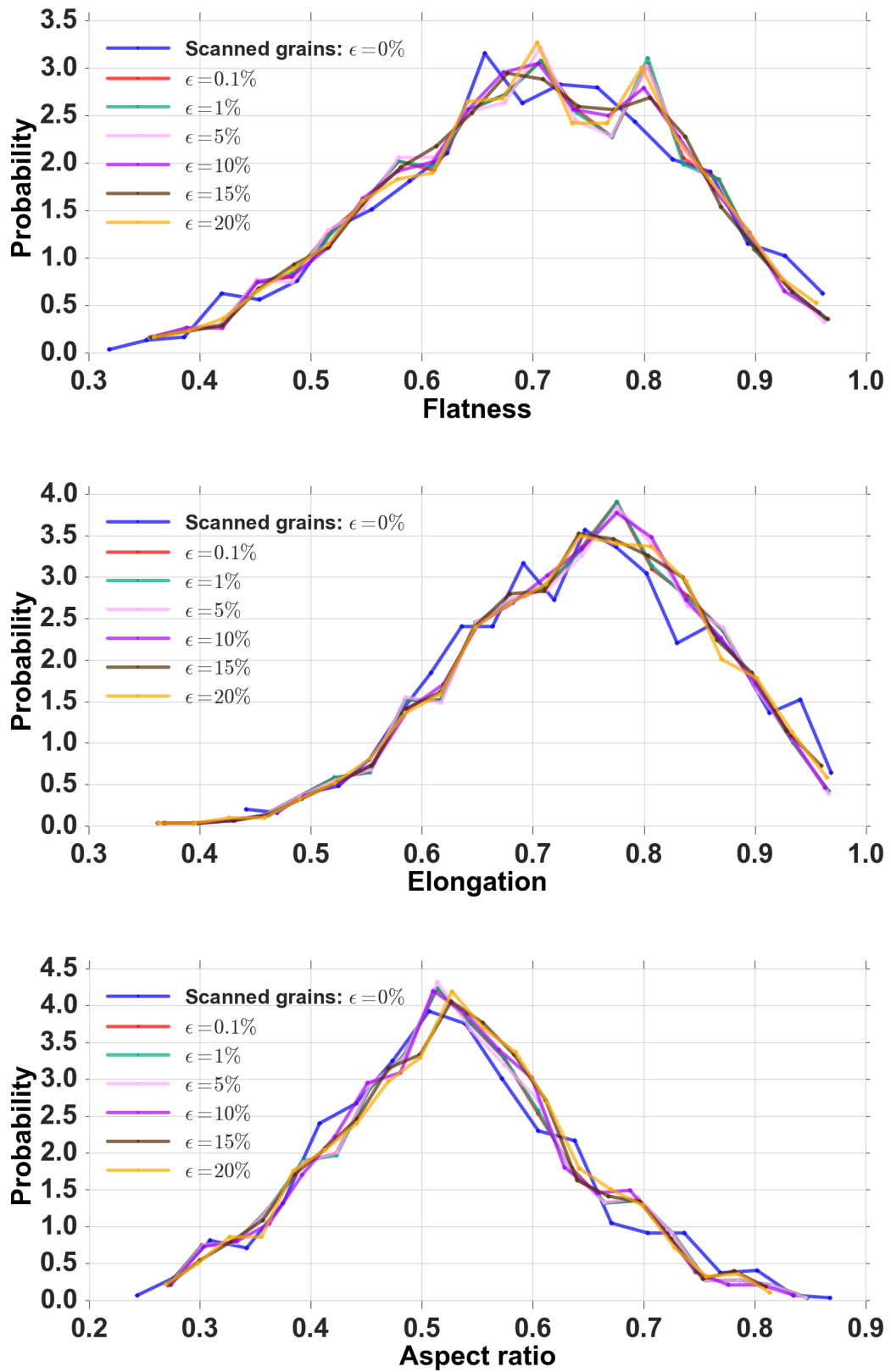


FIGURE 2.23 – PDFs de  $E$ ,  $F$  et  $a$  pour différents seuils d'erreur, comparées avec celle des grains scannés

sont présentées sur le Tab. 2.2, montrent qu'on reproduit très bien les propriétés de forme, même avec plus de 10% d'erreur sur la forme.

TABLE 2.2 – Résultats de Kolmogorov et Smirnov, pour différents ordres de troncature  $k$ , pour les distributions du volume  $V$ , de la surface  $\mathcal{S}$ , du rayon moyen  $\bar{R}$ , de l'aplatissement  $F$ , de l'élongation  $E$  et du rapport d'aspect  $a$ .

$k$	11	16	22	40	117	350
KS ( $\Omega$ )	0.866	0.879	0.984	0.997	0.997	0.997
KS ( $\mathcal{S}$ )	0.975	0.984	0.997	0.999	0.999	0.999
KS ( $\bar{R}$ )	0.725	0.737	0.764	0.785	0.849	0.947
KS ( $F$ )	0.926	0.943	0.965	0.973	0.987	0.998
KS ( $E$ )	0.934	0.941	0.945	0.981	0.986	0.996
KS ( $a$ )	0.887	0.888	0.919	0.928	0.965	0.983

Pour certains indicateurs scalaires comme le volume ou le rayon moyen, les Fig. 2.22(a) et Fig. 2.22(c) montrent une différence entre les grains scannés et reconstruits à partir de  $\epsilon = 10\%$  à  $\epsilon = 15\%$ . Pour la surface (Fig. 2.22(b)), nous commençons déjà à avoir une faible différence à partir de  $\epsilon = 1\%$ . Ensuite, pour les trois indicateurs scalaire  $E$ ,  $F$  et  $a$  (Fig. 2.23), la différence est déjà observable à partir de 0.1%. Ces observations confirment la nature des indicateurs calculés : Les indicateurs qu'on a choisi d'appeler "indicateurs simples" qui sont également des mesures permettent moins de détecter les variations de la forme, que les "indicateurs avancés" définis dans la section précédente.

Cette comparaison entre deux méthodes de caractérisation (scalaire vs par POD) permet de tirer deux conclusions essentielles. D'une part, le potentiel de la méthode à caractériser, à détecter finement les variations de la forme, et à reconstruire des grains à précision contrôlée sur la forme. D'une autre part, au vu des formes reconstruites avec peu de modes, la faiblesse des indicateurs scalaires devant cette méthode, notamment à décrire pertinemment les différences d'angularité et de texture. Afin d'assurer un traitement de ces aspects, nous avons alors besoin de caractériser l'angularité, les faces et les arêtes par des indicateurs plus pertinents, inexistant jusqu'ici dans la littérature. Cette partie est abordée dans le chapitre suivant.

Par analogie à la comparaison avec la caractérisation scalaire, nous comparons ensuite la caractérisation par POD à la caractérisation spectrale par harmoniques sphériques (SH).

### 3.6 La POD vs les harmoniques sphériques

Afin de faire une comparaison avec la même quantité de données à manipuler, nous faisons une décomposition en harmoniques sphériques en utilisant  $N_c = d = 800$  coefficients (SH), ce qui est équivalent à 100% d'information (au sens de la POD). Nous reconstruisons ensuite nos grains avec différents degré de précision, et nous comparons ces résultats avec ceux de la POD.

La Fig. 2.24 montre un exemple d'un grain reconstruit à l'aide de 800 coefficients POD, puis à l'aide de 800 coefficients SH. On peut très bien observer que pour un type de particule dont la forme est aussi complexe que le ballast, une décomposition du type (SH) nécessite une base de projection plus grande, puisqu'un nombre

maximal de 800 coefficients ne parvient pas à reconstruire la forme du grain scanné.

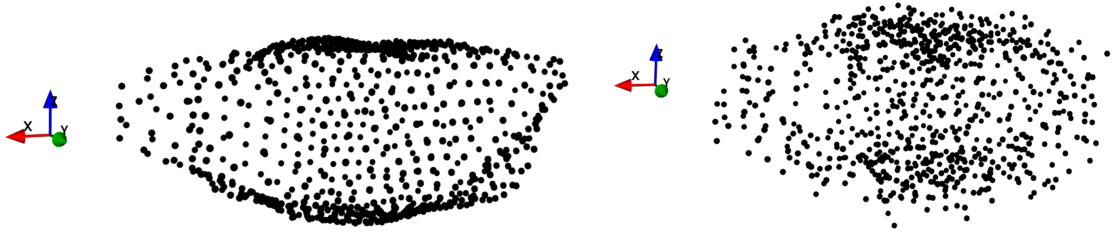


FIGURE 2.24 – Grains de ballast reconstruit par POD (gauche) puis par SH (droite) pour le même nombre  $N_c = 800$  de coefficients

Afin de comparer des quantités d'informations équivalentes, la décomposition a été faite avec des bases de dimensions différentes, et nous trouvons qu'une base de dimension 3600 coefficients permet d'avoir la précision voulue (100%) correspondant à la forme du grain scanné. Fig. 2.25 montre l'évolution de la même erreur  $\epsilon$  en fonction du nombre de coefficients retenus. On voit clairement l'apport de la POD en terme d'optimisation de la base de décomposition. En effet, l'avantage de la POD est d'avoir une base qui s'adapte à la morphologie des grains injectés dans la décomposition, contrairement aux harmoniques sphériques, dont la principale caractéristique est d'avoir une base imposée, définie quelque soit le type de données d'entrée. Il est alors logiquement plus difficile de reconstruire une forme d'autant plus qu'elle s'éloigne de la forme d'une sphère, ou d'une forme à texture arrondie. La comparaison des formes des modes dans les Fig. 2.16 et Fig. 2.17 permet de confirmer cela. La flexibilité de la POD permet de mieux apprécier le sens morphologique des modes, et permet surtout d'obtenir une base optimale.

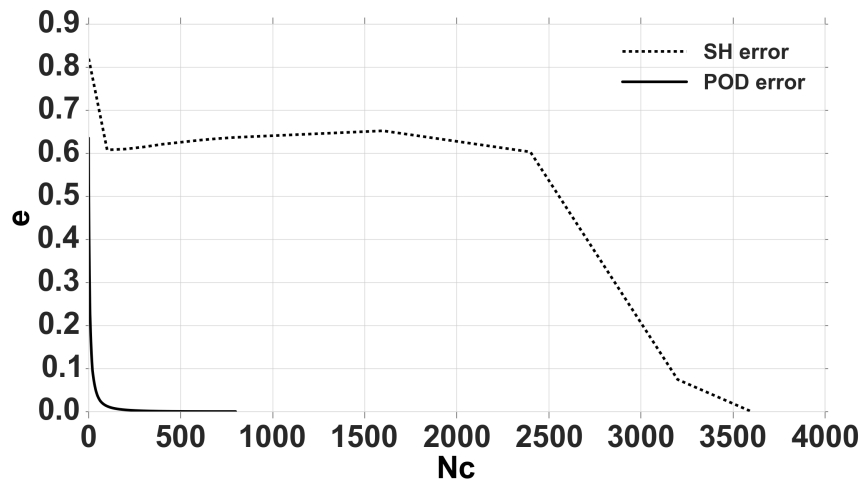


FIGURE 2.25 – Erreur en fonction du nombre de coefficients/modes pour les deux approches POD et SH

## 4 Conclusion

En conclusion, nous avons démontré par cette méthode, un grand potentiel de caractérisation, de reconstruction, et de contrôle de la forme, potentiel confirmé par les indicateurs scalaires simples ainsi que par les méthodes spectrales existantes. L'une des questions qui se posent est la comparaison avec des indicateurs de forme qui relèvent d'un degré plus avancé de la morphologie (faces, arêtes, angularité, etc.).

Autrement dit, la faculté de la méthode à optimiser l'information nous permet certes de réduire notre modèle de représentation de la forme tout en maintenant l'information essentielle, mais cette réduction aurait plus de sens si nous pouvions générer des grains équivalents aux grains scannés (en terme de forme), en ne gardant qu'un nombre réduit de modes et de coefficients.

Selon les résultats obtenus du paramètre (KS), on doit être capable de générer des grains équivalents en utilisant uniquement 117 modes au lieu de 800 modes, puisqu'on est à 99% de l'information globale.

Il est également important de garder en tête le besoin de relier l'erreur au sens POD à une erreur à sens physique, ou encore mieux, à sens mécanique. Ceci serait donné par une comparaison des résultats mécaniques donnés par les simulations numériques utilisant différentes configurations de forme dont l'erreur POD est connue.

Enfin, motivé par le besoin des calculs numériques d'un échantillon représentatif en termes de nombre, de variabilité et de forme, du ballast sur la voie ferroviaire, l'objectif du chapitre suivant est de générer des ensembles de grains équivalents aux grains scannés avec un modèle réduit, pour ensuite les injecter dans nos simulations numériques.

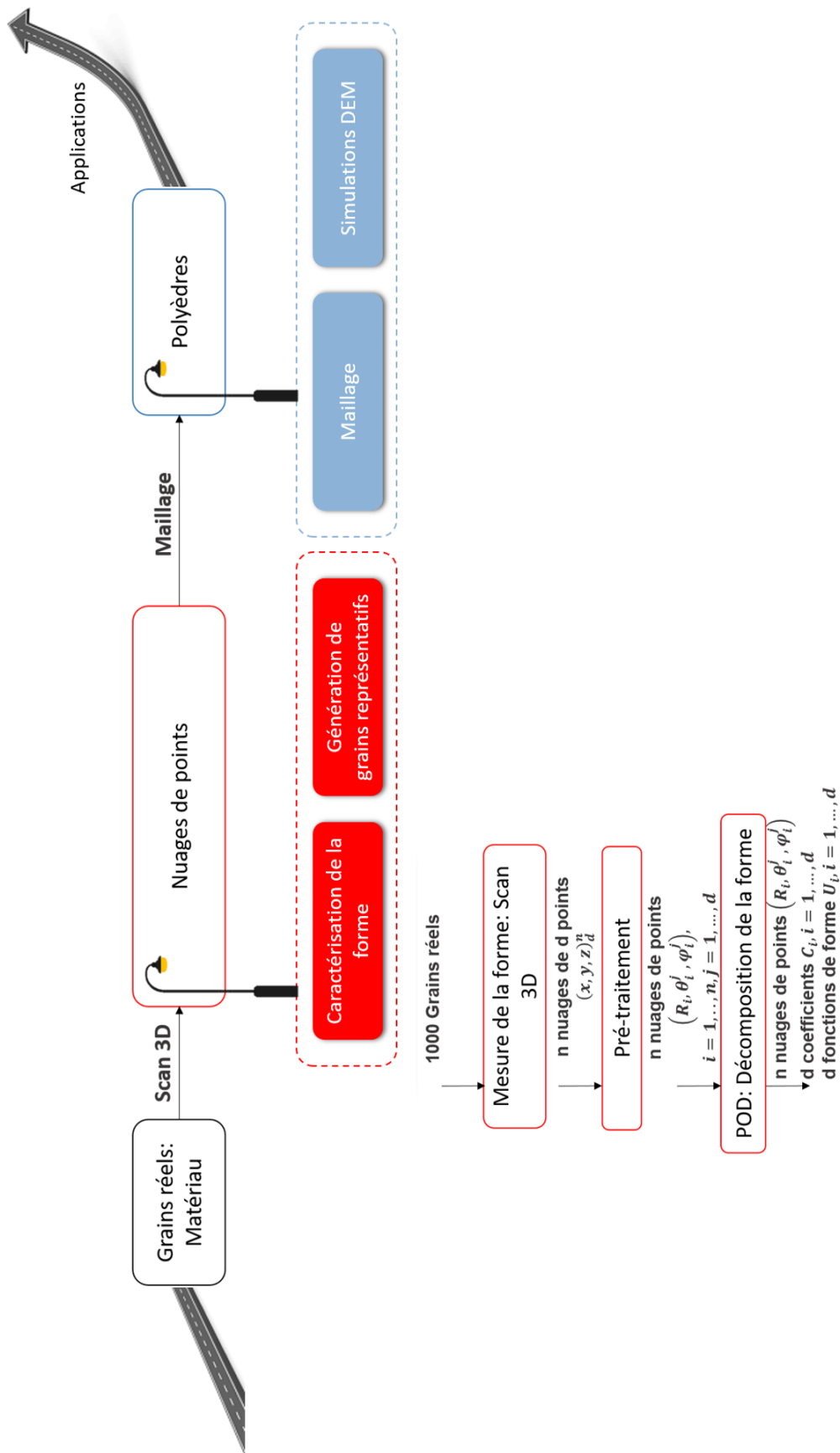


FIGURE 2.26 – Avancement sur le chemin d'étude du ballast



# Chapitre 3

## Génération de grains virtuels



Afin de garantir une bonne représentativité des simulations (DEM) dans les études de la forme, il est important de disposer d'une base de données de particules représentatives en terme de forme (reproduire la bonne primitive géométrique) et de variabilité (couvrir le spectre des variations possibles de la forme) : de grands assemblages de particules réalistes et distinctes. En 2D comme en 3D, plusieurs tentatives ont été faites, afin de générer les modèles les plus représentatifs de la réalité : des ellipses [101], des tiges [102], et des polygones [103] en 2D, ainsi que des formes 3D réalistes en utilisant la tomographie (mCT) [104]. Dans toutes ces études, nous remarquons un recours à la reconstruction de la morphologie particule-par-particule à partir d'un moyen de mesure donné (tomographie par exemple) [105]. Dans ces études parmi tant d'autres mettant en œuvre des simulations (DEM) avec des grains numérisés, deux limitations sont relevées :

1. Une limitation statistique en nombre et en variabilité des particules. En effet, à cause du nombre fini de particules scannées, la génération produit un grand nombre de particules identiques et répétées. Or, pour les sables et surtout pour le ballast, la micro-morphologie des grains est totalement aléatoire et distincte l'une de l'autre, et pourtant, celle-ci englobe plusieurs propriétés statistiques communes. Nous nous proposons alors de développer une méthode de génération de grands ensembles de grains ayant une morphologie équivalente à nos grains scannés.
2. Une limitation morphologique en forme des particules. En effet, comme nous le détaillerons dans la section suivante, les travaux actuels ne sont pas capables de générer des formes équivalentes aux formes réelles quand il s'agit de degré élevée d'irrégularité : nous soulignons notamment la difficulté de générer des particules anguleuses et facettées et le manque d'indicateurs prévus à cet effet, et nous nous proposons alors d'établir une méthode permettant de caractériser les faces et les courbures d'un nuage de points 3D.

Plusieurs études ont récemment été réalisées afin de relever ces défis. Ces méthodes proposées sont discutées dans les parties qui suivent, avec leurs avantages et limitations. Nous proposerons ensuite une méthode de caractérisation des faces et des courbures des nuages de points qui nous servira pour valider notre modèle de génération. Enfin, nous proposerons deux approches de génération appliquées au ballast, et nous discuterons les résultats obtenus et la validation du modèle par les indicateurs classiques et avancés.

## 1 Méthodes de génération de grains virtuels

### 1.1 Génération par indicateurs de Fourier

Comme expliqué dans le chapitre précédent, les indicateurs de Fourier permettent la décomposition de la forme et sa reconstruction à partir d'un spectre obtenu expérimentalement : à partir de la transformée de Fourier du signal discret  $r(\theta)$ , on obtient un spectre dont l'amplitude est utilisée pour caractériser la morphologie des particules, et dont les composantes permettent une reconstruction des particules à l'ordre voulu. Cette méthode a été utilisée dans plusieurs études notamment pour la

caractérisation de la forme des grains de sable. A partir de ce mécanisme, [69] propose d'inverser le concept : utiliser ce spectre afin de générer des grains aléatoires : ceci est fait en effectuant une transformation inverse d'un spectre de Fourier choisi. La Fig. 3.1 montre une réalisation du spectre généré aléatoirement avec l'approche détaillée dans [69] et le contour 2D correspondant.

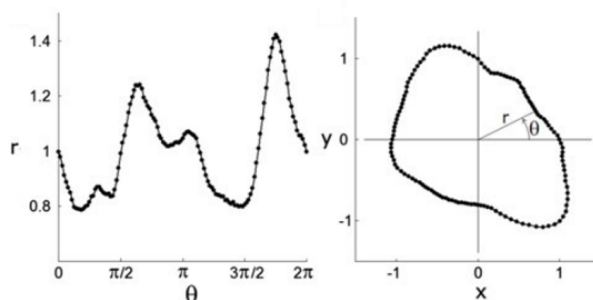


FIGURE 3.1 – Génération de grains virtuels 2D par transformation inverse de Fourier [69] : réalisation du spectre aléatoire généré (gauche) et contour 2D correspondant (droite)

Cette approche permet de générer des grains de sable en imposant leur taille, leur élongation, et leur orientation. Une extension directe des indicateurs de Fourier au cas tridimensionnel est proposée par [73, 74]. Elle consiste à choisir aléatoirement 3 contours 2D comme les trois sections orthogonales croisées d'un grain virtuel, puis à développant une méthode de révolution-dilatation de ces sections pour extrapoler la surface 3D (Fig. 3.2).

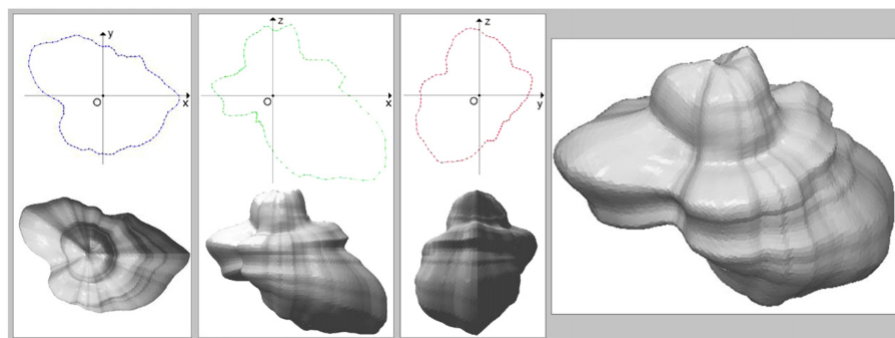


FIGURE 3.2 – Grain 3D généré par l'approche proposée dans [73]

Néanmoins, cette approche présente quelques limitations :

- La technique d'extrapolation ne produit pas de distribution isotrope des irrégularités : puisque la génération est basée uniquement sur 3 plans de coupe, les irrégularités du grain 3D obtenu par révolution sont basées sur ces sections au détriment des autres.
- La méthode ne parvient pas à générer des formes "facettées", alors que c'est le cas de l'échantillon initial des auteurs.

— Puisque les grains 3D sont obtenus à partir des 3 contours aléatoirement choisis, le résultat dépend fortement du choix de ces contours, d'où le risque élevé de générer des formes non voulues dans le cas où les 3 contours choisis donnent par exemple des grains d'un rapport d'aspect plus grande que celui de l'échantillon : nous notons une absence de corrélation/dépendance entre les contours choisis.

Une autre méthode plus adaptée au cas 3D classiquement utilisée permet de résoudre la 3<sup>me</sup> limitation, et d'améliorer la distribution des irrégularités (1<sup>re</sup> limitation).

## 1.2 Génération par harmoniques sphériques

Comme les indicateurs de Fourier, l'analyse en harmoniques sphériques permet non seulement la décomposition et la reconstruction de la forme, mais également la génération des particules.

Cette méthode a été présentée dans plusieurs études [30, 106]. On s'intéresse à l'approche présentée dans [29], puisqu'il s'agit de la modélisation de particules présentant une forme irrégulière et une angularité plus prononcée. Après une décomposition en harmoniques sphériques, la matrice contenant les coefficients  $c_n^m$  appelés indicateurs HS est exploitée pour générer des coefficients statistiquement équivalents. Cette étape est obligatoirement suivie par une PCA à des fins de réduction de modèle uniquement, car la modélisation de particules irrégulières nécessite un nombre total de  $2(n+1)^2$  indicateurs HS. La génération se fait ensuite en faisant l'hypothèse que les coefficients PCA issus de la réduction de modèle des coefficients SH, suivent tous une distribution gaussienne ce qui constitue une hypothèse forte mais à vérifier (cf. chapitre 2). En utilisant donc leur moyennes et leurs variances, on génère des distributions gaussiennes centrées réduites. Ensuite, les coefficients sont reconstruits comme suit :

$$\hat{c} = \bar{c}_n^m + p(x_1).c_v(x_1).x_2 \quad (3.1)$$

où  $\hat{c}$  est le coefficient SH généré,  $\bar{c}_n^m$  la moyenne du coefficient  $c_n^m$ ,  $p(x_1)$  la composante principale PCA  $n^o x_1$  où  $x_1 \in 1, \dots, K$ ,  $c_v(x_1)$  la variance correspondante, et  $x_2$  un nombre aléatoire qui suit une distribution gaussienne normalisée. Pour plus de détails, voir [29]. Le résultat de la génération est présenté sur la Fig. 3.3.

Visuellement, la méthode produit des ensembles de grains de formes très variables, mais avec des côtés arrondis et quasiment pas de faces. Les limitations de cette méthodes peuvent être résumées en ces quelques points :

1. L'absence de grains facettés, ce qui rend la méthode inappropriée pour des grains de ballast présentant des faces et des arêtes visibles à l'œil nu. A part l'aspect visuel, il n'existe pas, jusqu'à présent, de moyen pour caractériser la présence des faces et les courbures sur une particule représentée par un nuage de points 3D.
2. La performance de la méthode diminue d'autant plus que les grains sont irréguliers : ceci est montré par la validation du modèle en mesurant le volume et la surface pour deux populations de grains ayant deux niveaux d'irrégularité différents [29].
3. L'hypothèse du caractère gaussien des distributions de tous les coefficients produit des grains de forme irréaliste (Fig. 3.3) (nous remarquons cela également



FIGURE 3.3 – Grains générés avec l’approche des harmoniques sphériques(SH)

dans [74] en utilisant les indicateurs de Fourier combinés avec des vecteurs aléatoires gaussiens). Nous validerons l’effet de cette hypothèse dans le cas du ballast.

4. Plusieurs auteurs ont remarqué la présence de trous dans les pôles à cause de la construction de la méthode SH ainsi que par la méthode par indicateurs de Fourier [73, 107].

Afin d’éviter les limitations évoquées plus haut, nous proposons dans ce travail deux approches de génération de grains statistiquement équivalents à notre échantillon de grains de ballast scannés. La validation de ces méthodes pour ce qui est de la présence des faces sur les grains nécessite de construire un indicateur pour la caractérisation des faces/courbures que nous proposons d’abord dans la section suivante, à la hauteur des détails générés.

## 2 Caractérisation des faces/courbures des particules 3D

Nous nous proposons de définir une méthode qui permet de détecter la présence de faces et d’arêtes à partir d’un nuage de points 3D, forme sous laquelle se présenteront les grains scannés et générés. Cet indicateur est indispensable afin de vérifier la pertinence de la méthode de génération proposée dans la suite. L’idée globale consiste à identifier les faces sur chaque grain comme des groupes de sommets ayant des normales aux directions proches. Pour ce faire, nous expliquerons d’abord comment construire les ”vecteurs normaux aux sommets”, puis comment définir des faces et des arêtes en fonction de la variation locale de la surface.

### 2.1 Normales et variation de surface aux sommets

Pour l’approximation des normales par sommet, nous implémentons la méthode décrite dans [108] en utilisant les données issues du scan 3D des grains de ballast. Habituellement, un vecteur normal est défini par rapport à un plan. Or dans notre cas, le grain est représenté par un nuage de points. Pour obtenir une normale relative à un sommet, il est donc nécessaire d’obtenir des approximations de plans tangents

au grain à chaque sommet. Pour ce faire, nous utilisons l'algorithme KD-Tree [93] qui permet de déterminer un voisinage local  $N_p$  pour chaque point  $p$  (dans notre cas, nous avons choisi 9 voisins par sommet) comme présenté dans la Fig. 3.4.

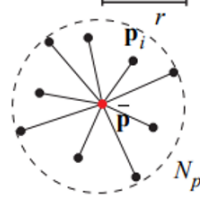


FIGURE 3.4 – Détection de voisinage pour un point 3D

A partir de ce voisinage défini pour chaque point (9 points les plus proches en distance), l'objectif est de déterminer un plan tangent et sa normale. La PCA (POD), expliquée dans le chapitre précédent, permet encore une fois de résoudre efficacement ce problème. En effet, la résolution du problème aux valeurs propres à partir de la matrice de covariance de  $N_p$ , permet de définir les axes portant le plus de variance. Dans ce cas, la variance se traduit par la dispersion dans l'espace : sur les trois axes définis par la PCA, les deux axes portant le plus de variance définissent le plan tangent au point  $p$ , et le troisième axe correspond donc à une estimation de la normale à ce plan.

La matrice de covariance  $C_v$  est définie comme suit :

$$\mathbf{C} = \begin{bmatrix} \mathbf{p}_{i_1} - \bar{\mathbf{p}} \\ \dots \\ \mathbf{p}_{i_k} - \bar{\mathbf{p}} \end{bmatrix}^T \cdot \begin{bmatrix} \mathbf{p}_{i_1} - \bar{\mathbf{p}} \\ \dots \\ \mathbf{p}_{i_k} - \bar{\mathbf{p}} \end{bmatrix}, i_j \in N_p \quad (3.2)$$

où  $\bar{\mathbf{p}}$  est le centroïde des voisins  $\mathbf{p}_{i_j}$  du point  $\mathbf{p}$ . La résolution aux valeurs propres de cette matrice de dimension  $3 \times 3$  nous donne :

$$\mathbf{C} \cdot \mathbf{v}_l = \lambda_l \cdot \mathbf{v}_l, l \in \{0, 1, 2\}. \quad (3.3)$$

$\lambda_l$  mesure la variation de  $\mathbf{p}_i, i \in N_p$  selon la direction du vecteur propre  $\mathbf{v}_l$  correspondant. En définissant les valeurs propres telles que  $\lambda_0 \leq \lambda_1 \leq \lambda_2$ , le plan  $T$  passant par  $\mathbf{p}$  :

$$T(\mathbf{x}) : (\mathbf{x} - \mathbf{p}) \cdot \mathbf{v}_0 = 0 \quad (3.4)$$

minimise la somme des carrés des distances aux voisins de  $\mathbf{p}$ . Par conséquent,  $\mathbf{v}_0$  approxime la normale à la surface en  $\mathbf{p}$  (Fig. 3.5).

De cette même analyse, nous notons que  $\lambda_0$  décrit quantitativement la variation de la direction de la normale à la surface. On définit également la variation de surface  $\sigma_m(\mathbf{p})$  au point  $\mathbf{p}$  dans un voisinage de taille  $m$  comme suit :

$$\sigma_m(\mathbf{p}) = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \quad (3.5)$$

Une valeur faible de la variation de surface en  $\mathbf{p}$  correspond alors à l'identification d'un plan tangentiel à l'échelle du voisinage de  $\mathbf{p}$  composé des  $m$  points. Tandis que

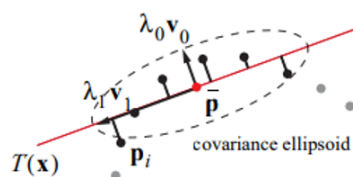


FIGURE 3.5 – Interprétation géométrique de l’analyse en covariance du voisinage du point  $p$ .

les valeurs élevées de  $\sigma_m(\mathbf{p})$ , elles, correspondent à des courbures élevées de surface. Cette notion se rapproche de la notion de courbure locale [109].

Finalement, avec un choix de  $m$  voisins ( $m = 9$ ), on attribue à chaque sommet un vecteur normal et une valeur de variation de surface. La Fig. 3.6 montre les normales calculées pour chacun des 800 sommets représentant un grain de ballast, tandis que la variation de surface est représentée sur la Fig. 3.7.

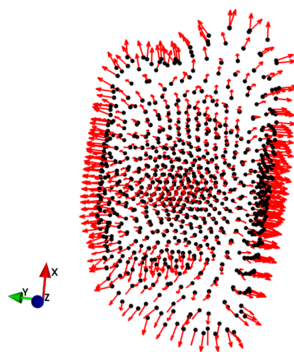


FIGURE 3.6 – Vecteurs normaux aux sommets d’un grain de ballast

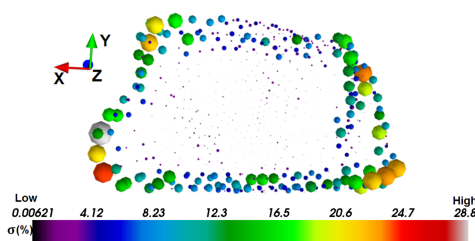


FIGURE 3.7 – La variation de surface calculée pour chacun des sommets d’un grain. La couleur et la taille du point varie en fonction de la valeur associée.

Dans la Fig. 3.7, la couleur et la taille du point représentant un sommet définit la valeur de sa variation de surface (qu’on appellera courbure) sous forme de pourcentage. Les valeurs sont comprises entre 0.006%. et 28.8%. Nous remarquons que cet indicateur a une valeur plus élevée pour les arêtes (plus de 12%) que pour les faces (moins de 1%), les faces et les arêtes étant pour l’instant celles pouvant être détectées par un observateur. Cet indicateur permet donc une caractérisation des courbures sur un nuage de points 3D.

## 2.2 Procédure de clustering & détection de faces et arêtes

Après avoir défini un vecteur normal ainsi qu'une valeur de variation de surface par sommet pour tous les grains (comparable à un rayon de courbure), nous nous proposons d'utiliser ces deux éléments afin de définir des faces et des arêtes sur nos nuages de points 3D. Pour ce faire, nous faisons les hypothèses suivantes :

- Les sommets composant une face ont des directions de normales proches, et une variation de surface faible.
- Les arêtes correspondent à des ensembles de sommets avec une plus grande variation de surface que celle des faces, et ne font donc pas partie des faces.

Afin de regrouper les sommets ayant des directions de normales proches, nous utilisons un algorithme d'apprentissage automatique dit de "clustering" [110]. Le clustering consiste à identifier des groupes d'objets  $\vec{X}$  ayant la même propriété choisie, à une tolérance près. Dans ce cas, nous choisissons de définir un cluster comme un groupe de sommets voisins ayant des directions de normales  $\vec{n}_i$  et  $\vec{n}_j$  proches vérifiant la condition suivante :

$$\|\vec{n}_i - \vec{n}_j\|^2 = \sum_{k=1}^3 (n_i^k - n_j^k)^2 \leq \epsilon^2 \quad (1 \leq i, j \leq d) \quad (3.6)$$

La valeur choisie de  $\epsilon$  a une influence significative sur la manière dont les points sont répartis dans les clusters. Un choix optimal, minimisant la sensibilité de l'ensemble des clusters obtenus aux variations de  $\epsilon$ , est déterminé en utilisant l'algorithme HDBSCAN [111]. La valeur de  $\epsilon = 0.116$  est trouvée, correspondant à un angle solide de 0.17 steradians.

Le résultat du clustering consiste à en ensemble "d'étiquettes" (labels) attribuées à chacun des sommets du grain analysé. Les sommets ayant la même étiquette forme un cluster, et vérifient donc le critère défini dans Eq. 3.6. Nous obtenons d'une part des clusters de tailles variables en terme de nombre de points, et d'une autre un ensemble de sommets non attribués (si l'une des deux conditions de voisinage (distance) et de proximité de direction de normales (Eq. 3.6) n'est pas vérifiée). Afin de caractériser la présence de faces sur les nuages de points, nous choisissons de nous intéresser uniquement aux "grandes faces" et aux arêtes, que nous définissons comme suit, en utilisant en plus des critères de courbure :

**F** Grandes faces, clusters de plus de 50 sommets : voir grandes faces 1, 2, 3 et 4 dans Fig. 3.8.

**sF** Petites faces, clusters de moins de 50 sommets.

**E** Sommets constituant les arêtes, définis comme les sommets non attribués avec une valeur élevée de variation de surface. (Nous choisissons le seuil de 5%, l'écart-type des variations de surfaces des sommets non attribués de moyenne 0.5%) (Fig. 3.8).

**R** Sommets restants (Fig. 3.8), qui n'appartiennent à aucune catégorie.

Testée sur quelques dizaines de grains comme dans la Fig. 3.8, la méthode permet une validation visuelle de la faculté de la méthode à détecter les grandes faces d'un grain, ainsi que de son pouvoir de détecter les arêtes parmi les sommets restants (Fig. 3.9). Deux grands avantages de cette approche : sa rapidité et son automatisme. Lors de l'usage de cet indicateur sur notre échantillon de grains scannés, nous accompagnerons cette validation visuelle par une analyse quantitative pour mieux

mettre en évidence le potentiel de la méthode à caractériser les faces/courbures sur un grain de ballast. L'analyse statistique de cet indicateur pourra également être testé sur les ensembles de grains générés dont la méthode est proposée par la suite.

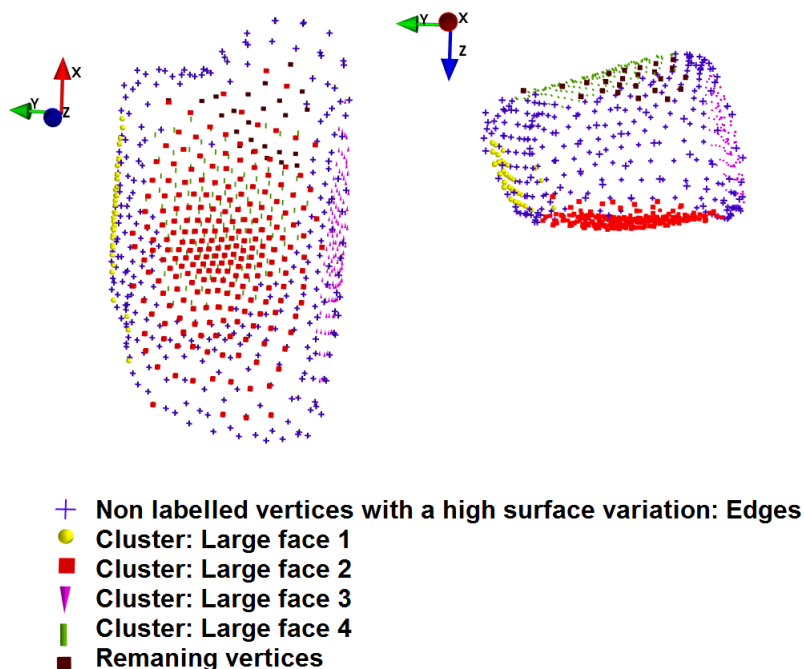


FIGURE 3.8 – Exemple de clustering effectué sur un grain (Les différentes couleurs représentent les différents clusters).

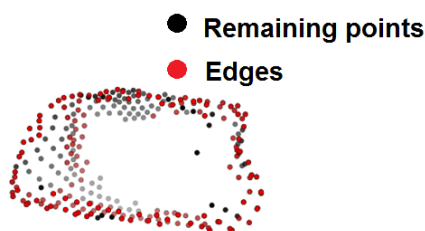


FIGURE 3.9 – Arêtes E et points restants R identifiés sur la surface d'un grain.

### 3 Méthodes proposées pour la génération des grains

Au vu des discussions sur les méthodes existantes de génération de la forme, nous choisissons d'aborder la génération de grands assemblages de grains à formes distinctes mais morphologiquement équivalentes aux grains scannés, d'un point de vue stochastique (aléatoire). C'est-à-dire de s'intéresser au comportement stochastique de l'information représentant les grains. A l'issue de la POD, nos grains sont représentés par des combinaisons linéaires de fonctions de forme et de coefficients.



Générer des grains statistiquement équivalents aux grains revient à générer des coefficients statistiquement équivalents aux premiers, et donc à proposer un modèle stochastique représentant l'échantillon scanné.

D'un point de vue stochastique, le modèle à estimer est une densité de probabilité multivariée, qui représente le vecteur aléatoire dont l'information disponible est un ensemble déterministe de  $n = 975$  réalisations  $C = (C_1, \dots, C_d)$  où  $C_i \in R^n$  est le  $i^{me}$  coefficient POD.

Théoriquement, si notre vecteur  $C$  provient d'une distribution  $f(x)$  inconnue, il suffit alors de déterminer complètement  $f(x)$  pour pouvoir générer des données qui suivent cette même distribution. Mais en pratique, il est très compliqué de déterminer la forme exacte de  $f(x)$ .

Pour estimer  $f(x)$ , on a classiquement recours aux méthodes paramétriques : on suppose que la distribution à simuler provient d'un modèle paramétrique (loi de probabilité). Si le modèle de l'hypothèse est correct, le problème est facilement résolu : si la loi d'un vecteur  $X$  est gaussienne par exemple, on peut facilement générer un vecteur statistiquement équivalent en utilisant la moyenne et la variance de  $X$ . Un test paramétrique sur la distribution de  $X$  peut très bien consister à calculer sa vraisemblance avec toutes les distributions dont les paramètres sont calculables (loi gaussienne, loi Gamma, loi de Bernoulli, etc.) et de conclure sur la nature de la distribution selon ce paramètre de vraisemblance.

Si toutefois la loi de l'hypothèse n'est pas correcte, et c'est souvent le cas pour des données spécifiques sur lesquelles nous n'avons pas de maîtrise, ceci peut fausser l'interprétation des données, et par conséquent empêcher de générer des particules ayant la forme voulue (un exemple est donné dans la suite).

Une autre approche pour estimer  $f(x)$  est d'appliquer une méthode non paramétrique, qui ne suppose donc aucune loi de probabilité préalable pour les données. La particularité de la statistique non paramétrique est que la fonction inconnue qu'on cherche à estimer ou à classifier n'est pas supposée appartenir à une famille indiquée par un petit nombre de paramètres réels, contrairement à l'estimation paramétrique.

On suppose donc que le nombre de paramètres qui décrivent la loi des observations est une fonction croissante du nombre d'observations, ou encore que le nombre de paramètres est infini. L'histogramme en est un exemple simple. En effet, l'histogramme peut approximer quasiment toutes les (PDF) pour un nombre d'observations  $n \rightarrow \infty$ .

La force de ces méthodes peut être résumée en 2 points :

- La complexité de la densité de probabilité estimée croît avec celle du vecteur stochastique à simuler.
- Plus la taille de l'échantillon est grande, plus il est possible d'estimer des modèles plus complexes. La flexibilité de l'estimation non paramétrique permet de considérer la structure réelle des données qui peut être négligée par un modèle paramétrique classique.

Naturellement, ces avantages sont accompagnés de quelques difficultés, dont voici une liste :

1. Le taux de convergence [112] est plus faible que les méthodes paramétriques :  $O(n^{\frac{-4}{d+4}})$  contre  $O(n^{\frac{-1}{2}})$  (où  $d$  est le nombre de variables,  $n$  le nombre de réalisations).
2. Les méthodes non paramétriques usuelles dépendent d'un paramètre de lissage

$h$  dit "bandwidth" qui contrôle la finesse de l'estimateur. Malheureusement, il n'existe pas de choix unique ou optimal de  $h$ , malgré plusieurs tentatives [113]. De plus, le choix de  $h$  a une influence significative sur les résultats.

3. La dépendance à la taille (nombre de réalisations) et de la dimension de l'échantillon (nombre de variables par réalisation) : puisque les méthodes non paramétriques sont asymptotiques, leur performance dépend fortement de la taille de l'échantillon. Le paramètre  $h$  en particulier est conditionné par la taille et la dimension de l'échantillon : selon quelques études [114,115], l'estimation non paramétrique donne de bons résultats pour une dimension allant jusqu'à  $d = 5$ . Or, pour des dimensions élevées comme c'est le cas pour nous ( $d = 800$ ), la valeur de  $h$  classiquement calculée est peu pertinente. Un exemple est donné lors de l'application de la méthode.

Nous choisissons d'essayer deux approches : une approche paramétrique, et une autre non paramétrique.

### 3.1 Méthode paramétrique : copule gaussienne

De manière générale, on dit que deux variables aléatoires  $C_i$  et  $C_j$  sont stochastiquement indépendantes si et seulement si pour tout  $c_i, c_j \in R$  :

$$P(C_i \leq c_i, C_j \leq c_j) = P(C_i \leq c_i)P(C_j \leq c_j) \quad (3.7)$$

Par conséquent, les variables sont dépendantes à chaque fois que Eq. 3.7 n'est pas vérifiée, ce qui fait de la dépendance stochastique un concept difficile à caractériser.

Pour  $d$  variables aléatoires à valeurs réelles  $C_1, \dots, C_d$ , la dépendance entre les variables est complètement décrite par leur fonction de répartition jointe  $F$  :

$$F(c_1, \dots, c_d) = P[C_1 \leq c_1, \dots, C_d \leq c_d]. \quad (3.8)$$

Leur marginales, quant à elles, sont décrites par les densités marginales  $f(c_i)$ ,  $i \in 1, \dots, d$  de chaque variable. Comment simuler des variables dépendantes provenant de lois différentes, en maintenant leur structure de dépendance ?

Notre premier choix est basé sur une hypothèse simple : la dépendance entre les variables aléatoires est entièrement contenue dans la matrice de corrélation. Dans ce cas, il est possible de passer par une copule gaussienne, un cas particulier des copules [116–118].

De manière générale, la copule  $H$  d'un vecteur aléatoire  $C = (C_1, \dots, C_d)$  est définie comme la fonction de distribution jointe du vecteur aléatoire  $Y = (F_1(C_1), \dots, F_d(C_d))$ , où  $F_i$  sont les distributions marginales des  $C_i$ ,  $i = 1, \dots, d$ . Le théorème de Sklar [119] permet d'établir l'existence et l'unicité (en cas de continuité des marginales) de la copule  $H$  telle que :

$$F(c_1, \dots, c_d) = H(F_1(C_1), \dots, F_d(C_d)) \quad (3.9)$$

La copule contient la structure de dépendance entre les marginales, et permet donc de caractériser/modéliser ces dépendances entre les différentes coordonnées d'un vecteur aléatoire à valeurs dans  $R^d$  sans se préoccuper de ses lois marginales.

Dans la même logique que le chapitre précédent, qui dit caractérisation, dit simulation et génération. En effet, les copules fournissent un moyen simple et efficace pour simuler des variables aléatoires dépendantes à marginales distinctes.

Dans le cas de la copule gaussienne, la procédure de génération se réduit à une génération de variables suivant la loi multivariée gaussienne, avec la moyenne et la matrice de corrélation voulues, puis à une transformation de marginales pour se ramener aux lois marginales voulues, selon le théorème suivant :

Soit  $X$  une variable aléatoire et soit  $F$  sa fonction de répartition. Soit  $F^{-1}$  la fonction quantile de  $F$ , i.e. :

$$F^{-1}(\alpha) = \inf \{x | F(x) \geq \alpha\}, \quad (3.10)$$

$\alpha \in (0, 1)$ . Alors :

1. Pour toute variable aléatoire uniforme  $U \sim U(0, 1)$ ,  $F^{-1}(U) \sim F$ . Ce qui permet de simuler des variables aléatoires provenant de fonction de distribution  $F$ .
2. Si  $F$  est continue, alors la variable aléatoire  $F(X)$  suit une loi uniforme :  $F(X) \sim U(0, 1)$ .

Pour un vecteur aléatoire  $C = (C_1, \dots, C_d)$  à marginales  $F_1, \dots, F_d$ , la fonction de répartition multivariée s'écrit :

$$F(c_1, \dots, c_d) = \Phi_{\Sigma}[\Phi^{-1}(F_1(c_1)), \dots, \Phi^{-1}(F_d(c_d))], \quad (3.11)$$

où  $\Phi_{\Sigma}$  est la fonction de distribution gaussienne multivariée, avec la matrice de corrélation  $\Sigma$ .

En partant du vecteur  $C = (C_1, \dots, C_d)$  où  $F_i$  est connue pour chaque  $C_i$ ,  $i = 1, \dots, d$ , et où le paramètre de corrélation  $\rho(C_i, C_j)$  est connu, il suffit de générer le vecteur gaussien  $Z$  de dimension  $d$ , avec la matrice de corrélation  $\Sigma$ . Ensuite, on transforme les variables  $Z_i$ ,  $i = 1, \dots, d$  en  $U_i = \Phi(Z_i)$ . Nous avons bien  $U_i \sim U(0, 1)$ . Enfin, on transforme les variables  $U_i$ ,  $i = 1, \dots, d$  en  $Y_i \equiv F_i^{-1}(U_i)$ . Nous avons donc  $Y_i \sim F_i$ , même lorsque  $F_i$  n'est pas continue.

Il est à noter que le choix de la copule gaussienne suppose donc que la seule dépendance existante entre ces variables est une dépendance linéaire exprimée dans leur matrice de corrélation. Nous verrons et discuterons les résultats de ce choix.

## 3.2 Application au ballast

### a. Résultats

Nous appliquons donc l'algorithme de génération à nos coefficients POD. Dans notre cas, les coefficients représentant nos variables aléatoires sont décorréliées au sens de Pearson (par construction suite à la POD) :  $\Sigma = [I_d]$  où  $[I_d]$  est la matrice identité. En se basant uniquement sur cette matrice, nous sommes exactement dans le cas de génération de  $d$  variables indépendantes.

Nous générons  $n$  réalisations des  $d$  coefficients POD, et nous reconstruisons  $n$  grains à l'aide de cette méthode. La Fig. 3.10 montre un exemple de quelques grains générés.



FIGURE 3.10 – Quelques grains générés par copule gaussienne

Les grains générés ont une taille équivalente, mais leur forme contient beaucoup de perturbations par rapport aux grains scannés. Nous avons également un pourcentage d'environ 5% de valeurs négatives dans les rayons. Sur les grains à rayons positifs, l'évaluation des indicateurs scalaires simples  $S$ ,  $V$  et  $\bar{R}$  donne toutefois une bonne correspondance (valeurs KS entre 0.5 et 0.65).

Les résultats montrent que l'approche paramétrique permet de couvrir l'aspect global de la forme. Cependant, à voir les résultats de vraisemblance (KS) pour l'élongation  $E$ , l'aplatissement  $F$  et la surface  $S$ , le résultat ne dépasse pas 0.07.

## b. Discussion et limitations

Comme expliqué dans [120], ces résultats sont dus aux hypothèses et aux limitations de ce choix paramétrique. D'abord, l'hypothèse de linéarité de la dépendance imposée par le choix de la copule gaussienne, combiné avec la nature décorrélée des coefficients POD. Ensuite, la transformation des distributions peut engendrer la perte de structure de dépendance. Il est donc possible que la corrélation soit différemment exprimée en transformant les variables. Afin de surmonter cette difficulté, une solution est proposée dans [121]. Enfin, la copule gaussienne, à cause de sa simplicité échoue à modéliser les dépendances dans les valeurs extrêmes des distributions.

Pour toutes ces raisons, nous choisissons d'aborder la question de la dépendance d'une manière différente. Nous faisons le choix d'une méthode non paramétrique. Dans la partie qui suit, une explication de la méthode, de son application et des ajustements nécessaires dans le cas du ballast sont détaillés.

## 3.3 Méthode non paramétrique : Multivariate Gaussian Kernel Density Estimation (Multivariate G-KDE)

### a. Introduction

Pour cette méthode non paramétrique, nous nous focalisons davantage sur la reproduction de la bonne structure locale du modèle, ce qui aura comme conséquence indirecte une bonne reproduction de la structure de dépendance.

Kernel Density Estimation (KDE) est une technique non paramétrique puissante pour estimer une distribution statistique. Le (KDE) peut être vu comme une généralisation de l'estimation de densité par histogramme. Cette technique a été d'abord

introduite pour le cas unidimensionnel dans les années 1950 et 1960 [122, 123], puis généralisée et validée pour le cas multidimensionnel. Depuis les années 2000, l'estimation par KDE multivarié a atteint un niveau de maturité suffisant pour être applicable à plusieurs domaines, et donner des résultats intéressants faisant de la méthode une contribution significative aux statistiques non paramétriques.

Dans le cas particulier du KDE gaussien (G-KDE), l'objectif est d'approximer une densité de probabilité (PDF) d'un vecteur aléatoire  $X$  (si elle existe), par une somme de  $n$  PDFs gaussiennes, toutes centrées autour de chaque observation de  $X$ .

Pour cette approche, l'optimisation des valeurs des matrices de covariance associées à ces  $n$  PDFs est une question importante, car ces matrices contrôlent l'influence de chaque observation de  $X$  sur l'approximation finale de la distribution de  $X$ . Malgré plusieurs contributions à ce sujet [124–128], des matrices de covariance paramétrées par un unique paramètre  $h$  (bandwidth) sont souvent considérées.

La règle de Silverman [129], en particulier, est largement utilisée pour déterminer la valeur de  $h$ , pour sa simplicité et son bon comportement asymptotique (quand  $n \rightarrow \infty$ ). Or, comme nous allons le montrer, pour une valeur fixée de  $n$  et une dimension  $d$  élevée, la règle de Silverman surestime la valeur de  $h$ , ce qui change la forme de la distribution, et empêche de concentrer correctement les réalisations générées de  $X$  autour de leurs régions de haute probabilité.

Dans le travail proposé par (Perrin et al. 2017) [130], un travail d'optimisation de  $h$  basé sur la maximisation de la vraisemblance vise à améliorer l'adaptation de la modélisation de la distribution de  $X$  à la structure de dépendance dans les composantes de  $X$ .

Une explication de la formulation mathématique est présentée dans la suite. Nous évoquerons brièvement les améliorations de la méthode, et notamment celles qui ont servi notre cas d'application. Nous choisissons d'expliquer le cas multidimensionnel..

## b. Cadre théorique classique

Soit  $\mathbf{X} := \{\mathbf{X}(w), w \in \Omega\}$  un vecteur aléatoire défini sur un espace de probabilité  $(\Omega, \Gamma, P)$  à valeurs dans  $R^d$ . Soit  $P_{\mathbf{X}}(d\mathbf{x})$  sa distribution inconnue, qu'on suppose concentrée sur un sous-espace de  $R^d$ , et notons  $[R_{\mathbf{X}}]$  sa matrice de covariance, inconnue elle aussi. On suppose également que la (PDF) de  $\mathbf{X}$  et l'inverse de la matrice  $[R_{\mathbf{X}}]$  existe, en les notant  $p_{\mathbf{X}}$  et  $[R_{\mathbf{X}}]^{-1}$  respectivement. Par définition,  $p_{\mathbf{X}}$  vérifie  $P_{\mathbf{X}}(d\mathbf{x}) = p_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$ , sachant que  $\int_{R^d} p_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = 1$  sur  $R^d$ .

Soit  $\Phi_{\mathbf{X}}$  la fonction caractéristique de  $\mathbf{X}$ , telle que pour tout  $t \in R^d$  :

$$\Phi_{\mathbf{X}}(t) := E[\exp(i \langle \mathbf{X}, t \rangle_d)] = \int_{R^d} \exp(i \langle \mathbf{X}, t \rangle_d) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (3.12)$$

où  $E$  est l'espérance mathématique,  $\langle \cdot \rangle_d$  le produit scalaire euclidien dans  $R^d$ . On suppose que l'information disponible sur  $p_{\mathbf{X}}$  est un ensemble déterministe de  $n > d$  réalisations distinctes et indépendantes de  $\mathbf{X}$ , qu'on représente par l'ensemble  $G(n) := \{\mathbf{X}(w_j), 1 \leq j \leq n\}$ . On définit ensuite 3 paramètres correspondants à cet ensemble :  $\hat{\mu}$ ,  $[\hat{R}_X]$  et  $\hat{\Phi}_{\mathbf{X}}$ , les estimateurs (non biaisés) de  $E$ , de  $[R_X]$  et de  $\Phi_{\mathbf{X}}$  respectivement. On définit ces paramètres comme suit :

$$E[\mathbf{X}] \approx \hat{\mu} := \frac{1}{n} \sum_{j=1}^n \mathbf{X}(w_j) \quad (3.13)$$

$$[R_X] \approx [\hat{R}_X] := \frac{1}{n-1} \sum_{j=1}^n (\mathbf{X}(w_j) - \hat{\mu}) \otimes (\mathbf{X}(w_j) - \hat{\mu}) \quad (3.14)$$

$$\Phi_{\mathbf{X}}(t) \approx \hat{\Phi}_{\mathbf{X}}(t) := \frac{1}{n} \sum_{j=1}^n \exp(i \langle \mathbf{X}(w_j), t \rangle_d), t \in R^d \quad (3.15)$$

Afin de déterminer une approximation satisfaisante de  $p_{\mathbf{X}}$ , l'idée du G-KDE multivarié est d'introduire par la suite, une transformée de Fourier inverse de  $\hat{\Phi}_{\mathbf{X}}$ .

Comme évoqué précédemment, définissons  $h \succ 0$  comme un paramètre de lissage (smoothing), tel que :

$$\forall \mathbf{x} \in R^d : p_{\mathbf{X}}(\mathbf{x}) \approx \hat{p}_{\mathbf{X}}(\mathbf{x}; h, G(n)) \quad (3.16)$$

C'est-à-dire que  $\hat{p}_{\mathbf{X}}$  approxime la fonction  $p_{\mathbf{X}}$  sur l'échantillon  $G(n)$ , avec le paramètre de lissage  $h$  qui contrôle la forme de la fonction.

En utilisant la définition de la fonction caractéristique, nous avons :

$$\hat{p}_{\mathbf{X}}(\mathbf{x}; h, G(n)) = \frac{1}{(2\pi)^d} \int_{R^d} \hat{\Phi}_{\mathbf{X}}(t) \exp\left(\frac{-1}{2} h^2 t^T [\hat{R}_{\mathbf{X}}] t\right) \exp(-i \langle \mathbf{x}, t \rangle_d) dt \quad (3.17)$$

$$= \frac{1}{n(2\pi)^d} \sum_{j=1}^n \int_{R^d} \exp\left(\frac{-1}{2} h^2 t^T [\hat{R}_{\mathbf{X}}] t + i \langle \mathbf{X}(w_j) - \mathbf{x}, t \rangle_d\right) dt \quad (3.18)$$

$$= \frac{1}{n} \sum_{j=1}^n \psi(\mathbf{x}; \mathbf{X}(w_j), h^2 [\hat{R}_{\mathbf{X}}]) \quad (3.19)$$

où  $\psi$  est la PDF de tout vecteur aléatoire gaussien dans  $R^d$  de moyenne  $\mu$  (vecteur dans  $R^d$ ) et matrice de covariance  $[Z]$  (matrice symétrique définie positive) définie comme suit :

$$\phi_Z(\mathbf{x}; \mu) = \psi(\mathbf{x}; \mu, [Z]) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det([Z])}} \exp\left(\frac{-1}{2} (\mathbf{x} - \mu)^T [Z]^{-1} (\mathbf{x} - \mu)\right), \mathbf{x} \in R^d \quad (3.20)$$

où  $\det(\cdot)$  est le déterminant. La formulation 3.19 montre qu'on arrive à approximer la PDF  $p_{\mathbf{X}}$  par une somme de  $n$  PDFs, centrées sur les réalisations (disponibles) de  $\mathbf{X}$  et dont les matrice de covariance son toutes égales à  $h^2 [\hat{R}_{\mathbf{X}}]$ .

Pour synthétiser, l'algorithme ci-dessous permet de générer  $M$  réalisations indépendantes d'un vecteur aléatoire  $\mathbf{X}$  :

1. Soit  $\eta(w'_1), \dots, \eta(w'_M)$   $M$  réalisations indépendantes suivant la loi discrète uniforme sur  $1, \dots, n$ .
2. On définit l'ensemble  $\bar{X} := [\mathbf{X}(w_{\eta(w'_1)}) \dots \mathbf{X}(w_{\eta(w'_M)})]$  contenant  $M$  variables aléatoires de dimension  $n$ , provenant de notre échantillon expérimental.
3. On définit ensuite la matrice  $[\xi]$  de dimension  $d \times M$  dont les composantes sont des réalisations indépendantes suivant la loi normale standard  $N(0, 1)$ .
4. On génère enfin nos  $M$  réalisations :  $[\hat{X}] = [\bar{X}] + h [\hat{R}_{\mathbf{X}}]^{1/2} [\xi]$ .

Après cette définition, intéressons-nous au paramètre  $h$  qui par construction, caractérise la contribution locale de chaque réalisation de  $\mathbf{X}$ . La difficulté est d'optimiser ce paramètre dans l'objectif de minimiser la différence entre  $p_{\mathbf{X}}$  (inconnue) et son estimateur  $\hat{p}_{\mathbf{X}}(\cdot; h, G(n))$ .

Pour quantifier cette différence, on utilise généralement l'erreur quadratique moyenne intégrée (mean integrated squared error) (MISE) définie comme suit :

$$MISE(h; d, n) = E_{G(n)} \left[ \int_{\mathbb{R}^d} (p_{\mathbf{X}}(\mathbf{x}) - \hat{p}_{\mathbf{X}}(\mathbf{x}; h, G(n)))^2 d\mathbf{x} \right] \quad (3.21)$$

L'erreur est définie de manière générale par rapport à n'importe quel ensemble de variables aléatoires indépendantes. Le choix de  $G(n)$  peut donc être aléatoire, même si dans notre cas, l'échantillon est déterministe, provenant du scan des grains.

Sans développer le calcul du  $h$  au sens de cette erreur, la formule donnée par Silverman, notée  $h^{silv}(d, n)$  semble un bon compromis entre complexité et précision :

$$h^{Silv}(d, n) := \left( \frac{1}{n} \frac{4}{(d+2)} \right)^{\frac{1}{(d+4)}} \quad (3.22)$$

On a effectivement :

$$\lim_{n \rightarrow +\infty} MISE(h^{Silv}(d, n); d, n) = 0 \quad (3.23)$$

Toutefois,  $h^{Silv}$  dépend uniquement de la dimension de l'espace et du nombre de réalisations de  $\mathbf{X}$  et peut donc pour des dimensions élevées, biaiser les résultats [?], un exemple sera montré sur l'application aux grains de ballast.

Afin de résoudre ce problème, plusieurs études récentes [131], basées sur les "diffusion maps" par exemple [132] ont tenté de biaiser la génération des réalisations afin d'améliorer le résultat. Dans [Guillaume, 2016], une autre approche basée sur la vraisemblance est proposée pour réévaluer la valeur de  $h$  appelée désormais  $h^{MLE}$  (pour maximum likelihood estimate). Ce même travail propose également des contraintes sur  $\hat{\mu}$  et  $[\hat{R}_{\mathbf{X}}]$  afin d'améliorer la performance de l'estimation des vrais paramètres. Par des calculs de vraisemblance, les auteurs ont montré la pertinence de ces méthodes (Perrin et al. 2017).

### 3.4 Application au ballast

Pour notre cas, nous avons appliqué la méthode classique G-KDE en utilisant la règle de Silverman, ainsi que toutes les autres méthodes proposées dans le travail de (Perrin et al. 2017).

Nous rappelons que l'on garde  $K = 117$  coefficients issus la PCA, avec 99% d'information sur la forme et  $n = 975$  grains (réalisations). Nous appliquons donc notre génération sur la matrice de coefficients de dimension  $n \times K$  représentant le vecteur aléatoire qu'on note  $\mathbf{X}$ , avec comme conséquence de la PCA :

$$E[\mathbf{X}] = \mathbf{0}, [\hat{R}_{\mathbf{X}}] = [I_K] \quad (3.24)$$

On suppose de plus que les réalisations sont indépendantes et font partie d'un processus aléatoire complexe.

A partir de la matrice des coefficients, nous avons donc  $n = 975$  réalisations de  $\mathbf{X}$  rassemblées dans l'échantillon noté  $G(n)$  dans la définition. La première tentative a été de générer  $n$  grains avec  $h^{Silv} \approx 0.98$ . Nous obtenons une bonne correspondance pour les indicateurs scalaires de "premier degré" (volume, rayon moyen, rapport d'aspect) et même pour l'élongation et l'aplatissement, sans aucune valeur négative de rayon. Toutefois, les détails de la texture de surface sont absents pour la plupart des grains, notamment dans les zones de non-convexité.

Ensuite, nous testons les approches basées sur les contraintes sur  $\hat{\mu}$  et  $[\hat{R}_{\mathbf{X}}]$ .

Afin d'évaluer la pertinence du résultat selon l'approche adoptée, le critère proposé par les auteurs, en plus de l'aspect visuel, est la comparaison des valeurs de la vraisemblance par Leave-One-Out [133]  $L^{LOO}(G(n) | h)$  utilisée à la fois pour l'estimation et pour l'évaluation de l'approximation :

$$L^{LOO}(G(n) | h) = \frac{1}{(n-1)^n} \prod_{i=1}^n \sum_{i=1, i \neq j}^n \psi(\mathbf{X}(w_i); \mathbf{X}(w_j), h^2 [\hat{R}_{\mathbf{X}}]) \quad (3.25)$$

La maximisation de cette fonction nous permet d'estimer une valeur de  $h = h^{MLE}$  telle que :

$$h^{MLE} = \operatorname{argmax}_{h>0} L^{LOO}(G(n) | h) \quad (3.26)$$

L'approximation de  $p_{\mathbf{X}}$  est d'autant plus pertinente que la valeur de  $L(G(n) | h)$  est élevée.

Les résultats du log-likelihood  $\log(L(G(n) | h, b))$  sont présentés dans le Tab. 3.1, où  $b = (b_1, \dots, b_p) \in 1, \dots, d^d$  correspond à une décomposition de la matrice en  $p$  blocs, au sens où si  $b_i = b_j$ , alors les variables de  $X$  contenus dans ces blocs sont considérés indépendantes. Alors :

- $b = (1, \dots, 1)$  correspond au cas où toutes les variables sont générées simultanément (dépendantes).
- $b = (1, 2, \dots, d)$  correspond au cas où toutes les variables sont générées séparément (indépendantes).

TABLE 3.1 – Influence de la valeur de  $h$ , de la présence de la contrainte sur les paramètres, et du choix de la décomposition par bloc sur la pertinence de l'estimateur de la PDF de  $\mathbf{X}$  par G-KDE (Perrin et al. 2017)

Valeur de $b$	$h$	Présence de contraintes	LOO log-likelihood
$(1, \dots, 1)$	$h^{Silv}(d, n)$	non	-179379
$(1, \dots, 1)$	$h^{MLE}(d, n)$	non	-176886
$(1, \dots, d)$	$h^{MLE}(d, n)$	non	-162398
$(1, \dots, 1)$	$h^{Silv}(d, n)$	oui	-161745
$(1, \dots, 1)$	$h^{MLE}(d, n)$	oui	-161262
$(1, \dots, d)$	$h^{MLE}(d, n)$	oui	-161775
$\mathbf{b}^{MLE}$	$h^{MLE}(d, n)$	oui	-160930

En terme de vraisemblance, nous remarquons l'intérêt de l'ajout de contraintes sur les estimateurs  $\hat{\mu}$  et  $[\hat{R}_{\mathbf{X}}]$ . En effet, pour une dimension élevée comme dans notre cas, une modélisation sans contraintes donne les modèles les plus pauvres (auxquels sont associés des valeurs faibles de vraisemblance).



Dans le cas précis où aucune contrainte n'est imposée, nous remarquons également que l'hypothèse de l'indépendance donne de meilleurs résultats que celle de la dépendance. L'explication peut être due au fait que dans le cas indépendant, nous imposons une structure diagonale de  $[\hat{R}_{\mathbf{X}}]$ , ce qui se rapproche du cas  $[\hat{R}_{\mathbf{X}}] = [I_K]$ . Par contre, dans le cas où nous imposons des contraintes de corrélation, nous remarquons que la valeur de vraisemblance augmente en prenant en compte les dépendances des coefficients.

Dans tous les cas, nous remarquons l'intérêt d'optimiser la valeur du paramètre de lissage  $h$  de  $h^{silv}$  à  $h^{MLE}$ . Enfin, la décomposition bloc-par-bloc de  $X$  proposée dans (Perrin et al. 2017) améliore davantage la valeur de vraisemblance.

Nous avons rajouté pour finir, une optimisation en dehors de la méthode : cette optimisation est basée sur la répartition des distances inter-points après génération, en fonction des valeurs avoisinantes du paramètre  $h^{MLE}$ , nous gardons la valeur qui minimise une erreur sur cet écart, et cette valeur est environ égale à  $h_{adapte}^{MLE} \approx 0.29$ .

La Fig. 3.11 montre un exemple de grains générés avec la méthode expliquée ci-dessus.

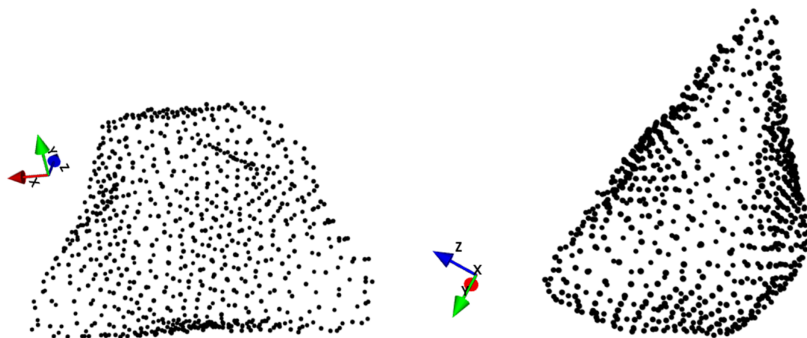


FIGURE 3.11 – Quelques grains générés par G-KDE optimisé

Comparés aux grains générés par les copules, nous remarquons l'amélioration significative du résultat de génération. Comme pour toutes les autres approches proposées dans ce travail, il est nécessaire d'évaluer la performance de la méthode, et de valider sa pertinence par des indicateurs quantitatifs.

## 4 Validation

### 4.1 Validation par indicateurs classiques

Nous calculons les indicateurs scalaires usuels  $S$ ,  $V$ ,  $\bar{R}$ ,  $E$ ,  $F$ , et  $a$  des  $n$  grains générés, et nous comparons leurs distributions statistiques à celles des grains scannés, déjà calculées précédemment. Les Fig. 3.12 et Fig. 3.13 présentent les PDFs de ces grandeurs pour les grains scannés et générés. Nous observons une très bonne correspondance pour tous les paramètres. En terme de valeurs (KS), toutes les valeurs calculées sont au-dessus de 0.8. Ce qui montre que les résultats sont très satisfaisants.

Comparée à la méthode paramétrique des copules, nous arrivons à des résultats meilleurs en terme de bonne reproduction des propriétés morphologiques des grains, sans oublier la réduction de modèle faite au préalable à l'issue de la POD. Ceci

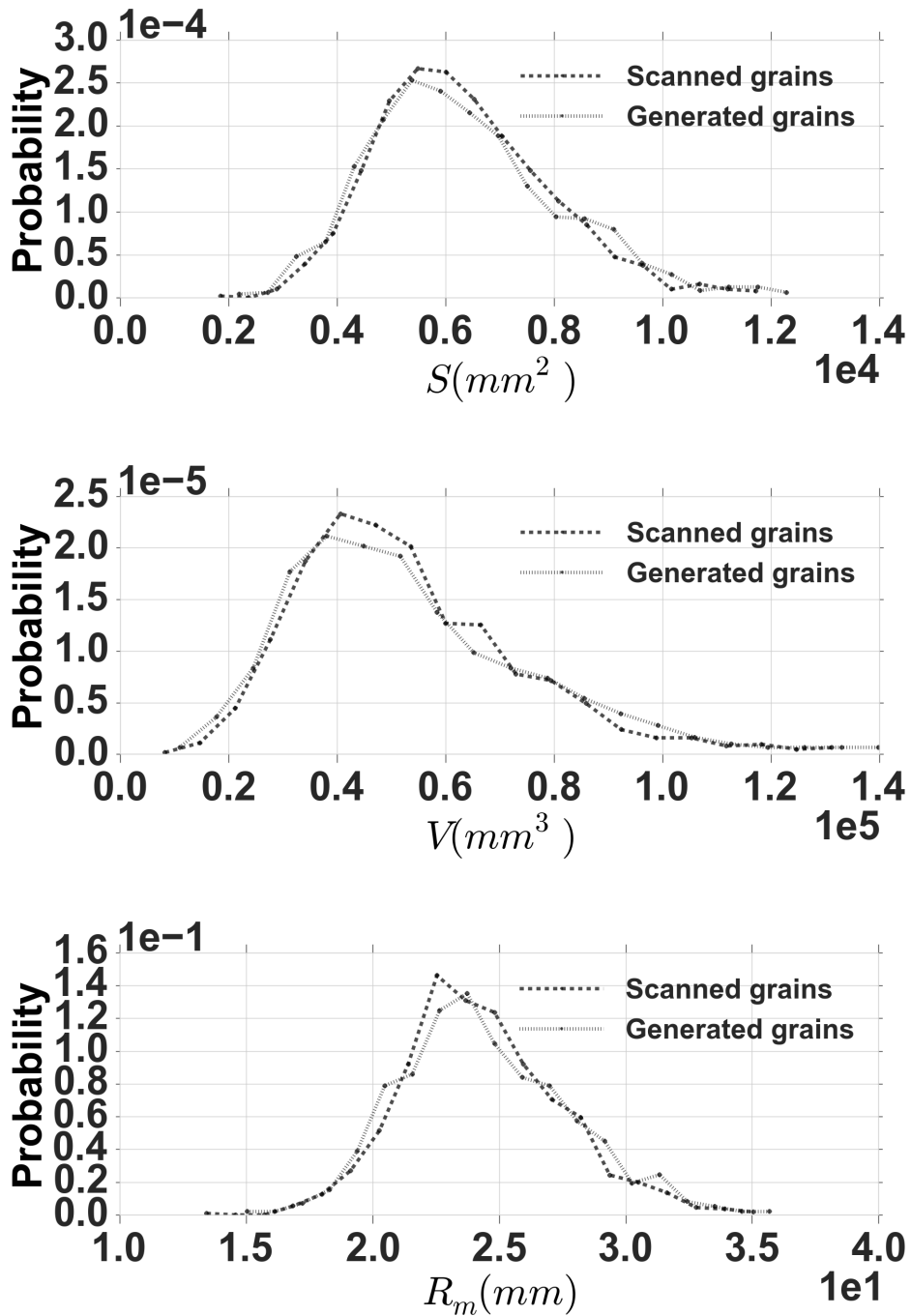


FIGURE 3.12 – Validation du modèle de génération par G-KDE en terme de  $S$ ,  $V$ ,  $\bar{R}$

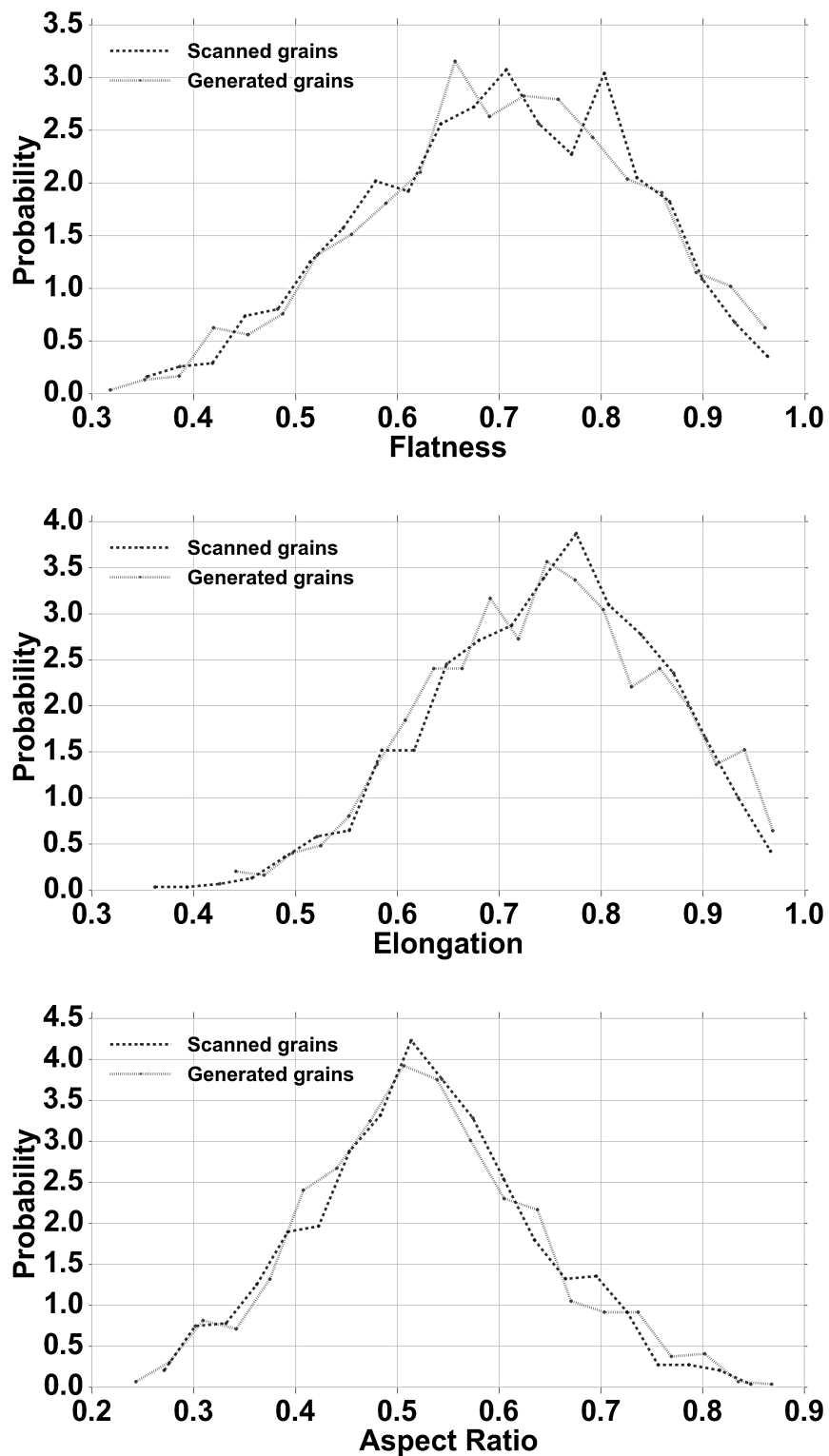


FIGURE 3.13 – Validation du modèle de génération par G-KDE en terme de  $F$ ,  $E$ ,  $a$

confirme non seulement le potentiel de la méthode de génération, mais également la capacité de la méthode de caractérisation à fournir une base de fonctions de forme qui permet une reconstruction pertinente des grains avec les propriétés voulues. Le besoin d'une méthode non paramétrique aussi avancée est également justifié par la complexité des dépendances entre les variables  $C_i$ ,  $i = 1, \dots, d$  (les positions des sommets dans l'espace).

Certes, la méthode de génération proposée plus haut a prouvé sa performance en terme de nombre, de variabilité et de reproduction des propriétés basiques de forme. Mais qu'en est-il des indicateurs plus avancés, comme les faces et l'angularité ?

## 4.2 Validation par les indicateurs de faces/courbures

Nous utilisons la méthode de détection de faces et d'arêtes proposée dans ce chapitre afin de comparer les distributions statistiques de ces indicateurs pour les deux populations de grains scannés et générés par G-KDE multivarié. La Fig. 3.14 montre la PDF du nombre de grandes faces (comme défini précédemment) identifiées sur les grains scannés et générés. La Fig. 3.15 montre également la distribution de la proportion  $\zeta_{LF_i}$  des sommets appartenant aux grandes faces, ce qui correspond à une estimation de la surface cumulée des grandes faces relativement à la surface totale. Nous remarquons clairement une bonne correspondance entre la population des grains scannés et générés.

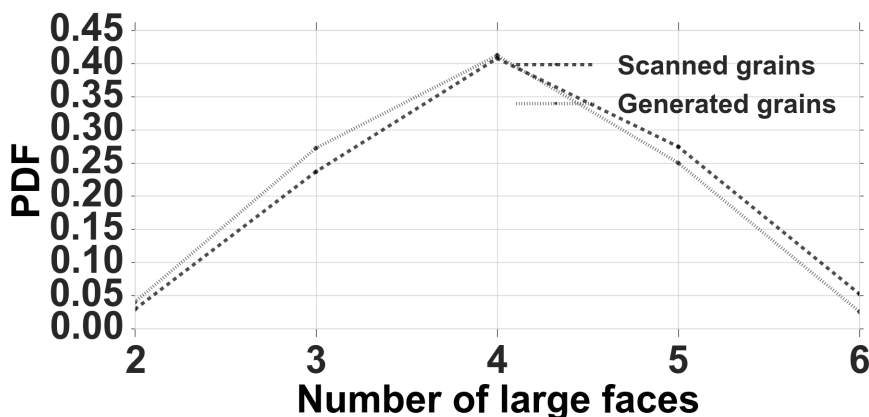


FIGURE 3.14 – PDFs du nombre de grandes faces, calculées pour les grains scannés et générés.

La distributions des variations de surface sont présentées dans les Figs. 3.16 et 3.17 pour les sommets appartenant aux grandes faces d'une part, et aux arêtes de l'autre. Nous remarquons encore un bon accord entre les grains scannés et générés.

Les résultats de la méthode de génération montre son excellent potentiel à générer des grains de nombre contrôlable, avec la variabilité souhaitée, et surtout avec les bonnes propriétés d'angularité, de faces et de texture de surface.

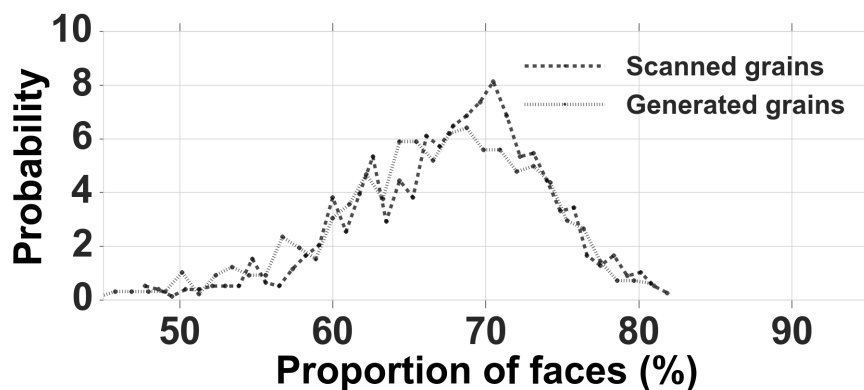


FIGURE 3.15 – PDFs de la proportion  $\zeta_{LF}$  calculées pour les grains scannés et générés.

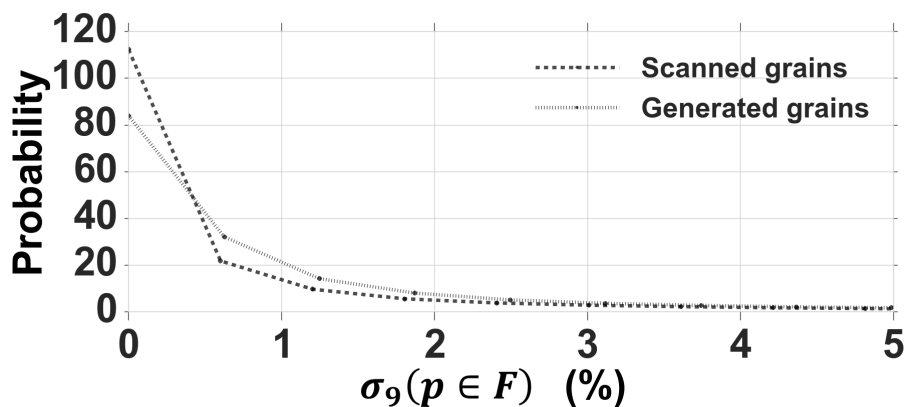


FIGURE 3.16 – PDFs de la variation de surface des sommets composant les grandes faces, calculées pour les grains scannés et générés.

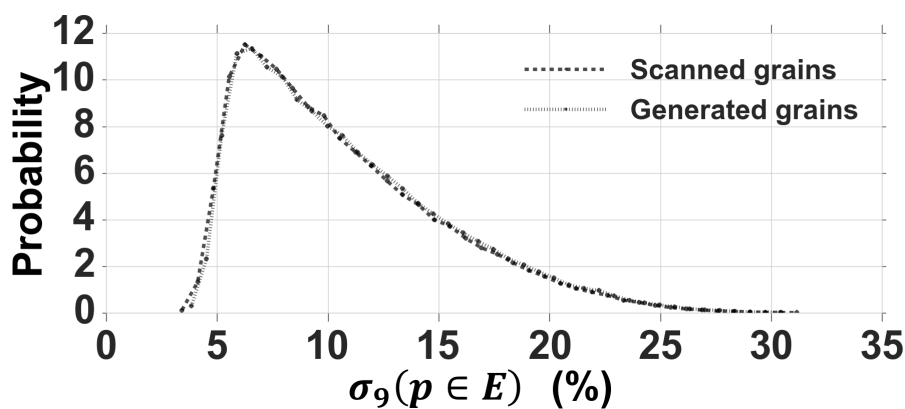


FIGURE 3.17 – PDFs de la variation de surface des sommets composant les arêtes, calculées pour les grains scannés et générés.

## 5 Conclusion

Dans ce chapitre, l'effort a été porté sur un objectif essentiel à la modélisation pertinente des grains de ballast : la génération de grains représentatifs de la forme variable de nos grains en voie.

Les méthodes de génération proposées ont permis de conclure sur une approche finale qui donne des résultats très prometteurs quant à la génération de grains à forme complexe et à degré élevé d'irrégularité. Nous avons proposé, pour valider ce modèle, de nouveaux indicateurs permettant de détecter les faces et les arêtes sur un grain de ballast numérisé. Ces indicateurs permettent de confirmer la bonne performance de l'algorithme de génération. Ce générateur de grains permet donc d'enrichir notre base de données construite à partir du scan *3D*, et servira lui aussi de base de données nouvelle de grains.

La performance de la méthode de génération permet également de prouver le potentiel de la méthode de caractérisation, à optimiser, contrôler et surtout reconstruire la forme à partir des coefficients générés.

Dans le chapitre suivant, nous passons à une autre configuration des grains de ballast, qui transforme les nuages de points anguleux et facettés en objets adaptés aux simulations numériques : les maillages. Notre objectif dans le chapitre suivant est donc d'évaluer (et de minimiser) la perte d'information relative à cette transformation. Nous ferons également une introduction aux outils numériques utilisés pour l'étude des propriétés mécaniques du ballast, en fonction de la variation de sa morphologie, présentée au chapitre d'après.

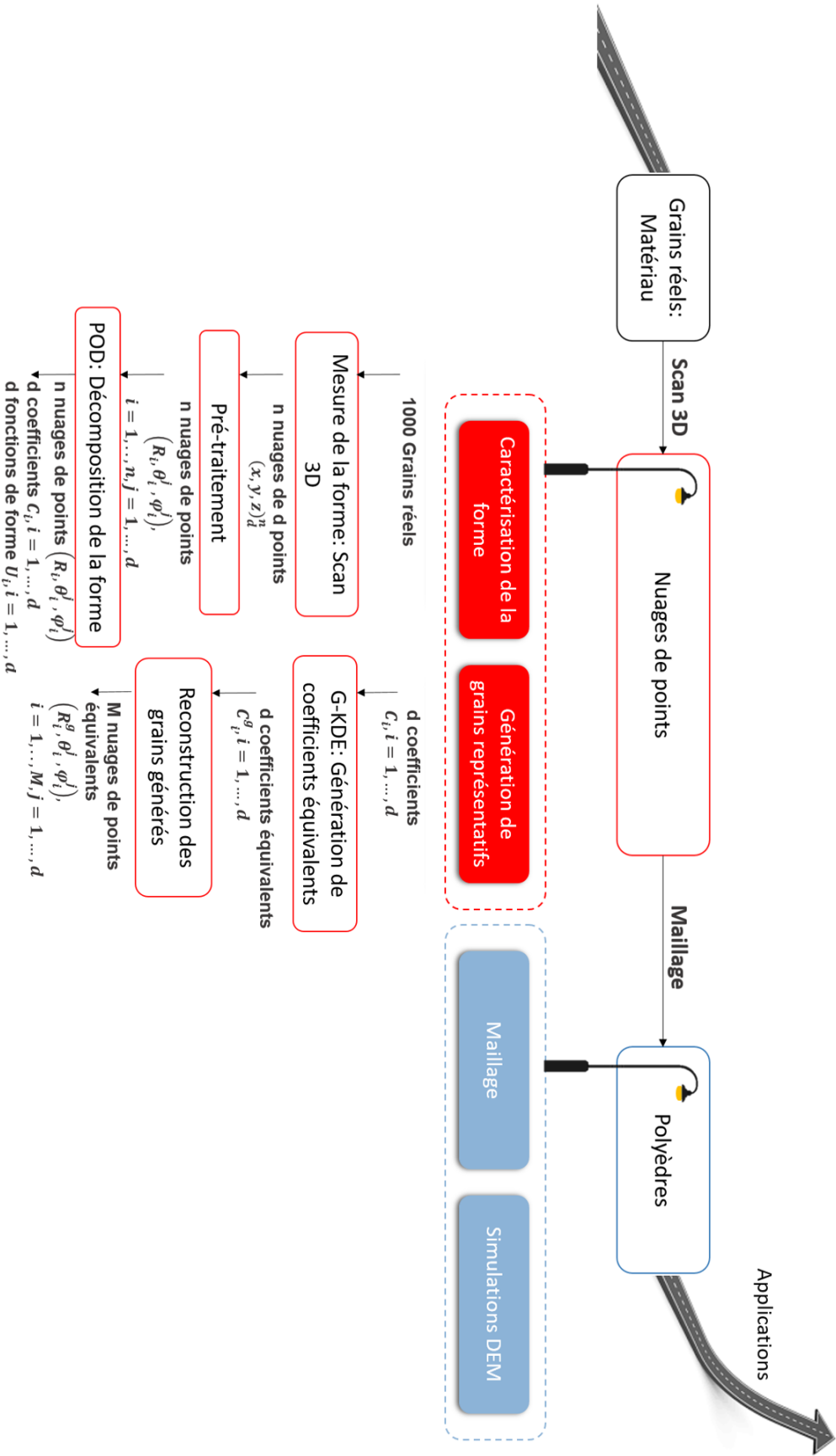


FIGURE 3.18 – Avancement sur le chemin d'étude du ballast

## Troisième partie

### Ballast : Les polyèdres aux mille facettes





# Chapitre 4

## Maillage & simulation

Comme expliqué au chapitre 1, afin d'étudier le lien entre la forme et le comportement mécanique par le biais de la simulation, il est important que les particules simulées soient représentatives de l'échantillon à étudier. Le scan 3D et les méthodes de caractérisation/génération assurent une première étape de transition entre le réel et le virtuel. Ensuite, une transformation nécessaire consiste à construire la primitive géométrique adaptée à la simulation : le maillage. Celui-ci dépend de la géométrie initiale, mais également du choix du maillage. Les résultats des études citées dans le chapitre 1 mettant en évidence la limitation des primitives simples (sphères, assemblages de sphères, etc.) à reproduire certains aspects du comportement mécanique (résistance au cisaillement). Nous faisons alors le choix de primitive polyédrique.

Dans ce chapitre, nous étudions l'étape du maillage qui consiste à transformer le nuage de points en primitive géométrique. Après une revue bibliographique rapide des différentes méthodes existantes utiles à ce travail, nous présentons une méthode de maillage adaptée à notre cas et dont nous évaluons l'effet sur la forme des grains de ballast. Enfin, nous introduisons la simulation numérique discrète par Dynamique des Contacts (CD).

Tout au long de ce chapitre, gardons en tête que nos attentes sur l'étape du maillage sont :

**Un maillage systématique** : un maillage robuste et automatisé.

**Un maillage contrôlable** : un maillage qui permet le traitement des surfaces convexes et non convexes, et qui permet le contrôle du nombre de faces.

**Un maillage rapide** : la vitesse n'est actuellement pas une priorité.

## 1 Maillage : du nuage de points au polyèdre

### 1.1 Revue bibliographique

#### a. Définition

Le maillage peut être défini comme une discrétisation spatiale d'un domaine continu, mais également comme une modélisation géométrique d'un domaine  $\Omega$  par des éléments finis dont la taille et la densité sont définies. L'objectif d'un maillage est de couvrir le domaine  $\Omega$  sans chevauchement ni trous. Le maillage est défini par son repère, les points qui le composent (sommets), et les éléments (polytopes) composés des sommets et des segments les reliant (arêtes). On caractérise un maillage par sa dimension (2D ou 3D), son volume (volume total de l'objet maillé), sa finesse (surface moyenne des éléments composant le maillage), et la forme des éléments (triangles, quadrilatères, polygones, etc. en 2D, tétraèdres, polyèdres, etc. en 3D). On le décrit également par le degré du polynôme décrivant les arêtes des éléments : le degré 1 par exemple correspondra à un côté rectiligne, et les arêtes seront d'autant plus courbés que le degré sera élevé.

## b. Types de maillage

En pratique, la distinction se fait d'abord entre deux grandes catégories de maillage : le maillage structuré, et le maillage non structuré.

Le maillage structuré/régulier [134] : il est composé de parallélogrammes (2D) ou de parallélépipèdes (3D). Son avantage principal est l'économie en temps de calcul et en mémoire. En effet, dans le cas d'un maillage structuré, la position des points et des cellules peut être déduite de manière implicite d'une règle de construction du maillage. Cependant, cette méthode ne convient que pour des géométries simples, car elle nécessite certaines conditions sur la forme de l'objet. Une possibilité, bien sûr, reste de décomposer les formes complexes en clusters de formes simples afin d'obtenir des maillages réguliers, mais cela risque d'être fastidieux.

D'une autre part, le maillage non structuré [134] n'exige aucune géométrie spécifique particulière du domaine, et offre donc plus de flexibilité, mais peut être plus compliqué à implémenter selon les objets. Grâce à son adaptabilité, ce type de maillage est de plus en plus utilisé. Parmi les différentes sous-familles de maillages non structurés, nous nous intéressons uniquement au cas des maillages triangulaires. Ensuite pour nos simulations discrètes, nous n'avons pas d'intérêt de mailler les volumes des grains, mais uniquement leurs surfaces. Par conséquent, nous utilisons un maillage surfacique triangulaire. Dans ce chapitre, nous détaillons uniquement la méthode développée dans cet objectif. Pour une présentation détaillée des différentes méthodes de maillage, voir par exemple [134].

Pour un objet 3D, le maillage surfacique peut être appliqué de deux manières : par un maillage 3D en rajoutant des contraintes (parmi lesquelles sur les intersections, les chevauchements des triangles, etc.) sur les éléments (polygones) : on l'appelle le maillage polygonal. D'un autre côté, le maillage surfacique peut être fait dans l'espace paramétrique : on exploite alors une représentation paramétrique de la surface, puis les nœuds sont projetés dans l'espace 3D de la surface (Fig. 4.1).

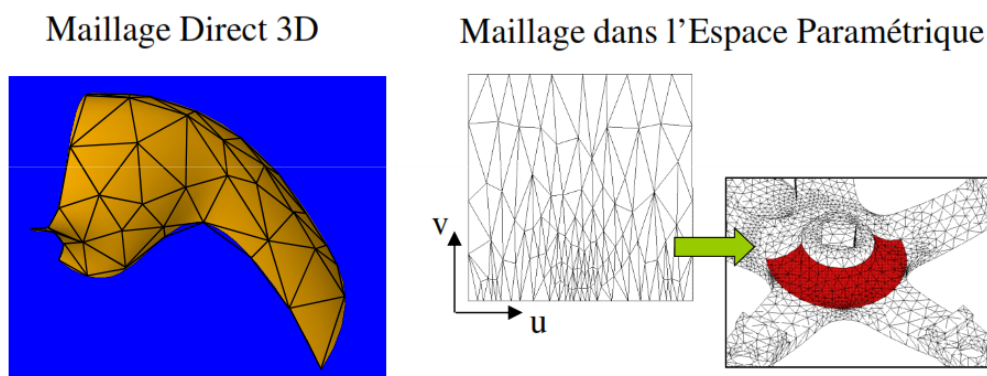


FIGURE 4.1 – Maillage surfacique : direct 3D (a) dans l'espace paramétrique (b)

Nous utilisons un maillage polygonal, contenant une information sur les position des sommets avec une information sur la connectivités de ces derniers. La qualité du maillage s'améliore d'autant plus que le nombre de polygones est élevé (Fig. 4.2).

Que ce soit pour les objets convexes ou non convexes, le développement des méthodes de maillage polygonal est en évolution incessante avec la performance des ou-

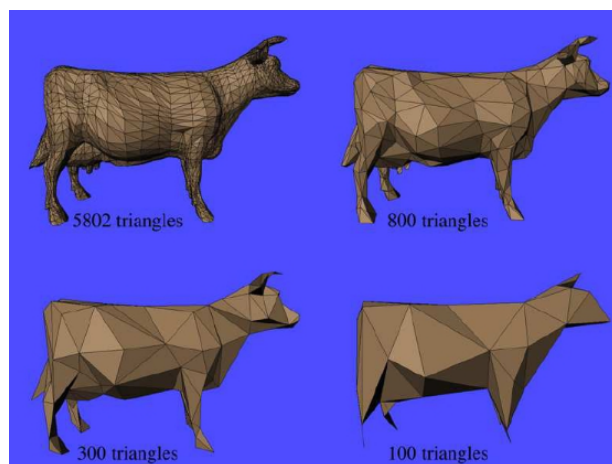


FIGURE 4.2 – Maillage polygonal

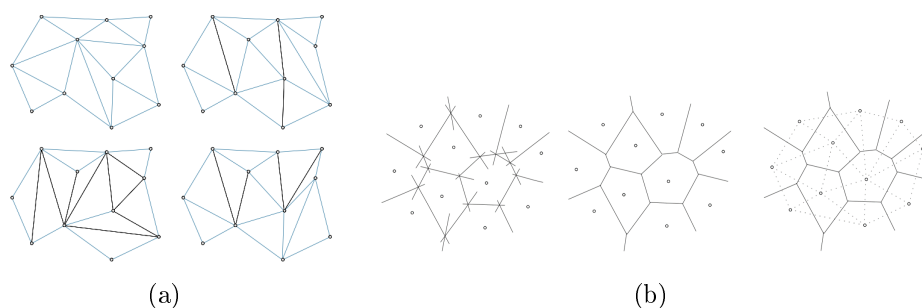


FIGURE 4.3 – Construction des maillages :(a) quelques triangulations possibles ; (b) construction à partir du diagramme de Voronoï.

tils numériques. En plus de la construction du maillage, d'autres opérations comme le lissage, la simplification et le raffinement font partie des techniques fréquemment utilisées pour le traitement des maillages polygonaux.

Nous nous intéressons maintenant à l'un des algorithmes les plus populaires pour les maillages triangulaires non structurés. Il s'agit de la triangulation de Delaunay [135,136]. Pour une présentation des autres méthodes, voir [134].

### c. La triangulation de Delaunay

Comme son nom l'indique, cet algorithme appliqué à un ensemble  $\Gamma = \{\Gamma_1, \dots, \Gamma_N\} \subset \mathbb{R}^2$  (ou  $\mathbb{R}^3$ ) de points permet de construire des triangles dont les sommets sont les points  $\Gamma_i$ . Comme montré sur la Fig. 4.3(a), plusieurs triangulations sont possibles, même en couvrant le même sous-espace de  $\mathbb{R}^2$  ( $\mathbb{R}^3$ ). Parmi toutes ces possibilités, il existe une triangulation basée sur le diagramme de Voronoï [137] construit à partir des intersections des médiatrices entre les points  $\Gamma_i$  et  $\Gamma_j$ .

La triangulation est construite comme suit :

1. Chaque centre de cellule de Voronoï est un point de la triangulation.
2. Chacun de ces centres est lié à ses voisins du diagramme de Voronoï.
3. Le résultat est une triangulation, puisque tous les points de Voronoï sont des

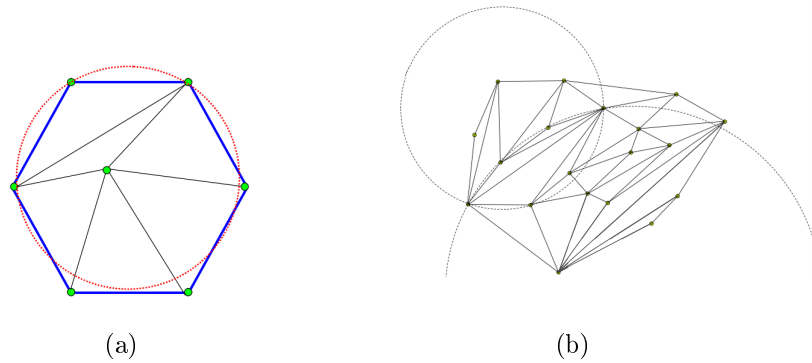


FIGURE 4.4 – Triangulations non-Delaunay :(a) exemple 1 ; (b) exemple 2.

points d'intersection de 3 médiateurs. (Fig. 4.3(b)).

Puisque le diagramme de Voronoï est unique, la triangulation de Delaunay qui en résulte est unique. En outre, cet algorithme possède par construction deux propriétés essentielles. La première est que les éléments de la triangulation sont des simplexes (triangles en 2D, tétraèdres en 3D). La seconde, est que ces éléments couvrent l'enveloppe convexe de la géométrie sans s'interpénétrer.

Il en résulte que tous les éléments de la triangulation vérifient la propriété du cercle vide : pour chaque simplexe de la triangulation, le cercle circonscrit ne contient aucun autre point de la triangulation. La Fig. 4.4 montre deux exemples d'une triangulation non-Delaunay (non-respect de la règle du cercle vide). Ce type de maillage n'est pas adapté à la simulation numérique, car quelques triangles sont trop allongés ou trop petits, ce qui pose problème pour la détection de contact (positionnement et calcul des coordonnées des points de contact). La triangulation de Delaunay fournit des triangles proches autant que possible de triangles équilatéraux.

Parmi les principaux avantages de cette méthode, sa capacité de construire un maillage indépendant de la géométrie et de l'orientation, qui présente une robustesse satisfaisante. Néanmoins, dans le cas du maillage surfacique des objets 3D, il est difficile d'obtenir des maillages de très bonne qualité, et surtout d'assurer les contraintes de bord. Il est à noter que cette méthode ne convient qu'aux formes convexes, et que lors de son application à un modèle géométrique non convexe, elle fournit l'enveloppe convexe de celui-ci, à moins que le bord du domaine soit défini au préalable (Fig. 4.5).

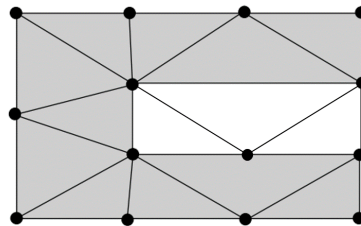


FIGURE 4.5 – Application de la triangulation de Delaunay au cas non convexe

Nous proposons dans la suite l'approche utilisée. Celle-ci s'appuie sur une triangulation de Delaunay, et est complétée pour correspondre à notre géométrie et à nos attentes.

## 1.2 Méthode proposée et résultats

De manière générale, l'évaluation de la qualité du maillage se fait lors de la simulation. Toutefois, l'optimisation préalable de la discrétisation par rapport à sa fidélité au modèle géométrique initial peut se faire en évaluant une erreur  $\epsilon$  (basée sur la différence de position des points avant et après maillage) et en la minimisant (Fig. 4.6).

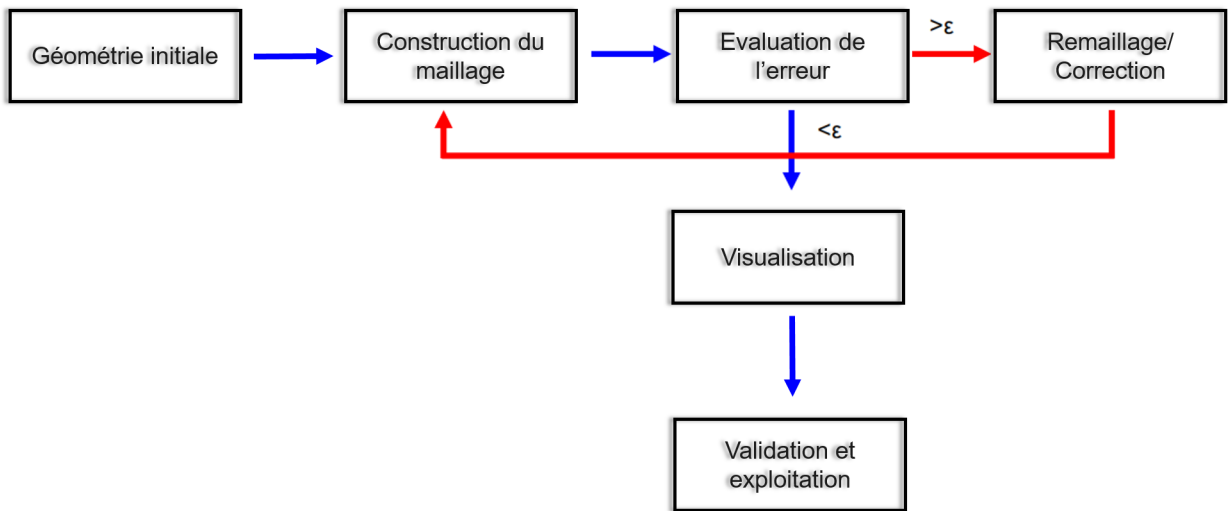


FIGURE 4.6 – Etapes de réalisation d'un maillage

Appliqué à notre population de grains, ce raisonnement est présenté dans le digramme de la Fig. 4.7.

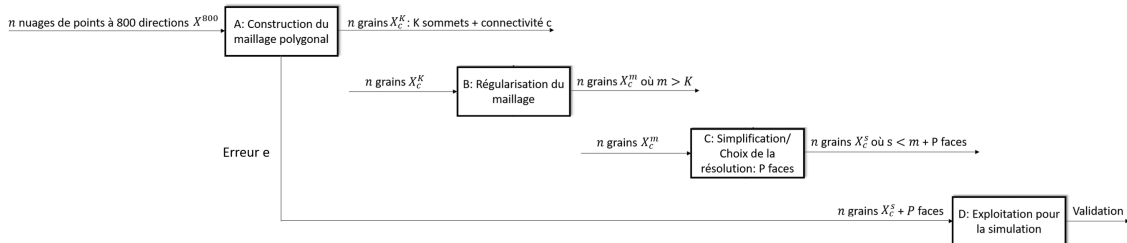


FIGURE 4.7 – Les étapes de la méthode proposée

Nous abordons ces étapes en détails dans la suite.

### a. Etape A : construction du maillage

Pour construire un maillage adapté à nos grains non convexes, nous avons d'abord recours à la méthode dite des " $\alpha$ -shape" [138] implémentée dans la librairie CGAL [139]. Soit  $S$  un ensemble dense et non structuré de points 2D ou en 3D. L'objectif des  $\alpha$ -shape est d'obtenir la forme "construite par ces points". Ils consistent à la reconstruction de la forme avec une information sur la frontière des cet ensemble.

Comme expliqué par [138], la notion d' $\alpha$ -shape peut être perçue comme suit : imaginons une énorme masse de glace, constituant l'espace  $R^3$ , et contenant des points représentés par des pépites de chocolat. Avec une cuillère de glace en forme de sphère, nous creusons toutes les parties du bloc de glace qu'on peut atteindre sans heurter les pépites de chocolat, ce qui permet de creuser des trous à l'intérieur. A l'issue de cette opération, nous obtenons un objet, pas nécessairement convexe, délimité par des arcs et des points. Si maintenant on projette en 2D toutes les zones creusées séparées par les arêtes et les pépites de chocolat, nous obtenons une description intuitive de ce qu'on appelle l' $\alpha$ -shape de l'ensemble de points  $S$ . En 2D, un exemple simple de ce processus peut être fait en remplaçant notre cuillère à glace par un cercle.

Les  $\alpha$ -shape dépendent d'un paramètre  $\alpha$ . Dans l'exemple de la glace,  $\alpha$  est simplement le carré du rayon de la cuillère. Une très faible valeur de  $\alpha$  permet de manger toute la glace, à l'exception des pépites de chocolat (points). L' $\alpha$ -shape converge donc vers le nuage de points pour  $\alpha \rightarrow 0$ . En revanche, une énorme valeur de  $\alpha$  nous empêcherait de déplacer la cuillère entre deux points, car celle-ci serait trop grande. Nous n'atteindrions jamais l'intérieur de l'enveloppe convexe de  $S$ , et par conséquent,  $\alpha \rightarrow \infty$  correspond à l'enveloppe convexe de  $S$  (Fig.4.8).

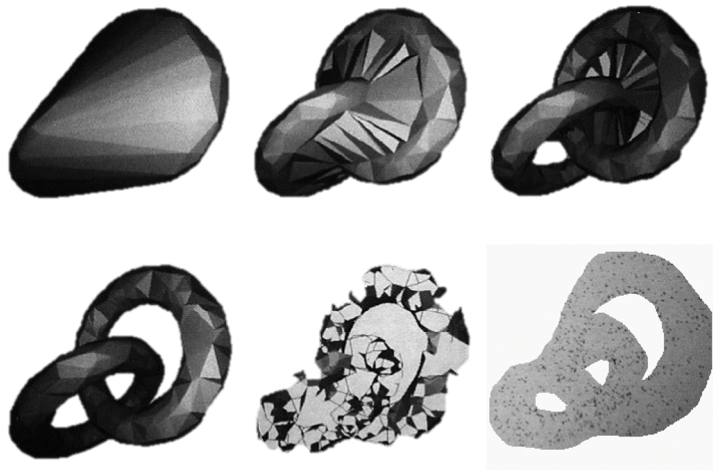


FIGURE 4.8 – Construction de l' $\alpha$ -shape pour différentes valeurs de  $\alpha$  ( $\alpha$  décroissant de  $\alpha \rightarrow \infty$  à  $\alpha = 0$ )

Le cas classique des  $\alpha$ -shape est basé sur une triangulation de Delaunay. Nous définissons d'abord le  $\alpha$ -complexe de l'ensemble  $S$  : c'est un sous-complexe de la triangulation Delaunay, c'est-à-dire que pour une valeur donnée de  $\alpha$ , le  $\alpha$ -complexe



comprend tous les simplexes de cette triangulation qui ont une sphère de circonscription vide de rayon  $r$  avec  $r^2 \leq \alpha$ . Dans ce cas "vide" signifie que la sphère concernée ne contient aucun point de  $S$ . L' $\alpha$ -shape est alors le domaine couvert par les simplexes du  $\alpha$ -complexe.

Les  $\alpha$ -shape d'un ensemble de points  $S$  forment une famille discrète, même s'ils sont définis pour tout  $\alpha \in \mathbb{R}$ . La famille des  $\alpha$ -shape peut-être représenté par la triangulation sous-jacente de  $S$ . Dans cette représentation, chaque simplexe (triangle en 2D, tétraèdre en 3D) de la triangulation est associé à un intervalle qui spécifie pour quelles valeurs de  $\alpha$  le simplexe appartient au  $\alpha$ -complexe. En appuyant sur cette définition, la famille des  $\alpha$ -shape peut être calculée efficacement et relativement facilement.

En outre, il est possible de sélectionner la valeur optimale de  $\alpha$  pour obtenir un  $\alpha$ -shape contenant le maximum de points du nuage de points et ayant un nombre donné de composantes connexes (dans ce cas 1 pour couvrir toute la surface du grain). Ce type de test nous donne une valeur de  $\alpha = 2 \text{ mm}^2$  pour le cas des grains de ballast scannés, ce qui correspond à un rayon de sphère d'environ  $r \approx 1.4 \text{ mm}$ . Par rapport à un grain de ballast de rayon moyen  $R_m \approx 2.5 \text{ mm}$ , ce rayon représente une proportion d'environ 5.6%. La Fig. 4.9 montre un exemple du maillage construit par  $\alpha$ -shape.

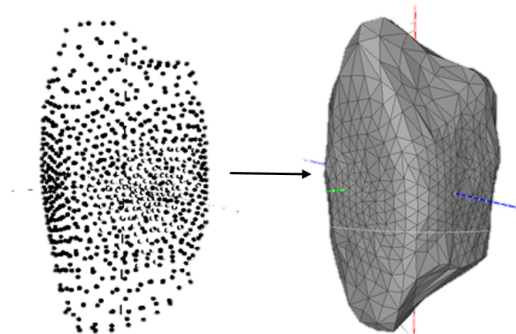


FIGURE 4.9 – maillage construit par  $\alpha$ -shape

A l'issue de cette étape, nous obtenons un premier maillage non convexe de nos  $n$  nuages de 800 points. On appelle  $X_c^K$  l'ensemble de ces grains maillés, décrits par leurs  $K$  sommets et leur connectivité  $c$ . Comme montré précédemment, nous obtenons, en plus de l'information sur les coordonnées des sommets, une information sur la connectivité de ces derniers. En visualisant tous nos grains maillés, nous obtenons un bon résultat lors de ces première construction. Néanmoins, nous remarquons que la forme et l'orientation des triangles reste à améliorer, notamment dans les zones où la variation de courbure est élevée (triangles difformes). Ceci est dû au fait que la valeur  $\alpha$  est fixée pour toutes les zones, en dépit de la texture de surface, et pour tous les grains.

Dans le but d'améliorer davantage la qualité de notre maillage et de contrôler le nombre de faces, nous procédons à l'augmentation de sa résolution pour ensuite le

simplifier en supprimant les arêtes et les sommets trop proches.

**b. Etape B : régularisation du maillage**

A partir des maillages construits par  $\alpha$ -shape, nous utilisons la méthode Least Squares Subdivision Surfaces ( $LS^3$ ) [140]. L'approche suit le même principe que les méthodes de subdivision de surfaces. Elle est composée de deux étapes principales :

- Fractionnement : consiste à raffiner le maillage en insérant de nouveaux sommets et en créant de nouvelles faces (Fig. 4.10(b)).
- Lissage : vise à produire des maillages plus lisses. Afin de produire des maillages localement uniformes, la méthode utilise des combinaisons affines pondérées des anciens sommets (relaxation).

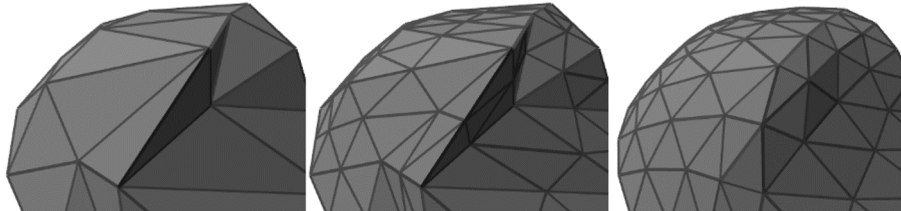


FIGURE 4.10 – Algorithme de subdivision de surfaces  $LS^3$

La méthode est implémentée sur le logiciel de maillage Meshlab [141]. A l'issue de cette étape, nous obtenons  $n$  grains maillés à  $m > K$  sommets  $X_c^m$ . L'application de cette méthode aboutit à un maillage est plus régulier et plus fin (environ 100000 triangles) (Fig. 4.11). Pour des raisons d'économie de mémoire et de temps de calcul dans les simulations DEM, il est nécessaire de simplifier notre maillage en diminuant le nombre de faces à un nombre raisonnable exploitable par nos machines de calcul.

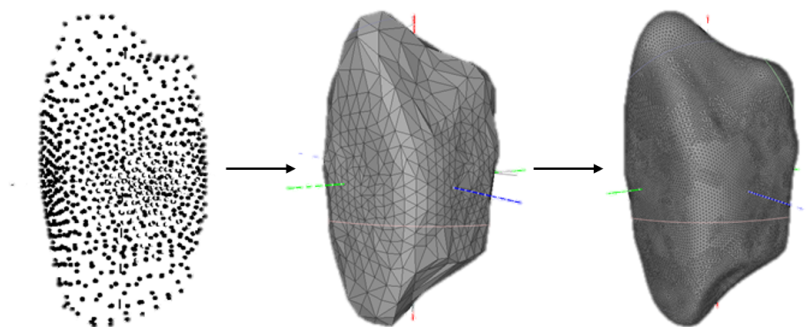


FIGURE 4.11 – Subdivision de surfaces  $LS^3$  appliquée au ballast

### c. Etape C : simplification (décimation) et choix de la résolution du maillage

Les algorithmes de simplification (décimation) [142] servent à réduire le nombre de faces/Sommets d'un maillage polygonal. Le processus itératif de simplification est souvent contrôlé par un ensemble de critères de qualité définis par l'utilisateur, afin d'assurer le maintien des propriétés spécifiques au maillage initial (distances géométriques, apparence visuelle, etc.), l'objectif final étant de réduire la complexité d'un maillage pour des fins de simulation.

On peut distinguer deux approches principales de simplification du maillage :

**Clustering de sommets** : rapide et robuste ; de complexité  $O(n)$  où  $n$  est le nombre de sommets ; la qualité n'est toutefois pas toujours satisfaisante : cet algorithme modifie significativement la topologie, et peut produire des maillages de type "non-manifold" (Fig. 4.12). L'algorithme est détaillé dans [143].

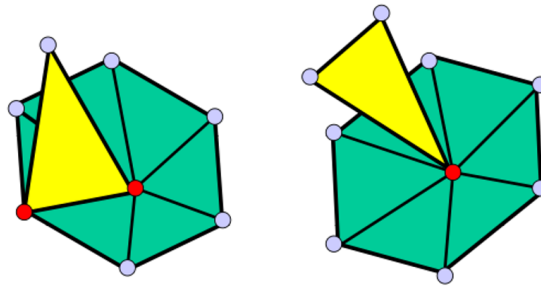


FIGURE 4.12 – Deux exemples d'un maillage de type non-manifold

**Décimation incrémentale** : donnant généralement des maillages de bonne qualité, cette méthode s'adapte aux critères de qualité imposés par l'utilisateur quant à la réalisation de chaque itération, ce qui permet un contrôle du résultat produit par le maillage. La complexité peut atteindre  $O(n \log_2 n)$  ou même  $O(n^2)$ , mais le compromis entre cette dernière et la qualité produite est assuré. Nous choisissons d'appliquer cette méthode.

Le principe de la décimation incrémentale est de supprimer un sommet/arête à la fois en se basant sur le critère fixé par l'utilisateur, et qui est révérifié à chaque changement de géométrie de la surface dans le voisinage du sommet/arête supprimé. On peut citer plusieurs opérateurs de décimation incrémentale, dont :

**"Vertex removal"** : après suppression du sommet ainsi que des arêtes et faces adjacents, la cavité est triangulée.

**"Edge collapse"** : une arête est sélectionnée puis transformée en sommet. Les deux triangles adjacents sont également transformés en deux arêtes. Il en résulte une réduction du nombre de sommets de 1, et du nombre de triangles de 2 à chaque itération.

Comme montré dans [143], la seconde approche est plus performante en terme de minimisation de l'erreur sur le maillage.

Mais la question se pose encore une fois quant au critère de sélection du sommet/arête à supprimer dans le cadre de la méthode "Edge collapse". [144] proposent une approche basée sur une erreur quadratique (Quadric Edge-Collapse) : on utilise une métrique calculée pour chaque sommet, qui permet, quand elle est minimale, de choisir le sommet à supprimer [143]. La méthode est implémentée sur le logiciel de maillage Meshlab.

La Fig. 4.13 montre l'application de cette méthode en partant du nuage de points. Dans ce cas, le nombre de faces choisi est de 100 faces.

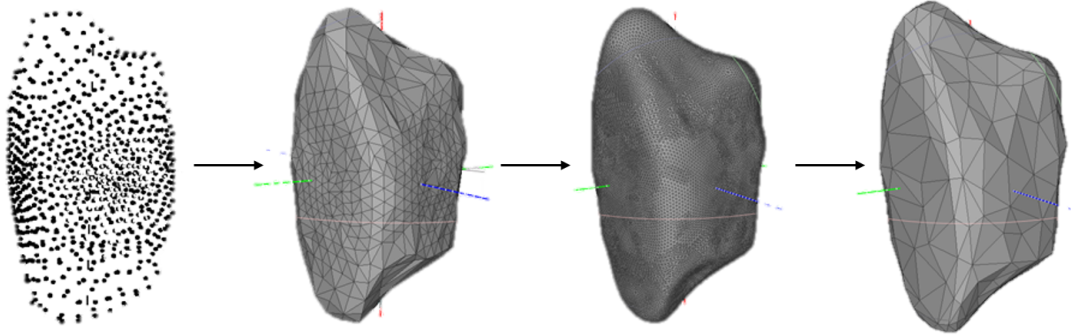


FIGURE 4.13 – La méthode "Quadric Edge Collapse" appliquée au ballast

Par rapport à la première construction par  $\alpha$ -*shape*, le résultat obtenu après décimation a permis de résoudre le problème d'orientation des triangles sur les zones à haute courbure. Nous obtenons également un maillage sans triangles difformes, qui se rapproche visuellement de la géométrie initiale, et dont nous contrôlons le degré de finesse (contrairement à la construction par  $\alpha$ -*shape*).

Nous procédons ainsi au maillage de nos  $n$  grains issus de l'étape  $B$ , pour récupérer  $n$  polyèdres non convexes à 100 faces  $X_c^s$  où  $s \ll m$ . Plus précisément, l'ensemble des arêtes  $e$ , des sommets  $v$  et des faces  $f$  obéissent au théorème de Descartes-Euler :  $v - e + f = 2$ , ce qui nous donne un nombre final de 52 sommets par grain à 100 faces.

### 1.3 Effet du maillage sur la forme

Suite à la procédure de maillage combinant un ensemble de transformations appliquées à la géométrie des grains scannés, il est nécessaire d'évaluer l'erreur relative à ce passage. Nous nous posons alors la question sur une méthode plus précise de quantifier l'erreur lié à ce maillage par rapport à la morphologie de chaque grain.

L'effet du maillage sur la forme peut être évalué de deux manières différentes :

**Les simulations numériques** : par un changement de la résolution du maillage, nous pouvons évaluer la variation de la réponse mécanique correspondante. Ceci suppose toutefois que la transition nuage de points-maillage engendre une erreur suffisamment faible.

**L'erreur au sens de la POD/PCA** : comme montré au chapitre sur la caractérisation de la forme, l'erreur PCA a prouvé son potentiel à détecter les variations

de la géométrie, par rapport aux indicateurs scalaires. Nous pouvons dans ce sens, évaluer la faculté des fonctions de formes (base PCA) déterminées précédemment pour les nuages de points à représenter les nuages de points post-maillage. Pour cela, les sommets constituant les maillages des grains à  $N = 800$  modes sont récupérés, et ceux-ci sont projetés sur la base utilisée pour représenter nos nuages de points initiaux (base  $(\theta_j, \varphi_j, j = 1, \dots, 800)$  obtenue par répulsion de points). Enfin, la même procédure PCA (cf. chapitre 2) est appliquée, et les grains sont reconstruits pour différents nombres de modes  $1 \leq K \leq d$  où  $d = 800$  en utilisant la base PCA précédemment obtenue. Nous sommes alors capables de représenter la courbe de l'erreur PCA en fonction du nombre de modes pour les nuages de points constituant tout maillage.

La Fig. 4.15 représente cette erreur pour trois configurations différentes :

1. Les nuages de points scannés (Scan 3D).
2. Les nuages de points issus du maillage non convexe présenté à la section précédente (non-convex mesh).
3. L'enveloppe convexe du maillage non convexe précédent (convex mesh) (Fig. 4.14).

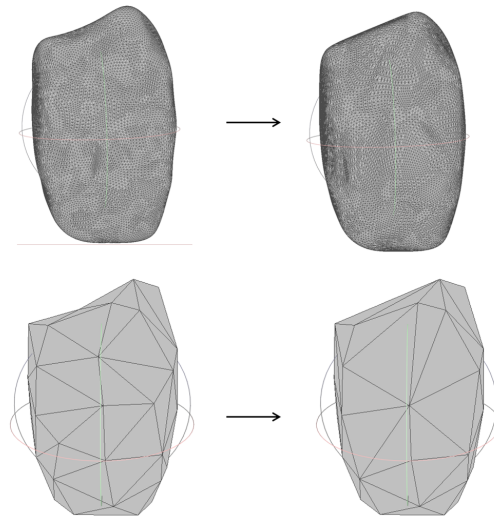


FIGURE 4.14 – Enveloppe convexe du maillage précédent pour 50000 faces puis 100 faces

L'analyse de cette figure peut être résumée dans les points suivants :

- Pour  $K = d = 800$  modes, l'erreur correspond à 0% dans le cas du scan, à moins de 5% dans le cas du maillage non convexe, et à plus de 20% dans le cas du maillage convexe. Ces valeurs signifient qu'une perte d'information a bien lieu, même en gardant le maximum du nombre de modes. Il est également à noter que suite à la contrainte du nombre de sommets ( $d = 800$ ) pour projeter les grains sur la base POD, les deux types de maillage ont été faits avec un nombre beaucoup plus élevé de faces (50000 au lieu de 100) mais qui conserve une valeur avoisinante de l'erreur puisqu'il s'agit simplement d'une subdivision de surface à partir de 100 faces (multiplication du nombre de triangles sans changement de la forme).

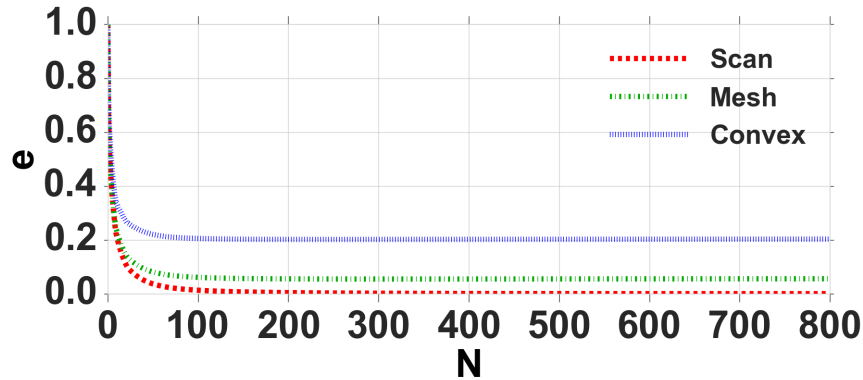


FIGURE 4.15 – L'erreur au sens POD en fonction du nombre de modes, calculée pour 3 configurations différentes

— Dans le cas convexe, l'erreur est constante à partir de  $K = 11$  modes environ. Ceci nous donne une prévision sur les résultats mécaniques. En effet, en introduisant plusieurs nombres de modes dans des études paramétriques sur la forme, il est possible que l'on remarque un effet plus significatif de la forme entre 1 et 11 modes, et moins d'effet entre 11 et 800 modes. Dans un cas où le maillage est de meilleure qualité (maillage non convexe), on ne remarquera plus beaucoup de différence de réponse à partir de  $K = 117$  modes.

Enfin, une question importante, à part le type de maillage choisi pour nos grains concerne le nombre de faces, car ce paramètre modifie nettement la morphologie (Fig. 4.16). Nous tenterons d'évaluer l'effet du nombre de faces du maillage sur le comportement mécanique dans le chapitre suivant.

## 2 Introduction aux simulations numériques discrètes

### 2.1 Méthodes de résolution

Afin de modéliser le comportement mécanique d'un matériau (granulaire ou autre), la simulation numérique est depuis des années un outil pertinent, permettant d'obtenir des résultats valides et représentatifs. Par rapport à l'expérience, elle offre une plus grande facilité pour le contrôle des paramètres et le traitement des résultats.

Dans le domaine ferroviaire, l'évolution des approches numériques a permis de mieux étudier le comportement des composantes de la voie ferrée. En effet, une approche connue dans la simulation est la modélisation par éléments finis (FEM) [145]. Cette méthode est performante pour représenter le comportement dynamique de systèmes physiques continus, sous certaines conditions. Cette approche a été par exemple utilisée pour le dimensionnement des couches d'assise de la voie ferrée en fonction de la qualité du sol support (en terme de densité et de module d'élasticité) [146, 147].

Néanmoins, considérer qu'un milieu est continu alors qu'il ne l'est pas est une approximation à faire avec prudence. En effet, dans un milieu composé de plusieurs particules, comme c'est cas pour la couche de ballast, cette approximation signifie

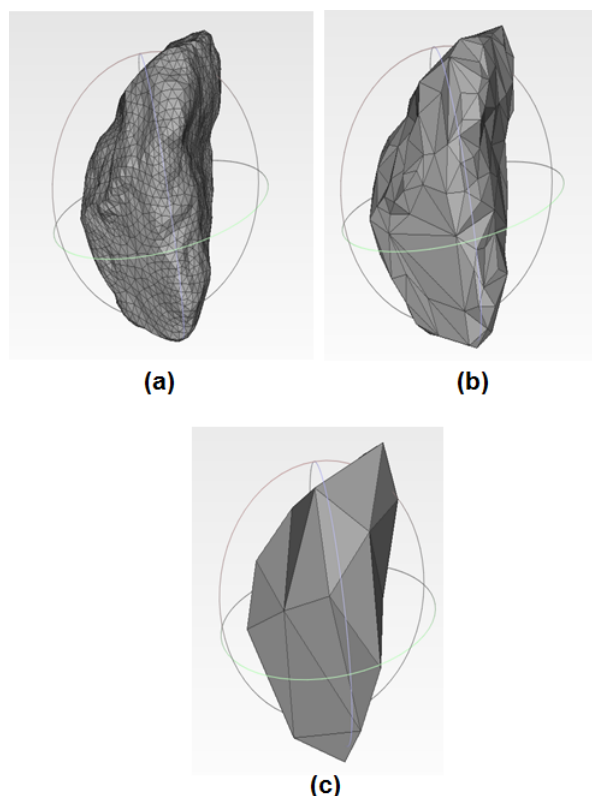


FIGURE 4.16 – Evolution de la forme en modifiant la résolution du maillage

qu'on ne tient compte ni de la nature granulaire du ballast, ni des caractéristiques propres des grains, et encore moins de leur disposition/arrangement.

L'approche discrète présente l'avantage de traiter chaque grain comme un corps (rigide ou déformable) distinct, avec un nombre défini et limité de degrés de liberté, et en interaction avec d'autres grains. Le traitement consiste simplement à intégrer les équations du mouvement simultanément pour toutes les particules, en considérant les forces de contact et les forces externes exercées sur elles. Ceci donne accès aux vitesses de chacun des grains et aux efforts de contacts durant la simulation. Ce niveau de précision nécessite des temps de calculs plus ou moins longs selon les cas.

Historiquement, l'approche en éléments distincts (DEM) fut introduite d'abord par [148] pour les géo-matériaux granulaires. Le terme "distinct" est venu distinguer la méthode de celle des éléments finis, et renvoie aux degrés de liberté portés par chaque particule. La méthode fut ensuite appelée Méthode des Éléments Discrets (DEM) en référence à la nature discrète du système.

La DEM compte deux catégories principales :

1. Les approches régulières dites de type "Smooth"  $S$  : qui autorisent l'interpénétration entre les particules.
2. Les approches non régulières dites de type "Non-Smooth"  $NS$  : qui considèrent les chocs sans interpénétration possible entre les particules.

Succinctement, les méthodes de type  $S$  [37] sont fondées sur l'approche de la dynamique moléculaire (MD) qui consiste à considérer que les grains sont rigides, mais que les contacts sont déformables et obéissent à une loi visco-élastique. C'est-à-dire qu'une déformation locale des grains, en l'occurrence l'interpénétration, est

possible et est intégrée aux lois de contact nécessaires pour la résolution numérique du système. Les liaisons entre les particules sont parfois modélisées par des ressorts. Ces approches ont été initialement développées pour les liquides, les cristaux et les atomes, puis étendues aux corps solides.

Les méthodes de type *NS* [37], quant à elles, sont fondées sur des lois de contact exprimant l'interdiction de recouvrement entre les particules et sur la loi de frottement de Coulomb. Les forces et les vitesses des particules sont déterminées simultanément par itérations sur l'ensemble des contacts. La solution vérifie à la fois la conservation de la quantité de mouvement pour toutes les particules et les inégalités exprimant les contraintes de non-interpénétration entre particules et la loi de frottement de Coulomb. Ces approches supposent la rigidité parfaite des contacts (contacts indéformables ; corps rigide ou déformable). Toute interpénétration pouvant apparaître avec une approche de type non régulière est considérée comme une « violation » et elle est donc indésirable. Ces méthodes s'appliquent à toute collection de corps rigides. Elles utilisent généralement des pas de temps plus grands, et prennent en compte l'unilatéralité au contact sans approximation ou régularisation. Il est important de préciser aussi que la résolution des équations du mouvement par les méthodes *NS* est indépendante de la forme des particules qui n'intervient que dans la détection de contact.

Il n'est pas question de comparer les avantages ou les inconvénients des deux types de méthodes. En effet, dans la réalité physique les contacts sont déformables, parce que le matériau qui constitue les grains est doté de propriétés élastiques, et qu'il y a une déformation, en général localisée au voisinage immédiat des points de contact. Le choix des méthodes non régulières est d'ignorer l'échelle de longueur associée à la déflexion élastique des contacts et de traiter les grains comme parfaitement rigides et indéformables. C'est un choix qui se défend bien pour les grandes déformations et les écoulements, parce qu'on peut espérer que les comportements sont peu affectés par les détails de très petite échelle. Un avantage de la Dynamique des contacts est donc de fonctionner avec moins de paramètres ( un coefficient de frottement et deux coefficients de restitution), alors qu'une méthode régulière fait appel à un modèle élastique pour les déflexions de contact. En revanche, le revers de cet avantage est qu'on doit s'inquiéter ultérieurement du niveau de violation, et les paramètres en moins dans le modèle physique apparaissent au moment de discuter de la précision numérique et de la qualité du calcul. Le pas de temps, effectivement, peut être beaucoup plus grand que dans les méthodes régulières, obligées de décrire le détail du mouvement lors d'une collision par exemple (dont la description demanderait de l'ordre de la centaine de pas de temps). Le choix du pas de temps reste un paramètre caché dans la dynamique des contacts, et il faut s'assurer qu'il n'a pas trop d'influence sur la simulation des phénomènes auxquels on s'intéresse. Supposons que l'on choisisse de résoudre un système composé d'un assemblage de grains de ballast en utilisant la méthode (MD). Comme un degré de déformabilité du contact est toléré, on serait contraint de choisir des pas de temps très petits (allant jusqu'à  $10^{-10}$ s) afin d'avoir une faible interpénétration, et de calibrer également la raideur des ressorts représentant les liaisons en fonction du pas de temps. La méthode (*CD*) permet d'éviter ces inconvénients et de prendre des pas de temps plus grands.

Dans le cas d'un système comptant plusieurs interactions simultanées comme c'est le cas pour un assemblage de grains de ballast, la méthode la mieux adaptée



est la Dynamique des Contacts (CD) [37, 149]. Nous présentons dans la suite les principes de la formulation de la méthode (CD) implémentée dans le code de calcul LMGC90 utilisé ensuite pour nos calculs mécaniques.

## 2.2 La dynamique des contacts (CD)

Le processus de résolution par la méthode (CD) [150] suit les étapes suivantes :

A l'instant  $t$ , le système est parfaitement déterminé en termes de positions, vitesses et forces appliquées. A la fin du pas de temps  $t + \delta t$ , on doit connaître l'évolution du système dans ces mêmes termes :

1. Détection des contacts.
2. Résolution globale (détermination des forces de contacts et des vitesses de tous les corps).
3. Actualisation des vitesses et des positions.

Pour une meilleure compréhension de la méthode, il est important de séparer ses différentes étapes. Il est important de noter que la détection de contacts dépend fortement de la forme des particules injectées dans la simulation. Le reste des étapes en est indépendant. Nous accordons donc une attention particulière à cette étape.

### a. Equation de la dynamique

L'équation de la dynamique s'écrit comme suit :

$$M\ddot{x}(t) = Q(x, \dot{x}, t) + F_{ext}(t) + r \quad (4.1)$$

où  $x, \dot{x}$  et  $\ddot{x}$  représentent respectivement la position, la vitesse et l'accélération des corps à un instant  $t$ ,  $M$  la matrice de masse du système,  $Q(x, \dot{x}, t)$  les efforts extérieurs ou intérieurs ainsi que des termes dépendant de la vitesse (que l'on appelle couramment centrifuges et gyroscopiques) et  $F_{ext}(t)$  les efforts extérieurs connus en l'absence de contact. Enfin,  $r$  représente les forces de contact.

Faisant l'hypothèse de l'unilatéralité (les particules ne peuvent s'interpénétrer), des chocs inter-particules se produisent. Ceci donne lieu à des discontinuité du champ de vitesse (sauts de vitesse), et l'accélération n'est pas définie. Ces sauts de vitesse relèvent de la non-régularité de la méthode. L'équation 4.1 doit donc être étendue au formalisme de la Dynamique Non Régulière.

En considérant que la fonction vitesse  $\dot{x}$  est une fonction à variation localement bornée, il est possible d'écrire le problème en terme de mesures différentielles comme le propose Moreau [151] :

$$M d\dot{x}(t) = Q(x, \dot{x}, t)dt + F_{ext}(t)dt + r d\nu \quad (4.2)$$

où  $dt$  est une mesure de Lebesgue,  $d\nu$  une mesure réelle positive, et  $dx$  la mesure différentielle de  $x$ .  $r$  représente la densité d'impulsions de contact (la contribution locale des impulsions exercées lors de la présence de contacts). La description du mouvement ne peut se faire sans conditions sur les contacts. Ces conditions sont d'abord géométriques, puis mécaniques.

## b. Détection géométrique des contacts

Comme évoqué précédemment, cette étape est fortement liée à la forme. En effet, dans une simulation discrète, les particules sont représentées par des corps maillés, dont la surface est discrétisée en plusieurs éléments appelés faces, où chaque face est composée d'un ensemble de sommets reliés par des arêtes. Tous ces éléments permettent de définir un contact entre deux ou plusieurs particules.

Dans le cas convexe, comme dans le cas non convexe, plusieurs méthodes permettent de répondre à ce besoin, dont 3 sont implémentées dans le code calcul LMGC90 [10, 152]. Dans tous les cas, afin de réduire le spectre de recherche, une première étape de détection "grossière" est effectuée, en se basant sur une distance d'alerte (proximité entre les grains). On arrive alors à éliminer rapidement plusieurs configurations où il n'y a pas de contact. Ensuite, une étape supplémentaire consiste à déterminer le type de contact selon le nombre de points choisis pour le représenter (Fig. 4.17) : les superpositions des parties de l'espace que les contacts occupent (sous l'hypothèse de la non-interpénétration) sont définies, et leur étendue définit le type du contact (un point, deux points, etc.). Enfin, les éléments nécessaires à la résolution du contact sont calculés (coordonnées dans l'espace des points de contact, repère local de contact, etc.).

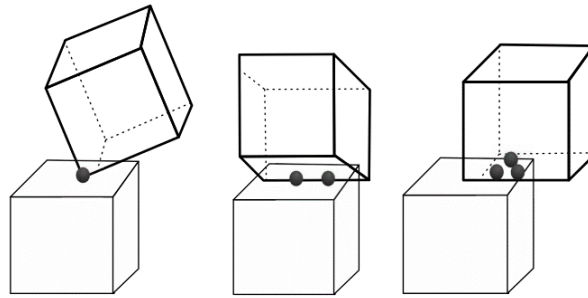


FIGURE 4.17 – Types de contacts : simple : sommet-face(a), double : arête-face(b), triple : face-face(c)

## c. Gestion de l'interaction et résolution

Après la détection géométrique des contacts, nous avons les outils nécessaires pour décrire correctement les interactions. Les particules sont considérées comme rigides et exercent des forces normales et tangentielles,  $f_n$  et  $f_t$  respectivement, l'une sur l'autre. Par convention, les forces normales de compression sont positives. De même, la vitesse relative normale  $v_n$  entre deux particules en contact est comptée positive lorsque les deux particules s'éloignent l'une de l'autre. Ainsi, la condition de contact entre deux particules s'exprime comme la réalisation de l'un des cas suivants :

$$f_n \geq 0 \quad \text{et} \quad v_n = 0 \quad (4.3)$$

$$f_n = 0 \quad \text{et} \quad v_n \geq 0 \quad (4.4)$$

Cette condition est appelée condition de Signorini-Coulomb (Fig. 4.18).

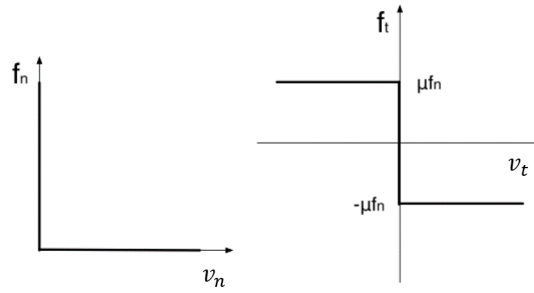


FIGURE 4.18 – Graphe de la condition de complémentarité en vitesse (Signorini) à gauche, et Graphe de la loi de Coulomb à droite.

De la même manière, la loi de frottement de Coulomb implique trois cas possibles :

$$f_t = -\mu f_n \quad \text{et} \quad v_t > 0, \quad (4.5)$$

$$-\mu f_n \leq f_t \leq \mu f_n \quad \text{et} \quad v_t = 0, \quad (4.6)$$

$$f_t = \mu f_n \quad \text{et} \quad v_t < 0 \quad (4.7)$$

où  $v_t$  est la vitesse de glissement au point de contact et  $\mu$  est le coefficient de frottement. Lors d'une collision, les vitesses  $v_n$  et  $v_t$  correspondent aux vitesses après collision. Autrement dit, la solution du problème à chaque instant  $t$  doit vérifier la condition de non-interpénétration à l'instant  $t + \delta t$  (où  $\delta t$  le pas de temps).

En outre, Moreau a montré qu'il est possible d'écrire  $v_n$  et  $v_t$  en fonction des vitesses avant et après le choc :

$$v_n = \frac{v_n^+ + \alpha_n v_n^-}{1 + \alpha_n} \quad (4.8)$$

$$v_t = \frac{v_t^+ + \alpha_t v_t^-}{1 + \alpha_t} \quad (4.9)$$

où  $v_n^+$ ,  $v_n^-$ ,  $v_t^+$  et  $v_t^-$  correspondent respectivement aux vitesses relatives normales et tangentielles des particules au contact au début et à la fin de chaque pas de temps. La condition de contact ou de collision correspond alors à  $v_n = 0$ , ce qui donne  $v_n^+ = -\alpha_n v_n^-$ . Ainsi, les paramètres  $\alpha_n$  et  $\alpha_t$  coïncident avec les coefficients de restitution (normal et tangentiel) définis habituellement pour les chocs binaires.

La méthode (CD) permet donc de généraliser le cas des collisions binaires au cas des collisions multiples ou aux systèmes multi-contacts, ce qui est le cas pour nous.

Pour ce qui est de la résolution, le code de calcul LMGC90 utilise un solveur basé sur la méthode de Gauss-Seidel non linéaire [37].

De manière générale, l'unicité de la solution n'est pas garantie pour les particules parfaitement rigides. Cependant, en initialisant chaque étape de calcul avec les forces calculées dans l'étape précédente, l'ensemble de solutions admissibles se réduit à des fluctuations autour de la solution numérique.

Enfin, on se pose la question de la qualité d'un calcul par la méthode CD. Quels sont les critères utilisés à cette fin ?

#### d. Critères de qualité d'un calcul

Afin d'évaluer la qualité et la propreté d'un calcul numérique, plusieurs indicateurs existent. Il est tout de même important de distinguer entre la qualité du calcul liée à l'algorithme de résolution (Gauss-Seidel) qui peut être mesurée par une erreur de convergence, et celle en lien avec la configuration choisie et les données d'entrée. Une manière de qualifier le deuxième aspect est de mesurer le taux d'interpénétration sur les grains, ainsi que le pourcentage de contacts où une interpénétration survient. Par définition de la méthode, aucun interpénétration n'est tolérée. Néanmoins dans toutes les simulations, avec la complexité et la taille des systèmes, des problèmes dans la résolution provoquent parfois des interpénétrations, qu'on cherche à quantifier.

Un taux d'interpénétration choisi est toléré. Nous choisissons de fixer ce seuil à 5% du rayon moyen des grains, c'est-à-dire à environ  $v_{max} = 2mm$ . Pour ce qui est du pourcentage d'interpénétration, nous représentons les distributions statistiques (PDF) de tous les taux d'interpénétration mesurés sur les contacts (données par le code de calcul LMGC90), et nous vérifions que  $p(v \geq v_{max}) \leq 10^{-2}$ .

Enfin, la convergence de l'algorithme de résolution est mesurée par une erreur dont la valeur est fournie dans les fichiers de résultats du code, et est de l'ordre de  $10^{-10}$ .

### 3 Conclusion

Dans ce chapitre, on se focalise sur le passage du nuage de points au grain maillé. L'idée est de minimiser la perte d'information sur la forme, tout en optimisant le maillage pour respecter des durées raisonnables de temps de calculs.

Après une vision globale des méthodes courantes de maillage, ainsi qu'une comparaison de leurs avantages et limitations, nous avons proposé la combinaison qui engendre le moins d'erreur possible sur la répartition initiale des nuages de points scannés. Néanmoins la méthode de caractérisation a permis de détecter et de quantifier une erreur sur la forme, engendrée par le maillage, jusque là non détectée avec les indicateurs classiques.

Ce résultat, intéressant pour la caractérisation et la mécanique à la fois, nous permet de conclure sur le potentiel de la POD à décrire précisément la morphologie des grains et de détecter les petites variations de forme. Il nous permet également de réduire, ou en tout cas de prévoir une réduction du spectre dans lequel nous nous attendons, lors des études paramétriques sur la forme, à remarquer un effet plus significatif de la forme sur la réponse mécanique, notamment dans le cas où le maillage se limite à une enveloppe convexe pour tous les grains.

Enfin, la présentation des méthodes numériques de résolution utilisées permettent d'introduire le chapitre suivant : nous nous intéressons à la réalisation de simulations numériques discrètes avec ces grains maillés, que l'on choisit de représenter par des polyèdres, et de quantifier l'effet de la variation de forme, ainsi que du frottement sur le comportement mécanique du ballast.

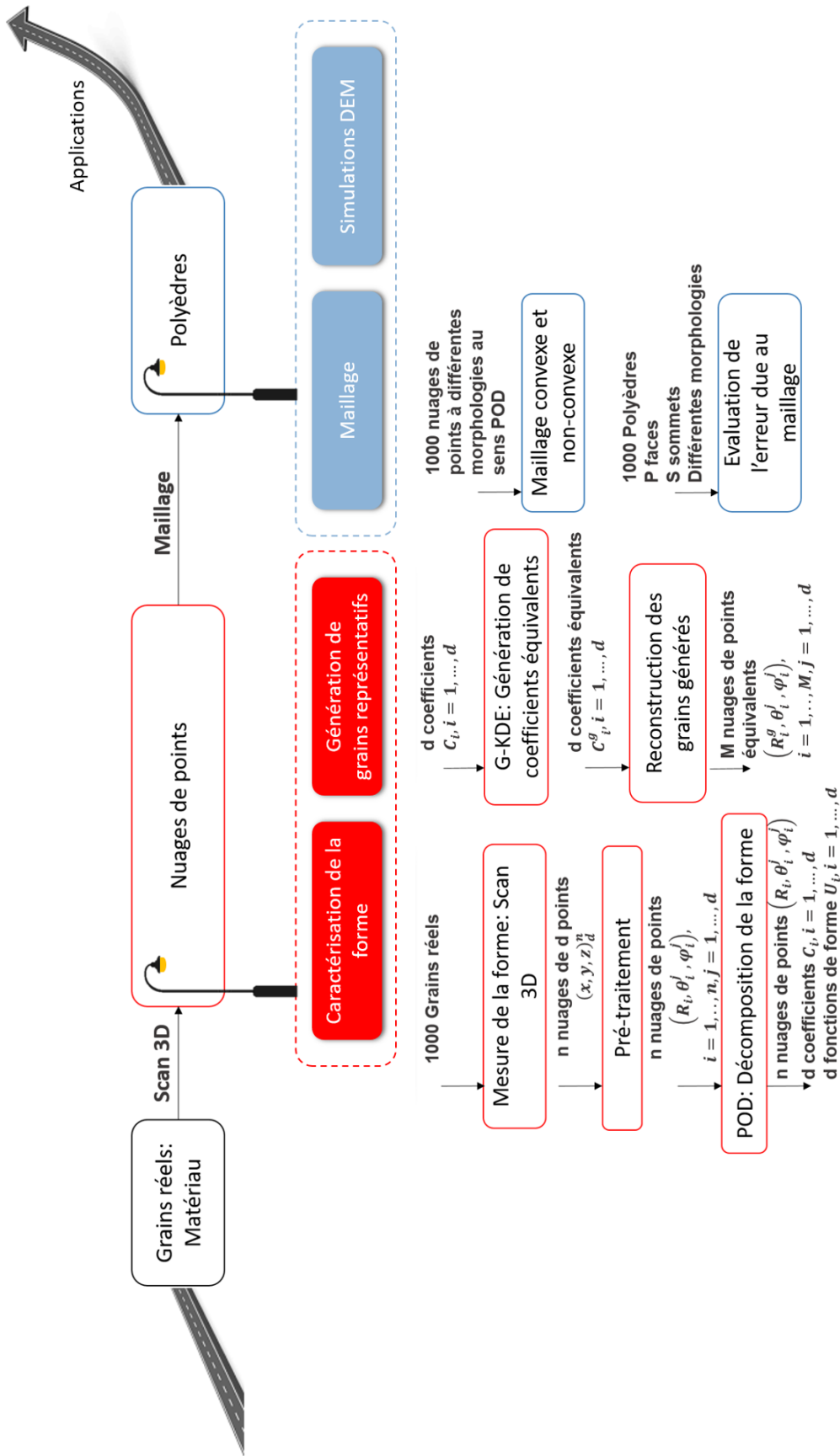


FIGURE 4.19 – Avancement sur le chemin d'étude du ballast



## Chapitre 5

### Etude de l'effet de la forme sur le comportement mécanique



Dans ce chapitre, l'intérêt est porté sur les simulations numériques par la méthode des éléments discrets (DEM). Le but global est d'explorer l'effet de la variation de la forme sur quelques aspects de la mécanique des grains de ballast. Pour ce faire, nous utilisons le logiciel LMGC90 [152] où la méthode de la Dynamique des Contacts est implémentée. Suite à la décomposition par POD expliquée dans le chapitre 2, nous avons en notre disposition  $n = 975$  grains avec des degrés différents de précision sur la forme définis par des seuils d'erreur que nous avons choisis (Tab. 5.1). Pour l'instant, nous n'utilisons pas les grains générés. Dans un premier temps, nous nous limitons à  $n_1 = 300$  grains tirés aléatoirement parmi les  $n$  grains.

TABLE 5.1 – Nombre de modes et erreur correspondante sur la forme, pour les ensembles de grains choisis

N échantillon	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	$S_7$	$S_8$	$S_9$	$S_{10}$	$S_{11}$
Erreur POD	63.5%	46%	35.7%	28%	20%	15 %	10%	5%	1%	0.1%	0%
Nombre de modes	1	3	5	7	11	16	22	40	117	350	800

Nous avons choisi d'explorer l'effet de la variation de la morphologie de la même primitive géométrique que représentent les polyèdres irréguliers facettés obtenus après maillage des différents ensembles de grains, sur le comportement mécanique du ballast. Notre choix se porte sur un nombre de faces de maillage de 100 pour tous les grains (52 sommets). La Fig. 5.1 montre les 11 configurations de grains maillés par un maillage convexe uniquement (cf. chapitre 4), sachant que pour des contraintes relatives au code de détection de contacts dans le logiciel, notre choix se limite au cas convexe.

Dans ce chapitre, nous présentons une étude du comportement mécanique du ballast par des tests de cisaillement simple maintenu, en régime quasi-statique. Nous nous intéressons alors à l'état critique du matériau sous cisaillement, qui ne dépend pas de son état initial (au-delà d'une déformation minimal de l'ordre de  $\gamma \approx 0.5$ ). Le comportement est caractérisé par l'angle de frottement interne, par la compacité, ainsi que par les différentes contraintes normales. Une attention particulière est prêtée à l'angle de frottement interne à l'état critique, proche de l'angle de talus expérimentalement obtenu pour le ballast (environ  $34^\circ$ ), même si ce dernier dépend un peu de la façon d'assembler le talus. Cette étude est précédée par une étape de préparation et d'assemblage de nos échantillons que nous décrivons dans la section qui suit.

## 1 Etat d'équilibre statique

### 1.1 Paramètres numériques

Nous disposons donc de 11 échantillons numériques  $S_i$ ,  $1 \leq i \leq 7$  dont nous utilisons  $n_1 = 300$  grains à niveau différent de précision sur la forme. Nous déposons d'abord chacun de ces ensembles de grains de manière géométrique [10,41] dans une boîte cubique de côté  $a = 0.7\text{m}$  (Fig. 5.2). Nous exerçons ensuite une compression verticale entre deux plans en appliquant une charge verticale de  $F_0 = 500\text{N}$  sur la paroi supérieure, la paroi inférieure étant immobile.

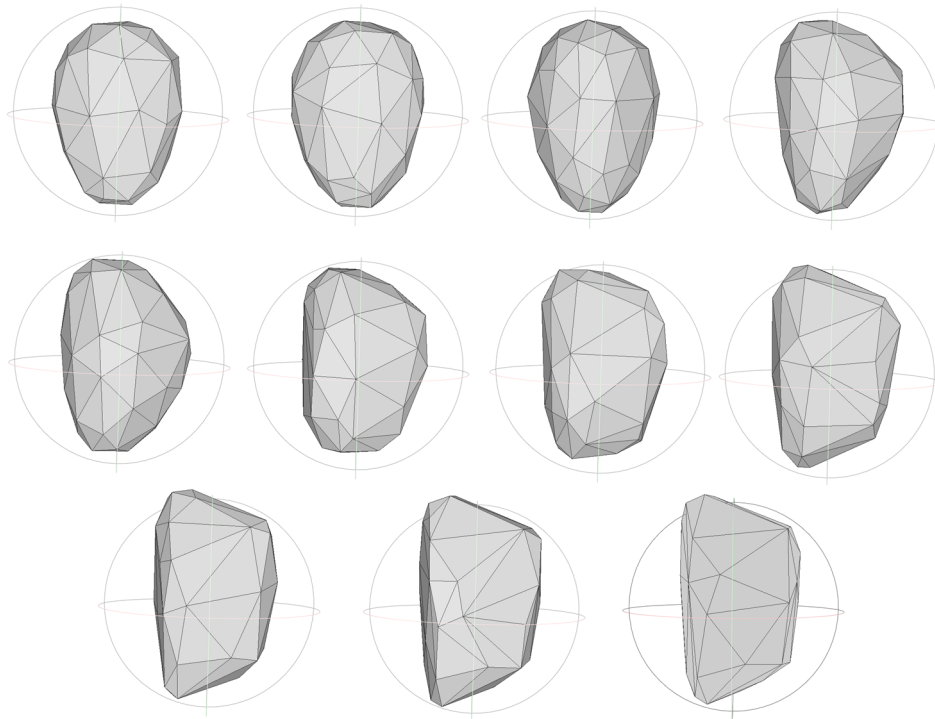


FIGURE 5.1 – Les 11 configurations de forme correspondant aux 11 troncatures du Tab. 5.1 : évolution de gauche à droite et de haut en bas.

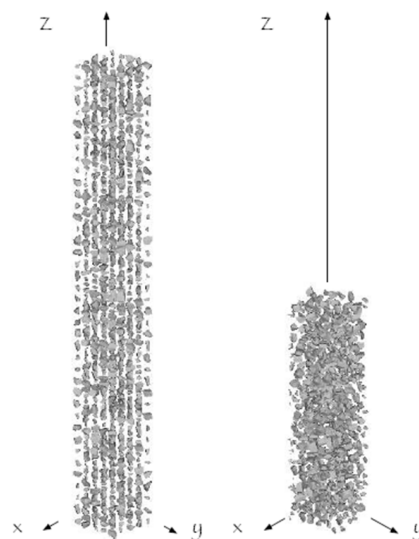


FIGURE 5.2 – Dépôt des grains dans une boîte : à gauche, dépôt selon une grille régulière ; à droite, dépôt géométrique.

La gravité est prise égale à 0 pour toutes les simulations afin d'avoir un état de contrainte homogène et éviter les gradients de force dans les échantillons. Le coefficient de frottement entre les grains est fixé à  $\mu_{gg} = 0.8$ , et à  $\mu_{gp} = 0$  avec les deux parois. La périodicité dans les deux directions horizontales est fixée à  $dx = dy = 0.3 m$ . Le pas de temps choisi est de  $dt = 5.10^{-4} s$ . Nous fixons un critère sur la vitesse de la paroi supérieure pour mettre fin à la simulation :  $v_{sup} \leq 10^{-4} m/s$ . On rappelle également que dans le cas de grains polyédriques, les contacts doubles (arête-face) et triples (face-face) seront comptés comme un seul contact dans les études statistiques bien que numériquement les contacts doubles et triples soient traités avec deux points et trois points de contacts appartenant à un même segment (face) (voir chapitre 4).

## 1.2 Vérifications de l'équilibre mécanique

La Fig. 5.3 montre un exemple de l'état final de la compression verticale.

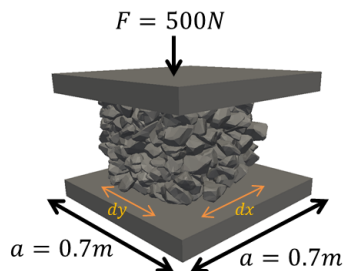


FIGURE 5.3 – Etat final de la compression : 800 Modes

Trois principales vérifications de l'équilibre mécanique sont faites à la suite de ces simulations :

**Le taux d'interpénétration** : comme expliqué au chapitre 4, la distribution statistique des interpénétrations permet de vérifier que le taux d'interpénétration moyen ne dépasse pas le seuil fixé (5% du rayon moyen des grains) (Fig. 5.4). La Figure montre la distribution statistique des distances entre les grains définis comme potentiellement en contact (selon une distance d'alerte choisie, cf. chapitre 4). Les valeurs négatives correspondent à une interpénétration. Nous tolérons une valeur inférieure à 5% du rayon moyen (2 mm), et la probabilité correspondante est à gauche du trait rouge, et sa valeur moyenne ne dépasse pas  $p = 8.25 \cdot 10^{-3}$  dans toutes les simulations.

**Le rapport K** : le rapport sans dimension :  $K = \frac{E_c}{\sigma D^3}$  où  $E_c = \frac{1}{2n_g} \sum_{i=1}^{n_g} m_i \vec{v}_i^2 + \vec{w}_i \mathbf{I}_i \vec{w}_i$  est l'énergie cinétique moyenne par grain, et  $\sigma = \frac{F_0}{S}$  la pression exercée sur la paroi supérieure, et  $D$  le diamètre moyen des grains, donne une valeur finale d'environ  $K = 3.9510^{-6}$ . Ce rapport très faible et sans décrit la stabilité mécanique du système.

**La force  $F_{inf}$**  : la force appliquée sur la paroi inférieure augmente jusqu'à atteindre une valeur constante de  $F_{inf} = 498.63N \approx F_0$  (Fig. 5.5).

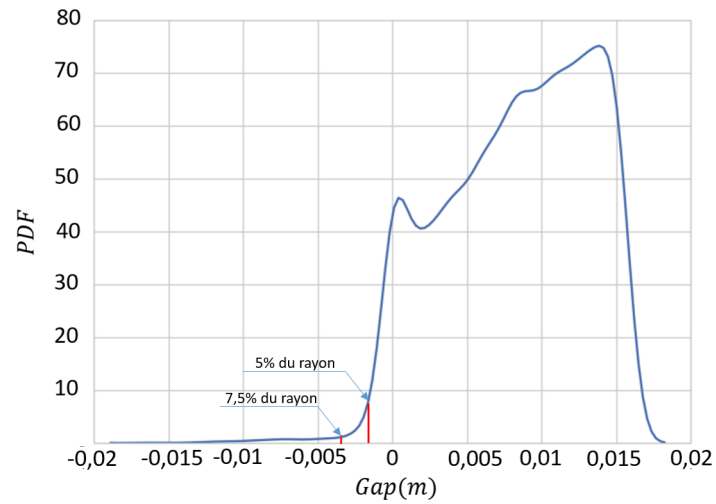


FIGURE 5.4 – Distribution statistique des distances entre les grains potentiellement en contact.

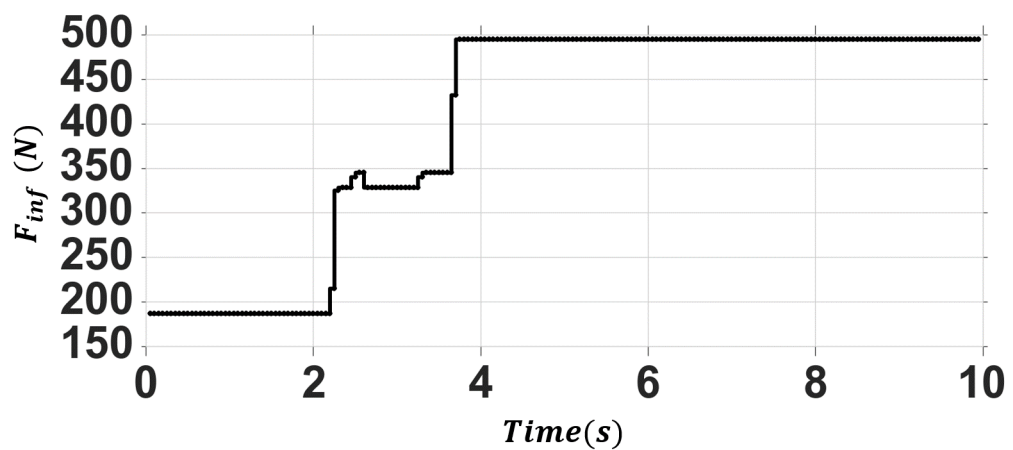


FIGURE 5.5 – Evolution de la force exercée sur la paroi inférieure.

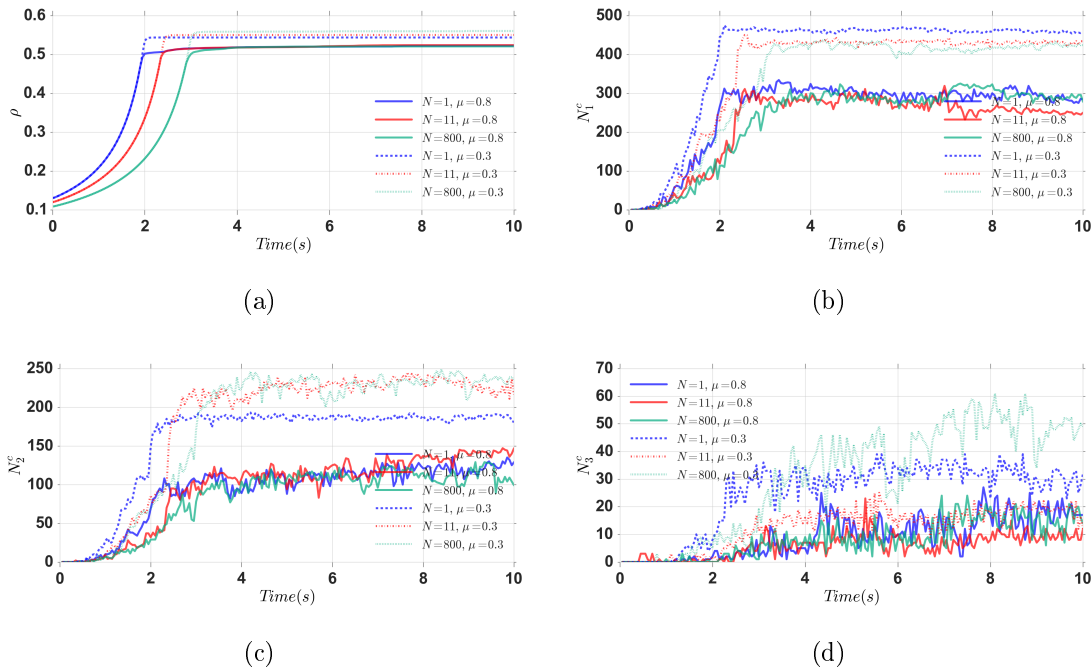


FIGURE 5.6 – Influence du nombre de modes sur quelques paramètres mécaniques pour  $\mu_{gg} = 0.8$  et pour  $\mu_{gg} = 0.3$  : (a)  $\rho$ ; (b)  $N_1^c$ ; (c)  $N_2^c$ ; et, (d)  $N_3^c$ .

Ces vérifications montrent que l'on atteint un état d'équilibre statique de qualité, qui permet d'assurer un bon état initial pour les simulations de cisaillement quasi-statique.

### 1.3 Etude préliminaire des effets de forme et du frottement

Nous nous proposons de réaliser une première évaluation de l'effet de la forme à celui du frottement sur le comportement mécanique. Nous choisissons alors deux valeurs pour le coefficient de frottement entre les grains :  $\mu_{gg} = 0.3$  et  $\mu_{gg} = 0.8$ . Nous évaluons l'effet de forme pour 3 valeurs de modes  $N = 1$ ,  $N = 11$  et  $N = 800$  modes, en termes de compacité (comportement macroscopique) (Fig. 5.6(a)) et de nombre de contacts simples, doubles et triples (comportement microscopique) (Fig. 5.6(b), Fig. 5.6(c) et Fig. 5.6(d)).

La Fig. 5.6 montre que le frottement pilote la réponse mécanique. A coefficient de frottement constant, nous remarquons un faible effet de forme. Cependant, avec un frottement de  $\mu_{gg} = 0.3$  entre les grains, l'influence de la forme est plus significative que dans le cas  $\mu_{gg} = 0.8$ . Pour confirmer cette forte dépendance au frottement, nous réalisons une étude paramétrique de l'effet du frottement à  $N = 800$  modes fixé et réciproquement à  $\mu_{gg} = 0.8$ . Le résultat de cette étude est présenté dans la Fig. 5.7.

La Fig. 5.7 met en évidence la dominance du frottement par rapport à la forme quant au pilotage de la réponse mécanique, lorsqu'il est comparé à l'effet de forme pour un frottement  $\mu_{gg}$ . En tenant compte des résultats donnés pour un frottement plus faible, et afin de mieux comprendre le lien forme-frottement, nous nous proposons alors d'explorer plusieurs valeur de frottement dans nos simulations de

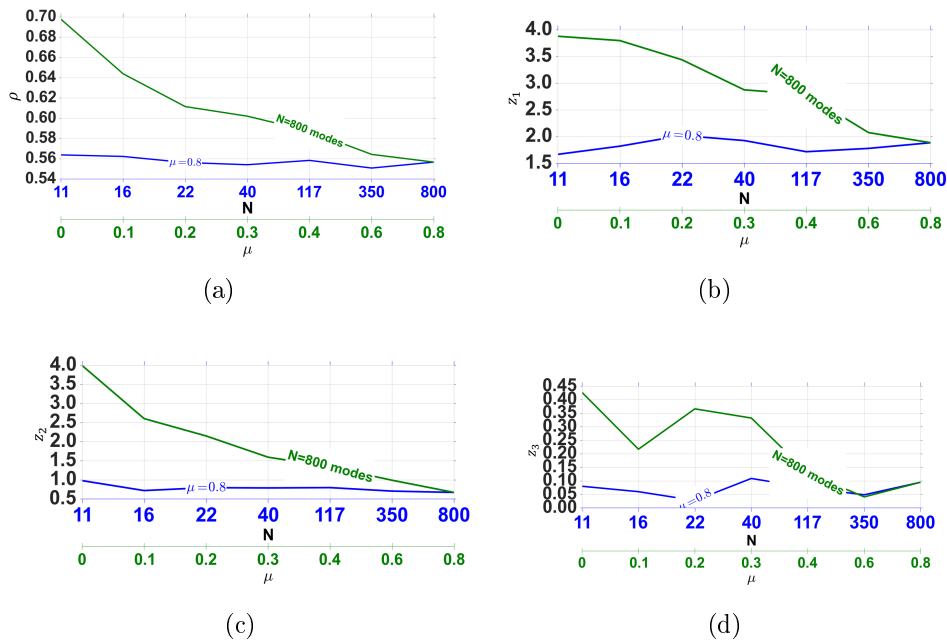


FIGURE 5.7 – Evolution de la valeur moyenne de quelques paramètres mécaniques à l'état stationnaire en fonction de la forme et du frottement : (a)  $\rho$  ; (b)  $z_1$  ; (c)  $z_2$  ; et, (d)  $z_3$ .

cisaillement, en plus des différentes morphologies de grains. Le choix du coefficient de frottement  $\mu_{gg} = 0.8$  est le résultat d'un ajustement du coefficient de frottement effectué dans [10] en utilisant un ensemble de grains scannés à forme pauvre (environ 15 sommets et 20 faces), mais permettant de reproduire l'angle de frottement interne du ballast. Nous nous posons alors la question sur une équivalence existant entre le frottement et la morphologie : pour des grains se rapprochant davantage de la morphologie réelle du ballast, quel coefficient de frottement est nécessaire afin de reproduire son comportement par la simulation ?

## 1.4 Synthèse

Cette partie consacrée aux états d'équilibre statique met en évidence l'influence couplée de la forme-frottement sur la réponse mécanique. Nous effectuons les mêmes simulations de préparation des échantillons avec la totalité des grains ( $n = 975$ ), que nous menons jusqu'à un état d'équilibre statique, afin de réaliser des simulations de cisaillement où nous prêtons une attention particulière à la dualité forme-frottement, et où nous nous intéressons à différents aspects du comportement mécanique.

# 2 Cisaillement simple maintenu

## 2.1 Paramètres numériques

Nous réalisons des tests de cisaillement direct sur nos échantillons compactés de  $n = 975$  grains chacun. En maintenant une force verticale sur la paroi supérieure de  $F_0 = 500N$ , nous appliquons une vitesse horizontale  $v_{shear} = 1 \text{ cm/s}$  dans la

direction  $y$  (Fig. 5.8). Pour simuler une paroi rugueuse, nous fixons le frottement entre les grains et la paroi à  $\mu_{gp} = 0.8$ . Nous limitons notre étude aux échantillons  $S_i$ ,  $i = 1, \dots, 9$ . En effet, comme résultat du chapitre 2, 3 et 4 ainsi que par l'étude des états d'équilibre statique,  $d = 117$  modes sont suffisamment représentatifs de l'ensemble des modes en terme de morphologie. Pour chacun de ces échantillons, nous menons des simulations pour 3 valeurs de frottement inter-granulaire :  $\mu_{gg} = 0.1; 0.3; 0.8$ . Nous maintenons les mêmes conditions de périodicité, et nous menons nos simulations jusqu'à obtenir une déformation de cisaillement cumulée  $\gamma = 3$ . Le pas de temps est maintenu à  $dt = 5.10^{-4}$  s. La Fig. 5.8 montre un exemple de l'état final de la simulation ( $N = 11$  modes). Tous les grains ont un nombre de faces fixé à  $N_f = 100$ .

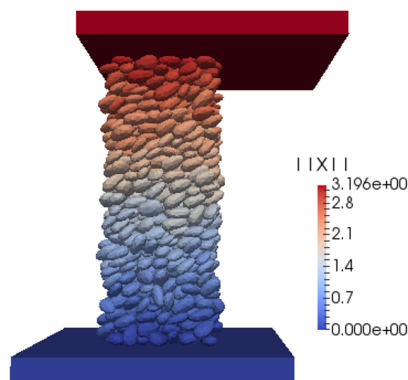


FIGURE 5.8 – Etat final de l'essai de cisaillement(11 modes) : la coloration correspond au déplacement

## 2.2 Vérifications de la qualité des calculs numériques

Puisque nous nous intéressons au comportement quasi-statique, le taux de déformation doit être assez faible pour que l'énergie cinétique induite par la déformation soit négligeable comparée à la pression statique. En d'autres termes, le paramètre d'inertie  $I$  [153–157] doit être petit devant 1 :

$$I = \dot{\gamma} \sqrt{\frac{m}{Dp}} \quad (5.1)$$

où :

$$\dot{\gamma} = \frac{v_{shear}}{h} \quad (5.2)$$

est le taux de cisaillement, avec  $h$  la hauteur de l'échantillon.  $m$  est la masse moyenne d'un grain ( $m \approx 100g$ ),  $D$  le diamètre moyen des grains  $D \approx 0.06m$  et  $p$  la pression moyenne ( $p = \frac{F_0}{S} \approx 5555$  pa). La Fig. 5.9(a) montre l'évolution du paramètre d'inertie. Globalement, dans l'ensemble des simulations,  $I < 2.5 \times 10^{-4}$ . Pour ce qui est d'interpénétrations, la probabilité d'obtenir une interpénétration de plus de 3 mm ne dépasse pas  $2.16 \cdot 10^{-6}$ . Le rapport  $K$  décroît jusqu'à atteindre

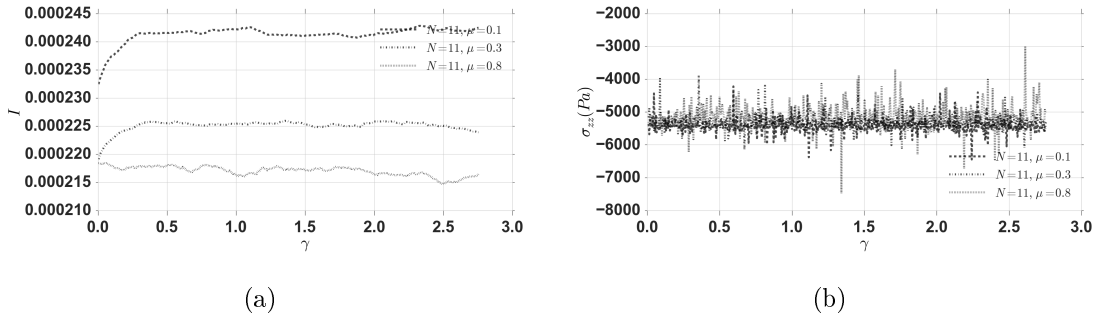


FIGURE 5.9 – Evolution du paramètre d'inertie  $I$  et de la contrainte verticale  $\sigma_{zz}$  au long de la simulation :(a)  $I$ ; (b)  $\sigma_{zz}$ .

une valeur de  $K_f = 3.10^{-7}$ , ce qui témoigne de l'état d'équilibre quasi-statique du système. Nous obtenons également des profils de vitesse linéaires, ce qui justifie de la bonne qualité du test de cisaillement réalisé. La valeur finale de la résultante de forces exercées sur la paroi inférieure est d'environ 499,27 N. Enfin, la Fig. 5.9(b) montre l'évolution de la contrainte verticale  $\sigma_{zz}$  au long de la simulation, et qui se stabilise vers la valeur de la pression  $p$  appliquée ( $\sigma_{zz} \approx 5432 Pa$ ).

Grâce à ces indicateurs simples, nous estimons avoir simulé des essais de cisaillement quasi-statique de qualité, permettant l'étude des états stationnaires de cisaillement. Nous étudions maintenant l'effet de la forme et du frottement sur le comportement mécanique du ballast à différentes échelles.

## 2.3 Effets de la forme et du frottement

### a. Comportement macroscopique

La Fig. 5.10(a) montre l'évolution de la compacité en fonction de la déformation de cisaillement cumulée  $\gamma$  pour les différents modes à  $\mu_{gg} = 0.3$ . On observe qu'un état stationnaire est atteint pour  $\gamma \geq 1$ . Nous remarquons également l'effet significatif du frottement sur la valeur de la compacité. La Fig. 5.10(b) résume les valeurs de la compacité à l'état stationnaire pour l'ensemble des modes et pour les trois valeurs de  $\mu_{gg}$  choisies. D'abord, une tendance de la compacité est claire d'autant plus que le frottement est faible. Pour un nombre de modes  $N \leq 11$ , nous remarquons davantage d'effet de la forme sur la compacité, puis une stabilisation au-delà de 11 modes. On note la différence significative entre le cas du frottement  $\mu_{gg} = 0.1$  et  $\mu_{gg} = 0.3$ , par rapport au cas  $\mu_{gg} = 0.3$  et  $\mu_{gg} = 0.8$ . Les cas  $\mu_{gg} = 0.1$  et  $\mu_{gg} = 0.3$  montrent un comportement contractant, tandis que le cas  $\mu_{gg} = 0.8$  montre un comportement plus dilatant. Cela s'explique par la valeur du frottement durant l'étape de préparation de l'échantillon ( $\mu_{gg} = 0.3$ ).

Nous étudions également le comportement macroscopique en terme de résistance. Nous rappelons d'abord la définition du tenseur de contrainte  $\sigma$  à partir des moments tensoriels  $M_i$ ,  $i = 1, \dots, n$  de chaque grain comme suit :

$$\sigma = \frac{1}{V} \sum_{i \in V} M^i \quad (5.3)$$



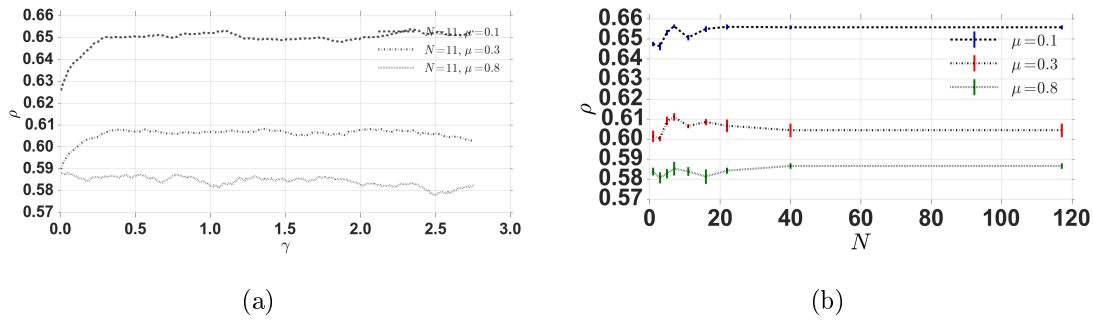


FIGURE 5.10 – Evolution de la compacité pour différents coefficients de frottement pour 11 modes en fonction de la déformation de cisaillement  $\gamma$ , et son évolution à l'état stationnaire en fonction du nombre de modes pour différents coefficients de frottement. Les barres représentent l'écart-type de chaque valeur moyenne. : (a)  $\rho(\gamma)$ ; (b)  $\rho_M(N)$ .

où :

$$M_{\alpha\beta}^i = \sum_{c \in i} f_{\alpha}^c r_{\beta}^c \quad (5.4)$$

$M_i$  est le moment tensoriel relatif au grain  $i$ ,  $f_{\alpha}^c$  est la composante de la force exercée sur le grain  $i$  suivant la direction de l'espace  $\alpha$  ( $\alpha = 1, 2 \text{ ou } 3$ ) et  $r_{\beta}^c$  est la composante  $\beta$  ( $\alpha = 1, 2, 3$ ) du vecteur position sur le même contact  $c$ . Le tenseur de contrainte est donc une moyenne statistique sur le volume total  $V$  formé par l'ensemble des grains. Comme tout paramètre statistique, sa pertinence s'améliore d'autant plus que la valeur de  $V$  est grande.

Comme montré dans [41], ce tenseur peut être exprimé en utilisant les vecteurs "branches"  $l^c$  au niveau de chaque contact, ou encore les vecteurs inter-centres des grains ayant un contact  $c$  comme suit, pour chaque deux directions  $\alpha$  et  $\beta$  :

$$\sigma_{\alpha\beta} = \frac{1}{V} \sum_{c=1}^{N_c} f_{\alpha}^c l_{\beta}^c \quad (5.5)$$

où  $N_c$  est le nombre total de contact. le calcul du tenseur de contraintes permet l'évaluation des trois indicateurs suivants :

$$N_1 = \frac{\sigma_{zz} - \sigma_{yy}}{\sigma_{zz}} \quad (5.6)$$

$$N_2 = \frac{\sigma_{xx} - \sigma_{zz}}{\sigma_{xx}} \quad (5.7)$$

$$\tan \varphi = \frac{|\sigma_{yz}|}{|\sigma_{zz}|} \quad (5.8)$$

$N_1$  et  $N_2$  sont appelées différences des contraintes normales. Elles permettent de décrire l'état du matériau, et contrôlent notamment les orientations des directions principales des contraintes. La Fig. 5.11 résume l'ensemble des valeurs calculées de  $N_1$  et  $N_2$  pour différents échantillons et différents coefficients de frottement  $\mu_{gg}$ .

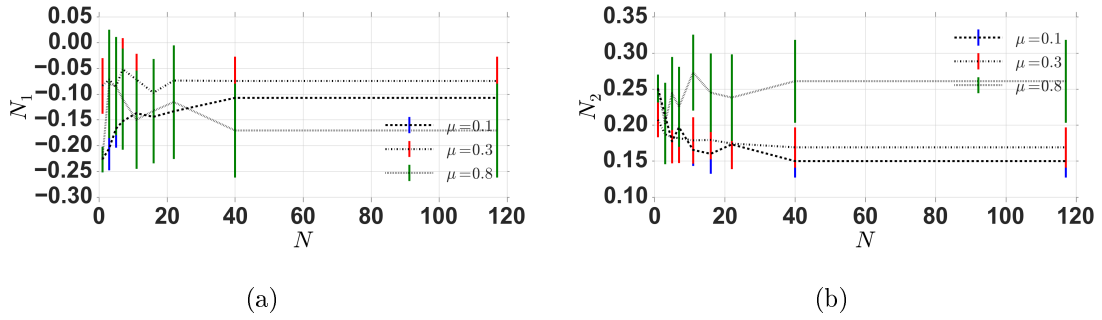


FIGURE 5.11 – Evolution des différences de contraintes normales à l'état stationnaire en fonction du nombre de modes pour différents coefficients de frottement : (a)  $N_1$  ; (b)  $N_2$ .

Dans le cas de simulations réalisées avec des billes par exemple, les valeurs observées de  $N_1$  et  $N_2$  sont très faibles, ce qui n'est pas le cas ici.

$\varphi$  est l'angle de frottement interne du matériau. Il permet d'évaluer la résistance au cisaillement : celle-ci est d'autant plus élevée que la valeur de  $\varphi$  est grande. Comme montré par plusieurs études [10, 38, 41], les grains modèles de ballast possèdent une valeur  $\varphi \approx 34^\circ$  (contre  $\varphi \approx 18^\circ$  pour des sphères de même granulométrie). Nous définissons également la contrainte moyenne  $p = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3}$  ainsi que la contrainte déviatorique  $q = \frac{\sigma_1 - \sigma_3}{3}$ , où  $\sigma_i$ ,  $i = 1, 2, 3$  les contraintes principales. Le rapport  $\frac{q}{p}$  est appelé déviateur normalisé des contraintes. Il permet de caractériser l'état moyen des contraintes dans le matériau, et la courbe contrainte-déformation correspondante croît jusqu'à un seuil pour se stabiliser à l'état stationnaire. Selon une approche utilisée par l'équipe granulaire du LMGC, ce rapport est étroitement lié à l'angle de frottement interne, puisque l'angle de frottement interne est défini à partir des valeurs propres du tenseur de contraintes en utilisant le critère de Mohr-Coulomb :

$$\sin \varphi_M = \frac{q_M}{p} = \frac{\sigma_1 - \sigma_3}{\sigma_1 + \sigma_3} \quad (5.9)$$

De manière générale,  $\varphi_M \neq \varphi$ . Dans le cas particulier de matériaux obéissant exactement au critère de Mohr-Coulomb, les deux voies de détermination de  $\varphi$  produisent des résultats identiques. Chacune des deux définitions est alors purement conventionnelle. Cependant, n'étant pas le cas général de tous les matériaux [158–160], il faut alors bien préciser la définition de l'angle de frottement interne avant de calculer. Dans notre cas, nous utiliserons la définition de l'Eq. 5.8, jugée comme présentant l'avantage d'une plus grande simplicité. La Fig. 5.12(a) montre l'évolution de  $\tan \varphi$  en fonction de la déformation de cisaillement cumulée  $\gamma$  pour l'échantillon  $S_5$ .

La Fig. 5.12(a) permet d'observer l'effet marqué du frottement pour quelques simulations, mais également une atteinte rapide de l'état stationnaire pour un coefficient de frottement de  $\mu_{gg} = 0.8$  par rapport aux deux autres valeurs. L'évolution de  $\tan \varphi$  pour les trois valeurs de frottement confirme le comportement contractant relevé pour  $\mu_{gg} = 0.3$  et  $\mu_{gg} = 0.1$ , et le comportement dilatant observé pour  $\mu_{gg} = 0.8$ . La Fig. 5.12(b) résume l'ensemble des valeurs calculées de  $\tan \varphi$  pour différents échantillons et différents coefficients de frottement  $\mu_{gg}$ . De la même ma-

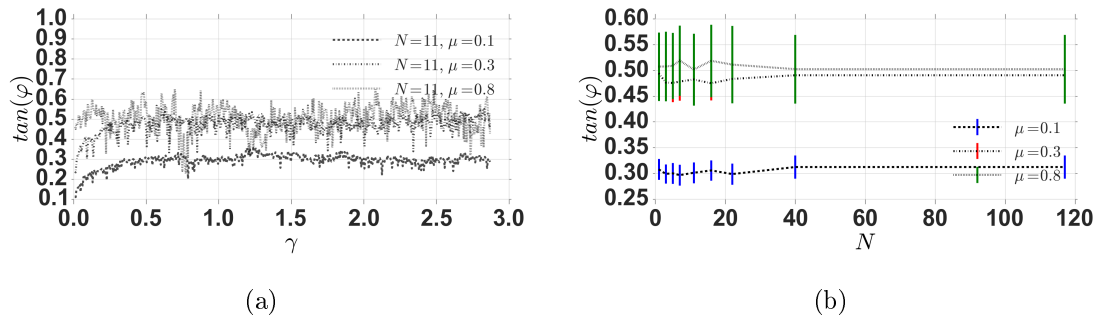


FIGURE 5.12 – Evolution de  $\tan \varphi$  pour différents coefficients de frottement pour 11 modes en fonction de la déformation de cisaillement  $\gamma$ , et son évolution à l'état stationnaire en fonction du nombre de modes pour différents coefficients de frottement. Les barres représentent l'écart-type de chaque valeur moyenne. :(a)  $\tan \varphi(\gamma)$ ; (b)  $\tan \varphi_M(N)$ .

nière que pour la compacité, les différences les plus notables concernent un nombre de modes inférieur à  $N = 11$  modes. Au-delà de cette configuration, cette valeur se stabilise à environ 0.52, ce qui correspond à un angle de frottement interne d'environ  $27.5^\circ$ . Nous discuterons plus en détail cette valeur dans la conclusion.

## b. Connectivité

À plus petite échelle, la connectivité des grains est décrite par le nombre de coordination  $z$  représentant le nombre moyen de contacts qui transmettent des forces non négligeables. Notons qu'un grain voisin en contact avec un autre grain ne transmet pas forcément une force ni un moment. Il peut en effet s'agir d'un grain dit flottant. La proportion  $x_0$  de ces grains est mesurée pour tous les échantillons et vaut en moyenne 2%, contre une valeur moyenne de 10% à la fin de la préparation de l'échantillon. De manière analogue, les grains voisins effectivement impliqués dans la transmission des efforts sont appelés grains actifs et représentent une proportion  $n(1-x_0)$  où  $n$  est le nombre total de grains. Dans notre description de la connectivité des grains, nous nous intéressons uniquement aux grains actifs.

Nous définissons le nombre de coordination à partir des  $z_i$ ,  $i = 1, 2, 3$  nombres de coordinations selon le type de contact :

$$z_i = \frac{2N_i^c + N_i^p}{n(1-x_0)} \quad (5.10)$$

$$z = z_1 + z_2 + z_3 \quad (5.11)$$

La Fig. 5.13(a) montre l'évolution de  $z$  en fonction de la déformation de cisaillement cumulée  $\gamma$  pour l'échantillon  $S_5$ . L'effet de frottement est très marqué pour le nombre de coordination. Nous atteignons une valeur stationnaire d'environ  $z \approx 4.5$  pour  $\mu_{gg} = 0.1$ ,  $z \approx 7$  pour  $\mu_{gg} = 0.3$  et  $z \approx 8.5$  pour  $\mu_{gg} = 0.8$ . La Fig. 5.13(b) résume l'ensemble des valeurs calculées de  $z$  pour différents échantillons et différents coefficients de frottement  $\mu_{gg}$ . Dans cette figure, l'effet de forme est plus prononcé par rapport aux résultats fournis par les paramètres mécaniques macroscopiques. Pour un coefficient de frottement de  $\mu_{gg} = 0.1$ ,  $z$  dépend de la forme (au nombre de

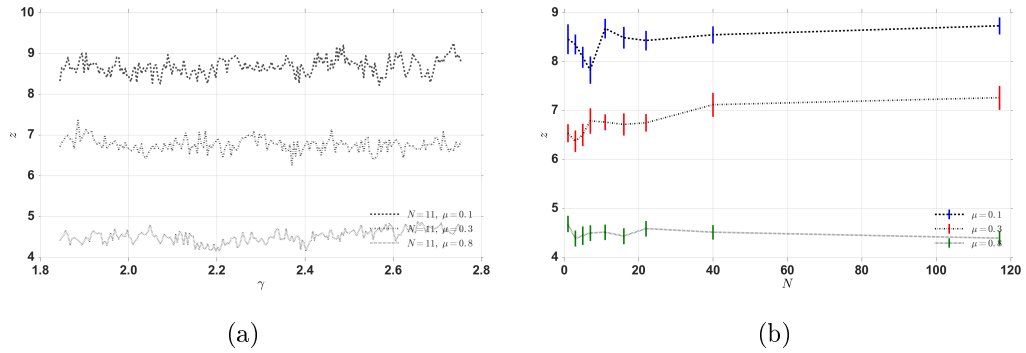


FIGURE 5.13 – Evolution de  $z$  pour différents coefficients de frottement pour 11 modes en fonction de la déformation de cisaillement  $\gamma$ , et son évolution à l'état stationnaire en fonction du nombre de modes pour différents coefficients de frottement. Les barres représentent l'écart-type de chaque valeur moyenne : (a)  $z(\gamma)$ ; (b)  $z_M(N)$ .

modes) s'exprime plus clairement. La tendance reste claire même au-delà de  $N = 11$  modes. Cela suggère une plus grande sensibilité à la forme à l'échelle microscopique.

La Fig. 5.14 résume l'ensemble des valeurs calculées de  $z_1$ ,  $z_2$  et  $z_3$  à l'état stationnaire pour différents échantillons et différents coefficients de frottement  $\mu_{gg}$ .

La variation de ces 3 indicateurs semble dépendre du frottement, mais également de la forme, et leur dépendance à la forme est d'autant plus visible que le frottement est faible. Pour  $z_1$  et  $z_2$ , nous notons surtout un comportement non monotone. Nous remarquons une décroissance de leur valeur moyenne entre  $N = 1$  mode et  $N = 11$  modes, au profit de  $z_3$ . Cette évolution est justifiée par la forme des grains relatifs à cet intervalle de modes (Fig. 5.1) qui favorise davantage de contacts face-face, puisqu'elle tend vers une forme ellipsoïdale à courbures faibles. Au fur et à mesure de l'augmentation du nombre de modes, les contacts face-face sont remplacés par des contacts simples et doubles suite à l'apparition progressive de l'angularité et des grandes faces (cf. chapitre 3). A partir de  $N = 40$  modes, où la forme est déjà suffisamment bien décrite, la dépendance des trois indicateurs à la forme est moins prononcée que pour les modes inférieurs.

## 2.4 Comparaison aux grains historiques de ballast

Les grains historiques de ballast proviennent d'une campagne de scan 3D réalisée dans [10]. Ce sont les grains traditionnellement utilisés pour les simulations numériques à SNCF. Comparés à l'ensemble des grains scannés et utilisés dans les simulations numériques de ce travail, ces grains comportent une description moins riche de la forme (Fig. 5.15).

Comme expliqué précédemment, cette description très anguleuse de la forme a permis, en ajustant le coefficient de frottement inter-granulaire  $\mu_{gg}$  à une valeur de 0.8, de reproduire un angle de frottement interne égal à l'angle de repos du ballast évalué expérimentalement [41], égal à environ  $34^\circ$ . Dans notre cas, avec une paramétrisation plus riche de la forme où nous trouvons une valeur de frottement de  $\varphi \approx 27.5^\circ$ , nous disposons de grains se rapprochant davantage de la forme réelle du ballast, mais pour lesquels nous ne possédons pas d'angularité "artificielle" créée par l'effet d'un nombre faible de sommets et de faces comme pour les grains

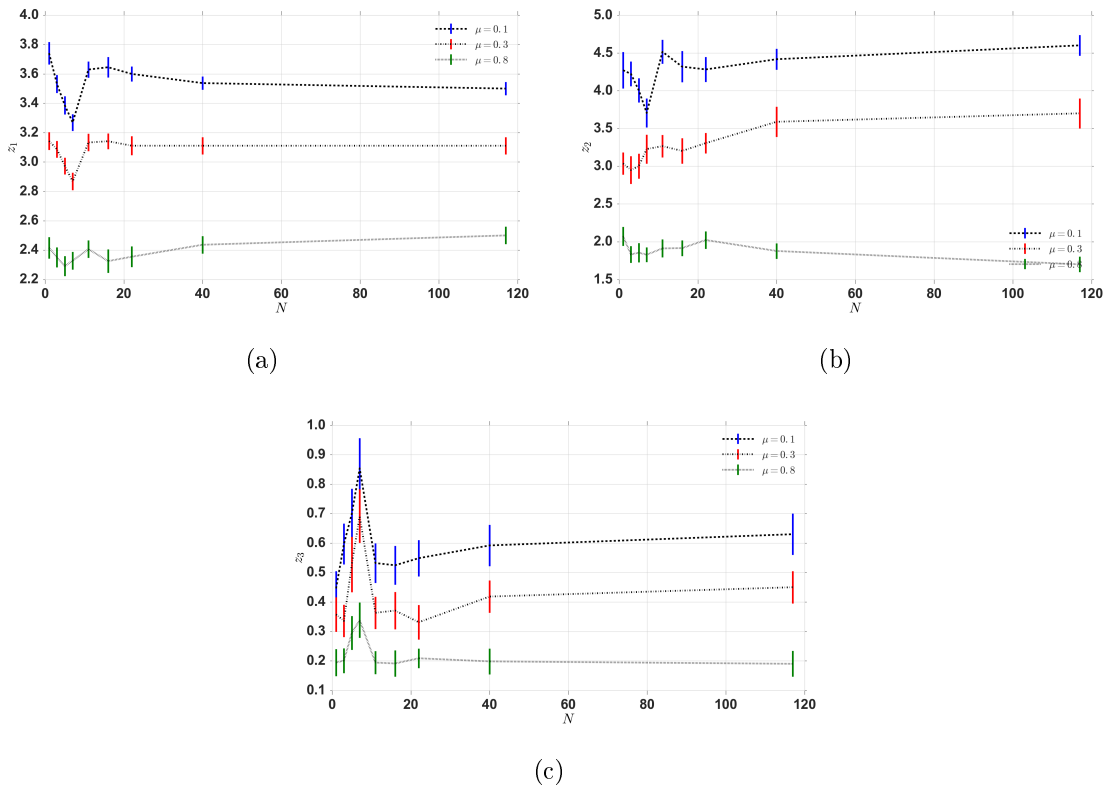


FIGURE 5.14 – Evolution de la valeur moyenne des  $z_i$  à l'état stationnaire en fonction du nombre de modes pour différents coefficients de frottement : (a)  $z_1$  ; (b)  $z_2$  ; et, (c)  $N_2$ .

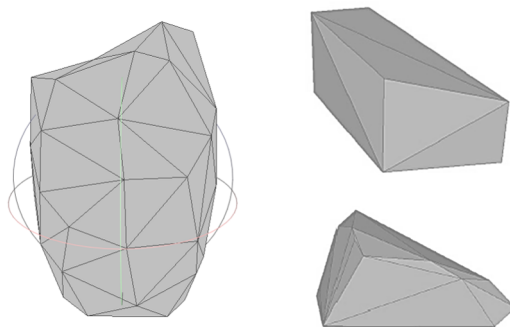


FIGURE 5.15 – Grain classiquement utilisé (à gauche) vs grain de notre échantillon scanné (à droite)

historiques (effet du maillage). Pour mieux illustrer cette observation, la Fig. 5.16 montre les différentes configurations de forme obtenues dans notre cas pour différents modes, comparés aux grains utilisés historiquement à SNCF.

La ligne horizontale correspond aux grains obtenus pour différents nombres de modes après maillage, qui sont ensuite utilisés pour nos simulations ( $N_f = 100$  faces). Les deux colonnes correspondent aux configurations de forme respectives relatives à  $N = 1$  mode et  $N = 800$  modes pour différents nombre de faces. Enfin, les trois grains sur la colonne de droite montrent un exemple de grains historiquement utilisés pour 3 valeurs de nombre de faces ( $N_f = 10$ ,  $N_f = 20$ ,  $N_f = 30$  faces).

De manière générale, pour un maillage convexe, les grains utilisés dans nos simulations pour tous les nombres de modes sont plus "ronds" que les grains historiquement utilisés dans les études précédentes [10, 41], ce qui peut expliquer la réduction de la valeur de l'angle de frottement interne par rapport au cas classique. De plus, nous remarquons que les différences initiales de géométrie sont conservées pour un nombre élevé de faces (2000 faces par exemple), mais se perdent pour un nombre de faces plus faibles : pour 100 faces, les modes de  $N = 1$  à  $N = 5$ , de  $N = 7$  à  $N = 22$  et de  $N = 40$  à  $N = 800$  modes semblent marquer une faible différence dans la forme, ce qui peut expliquer le faible effet de forme observé dans nos simulations de cisaillement, et une stabilisation des paramètres mécaniques macroscopiques à partir de  $N = 40$  modes. Ensuite, pour un faible nombre de faces (20,30 faces), nous remarquons bien que le résultat obtenu par  $N = 1$  mode et  $N = 800$  modes est semblable. Afin de conserver la différence de la géométrie, quel que soit le nombre de modes, le maillage est capital : il est alors souhaitable de réaliser des simulations avec un nombre de faces d'au moins  $N_f = 1000$  faces, mais qui nécessite des temps de calcul plus longs.

La valeur de frottement interne obtenue pour les grains historiquement utilisés à SNCF a permis de calibrer un coefficient de frottement conventionnel de  $\mu_{gg} = 0.8$  à partir d'essais en tambour tournant [10]. Comme le montre l'étude faite par l'auteur sur l'effet du coefficient de frottement sur l'angle de frottement interne, la seule variation du coefficient de frottement permet de gagner jusqu'à 7 sur l'angle de frottement. Ceci est valable pour les grains très anguleux en question. Dans notre cas, où nous disposons d'une paramétrisation plus riche de la forme, la combinaison forme-frottement gagne en complexité, d'autant plus qu'un effort sur la fidélité du maillage à la géométrie initiale montre que le maillage se rajoute à la dualité forme-frottement. Il devient alors plus difficile de contrôler le frottement interne uniquement à l'aide le coefficient de frottement. Néanmoins, en dépit de sa complexité, l'avantage de cette paramétrisation plus riche est le contrôle des erreurs relatives à chaque étape (scan - caractérisation - maillage). La reproduction du frottement interne reste alors une question intéressante à creuser.

Les Fig. 5.17(a) et Fig. 5.17(b) permettent de mettre en évidence la différence obtenue entre les grains historiquement utilisés avec 20-30 faces et nos grains à 100 faces, en termes de compacité et de frottement interne. Nous remarquons bien que l'on atteint une valeur de  $\tan \varphi \approx 0.64$  pour les grains très anguleux, ce qui correspond à environ  $33^\circ$ , contre une valeur de  $27.5^\circ$  pour nos grains. La compacité montre également une différence notable, et met en évidence un comportement dilatatant pour les deux cas, encore plus prononcé pour les grains historiquement utilisés. D'une autre part, les Fig. 5.18(a) et Fig. 5.18(b) montrent les valeurs de ces deux

indicateurs pour 20,30,50, et 100 faces entre nos échantillons de grains à  $N = 800$  modes. Nous remarquons une augmentation du frottement pour un faible nombre de faces, augmentation qui sera d'autant plus prononcée pour un frottement plus élevé, de 0.8 par exemple. Les simulations relatives à ces cas sont actuellement en cours pour mieux mettre en évidence le mécanisme maillage-frottement. Un autre point consiste à réaliser des simulations avec le cas non convexe. En effet, la suppression de la non-convexité (pour des raisons de détection de contact non convexe encore en cours d'amélioration par l'équipe LMGC) peut expliquer une réduction du frottement interne. Enfin, le réel intérêt ici est porté sur l'effet relatif de la forme sur le comportement mécanique, objectif atteint avec les simulations effectuées pour les formes issues du maillage, quelles qu'elles soient. Les résultats obtenus permettent d'ouvrir une piste intéressante pour tenir compte du rôle du maillage, de la non-convexité, et du frottement dans les simulations numériques effectuées.

### 3 Conclusion

Dans ce chapitre, l'effort est porté sur la simulation numérique de l'effet de la forme et du frottement, en utilisant les grains issus de la décomposition de la forme par POD, après maillage. Par des essais de cisaillement simple maintenu menés jusqu'à l'état stationnaire, nous mettons en évidence, avec la description de la forme choisie, la forte dépendance au maillage, au-delà d'une preuve de faisabilité de l'étude détaillée de l'effet de forme. Il est alors important d'avoir une vision prédictive du rapport forme (nuage de points)-maillage-frottement, qui devient plus complexe pour des formes se rapprochant davantage de la forme réelle du ballast ferroviaire.

Afin de maintenir plus de contrôle sur la forme des grains, deux points importants semblent influencer les résultats : le maillage, et la nature convexe ou non convexe des grains. Comme montré dans ce chapitre, le maillage influe significativement sur des indicateurs mécaniques fondamentaux comme l'angle de frottement interne, et il est alors souhaitable de disposer d'un maillage le plus raffiné possible afin de maintenir la différence de forme observée à l'étape des nuages de points. De plus, il est important de remarquer la différence entre l'angularité artificielle créée par le maillage, laquelle, accompagnée par un ajustement du frottement, permet plus facilement de reproduire l'angle de frottement souhaité, contrairement à l'angularité réelle des grains, qui elle, dépend fortement du maillage et dont le lien avec le frottement est moins évident à définir. Enfin, la simplification du cas non convexe vers un cas convexe induit une erreur non négligeable sur la forme, dont la conséquence est de rapprocher des configurations de forme initialement différentes et accélère la convergence des paramètres mécaniques au fur et à mesure du raffinement de la forme.

L'étude réalisée ne permet certes pas de conclure quant à l'effet de la forme réelle du ballast sur ses performances mécaniques, mais ouvre tout de même des perspectives intéressantes quant à la compréhension du lien forme-maillage-frottement pour des grains réalistes, et attire l'attention sur l'intérêt d'améliorer les outils actuels de simulation (maillage et code de calcul non convexe).

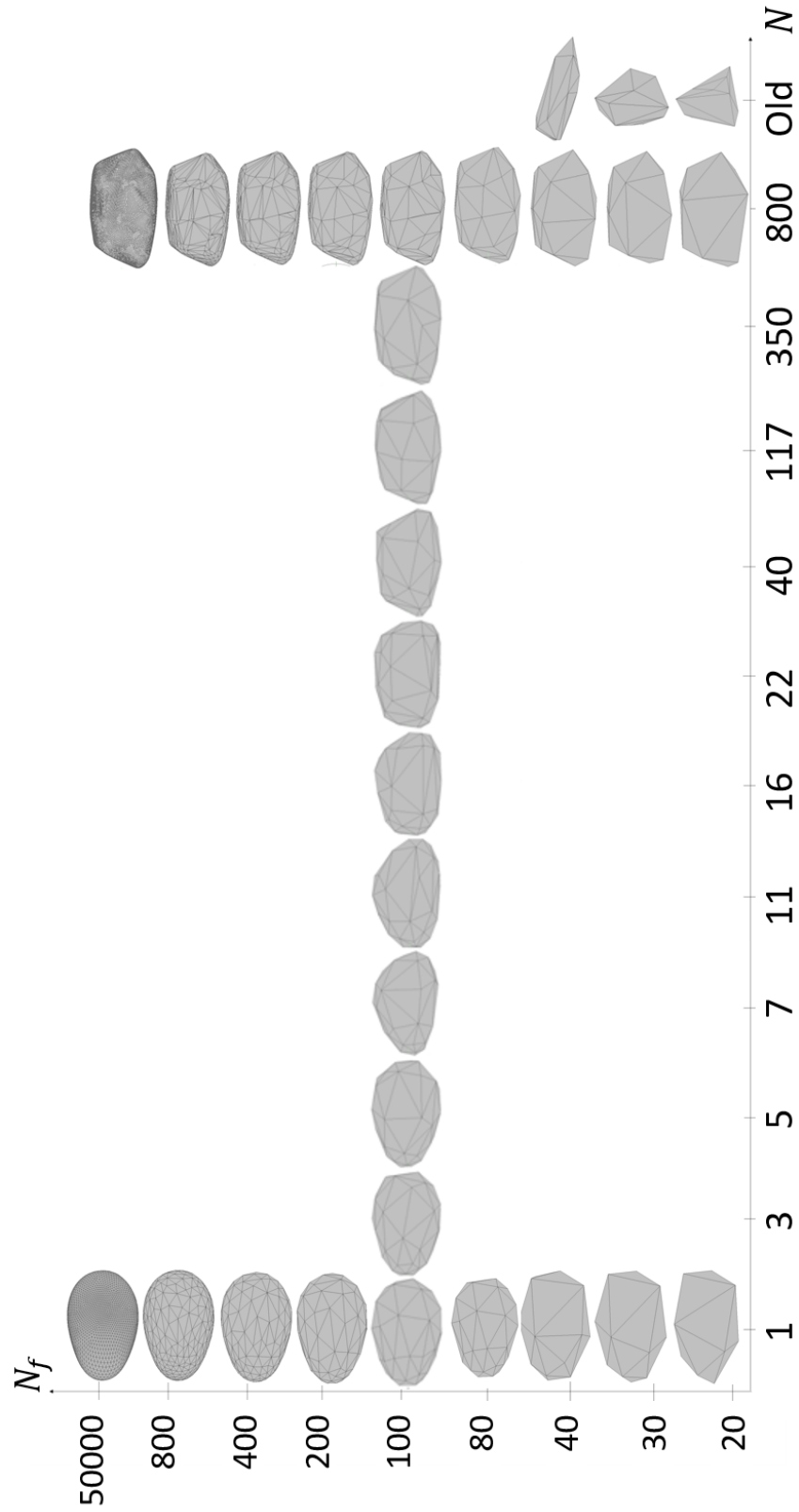


FIGURE 5.16 – Différentes configurations de forme obtenues selon le nombre de modes et le nombre de faces du maillage



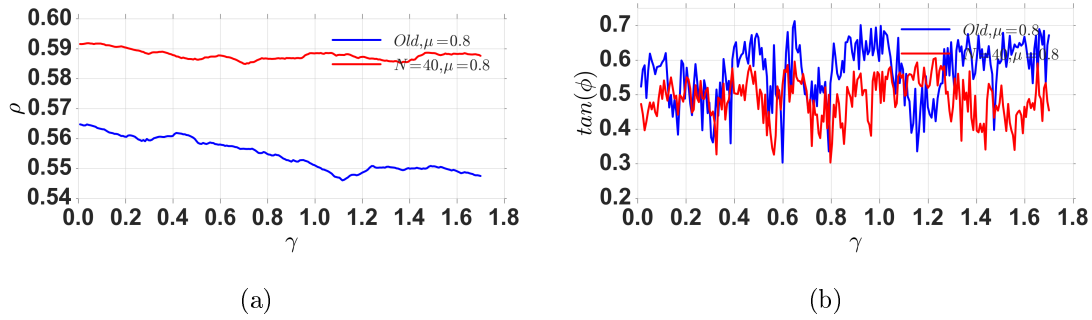


FIGURE 5.17 – Evolution de la compacité  $\rho$  et de l'angle de frottement interne  $\tan \varphi$  pour les grains historiquement utilisés pour modéliser le ballast, en comparaison avec les grains à  $N = 40$  modes pour un coefficient de frottement  $\mu_{gg} = 0.8$  : (a)  $\rho$  ; (b)  $\varphi$ .

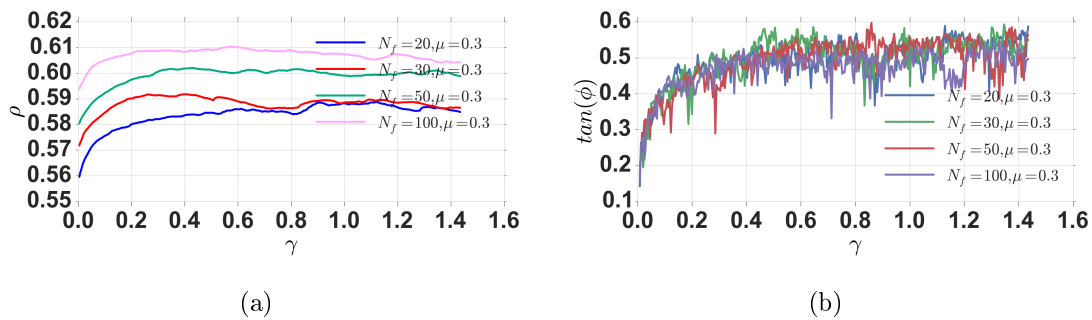


FIGURE 5.18 – Evolution de la compacité  $\rho$  et de l'angle de frottement interne  $\tan \varphi$  pour les grains les grains à  $N = 117$  modes pour un coefficient de frottement  $\mu_{gg} = 0.3$  et différents nombre de faces : (a)  $\rho$  ; (b)  $\varphi$ .

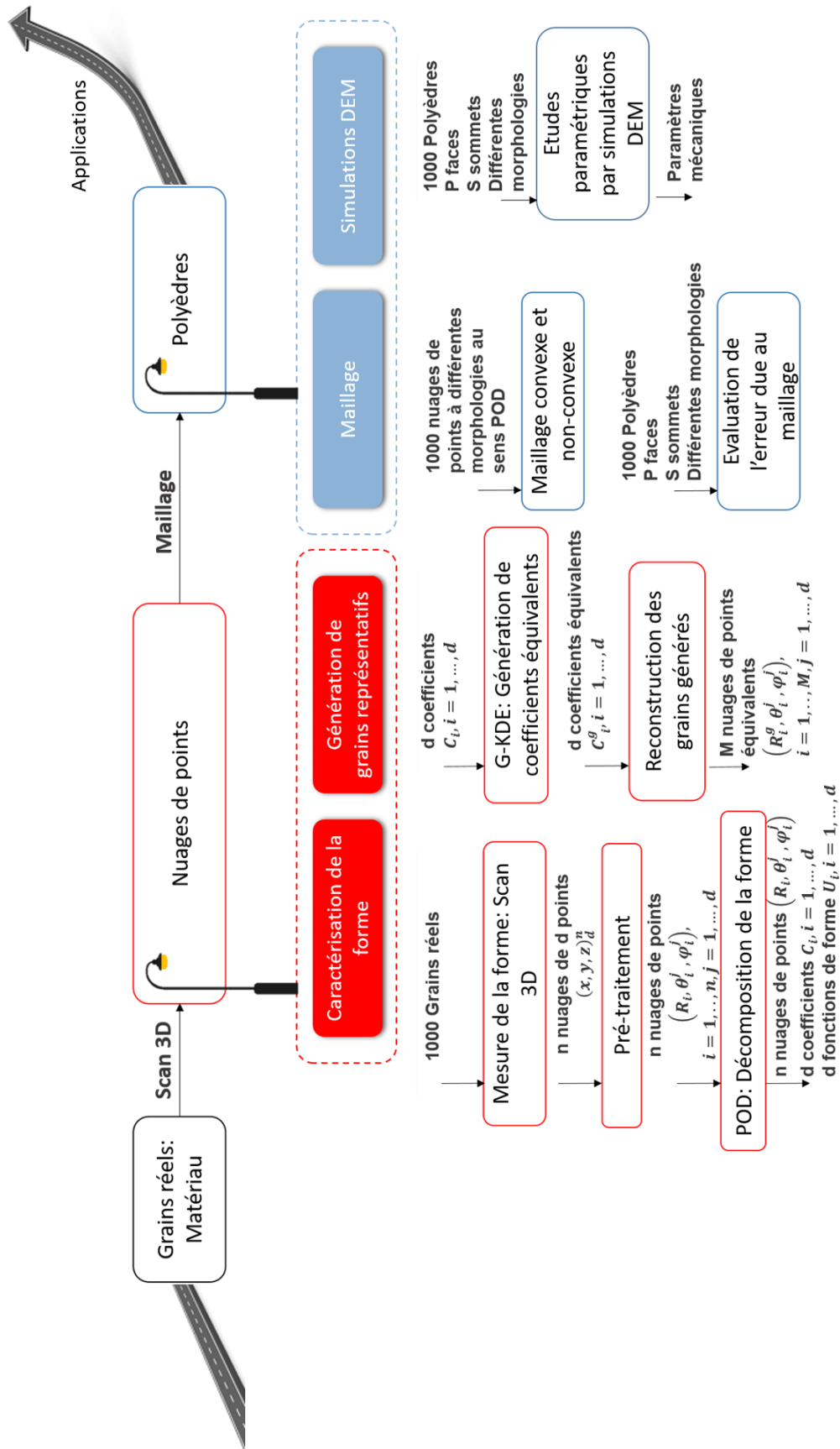


FIGURE 5.19 – Avancement sur le chemin d'étude du ballast



Quatrième partie  
Ballast : Les applications



# Chapitre 6

## Applications

Dans ce dernier chapitre, l'objectif est d'explorer une application des méthodes de caractérisation présentées précédemment. Comme expliqué dans le premier chapitre, l'optimisation de la durée de vie de la couche de ballast sur la voie ferrée est un réel enjeu pour SNCF. Avec les seuils fixés actuellement sur la forme, les opérations de remplacement de ballast entraînent un renouvellement total ou partiel de la couche à l'aide du criblage (tamisage). Si le ballast est considéré comme bon en terme de granulométrie, il est restitué en voie. Dans le cas contraire, il est retiré et remplacé par du ballast neuf. Néanmoins, la nature limitée de tri que permet le criblage est sous optimale. En conséquence, un nombre de grains conformes est retiré, et en même temps, un nombre de grains non conformes est maintenu en voie, la conformité étant définie par la norme. A partir du premier passage de train, le grain neuf devient un grain usé (Fig. 6.1), et le degré d'usure augmente au fur et à mesure des passages de trains. La question à laquelle on se propose de répondre consiste à la possibilité d'améliorer/ compléter l'opération de criblage afin de permettre de discerner des grains conformes ou non conformes avec différents degrés d'usure.



FIGURE 6.1 – Grain de ballast usé (gauche) et neuf (droite).

Nous nous posons alors la question suivante : à partir du scan  $3D$  d'un grain, comment dire avec certitude s'il est usé ? Pour tenter de répondre à cette question, ce chapitre met en évidence la faculté des méthodes de caractérisation à répondre au besoin opérationnel de recyclage du ballast, et illustre les limitations des méthodes actuelles de caractérisation utilisées (analyse granulométrique). Pour une bonne compréhension de l'approche proposée dans ce chapitre, il est important de distinguer la caractérisation de la classification (tri des grains) : dans un premier temps, nous observons les résultats donnés par les différentes approches de caractérisation (distributions statistiques, coefficients POD, etc.) entre des échantillons de nature différente. Dans un deuxième temps, nous évaluons la performance de la classification à partir de ces approches, en se basant sur la notion de vrai positif et la matrice de confusion [161]. Ainsi, pour définir la performance, nous définissons quatre cas de figure : un grain neuf classé comme usé, un grain usé classé comme neuf, un grain neuf classé comme neuf, et un grain usé classé comme usé par l'algorithme de prédiction. Le taux de vrai positif, correspond à la superposition du cas réel et de la prédiction (grain neuf prédit comme neuf ou grain usé prédit comme usé). Le taux de bonnes prédictions par rapport au nombre total de prédictions donne la performance, évaluée tout au long de ce chapitre pour les différentes méthodes de caractérisation. En plus de l'analyse granulométrique (évaluation numérique de la granulométrie par tamisage), nous étudions le potentiel de classification de 3 méthodes de caractérisation : l'évaluation statistique des indicateurs scalaires, l'analyse par POD, et l'analyse de la morphologie avancée (faces/courbures).

Après une présentation succincte du processus de classification, nous tentons de classifier 2 types de données :

1. Populations de grains à différents degrés d'usure.
2. Populations de grains provenant de carrières différentes.

## 1 La classification

Dans notre contexte, la classification est une procédure consistant à attribuer une observation (ou un échantillon) à une catégorie donnée. Il existe trois approches pour la classification : les approches statistiques, la classification par apprentissage automatique [162], ou par réseaux neuronaux. Ces approches au formalisme différent partagent trois principaux objectifs communs. D'abord, la volonté d'égaliser (voire dépasser) le comportement d'un décideur humain en maintenant un niveau satisfaisant de cohérence. Ensuite, la faculté à traiter une grande variété de problèmes et d'être applicables avec succès à différents contextes.

Dans ce travail de thèse, nous utilisons une combinaison des approches statistiques, et de classification par apprentissage automatique. L'utilisation des approches statistiques consiste au calcul et à la comparaison visuelle des distributions statistiques des indicateurs scalaires sur la forme. La classification par apprentissage automatique consiste, elle, soit à attribuer des observations/grains/échantillons à des catégories connues au préalable ou non connues en se basant sur des mesures de similarité ou de distance. Le diagramme de la Fig. 6.2 présente l'ensemble des méthodes utilisées et comparées dans la suite.

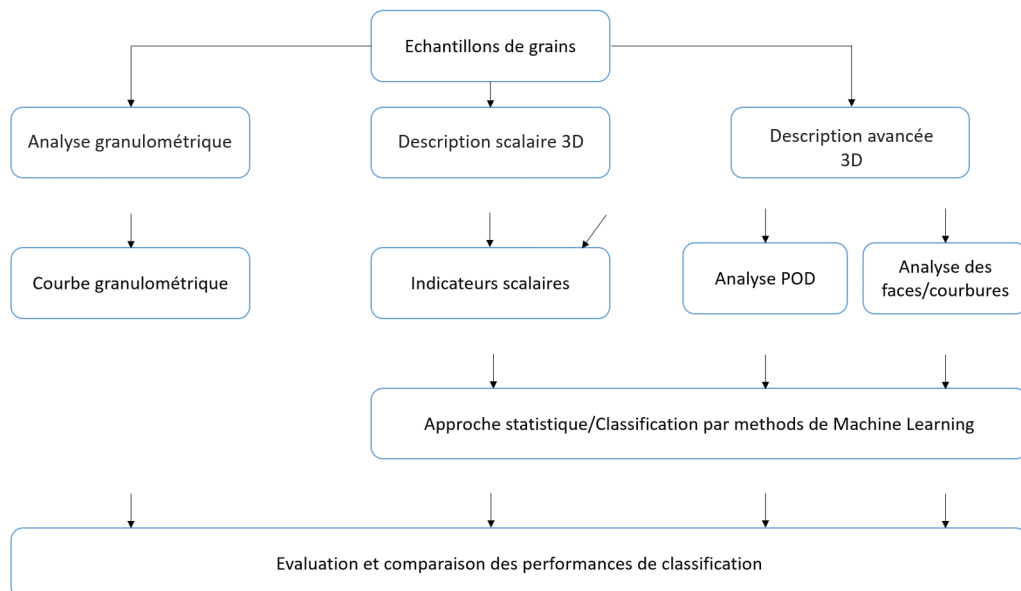


FIGURE 6.2 – 3 méthodes de caractérisation choisies pour la classification des grains

Nous comparons alors 3 approches :

1. L'analyse granulométrique (courbe granulométrique).



2. La description scalaire 3D (surfaces, volumes, rayons moyens).
3. La description avancée 3D (caractérisation par POD, et analyse des faces/courbures).

L'évaluation des performances nécessite la construction de 3 matrices distinctes, relatives aux trois méthodes de caractérisation à comparer : pour la description scalaire, nous construisons une matrice contenant la surface, le volume et le rayon moyen de chaque grain pour tous les échantillons. Ensuite, pour la description avancée par POD, la matrice contient les 100 premiers coefficients POD obtenus après projection sur la base POD à disposition (cf. chapitre 2). Enfin, pour la caractérisation des faces/courbures, la matrice construite contient les valeurs de courbures des 800 sommets et le nombre de grandes faces par grain pour tous les échantillons. Pour l'analyse granulométrique, nous nous limitons à l'analyse qualitative des courbes granulométriques.

Tous les échantillons de grains étudiés sont obtenus à partir du scan 3D. Le cas particulier de la caractérisation par POD nécessite un nombre conséquent de grains (cf. Chapitre 2). Une étude de la convergence des matrices de covariance utilisées pour appliquer la POD, donne une valeur de 400 grains. C'est idéalement le nombre de grains que chacun de nos échantillons doit contenir si on cherche à appliquer une POD. Néanmoins, possédant déjà une base POD grâce à l'échantillon de  $n = 975$  grains neufs (cf. chapitre 2), il suffit de projeter les grains numérisés (quel que soit leur nombre) sur cette base pour récupérer les coefficients POD qu'on souhaite utiliser pour la classification. Le Tab. 6.1 résume le nombre total des grains scannés par catégorie. Le sens des degrés d'usure est expliqué en détail dans la suite.

TABLE 6.1 – Nombre de grains scannés en fonction des catégories

N échantillon	Nature des grains	Nombre de grains
$U_0$	Neufs (usure nulle)	160
$U_1$	usés (Mi-usure)	160
$U_2$	usés (Usure maximale)	160
$Q_2$	Neufs (carrière 1)	300
$Q_1$	Neufs (carrière 2)	300
$Q_2$	Neufs (carrière 3)	300

## 2 Tri de populations de grains de ballast selon le degré d'usure

Pour ce qui est de l'usure, une procédure convenable de tri des grains lors du dégarnissage du ballast (Fig. 6.3) permettrait d'améliorer les opérations de renouvellement de ballast.

Ce tri peut être fait à l'aide de la caractérisation numérique de la forme en comparant des échantillons de grains neufs et usés. Pour discerner des populations de grains à différents degrés d'usure, une première étape consiste à déterminer les degrés d'usure à étudier. Il est possible de prélever des grains usés sur la voie, mais leur degré d'usure ne sera pas connu ni contrôlé. Nous choisissons alors de récupérer des grains usés par Micro Deval [5] (Fig. 6.4). Cette procédure, servant également à déterminer la résistance à l'usure par frottement réciproque des grains de ballast

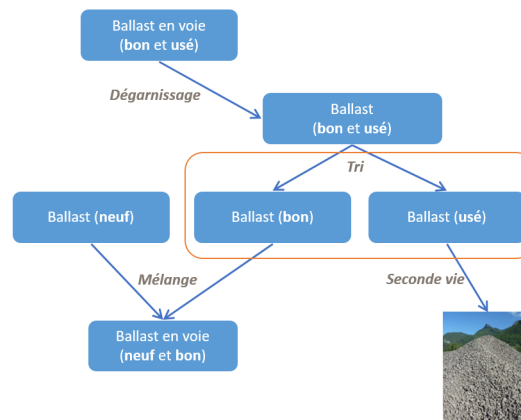


FIGURE 6.3 – Proposition d’une procédure de tri des grains usés en voie

(cf. chapitre 1), consiste à faire tourner  $10Kg$  de grains dans un cylindre creux en acier de longueur  $l = 40cm$  et de diamètre  $D = 20\text{ cm}$  à  $100\text{ tours/min}$  pendant une durée maximale de  $140\text{ min}$ . Les contacts inter-granulaires ainsi qu’avec le cylindre entraînent un processus accéléré d’usure des grains, et engendrent des changements dans leur morphologie. On précise qu’il n’existe actuellement aucun lien établi entre l’usure au Micro-Deval et l’usure sur la voie, puisque les phénomènes physiques mis en jeu dans chacune sont différents. Néanmoins, le test de Micro-Deval reste un essai normatif utilisé pour évaluer la dureté, son but étant (en mesurant la masse des particules fines produites de  $d < 1.6\text{ mm}$ ), de calculer le coefficient de Micro-Deval (MD) [5]. Dans le but de mieux comprendre le mécanisme de l’usure du ballast à l’aide du Micro-Deval, un travail de collaboration est établi avec une étude concentrée sur l’étude de l’usure du ballast, toujours en cours à SNCF [163].



FIGURE 6.4 – Dispositif du test de Micro Deval : appareil avec un seul cylindre. il est possible de faire tourner jusqu’à 4 cylindres simultanément

Pour notre étude, nous nous intéressons davantage à l’évolution de la masse des grains eux-mêmes, puisqu’une perte significative de masse est plus à même de correspondre à un changement de morphologie. La Fig. 6.5 montre l’évolution de la masse des grains pendant  $140\text{ min}$  de test.

Pour nos échantillons, nous choisissons deux degrés d’usure différents, à  $70\text{ min}$  (mi-usure) et à  $140\text{ min}$  (usure maximale) du Micro-Deval, à comparer avec l’échantillon de grains neufs déjà à disposition. La Fig. 6.6 montre en exemple de trois grains

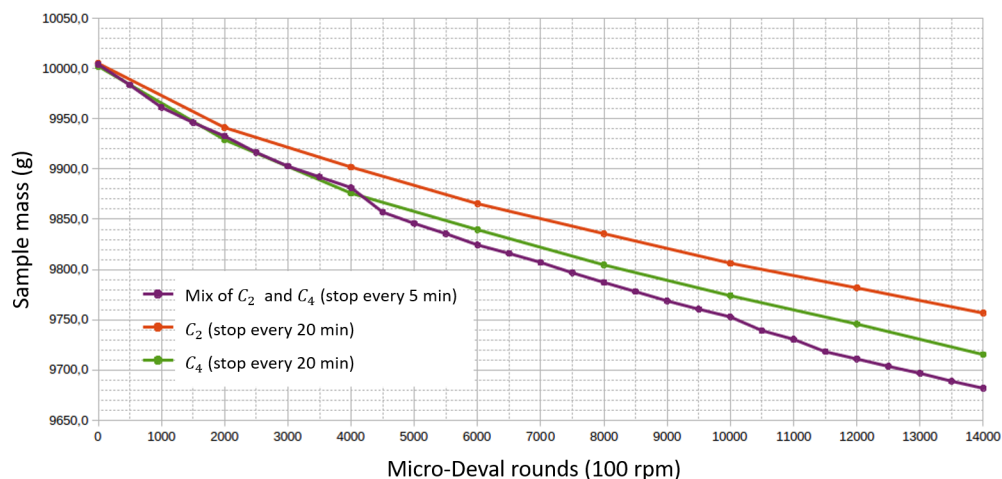


FIGURE 6.5 – Evolution de la masse des grains de ballast : les courbes  $C_2$  et  $C_4$  proviennent de carrières de ballast différentes, et la troisième courbe présente un mélange des deux pour une mesure plus fréquente de la masse

à différents degrés d'usure. Malgré la différence visuelle qui apparaît entre un grain neuf et un grain usé, la difficulté reste d'automatiser cette différenciation sur des critères de morphologie, et de rendre sa quantification précise possible même pour des degrés d'usure faibles et similaires.



FIGURE 6.6 – Exemple de trois grains obtenus par scan 3D de grains à différents degrés d'usure : neuf(a) ; usé à 70 min au (MD) (b) ; usé à 140 min au (MD)

Après une première étape de projection sur la base choisie dans le chapitre 2 et composée de  $d = 800$  directions, nous procédons à la comparaison de ces échantillons.

## 2.1 Méthodes de caractérisation

### a. Indicateurs scalaires

Une première approche consiste à représenter les distributions statistiques de quelques indicateurs classiques pour les différents échantillons. La Fig. 6.7 montre ces distributions pour les 3 échantillons de grains.

La Fig. 6.7(a) montre un déplacement du pic de la distribution de la surface d'environ  $3000 \text{ mm}^2$  entre les grains neufs et les grains à mi-usure. En terme de particule fines, cette perte correspond à environ 150 g de fines sur 10 Kgs de grains de ballast ( $\delta m \approx 600 \text{ g}$  sur 400 grains), donc environ à 1.5% de la masse totale. A partir de la mi-usure jusqu'à l'usure maximale, la probabilité est plus concentrée sur une moyenne de  $S = 4000 \text{ mm}^2$ , ce qui correspond à une masse de fines produites d'environ 80 g sur 10 Kgs de grains (production cumulée d'environ 230 g sur 10 Kgs de ballast), ce qui revient à environ 0.8% de la masse totale des grains, ou d'environ 2.3% au total depuis le début du test. Ces valeurs donnent un coefficient (MD) de 2.3 (Bonne résistance à l'attrition selon la norme).

De manière analogue, nous remarquons un déplacement du pic de la distribution des volumes (Fig. 6.7(b)) d'environ  $25000 \text{ mm}^3$ , puis une stabilisation des valeurs avec une concentration autour du pic pour l'usure maximale. La perte de volume totale est d'environ  $\delta V \approx 2.3 \cdot 10^{-4} \text{ m}^3$ , ce qui est en très bon accord avec la perte totale de masse  $\delta m = \rho \delta V = 2600 \cdot 2.3 \cdot 10^{-4} = 598 \text{ g}$ . Pour ce qui est des rayons moyens (Fig. 6.7(c)), nous constatons un déplacement du pic d'environ 4 mm entre les grains neufs et les grains à mi-usure, puis d'environ 2 mm de la mi-usure à l'usure maximale. Pour les trois mesures de forme, nous remarquons un changement brusque jusqu'à la mi-usure, puis un changement moins significatif à partir de celle-ci.

Cette analyse montre que l'approche statistique sur la morphologie 3D permet de discerner ces populations, jusqu'à un certain point qui sera donné par l'évaluation de sa performance dans la suite, par rapport aux autres approches possibles.

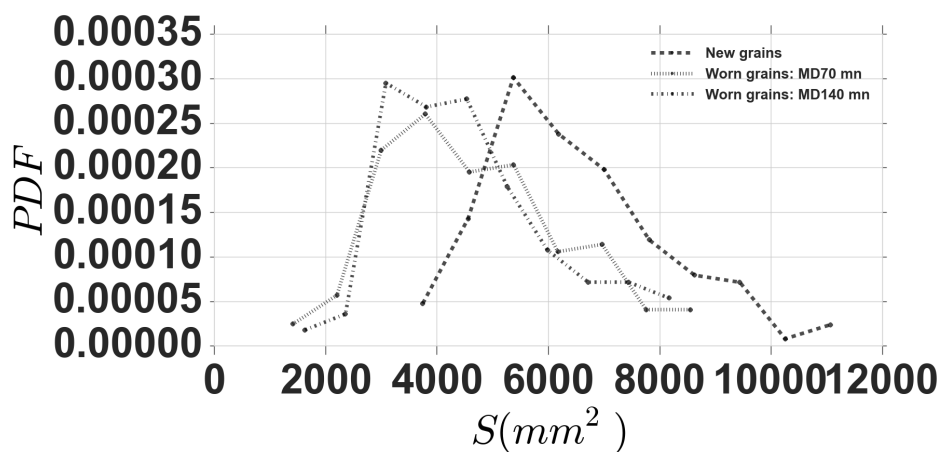
### b. Analyse POD

Avant de nous intéresser à la projection POD qui traite uniquement de l'écart au grain moyen, nous nous intéressons à la comparaison du grain moyen  $R_m^{new}$  des grains neufs à celui des grains usés  $R_m^{MD70}$  et  $R_m^{MD140}$ . La Fig. 6.8 montre le grain moyen extrait des trois échantillons. Visuellement, nous remarquons une diminution marquante de la taille du grain moyen entre l'échantillon neuf et l'échantillon à mi-usure, puis une petite variation entre ce dernier et l'échantillon à usure maximale. Quantitativement, nous avons :

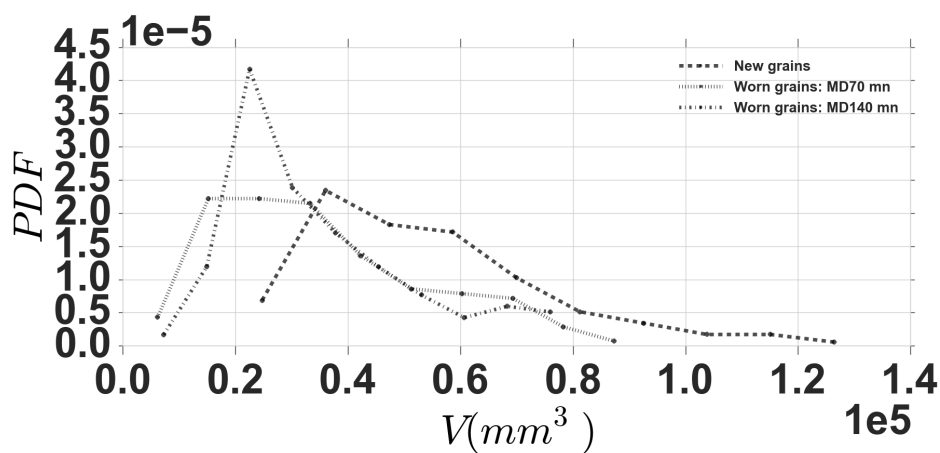
$$100. < \frac{R_m^{MD70}}{R_m^{new}} > \approx 84\% \quad (6.1)$$

et :

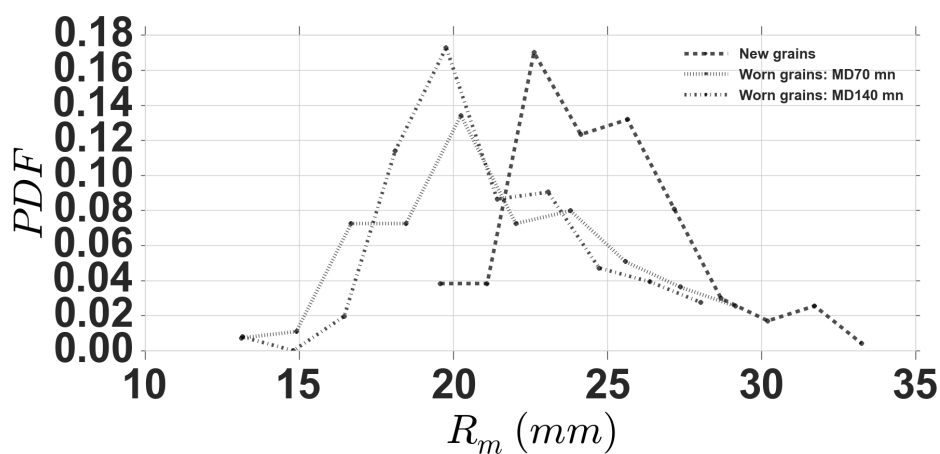
$$100. < \frac{R_m^{MD140}}{R_m^{new}} > \approx 86\% \quad (6.2)$$



(a)



(b)



(c)

FIGURE 6.7 – Distributions statistiques de quelques indicateurs scalaires pour les 3 échantillons de grains usés :(a)  $S$ ; (b)  $V$ ; (c)  $R_m$ .

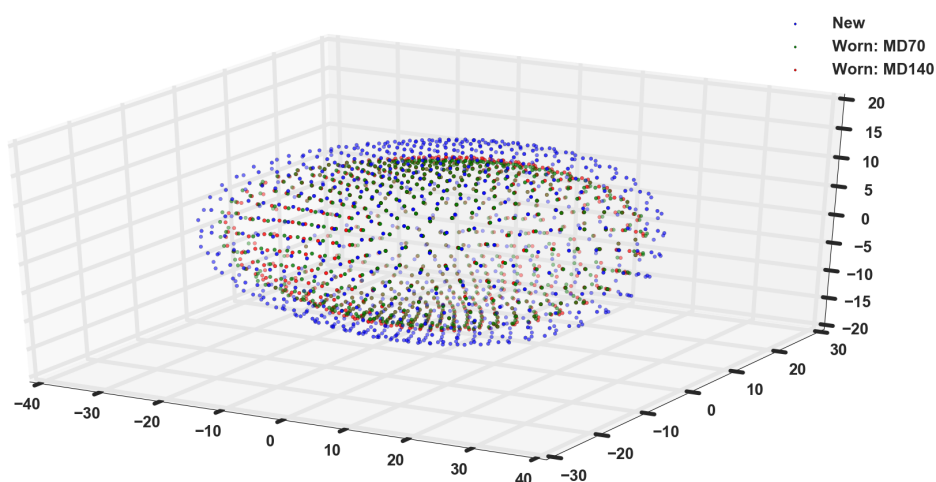


FIGURE 6.8 – Les trois grains moyens extraits des populations de grains à 3 niveaux d’usure en utilisant la base POD.

Par projection de nos échantillons usés sur la base POD obtenue au chapitre 2, on extrait  $d = 800$  coefficients, dont nous représentons les 4 premiers dans la Fig. 6.9.

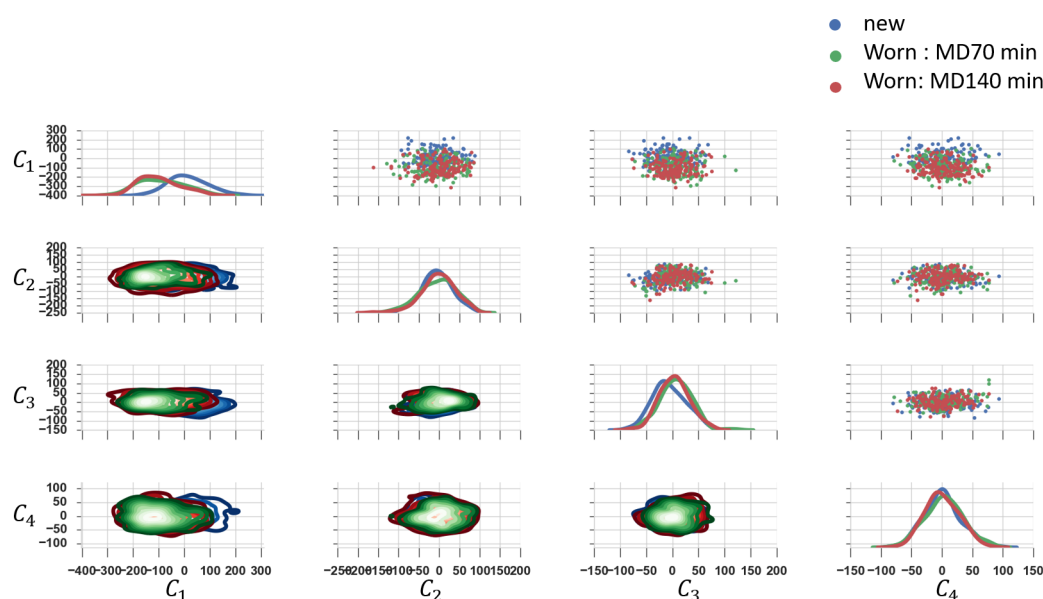


FIGURE 6.9 – Les quatre premiers coefficients  $C_1$ ,  $C_2$ ,  $C_3$  et  $C_4$  extraits des populations de grains à 3 degrés d’usure en utilisant la base POD. Sur la diagonale les distributions statistiques, sur la zone triangulaire supérieure les nuages de points et sur la zone triangulaire inférieure les distributions jointes bivariées.

La Fig. 6.9 montre une distinction visuelle et qualitative possible entre les trois échantillons, en termes de valeur, et distribution des 4 premiers coefficients. En effet, on note un décalage dans la distribution des premiers coefficients avec l’augmentation du degré d’usure (décalage plus marqué des grains neufs aux grains à mi-usure). De

même, pour les distributions jointes du premier coefficient avec les 3 coefficients suivants, une différence est visuellement constatée entre la population des grains neufs et les deux autres populations usées. Ce résultat montre qu'il est alors possible de différencier les échantillons à différents degrés d'usure avec cette méthode de caractérisation. La performance de discernement sera d'autant plus élevée que l'on intègre un plus grand nombre de coefficients dans le processus de classification, car au fur et à mesure de l'augmentation du nombre de modes POD, plus de détails sur la morphologie apparaissent, ce qui augmente les différences entre les grains neufs et usés. Un point important à souligner est que la base POD utilisée pour projeter les populations de grains usés n'est pas optimale puisqu'elle a été établie à partir des grains neufs uniquement. Il serait alors intéressant de la comparer à la base POD obtenue à partir de l'ensemble des grains neufs et usés. Néanmoins, comme nous le verrons par la suite, la base à disposition est suffisante pour fournir une classification satisfaisante. Il reste finalement à comparer quantitativement la performance de cette méthode avec les autres approches.

### c. Analyse des faces/courbures

A partir de la méthode développée dans le chapitre 3 sur la caractérisation des faces et des courbures d'un nuage de points, il est possible de l'appliquer à nos trois échantillons pour détecter la variation de ces indicateurs sur des populations de différents degrés d'usure. L'application de cet algorithme donne les Figs. 6.10, 6.11, 6.12, 6.13.

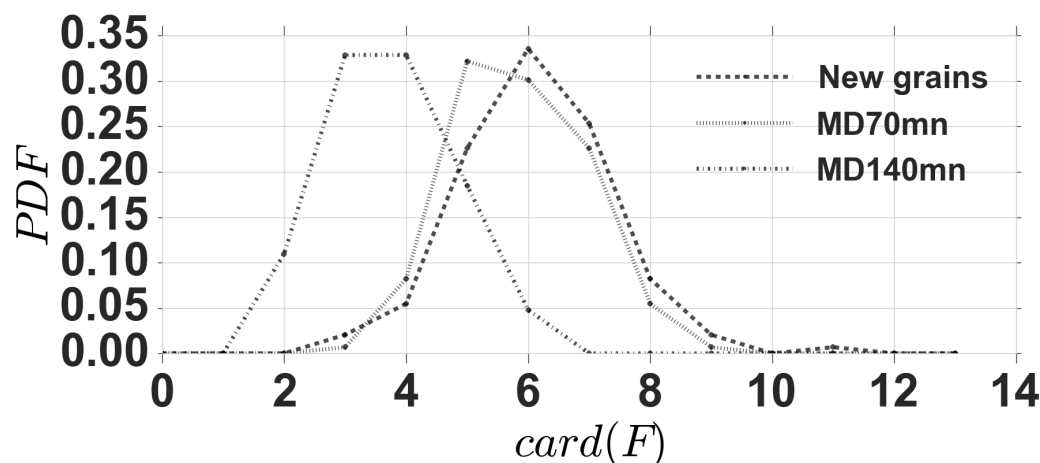


FIGURE 6.10 – Distribution du nombre de grandes faces pour les trois échantillons à degré d'usure différent

La Fig. 6.10 montre l'évolution du nombre de grandes faces (cf. Chapitre 3). Nous observons un déplacement du pic de 6 grandes faces à 5 entre les grains neufs et les grains à mi-usure. Ensuite, un décalage de 5 grandes faces à 3-4 grandes faces à partir de la mi-usure. Cette décroissance du nombre de grandes faces ne montre pas de grande variation dans leurs courbures (Fig. 6.11), mais montre un changement dans les courbures des petites faces. En effet, comme le montre la Fig. 6.10, dans la zone

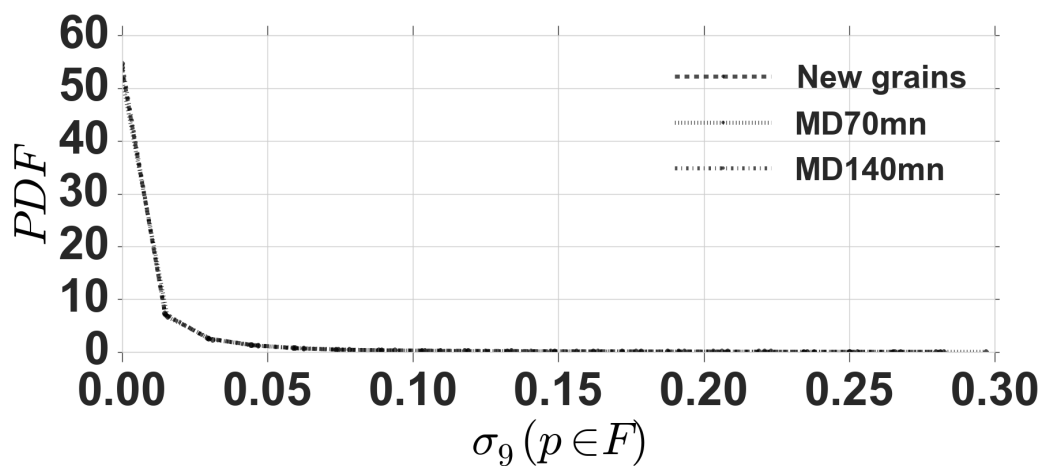


FIGURE 6.11 – Distribution des courbures des grandes faces pour les trois échantillons à degré d'usure différent

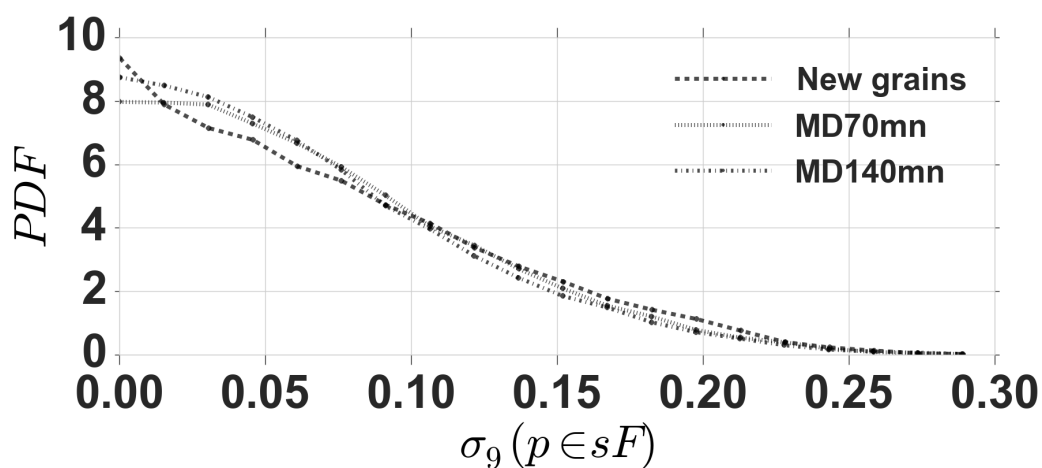


FIGURE 6.12 – Distribution des courbures des petites faces pour les trois échantillons à degré d'usure différent



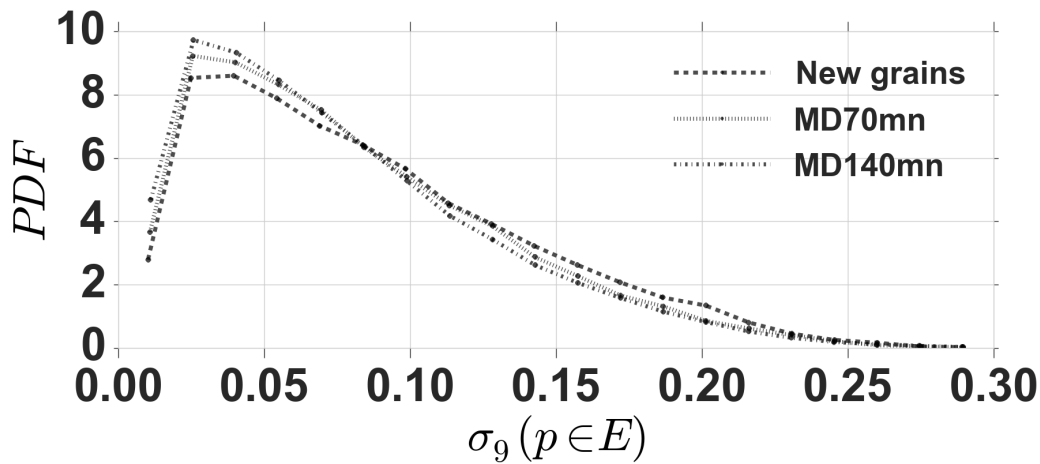


FIGURE 6.13 – Distribution des courbures des arêtes pour les trois échantillons à degré d’usure différent

de courbure entre 0 et 11%, la population des grains usés présente une probabilité plus élevée, ce qui signifie que les grandes faces sont en partie remplacées par des faces plus petites. En revanche, à partir de 11%, les grains neufs présentent une probabilité plus élevée. Ceci est en lien direct avec les courbures élevées des arêtes (Fig. 6.13) qui montrent la même tendance : pour les valeurs les plus faibles (entre 2.5% et 8%), les grains usés présentent des probabilités plus élevées, puis laissent la place aux grains neufs à partir de 12%.

La méthode de caractérisation des faces et des courbures montrent également un potentiel intéressant quant au discernement de populations de grains neufs et usés. Au-delà de la classification, les résultats permettent également d’expliquer un certain mécanisme de l’usure (Fig. 6.14).



FIGURE 6.14 – scan 3D de 5 degrés d’usures du même grain de ballast. De gauche à droite, une usure croissante au Micro Deval : 0 min, 10 min, 30 min, 70 min et 140 min.

Après avoir présenté différentes méthodes de caractérisation permettant la classification, nous nous intéressons maintenant à la comparaison de leur performance à discerner les populations de grains de nature différente.

## 2.2 Comparaison des performances

Afin de comparer les performances des différentes méthodes de caractérisation dans la classification des populations de grains, à commencer par la granulométrie classique, nous créons un nouvel ensemble de grains par usure "numérique" de notre échantillon neuf à l'aide d'un algorithme, afin de créer des échantillons à état d'usure intermédiaire (Fig. 6.15). Le fonctionnement de cet algorithme peut être résumé comme suit :

1. Détermination des  $N$  voisins ( $N=4$ ) de chacun des 800 sommets sur la sphère unitaire discrétisée selon les mêmes directions  $(\theta_j, \varphi_j)$  que les grains.
2. Une itération d'usure numérique d'un grain consiste en :
  - Pour chaque sommet issu de l'itération précédente, on calcule la moyenne des rayons relatifs aux sommet et aux  $N$  voisins.
  - Si cette valeur est inférieure au rayon relatif au sommet en question, elle est attribuée à celui-ci. Sinon, on maintient le rayon initial. De cette manière, à chaque itération, on "rabote" la surface du grain (Fig. 6.15).

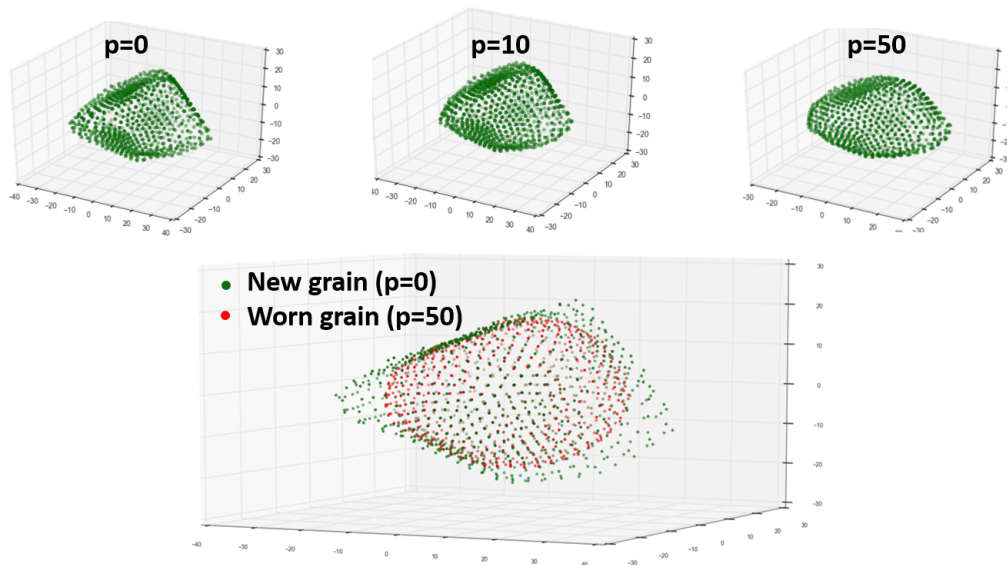


FIGURE 6.15 – Création d'états intermédiaires d'usure en appliquant 50 fois un algorithme d'usure numérique

L'analyse granulométrique de ces échantillons est présentée dans la Fig. 6.16. Nous remarquons que les états d'usure numériques sont tous considérés comme du bon ballast, puisqu'ils appartiennent tous au fuseau de spécification à des tolérances près. D'autre part, les deux courbes granulométriques correspondant à la mi-usure et à l'usure maximale sont presque superposées et sont hors du fuseau de spécification. La classification par granulométrie est alors peu performante pour discerner des grains à différents degrés d'usure, puisqu'elle est basée uniquement sur une taille moyenne des grains.

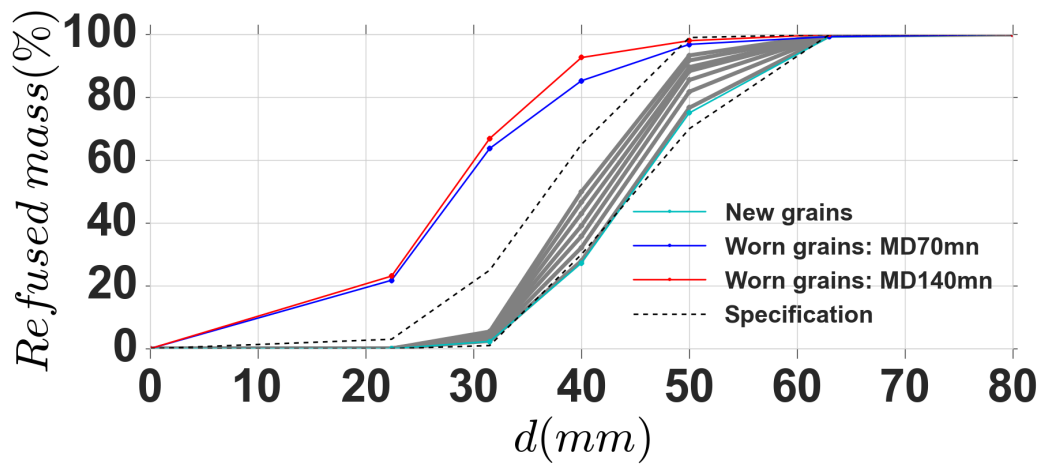


FIGURE 6.16 – Courbes granulométriques correspondant aux différents degrés d'usure : les courbes grises représentent des états intermédiaires d'usure numérique

Nous comparons ensuite la caractérisation par indicateurs scalaires à la caractérisation par POD, et la morphologie avancée des faces/courbures. Pour cela, nous calculons, à l'aide d'un algorithme, le taux de bonne prédiction de l'algorithme avec l'augmentation du degré d'usure. Les résultats sont présentés dans la Fig. 6.17. La classification par POD a été faite à l'aide des 100 premiers coefficients.

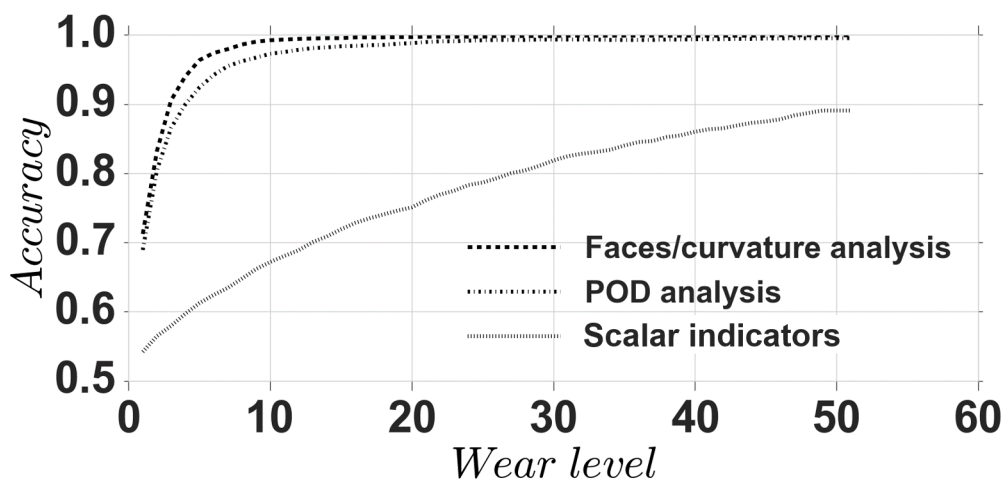


FIGURE 6.17 – Comparaison de la performance de classification des méthodes de caractérisation

Sur la Fig. 6.17, l'axe des abscisses correspond à un degré d'usure : de  $p = 0$  à  $p = 50$ , l'indice des états d'usure intermédiaires à partir du grain neuf, obtenus à l'aide de l'algorithme d'usure numérique. Ensuite  $p = 51$  et  $52$  correspondent respectivement à l'état de mi-usure ( $70 \text{ min}$  au MD) et d'usure maximale ( $140 \text{ min}$  au MD). L'axe des ordonnées correspond au taux de bonne prédiction. Les résultats montrent une performance nettement plus élevée de la caractérisation par POD et de l'analyse des

faces et des courbures d'un nuage de points par rapport aux approches statistiques des indicateurs scalaires 3D, puisqu'elles permettent d'atteindre plus rapidement une performance maximale. La granulométrie, comme montré dans la courbe 6.16 échoue à différencier des états d'usure intermédiaires ou des états d'usure élevés. Le calcul d'indicateurs scalaires simples, quant à lui, ne permet qu'avec une performance inférieure à 0.7, de discerner des grains à degrés d'usure intermédiaires, considérés comme du bon ballast, comme montré par la première phase de la courbe (degré d'usure inférieur à 20).

Dans la suite, nous explorons davantage le potentiel de la POD et par l'analyse des faces/courbures à différencier des grains de ballast neufs provenant de différentes carrières.

### 3 Tri de populations de grains de ballast provenant de différentes carrières

Nous disposons du résultat de scan 3D de 900 grains provenant de 3 différentes carrières (par lot de 300 grains), le but étant d'étudier leurs caractéristiques de forme et de tester la faculté des différentes méthodes de caractérisation à différencier ces échantillons. Nous employons les mêmes approches que pour l'usure, mais l'enjeu ici est différent : une valeur faible de la performance de la classification permettrait de conclure que les grains sont indiscernables et que les fournisseurs des grains produisent un ballast similaire en forme. En revanche, une valeur élevée signifierait probablement un besoin de spécifier davantage d'indications sur la forme dans les normes définissant les caractéristiques des grains neufs.

#### 3.1 Méthodes de caractérisation

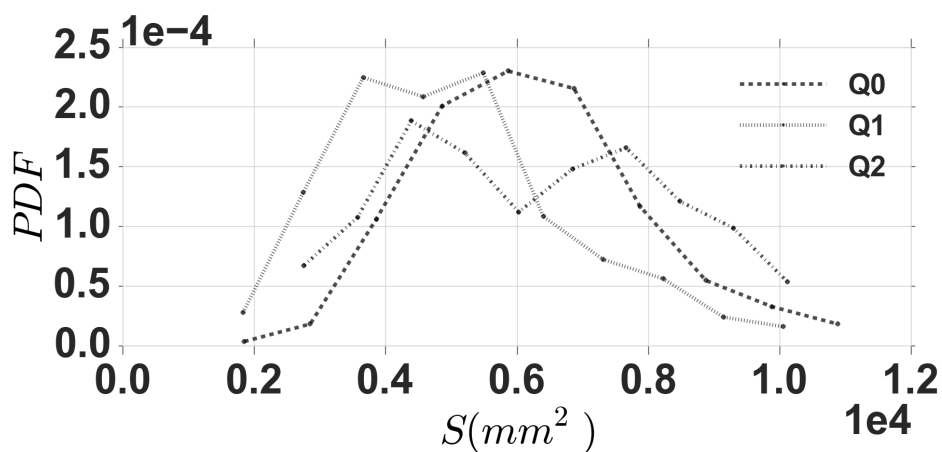
##### a. Indicateurs scalaires

La Figs. 6.18(a), 6.18(b) et 6.18(c) présentent les distributions statistiques des trois échantillons. Comme on est face à trois échantillons de grains neufs, il n'est pas étonnant de voir que les grains des différentes carrières couvrent en partie (ou intégralement) les mêmes spectres de valeurs de surface, de volume ou de rayon moyen.

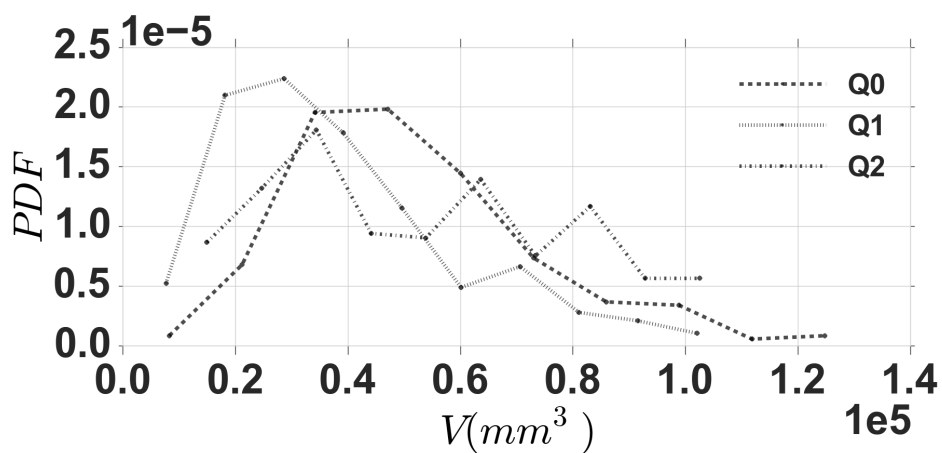
##### b. Analyse POD

De la même manière que pour le cas de l'usure, nous représentons les 4 premiers coefficients POD (Fig. 6.19). Par rapport à la même représentation obtenue dans le cas d'échantillons à différents degrés d'usure, nous observons une certaine zone de recouvrement entre les échantillons au niveau des distributions des marginales et des distributions jointes, avec une légère démarcation de l'échantillon  $Q_1$ , mais généralement, les trois échantillons couvrent les mêmes valeurs de coefficients.

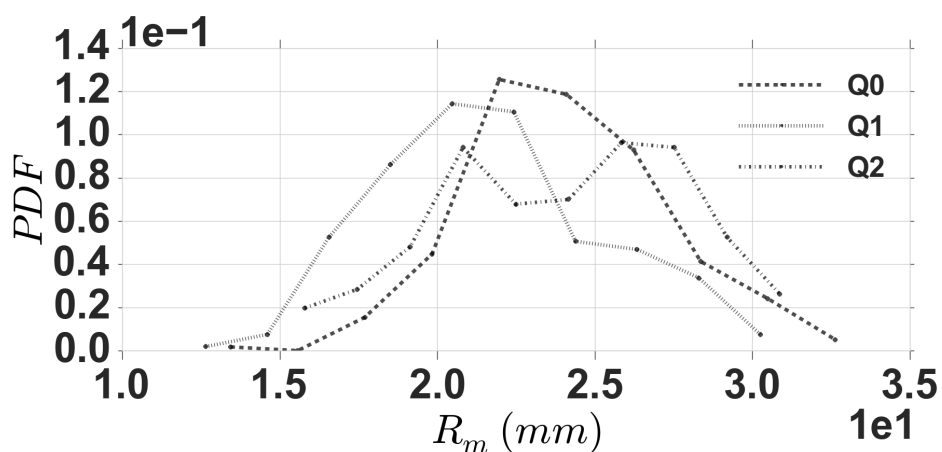
Le résultat de cette analyse montre une performance inférieure à celle donnée par des échantillons à différents degrés d'usure, ce qui est en accord avec les résultats de



(a)



(b)



(c)

FIGURE 6.18 – Distributions statistiques de quelques indicateurs scalaires pour les 3 échantillons de grains usés : (a)  $S$  ; (b)  $V$  ; (c)  $R_m$ .

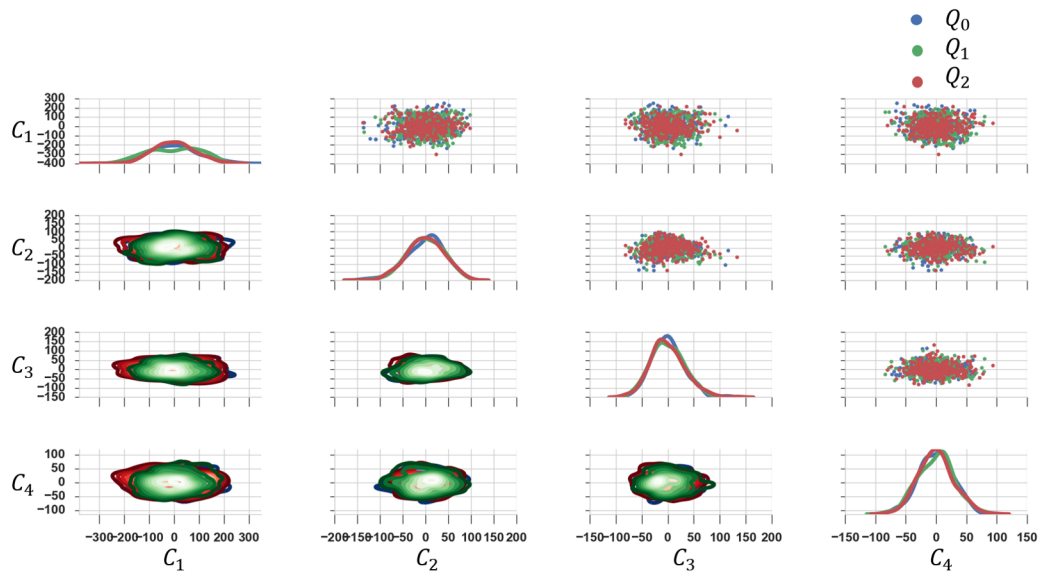


FIGURE 6.19 – Les quatre premiers coefficients  $C_1$ ,  $C_2$ ,  $C_3$  et  $C_4$  extraits des populations de grains provenant de 3 carrières différentes en utilisant la base POD. Sur la diagonale les distributions statistiques, sur la zone triangulaire supérieure les nuages de points et sur la zone triangulaire inférieure les distributions jointes bivariées.

l'analyse granulométrique.

### c. Analyse des faces/courbures

Le résultat de détection de faces/courbures des trois échantillons est présentée dans les Figs. 6.20, 6.21, 6.22 et 6.23.

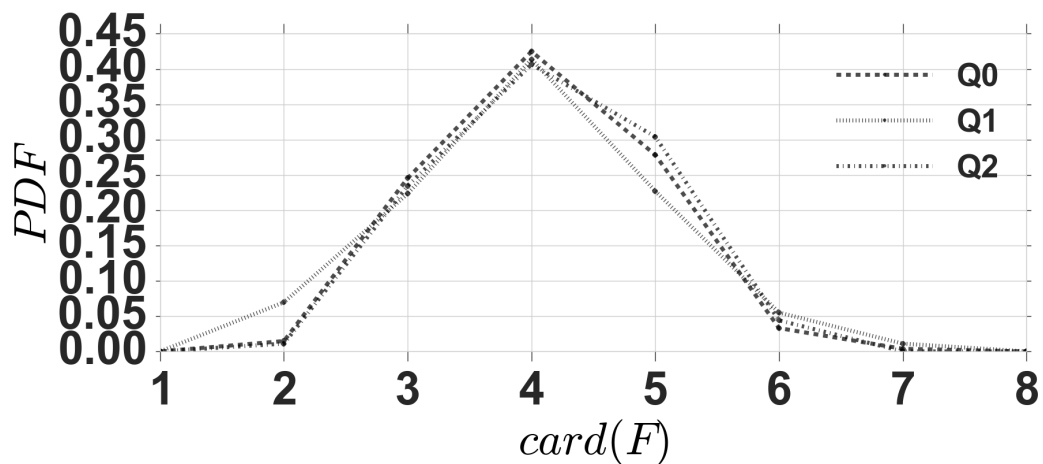


FIGURE 6.20 – Distribution du nombre de grandes faces pour les échantillons provenant de 3 différentes carrières

Les Figs. 6.20, 6.21, 6.22 et 6.23 montrent le résultat de ce calcul pour les trois échantillons. Pour tous les indicateurs, nous remarquons quelques différences entre

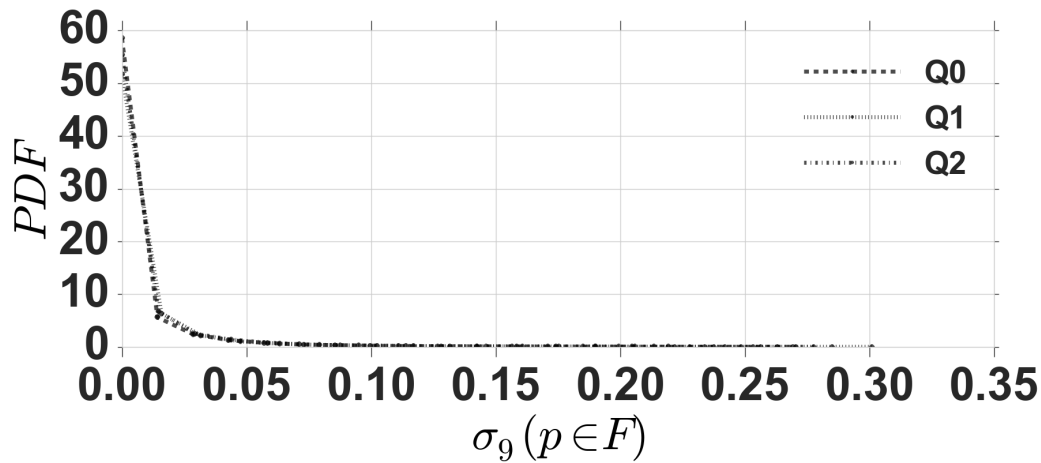


FIGURE 6.21 – Distribution des courbures des grandes faces pour les échantillons provenant de 3 différentes carrières

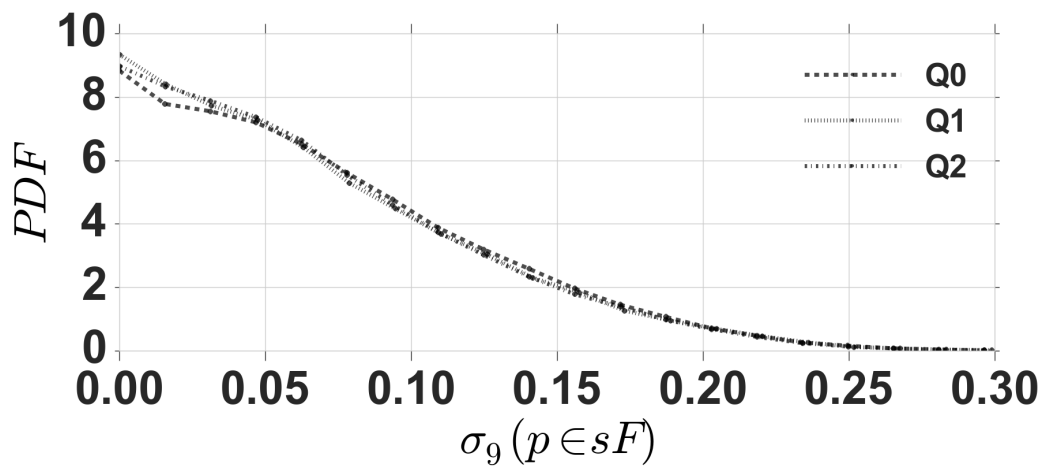


FIGURE 6.22 – Distribution des courbures des petites faces pour les échantillons provenant de 3 différentes carrières

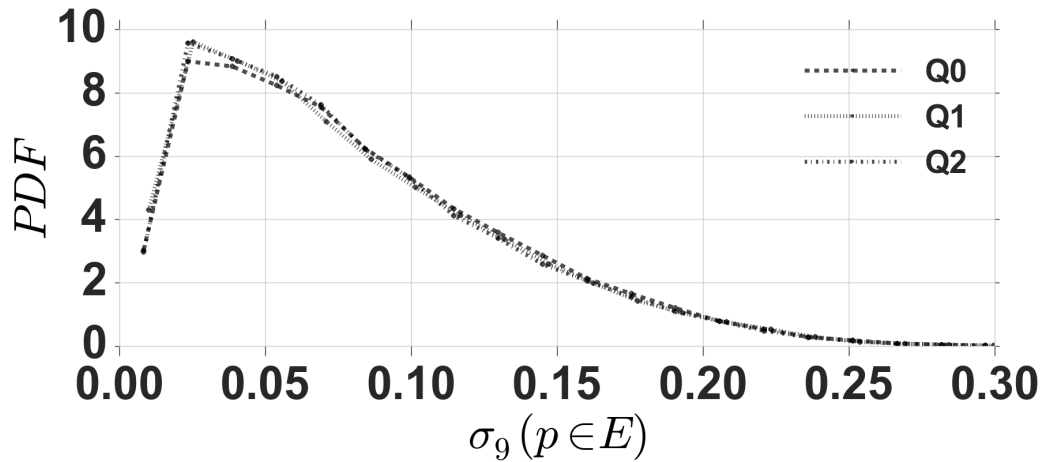


FIGURE 6.23 – Distribution des des courbures des arêtes pour les échantillons provenant de 3 différentes carrières

ces grains, mais qui restent minimales devant les résultats donnés par les grains usés. Il serait alors intéressant de compléter ces résultats par des simulations numériques pour explorer l'équivalence entre ces résultats et le comportement mécanique.

### 3.2 Comparaison des performances

L'analyse granulométrique de ces échantillons est présentée sur la Fig. 6.24. Nous remarquons que les 3 échantillons sont considérés comme du bon ballast (à part un petit dépassement pour les grains de la carrière  $Q_0$ ), puisqu'ils appartiennent tous au fuseau de spécification à des tolérances près. Nous remarquons une petite différence de l'échantillon de la carrière  $Q_1$  par rapport aux deux autres échantillons, mais qui ne permet toujours pas de les discerner. Ce résultat n'est pas étonnant, puisque les trois échantillons sont définis comme du bon ballast par les normes.

Comparons alors les trois méthodes de caractérisation pour ce qui est de leur potentiel de classification. La Fig. 6.25 montre le résultat de cette évaluation.

D'abord, nous constatons une performance inférieure à celle donnée par un mélange d'échantillons neufs et usés. En effet, dans ce cas, nous atteignons une valeur maximale de 0.68, contre une valeur qui approche 1 pour les populations de grains usés. Ensuite, nous remarquons la même tendance que pour la première classification : la meilleure performance est donnée par la caractérisation par faces/courbures, suivie par l'analyse POD, elle-même suivie par la caractérisation par indicateurs scalaires. Enfin, nous constatons que la performance de classification relative à la carrière  $Q_1$  est légèrement plus élevée que pour la carrière  $Q_2$ , ce qui est du à l'observation faite sur la courbe granulométrique de la (Fig. 6.25) montrant une légère démarcation de la carrière  $Q_1$  par rapport aux deux autres. Mais globalement, les grains sont indiscernables avec les outils proposés.



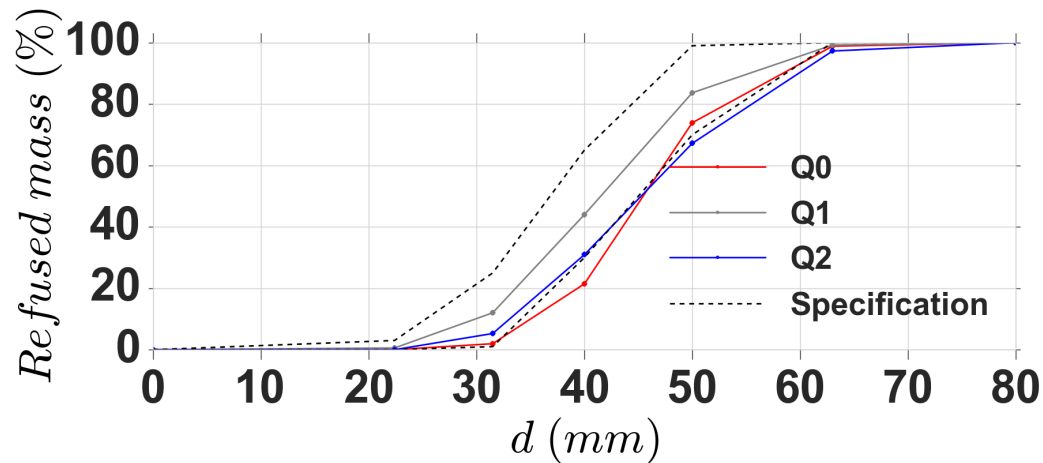


FIGURE 6.24 – Courbes granulométriques correspondant aux trois carrières de ballast

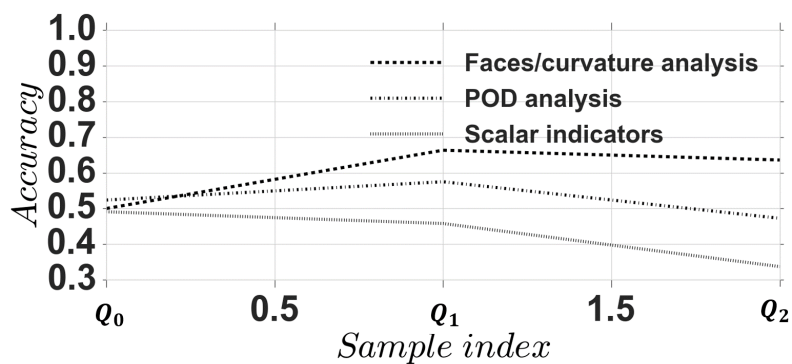


FIGURE 6.25 – Performance de classification évaluée pour les 3 approches de caractérisation dans le cas d'échantillons provenant de trois carrières de ballast.

## 4 Conclusion

Dans ce chapitre, nous nous sommes intéressés à la question du tri des populations de grains sur des critères de morphologie. Pour ce faire, plusieurs méthodes de caractérisation allant du plus classique (granulométrie) au plus complexe (analyse des faces/courbures) ont été testées et comparées dans deux cas de figures : l'usure des grains et leur carrière de provenance. Les résultats montrent un potentiel intéressant des méthodes de caractérisation par POD et de l'analyse des faces/courbures par rapport aux indicateurs scalaires et à la granulométrie classique, quant au discernement de grains à états d'usure similaires. La comparaison d'échantillons de grains provenant de différentes carrières montrent des différences moins notables pour toutes les méthodes de caractérisation, mais ceci ne va pas sans préciser que l'algorithme de classification peut être optimisé, amélioré. Entre la caractérisation par POD, et l'analyse des faces/courbures, cette dernière démontre une meilleure performance. Néanmoins, elle nécessite une plus grande quantité de données (valeurs de courbures des  $d = 800$  sommets ainsi que le nombre de grandes faces), contrairement à la POD qui nécessite uniquement les  $K = 100$  premiers coefficients POD. La classification par POD reste alors la plus performante vu son optimalité.

Pour conclure, un réel potentiel des méthodes de caractérisation développées dans les chapitres précédents est relevé quant à leur capacité à discerner des grains ayant une morphologie différente. Pour les grains provenant de différentes carrières, d'autres améliorations des algorithmes de classification peuvent être envisagées. L'ensemble des résultats obtenus dans ce chapitre, notamment sur l'usure du ballast, ouvre une perspective intéressante sur le recyclage du ballast sur la voie.



# Conclusion

Dans ce travail de thèse, l'effort a été porté sur le développement d'une méthode de caractérisation précise de la forme des grains de ballast, dans l'objectif d'explorer l'effet de la forme sur le comportement mécanique de ce matériau crucial pour la voie ferrée.

Pour répondre à l'ensemble des questions posées sur la forme, la première étape a consisté en une analyse approfondie des différentes approches possibles de caractérisation de la forme, suivie de la proposition d'une méthode innovante de caractérisation basée sur la POD appliquée à un ensemble de grains numérisés obtenus expérimentalement par scan 3D. La méthode proposée a montré de plus grandes capacités que celles fournies par une caractérisation scalaire ou spectrale : à l'aide de la reconstruction contrôlée de la forme, nous disposons de plusieurs ensembles de grains à différents degrés de raffinement sur la forme, permettant une étude paramétrique de l'effet de la forme sur le comportement mécanique. Ensuite, la méthode permet une réduction optimale des données par rapport à la quantité d'information conservée. Enfin, un avantage considérable de la méthode est la simplicité de sa mise en œuvre.

A partir de la description de la morphologie des grains par POD, nous avons proposé une méthode de génération d'ensembles de grains possédant des propriétés de forme équivalentes à notre ensemble de grains scannés. Cette méthode stochastique basée sur la G-KDE (Gaussian Kernel Density Estimation), et appliquée aux coefficients POD, permet de prendre en compte les dépendances complexes entre les différentes directions de représentation des grains de ballast, et de reproduire des grains ayant une dispersion de points statistiquement équivalente. Ce résultat permet d'apporter une solution à un problème non satisfait jusque là dans la littérature. L'approche de génération proposée a démontré une performance satisfaisante en observant des indicateurs simples de la forme, ainsi que des indicateurs plus avancés, dont nous proposons une méthode de description : L'analyse des faces et des courbures. En effet, la plupart des études menées sur la génération de particules 3D montre une limitation quant à la génération de grains facettés. La proposition d'une méthode de caractérisation des grandes faces et des courbures des sommets des grains, nous avons validé la performance de la méthode de génération à générer des grains à la forme voulue.

La simulation numérique discrète nécessite la modélisation des grains par des objets maillés. Une procédure optimisée de maillage a permis de passer de nuages de points à des polyèdres dont le degré de discrétisation est contrôlé. La nature du maillage réalisé (convexe ou non convexe), ainsi que la finesse de celui-ci (nombre de faces) sont particulièrement capitales.

Un ensemble d'études paramétriques sur l'effet de la forme et du frottement

est mené par le biais de simulations numériques discrètes d'essais de cisaillement quasi-statique avec des conditions périodiques. Les résultats de ces études, évalués à l'échelle macroscopique en termes de résistance au cisaillement et de dilatance, ainsi qu'à l'échelle microscopique en terme de connectivité, a permis de mettre en évidence plusieurs points intéressants. D'abord, à toutes les échelles d'analyse, la réponse mécanique est pilotée davantage par le frottement que par la forme. Nous observons également une dépendance plus visible à la forme d'autant plus que le frottement est faible. Une plus grande variation des paramètres mécaniques est remarquée dans les premières valeurs de modes (en dessous de 11 modes), par rapport aux modes supérieurs où l'on remarque une convergence de la plupart des paramètres mécaniques, complètement justifiée par les géométries obtenues après le maillage. Sur les valeurs obtenues, une remarque particulière est que l'on obtient un angle de frottement interne inférieur à celui de l'angle de repos du ballast mesuré expérimentalement. Cet angle de frottement interne est reproduit historiquement par des grains à faible degré de discrétisation avec un ajustement du coefficient de frottement. Le faible nombre de faces dont dispose les grains historiquement utilisés en simulation du ballast crée une angularité artificielle qui réduit la dépendance de l'angle de frottement au coefficient de frottement. Dans notre cas, la complexité de la géométrie, la convexité, et l'angularité se retrouvent de manière couplée dans les maillages, en plus de la finesse de celui-ci. Ceci illustre donc le besoin de mener des études détaillées sur la sensibilité des résultats à chacun de ces paramètres, pour espérer faire une analyse prédictive avec des formes réalistes de grains de ballast.

Au-delà de la preuve de faisabilité de l'étude détaillée de l'effet de forme, notre analyse permet de mettre en évidence l'effet du maillage, et ouvre des portes vers l'amélioration et l'optimisation de cette étape afin de reproduire au mieux les paramètres mécaniques voulus.

Enfin, nous évaluons les performances des différentes méthodes de caractérisation de ballast existantes et proposées, dans le but de discerner des grains neufs et usés, ou encore des grains provenant de différentes carrières de ballast. L'analyse par POD, ainsi que la caractérisation avancées de la forme (faces et courbures) ont donné des résultats satisfaisants, d'autant plus que le degré d'usure est important, en comparaison avec l'analyse granulométrique classique. Par rapport à la POD, l'analyse avancée de la morphologie, malgré une performance plus élevée, nécessite un effort plus significatif de calcul. L'analyse par POD demeure alors une méthode de caractérisation optimale pour la décomposition de la forme, ainsi que pour la classification de populations différentes de grains.

L'ensemble des études menées dans ce travail de thèse suggèrent plusieurs ouvertures, selon "l'état" de représentation des grains : nuages de points, ou polyèdres. En effet, le traitement des nuages de points interroge sur l'éventuelle influence de la taille de la base de projection, ainsi que sur la nature des données sujettes à l'application de la POD, qui peuvent aller des simples coordonnées des sommets, aux valeurs des courbures par sommets. Ce changement pourrait mettre en relief d'autres propriétés communes dominantes au sein de nos échantillons de grains, et permettre une description/un contrôle plus avancé de la forme. Une analyse POD appliquée à la suite du maillage permettrait également un contrôle direct de la morphologie avant la simulation, mais pose le problème de la base commune de projection, qui risque de changer la forme initiale. A l'issue du scan 3D des grains neufs et usés, il se

rait intéressant de construire une base POD à partir de la totalité des grains, afin de couvrir toutes les configurations possibles des grains. Ceci offrirait la possibilité, par la suite, en utilisant la méthode de génération proposée, de générer des ensembles de grains neufs et usés, par catégories d'usure.

Une autre question qui se pose est l'équivalence entre les faces/courbures propres aux grains, et celles obtenues ou créées à la suite du maillage. La possibilité d'évaluer cette différence pourrait nous éclairer sur la notion d'angularité intrinsèque à la géométrie du grain, confrontée à celle créée par le maillage, et permettrait alors de comprendre la contribution de l'une ou l'autre au frottement interne. Sur le plan mécanique, des simulations numériques menées avec des valeurs de frottement intergranulaire plus élevées est à envisager pour nos échantillons de grains, afin d'établir une équivalence forme-frottement dans le même sens que pour les grains historiquement utilisées pour représenter le ballast, mais tout en faisant varier le nombre de faces des grains, qui donnerait une équivalence forme-maillage. La non-convexité constitue également un onglet intéressant à explorer, afin d'évaluer sa contribution aux paramètres mécaniques obtenus, étant donnée l'erreur plus faible retenue sur la forme par rapport au cas convexe. Globalement, un objectif ambitieux mais intéressant consisterait à établir une équivalence forme-frottement-maillage (nature et finesse) pour différentes configurations possibles. La forme pourrait également englober le degré d'usure, en établissant une équivalence usure-nombre de modes correspondant. Enfin, au-delà de la mécanique, les performances obtenues dans la classification des populations de grains par les méthodes de caractérisation développées suggère une mise en œuvre prometteuse d'une procédure de tri des grains, qui permettrait de maximiser l'exploitation des populations de grains pour la voie ferrée lors des opérations de dégarnissage.



# Bibliographie

- [1] F. Lamas-Lopez, *Etude 'in-situ' et en laboratoire sur le comportement dynamique des matériaux constitutifs des plateformes ferroviaires classiques dans le contexte d'augmentation du trafic*. PhD thesis, Université Paris-Est, 2016.
- [2] M. Hosseingholian, D. Levacher, and M. Khay, "Mesure en continu de la raideur dynamique d'une voie ferrée," *Canadian Geotechnical Journal*, vol. 48, no. 3, pp. 439–450, 2011.
- [3] "Granulats pour ballasts de voies ferrées," standard, AFNOR, France, 2003.
- [4] B. Aursudkij, *A laboratory study of railway ballast behaviour under traffic loading and tamping maintenance*. PhD thesis, University of Nottingham, 2007.
- [5] "Essais pour déterminer les caractéristiques mécaniques et physiques des granulats - partie 1 : détermination de la résistance à l'usure (microdeval)," standard, AFNOR, France, 2011.
- [6] "Essais pour déterminer les caractéristiques mécaniques et physiques des granulats - partie 2 : méthodes pour la détermination de la résistance à la fragmentation," standard, AFNOR, France, 2010.
- [7] E. T. Selig and J. M. Waters, *Track geotechnology and substructure management*. Thomas Telford, 1994.
- [8] J. N. Varandas, P. Hölscher, and M. A. Silva, "Settlement of ballasted track under traffic loading : application to transition zones," *Proceedings of the Institution of Mechanical Engineers, Part F : Journal of Rail and Rapid Transit*, vol. 228, no. 3, pp. 242–259, 2014.
- [9] J. Alias and P. Gentil, *La voie ferrée : techniques de construction et d'entretien*. Eyrolles, 1984.
- [10] G. Saussine, *Contribution à la modélisation de granulats tridimensionnels : application au ballast*. PhD thesis, Université Montpellier II-Sciences et Techniques du Languedoc, 2004.
- [11] V. Janoo, "Quantification of shape, angularity, and surface texture of base course materials," tech. rep., COLD REGIONS RESEARCH AND ENGINEERING LAB HANOVER NH, 1998.
- [12] E. Bergliv and T. Edeskär, "Laboratory study on two-dimensional image analysis as a tool to evaluate degradation of granular fill materials," in *Nordic Geotechnical Meeting : Challenges in Nordic Geotechnics 25/05/2016-27/05/2016*, pp. 461–470, 2016.



- [13] L. Le Pen, W. Powrie, A. Zervos, S. Ahmed, and S. Aingaran, "Dependence of shape on particle size for a crushed rock railway ballast," *Granular Matter*, vol. 15, no. 6, pp. 849–861, 2013.
- [14] M. Koohmishi and M. Palassi, "Evaluation of morphological properties of railway ballast particles by image processing method," *Transportation Geotechnics*, vol. 12, pp. 15–25, 2017.
- [15] E. Masad and J. W. Button, "Unified imaging approach for measuring aggregate angularity and texture," *Computer-Aided Civil and Infrastructure Engineering*, vol. 15, no. 4, pp. 273–280, 2000.
- [16] D. Zhang, X. Huang, and Y. Zhao, "Investigation of the shape, size, angularity and surface texture properties of coarse aggregates," *Construction and Building Materials*, vol. 34, pp. 330–336, 2012.
- [17] Y. Descantes, F. Russo, and J. M. Balabaud, "Angularity assessment of several railroad ballast sources using image processing," in *Advanced Characterisation of Pavement and Soil Engineering Materials - Proceedings of the International Conference on Advanced Characterisation of Pavement and Soil Engineering Materials*, vol. 2, pp. 1809–1816, 2007.
- [18] Y. Descantes, Y. Fosse, and F. Milcent, "Automated measurement of railway ballast angularity," *Journal of materials in civil engineering*, vol. 18, no. 4, pp. 612–618, 2006.
- [19] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [20] J. Lee, M. Smith, and L. Smith, "A new approach to the three-dimensional quantification of angularity using image analysis of the size and form of coarse aggregates," *Engineering Geology*, vol. 91, no. 2, pp. 254–264, 2007.
- [21] M. R. Muszynski and S. J. Vitton, "Particle shape estimates of uniform sands : visual and automated methods comparison," *J. Mater. Civ. Eng.*, vol. 24, no. 2, pp. 194–206, 2012.
- [22] F. Altuhafi, C. O'sullivan, and I. Cavarretta, "Analysis of an image-based method to quantify the size and shape of sand particles," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 139, no. 8, pp. 1290–1307, 2012.
- [23] G.-C. Cho, J. Dodds, and J. C. Santamarina, "Particle shape effects on packing density, stiffness, and strength : Natural and crushed sands," *Journal of geotechnical and geoenvironmental engineering*, vol. 132, no. 5, pp. 591–602, 2006.
- [24] P. C. Rousé, R. J. Fannin, and D. A. Shuttle, "Influence of roundness on the void ratio and strength of uniform sand," *Géotechnique*, vol. 58, no. 3, pp. 227–231, 2008.
- [25] J. Fonseca, C. O'Sullivan, M. R. Coop, and P. Lee, "Non-invasive characterization of particle morphology of natural sands," *Soils and Foundations*, vol. 52, no. 4, pp. 712–722, 2012.
- [26] Y. Sun, B. Indraratna, and S. Nimbalkar Dr, "Three-dimensional characterisation of particle size and shape for ballast," 2014.

- [27] B. Zhou and J. Wang, "Random generation of natural sand assembly using micro x-ray tomography and spherical harmonics," *Géotechnique Letters*, vol. 5, no. 1, pp. 6–11, 2015.
- [28] J. Anochie-Boateng, J. J. Komba, N. Makulenga, and A. Maharaj, "Evaluation of 3d laser device for characterizing shape and surface properties of aggregates used in pavements," 2010.
- [29] B. Zhou, J. Wang, and B. Zhao, "Micromorphology characterization and reconstruction of sand particles using micro x-ray tomography and spherical harmonics," *Engineering Geology*, vol. 184, pp. 126–137, 2015.
- [30] E. J. Garboczi, "Three-dimensional mathematical analysis of particle shape using x-ray tomography and spherical harmonics : Application to aggregates used in concrete," *Cement and concrete research*, vol. 32, no. 10, pp. 1621–1638, 2002.
- [31] J. M. Fernlund, "The effect of particle form on sieve analysis : a test by image analysis," *Engineering Geology*, vol. 50, no. 1, pp. 111–124, 1998.
- [32] C. K. Ballantyne, "Aggregate clast form characteristics of deposits near the margins of four glaciers in the jotunheimen massif, norway," 1982.
- [33] P. Barrett, "The shape of rock particles, a critical review," *Sedimentology*, vol. 27, no. 3, pp. 291–303, 1980.
- [34] H. Wadell, "Volume, shape, and roundness of quartz particles," *The Journal of Geology*, vol. 43, no. 3, pp. 250–280, 1935.
- [35] A. G. Athanassiadis, M. Z. Miskin, P. Kaplan, N. Rodenberg, S. H. Lee, J. Merritt, E. Brown, J. Amend, H. Lipson, and H. M. Jaeger, "Particle shape effects on the stress response of granular packings," *Soft Matter*, vol. 10, no. 1, pp. 48–59, 2014.
- [36] T. Börzsönyi, E. Somfai, B. Szabó, S. Wegner, A. Ashour, and R. Stannarius, "Elongated grains in a hopper," in *EPJ Web of Conferences*, vol. 140, p. 06017, EDP Sciences, 2017.
- [37] F. Radjaï and F. Dubois, *Discrete-element modeling of granular materials*. Wiley-Iste, 2011.
- [38] J.-C. Quezada, *Mécanismes de tassement du ballast et sa variabilité*. PhD thesis, Université Montpellier II-Sciences et Techniques du Languedoc, 2012.
- [39] M. Lu and G. McDowell, "The importance of modelling ballast particle shape in the discrete element method," *Granular matter*, vol. 9, no. 1-2, p. 69, 2007.
- [40] E. Azéma, F. Radjai, and G. Saussine, "Quasistatic rheology, force transmission and fabric properties of a packing of irregular polyhedral particles," *Mechanics of Materials*, vol. 41, no. 6, pp. 729–741, 2009.
- [41] E. Azéma, *Etude numérique de milieux granulaires à grains polyédriques : rhéologie quasi-statique, dynamique vibratoire et application au procédé de bourrage du ballast*. PhD thesis, Université Montpellier II-Sciences et Techniques du Languedoc, 2007.
- [42] E. Azéma and F. Radjaï, "Stress-strain behavior and geometrical properties of packings of elongated particles," *Physical Review E*, vol. 81, no. 5, p. 051304, 2010.

- [43] N. Estrada, E. Azéma, F. Radjai, and A. Taboada, “Identification of rolling resistance as a shape parameter in sheared granular media,” *Physical Review E*, vol. 84, no. 1, p. 011306, 2011.
- [44] E. Azéma, N. Estrada, and F. Radjai, “Nonlinear effects of particle shape angularity in sheared granular media,” *Physical Review E*, vol. 86, no. 4, p. 041301, 2012.
- [45] E. Azéma, F. Radjai, and F. Dubois, “Packings of irregular polyhedral particles : Strength, structure, and effects of angularity,” *Physical Review E*, vol. 87, no. 6, p. 062203, 2013.
- [46] É. Azéma, F. Radjai, and F. Dubois, “Numerical simulation of granular media composed with irregular polyhedral particles : effect of particles’ angularity,” in *II International Conference on Particle-based Methods-Fundamentals and Applications-PARTICLES 2011*, 2011.
- [47] E. Azéma, F. Radjai, B. Saint-Cyr, J.-Y. Delenne, and P. Sornay, “Rheology of three-dimensional packings of aggregates : Microstructure and effects of nonconvexity,” *Physical Review E*, vol. 87, no. 5, p. 052205, 2013.
- [48] G. M. Mvelase, P. Gräbe, and J. K. Anochie-Boateng, “The use of laser technology to investigate the effect of railway ballast roundness on shear strength,” *Transportation Geotechnics*, vol. 11, pp. 97–106, 2017.
- [49] N. Das, *Modeling three-dimensional shape of sand grains using discrete element method*. University of South Florida, 2007.
- [50] C. K. Wentworth, “The shapes of rock particles : a discussion,” *The Journal of Geology*, vol. 41, no. 3, pp. 306–309, 1933.
- [51] W. C. Krumbein, “Measurement and geological significance of shape and roundness of sedimentary particles,” *Journal of Sedimentary Research*, vol. 11, no. 2, 1941.
- [52] S. J. Blott and K. Pye, “Particle shape : a review and new methods of characterization and classification,” *Sedimentology*, vol. 55, no. 1, pp. 31–63, 2008.
- [53] H. Wadell, “Sphericity and roundness of rock particles,” *The Journal of Geology*, vol. 41, no. 3, pp. 310–331, 1933.
- [54] G. Lees, “A new method for determining the angularity of particles,” *Sedimentology*, vol. 3, no. 1, pp. 2–21, 1964.
- [55] B. C. Aschenbrenner, “A new method of expressing particle sphericity,” *Journal of Sedimentary Research*, vol. 26, no. 1, 1956.
- [56] D. I. Benn and C. K. Ballantyne, “The description and representation of particle shape,” *Earth Surface Processes and Landforms*, vol. 18, no. 7, pp. 665–672, 1993.
- [57] E. D. Sneed and R. L. Folk, “Pebbles in the lower colorado river, texas a study in particle morphogenesis,” *The Journal of Geology*, vol. 66, no. 2, pp. 114–150, 1958.
- [58] W. K. Illenberger, “Pebble shape (and size!),” *Journal of Sedimentary Research*, vol. 61, no. 5, 1991.

- [59] W. K. Illenberger and J. S. Reddering, "An evaluation of shape indices as palaeoenvironmental indicators using quartzite and metavolcanic clasts in upper cretaceous to palaeogene beach, river and submarine fan conglomerates," *Sedimentology*, vol. 40, no. 5, pp. 1019–1020, 1993.
- [60] J. L. Howard, "An evaluation of shape indices as palaeoenvironmental indicators using quartzite and metavolcanic clasts in upper cretaceous to palaeogene beach, river and submarine fan conglomerates," *Sedimentology*, vol. 39, no. 3, pp. 471–486, 1992.
- [61] R. J. Oakey, M. Green, P. A. Carling, M. W. Lee, D. A. Sear, and J. Warburton, "Grain-shape analysis—a new method for determining representative particle shapes for populations of natural grains," *Journal of Sedimentary Research*, vol. 75, no. 6, pp. 1065–1073, 2005.
- [62] G. Bagheri, C. Bonadonna, I. Manzella, and P. Vonlanthen, "On the characterization of size and shape of irregular particles," *Powder Technology*, vol. 270, pp. 141–153, 2015.
- [63] H. Wadell, "Volume, shape, and roundness of rock particles," *The Journal of Geology*, vol. 40, no. 5, pp. 443–451, 1932.
- [64] J.-F. Ferrellec and G. R. McDowell, "A method to model realistic particle shape and inertia in dem," *Granular Matter*, vol. 12, no. 5, pp. 459–467, 2010.
- [65] M. Frederick, "Notes on the shape of particles and its influence on the properties of sands," *Proc. of the Midland Soil Mechanics and Foundation Engineering Society*, pp. 157–162, 1961.
- [66] B. Sukumaran and A. Ashmawy, "Quantitative characterisation of the geometry of discret particles," *Geotechnique*, vol. 51, no. 7, pp. 619–627, 2001.
- [67] C. Rao, E. Tutumluer, and I. T. Kim, "Quantification of coarse aggregate angularity based on image analysis," *Transportation Research Record : Journal of the Transportation Research Board*, no. 1787, pp. 117–124, 2002.
- [68] B. B. Mandelbrot, *Fractals*. Wiley Online Library, 1977.
- [69] G. Mollon and J. Zhao, "Fourier–voronoi-based generation of realistic samples for discrete modelling of granular materials," *Granular Matter*, vol. 14, no. 5, pp. 621–638, 2012.
- [70] R. Ehrlich and B. Weinberg, "An exact method for characterization of grain shape," *Journal of sedimentary research*, vol. 40, no. 1, 1970.
- [71] T. Meloy, "Fast fourier transforms applied to shape analysis of particle silhouettes to obtain morphological data," *Powder Technology*, vol. 17, no. 1, pp. 27–35, 1977.
- [72] E. T. Bowman, K. Soga, and T. W. Drummond, *Particle shape characterisation using Fourier analysis*. University of Cambridge, Department of Engineering, 2000.
- [73] G. Mollon and J. Zhao, "Generating realistic 3d sand particles using fourier descriptors," *Granular Matter*, vol. 15, no. 1, pp. 95–108, 2013.
- [74] G. Mollon and J. Zhao, "3d generation of realistic granular samples based on random fields theory and fourier shape descriptors," *Computer Methods in Applied Mechanics and Engineering*, vol. 279, pp. 46–65, 2014.

- [75] M. A. Taylor, E. J. Garboczi, S. Erdogan, and D. Fowler, “Some properties of irregular 3-d particles,” *Powder Technology*, vol. 162, no. 1, pp. 1–15, 2006.
- [76] L. Shen, H. Farid, and M. A. McPeck, “Modeling three-dimensional morphological structures using spherical harmonics,” *Evolution*, vol. 63, no. 4, pp. 1003–1016, 2009.
- [77] C. Brechbühler, G. Gerig, and O. Kübler, “Parametrization of closed surfaces for 3-d shape description,” *Computer vision and image understanding*, vol. 61, no. 2, pp. 154–170, 1995.
- [78] Wikipédia, “Corrélation (statistiques) — wikipédia, l’encyclopédie libre,” 2017.
- [79] A. C. Antoulas, D. C. Sorensen, and S. Gugercin, “A survey of model reduction methods for large-scale systems,” *Contemporary mathematics*, vol. 280, pp. 193–220, 2001.
- [80] A. J. Newman, “Model reduction via the karhunen-loeve expansion part i : An exposition,” tech. rep., 1996.
- [81] H. Hotelling, “Analysis of a complex of statistical variables into principal components.,” *Journal of educational psychology*, vol. 24, no. 6, p. 417, 1933.
- [82] I. T. Jolliffe, “Principal component analysis and factor analysis,” in *Principal component analysis*, pp. 115–128, Springer, 1986.
- [83] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, “Singular value decomposition and principal component analysis,” in *A practical approach to microarray data analysis*, pp. 91–109, Springer, 2003.
- [84] Y. Liang, H. Lee, S. Lim, W. Lin, K. Lee, and C. Wu, “Proper orthogonal decomposition and its applications—part i : Theory,” *Journal of Sound and vibration*, vol. 252, no. 3, pp. 527–544, 2002.
- [85] Wikipedia, “Machine learning — wikipedia, the free encyclopedia,” 2017.
- [86] J. Shlens, “A tutorial on principal component analysis derivation,” *Discussion and Singular Value Decomposition*, vol. 25, 2003.
- [87] S. Volkwein, “Proper orthogonal decomposition : Theory and reduced-order modelling. university of konstanz, 2013. lecture notes.”
- [88] S. Volkwein, “Proper orthogonal decomposition : Theory and reduced-order modelling,” *Lecture Notes, University of Konstanz*, vol. 4, no. 4, 2013.
- [89] J. Holmes, “Lumley, and g. berkooz,” *Turbulence, coherent structures, dynamical systems and symmetry*, 1996.
- [90] B. Feeny and R. Kappagantu, “On the physical interpretation of proper orthogonal modes in vibrations,” *Journal of sound and vibration*, vol. 211, no. 4, pp. 607–616, 1998.
- [91] S. Mika, B. Schölkopf, A. J. Smola, K.-R. Müller, M. Scholz, and G. Rätsch, “Kernel pca and de-noising in feature spaces,” in *Advances in neural information processing systems*, pp. 536–542, 1999.
- [92] R. Bellman, *Dynamic programming*. Courier Corporation, 2013.
- [93] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, “Scikit-learn : Machine learning in python,” *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

- [94] Wikipedia, “3d scanner — wikipedia, the free encyclopedia,” 2017.
- [95] G. Turk, “Re-tiling polygonal surfaces,” *ACM SIGGRAPH Computer Graphics*, vol. 26, no. 2, pp. 55–64, 1992.
- [96] S. Geisser, *Predictive inference*, vol. 55. CRC press, 1993.
- [97] R. Kohavi *et al.*, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Ijcai*, vol. 14, pp. 1137–1145, Stanford, CA, 1995.
- [98] Wikipedia, “Cross-validation (statistics) — wikipedia, the free encyclopedia,” 2017.
- [99] G. Monari, *Sélection de modèles non linéaires par "leave-one-out" : étude théorique et application des réseaux de neurones au procédé de soudage par points*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 1999.
- [100] Wikipedia, “Binomial coefficient — wikipedia, the free encyclopedia,” 2017.
- [101] Z. Mahmood and K. Iwashita, “Influence of inherent anisotropy on mechanical behavior of granular materials based on dem simulations,” *International journal for numerical and analytical methods in geomechanics*, vol. 34, no. 8, pp. 795–819, 2010.
- [102] E. Azéma and F. Radjaï, “Force chains and contact network topology in sheared packings of elongated particles,” *Physical review E*, vol. 85, no. 3, p. 031303, 2012.
- [103] W. Zhou, K. Xu, L. Yang, and G. Ma, “Influence of particle shape on mechanical behavior of granular materials,” in *Proceedings of the 7th International Conference on Discrete Element Methods*, pp. 245–252, Springer, 2017.
- [104] A. Hasan and K. Alshibli, “Three dimensional fabric evolution of sheared sand,” *Granular Matter*, pp. 1–14, 2012.
- [105] J. Ferrellec and G. McDowell, “Modelling realistic shape and particle inertia in dem,” *Géotechnique*, vol. 60, no. 3, pp. 227–232, 2010.
- [106] M. Grigoriu, E. Garboczi, and C. Kafali, “Spherical harmonic-based random fields for aggregates used in concrete,” *Powder Technology*, vol. 166, no. 3, pp. 123–138, 2006.
- [107] T. Gruber, “Validation concepts for gravity field models from new satellite missions,” in *Proc. Second International GOCE Users Workshop*, 2004.
- [108] L.-A. Albrecht, “Curvature-based analysis of point clouds,” Master’s thesis, 2014.
- [109] E. Abbena, S. Salamon, and A. Gray, *Modern differential geometry of curves and surfaces with Mathematica*. CRC press, 2006.
- [110] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering : a review,” *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.
- [111] A. Dockhorn, C. Braune, and R. Kruse, “An alternating optimization approach based on hierarchical adaptations of dbscan,” in *Computational Intelligence, 2015 IEEE Symposium Series on*, pp. 749–755, IEEE, 2015.
- [112] M. Schatzman and M. Schatzman, *Numerical analysis : a mathematical introduction*. Oxford University Press, 2002.

- [113] D. W. Scott and S. R. Sain, “9-multidimensional density estimation,” *Handbook of statistics*, vol. 24, pp. 229–261, 2005.
- [114] M. C. Jones, J. S. Marron, and S. J. Sheather, “A brief survey of bandwidth selection for density estimation,” *Journal of the American Statistical Association*, vol. 91, no. 433, pp. 401–407, 1996.
- [115] A. Baillo, “Advanced course in statistics : Multivariate nonparametric estimation,” 2014.
- [116] T. Schmidt, “Coping with copulas,” *Copulas-From theory to application in finance*, pp. 3–34, 2007.
- [117] R. Salinas-Gutiérrez, A. Hernández-Aguirre, M. J. Rivera-Meraz, and E. R. Villa-Diharce, “Using gaussian copulas in supervised probabilistic classification,” in *Soft Computing for Intelligent Control and Mobile Robotics*, pp. 355–372, Springer, 2010.
- [118] P. Embrechts, A. McNeil, and D. Straumann, “Correlation and dependence in risk management : properties and pitfalls,” *Risk management : value at risk and beyond*, vol. 176223, 2002.
- [119] M. Sklar, “Fonctions de repartition an dimensions et leurs marges,” *Publ. Inst. Statist. Univ. Paris*, vol. 8, pp. 229–231, 1959.
- [120] D. X. Li, “On default correlation : A copula function approach,” *The Journal of Fixed Income*, vol. 9, no. 4, pp. 43–54, 2000.
- [121] L. Madsen and D. Birkes, “Simulating dependent discrete data,” *Journal of Statistical Computation and Simulation*, vol. 83, no. 4, pp. 677–691, 2013.
- [122] M. Rosenblatt *et al.*, “Remarks on some nonparametric estimates of a density function,” *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.
- [123] E. Parzen, “On estimation of a probability density function and mode,” *The annals of mathematical statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.
- [124] T. Duong, A. Cowling, I. Koch, and M. P. Wand, “Feature significance for multivariate kernel density estimation,” *Computational Statistics & Data Analysis*, vol. 52, no. 9, pp. 4225–4242, 2008.
- [125] T. Duong and M. L. Hazelton, “Convergence rates for unconstrained bandwidth matrix selectors in multivariate kernel density estimation,” *Journal of Multivariate Analysis*, vol. 93, no. 2, pp. 417–433, 2005.
- [126] T. Duong and M. L. Hazelton, “Cross-validation bandwidth matrices for multivariate kernel density estimation,” *Scandinavian Journal of Statistics*, vol. 32, no. 3, pp. 485–506, 2005.
- [127] M. Filippone and G. Sanguinetti, “Approximate inference of the bandwidth in multivariate kernel density estimation,” *Computational Statistics & Data Analysis*, vol. 55, no. 12, pp. 3104–3122, 2011.
- [128] N. Zougab, S. Adjabi, and C. C. Kokonendji, “Bayesian estimation of adaptive bandwidth matrices in multivariate kernel density estimation,” *Computational Statistics & Data Analysis*, vol. 75, pp. 28–38, 2014.
- [129] B. W. Silverman, *Density estimation for statistics and data analysis*, vol. 26. CRC press, 1986.

- [130] G. Perrin, C. Soize, and N. Ouhbi, “Data-driven kernel representations for sampling with an unknown block dependence structure under correlation constraints,” *Computational Statistics & Data Analysis*, vol. 119, pp. 139–154, 2018.
- [131] C. Soize and R. Ghanem, “Data-driven probability concentration and sampling on manifold,” *Journal of Computational Physics*, vol. 321, pp. 242–258, 2016.
- [132] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, “Geometric diffusions as a tool for harmonic analysis and structure definition of data : Diffusion maps,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 21, pp. 7426–7431, 2005.
- [133] M. J. van der Laan, S. Dudoit, and S. Keles, “Asymptotic optimality of likelihood-based cross-validation,” *Statistical Applications in Genetics and Molecular Biology*, vol. 3, no. 1, pp. 1–23, 2004.
- [134] S. Owen, “An introduction to mesh generation algorithms,” *Proceedings of the 14th International Meshing Roundtable, San Diego, CA, USA*, pp. 11–14, 2005.
- [135] J.-D. Boissonnat and M. Yvinec, “Géométrie algorithmique,” 1995.
- [136] P. J. Frey and P.-L. George, “Delaunay-based mesh generation methods,” *Mesh Generation : Application to Finite Elements, Second Edition*, pp. 235–273, 2008.
- [137] F. Aurenhammer, R. Klein, and D.-T. Lee, *Voronoi diagrams and Delaunay triangulations*. World Scientific Publishing Co Inc, 2013.
- [138] H. Edelsbrunner and E. P. Mücke, “Three-dimensional alpha shapes,” *ACM Transactions on Graphics (TOG)*, vol. 13, no. 1, pp. 43–72, 1994.
- [139] E. Fogel and M. Teillaud, “Generic programming and the cgal library,” in *Effective Computational Geometry for Curves and Surfaces*, pp. 313–320, Springer, 2006.
- [140] S. Boyé, G. Guennebaud, and C. Schlick, “Least squares subdivision surfaces,” in *Computer Graphics Forum*, vol. 29, pp. 2021–2028, Wiley Online Library, 2010.
- [141] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, “Meshlab : an open-source mesh processing tool.,” in *Eurographics Italian Chapter Conference*, vol. 2008, pp. 129–136, 2008.
- [142] P. Cignoni, C. Montani, and R. Scopigno, “A comparison of mesh simplification algorithms,” *Computers & Graphics*, vol. 22, no. 1, pp. 37–54, 1998.
- [143] C.-K. Shene, “Mesh simplification,” 2010.
- [144] M. Garland and P. S. Heckbert, “Simplifying surfaces with color and texture using quadric error metrics,” in *Visualization’98. Proceedings*, pp. 263–269, IEEE, 1998.
- [145] G. Dhatt, G. Touzot, *et al.*, *Finite element method*. John Wiley & Sons, 2012.
- [146] V. Bodin, *Comportement du ballast des voies ferrées soumises à un chargement vertical et latéral*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 2001.
- [147] E. Arlaud, *Modèles dynamiques réduits de milieux périodiques par morceaux : application aux voies ferroviaires*. PhD thesis, Paris, ENSAM, 2016.



- [148] P. A. Cundall and O. D. Strack, “A discrete numerical model for granular assemblies,” *geotechnique*, vol. 29, no. 1, pp. 47–65, 1979.
- [149] M. Jean, “The non-smooth contact dynamics method,” *Computer methods in applied mechanics and engineering*, vol. 177, no. 3-4, pp. 235–257, 1999.
- [150] F. Radjai and V. Richefeu, “Contact dynamics as a nonsmooth discrete element method,” *Mechanics of Materials*, vol. 41, no. 6, pp. 715–728, 2009.
- [151] J. J. MOREAU, “Some numerical methods in multibody dynamics : application to granular materials,” *European journal of mechanics. A. Solids*, vol. 13, pp. 93–114, 1994.
- [152] F. Dubois and M. Jean, “Lmgc90 une plateforme de développement dédiée à la modélisation des problèmes d’interaction,” in *Actes du sixieme colloque national en calcul des structures*, vol. 1, pp. 111–118, 2003.
- [153] F. Da Cruz, S. Emam, M. Prochnow, J.-N. Roux, and F. Chevoir, “Rheophysics of dense granular materials : Discrete simulation of plane shear flows,” *Physical Review E*, vol. 72, no. 2, p. 021309, 2005.
- [154] G. MiDi, “On dense granular flows.,” *European Physical Journal E-Soft Matter*, vol. 14, no. 4, 2004.
- [155] T. Hatano, “Power-law friction in closely packed granular materials,” *Physical Review E*, vol. 75, no. 6, p. 060301, 2007.
- [156] P.-E. Peyneau and J.-N. Roux, “Frictionless bead packs have macroscopic friction, but no dilatancy,” *Physical review E*, vol. 78, no. 1, p. 011307, 2008.
- [157] E. Azéma and F. Radjai, “Internal structure of inertial granular flows,” *Physical review letters*, vol. 112, no. 7, p. 078001, 2014.
- [158] C. Thornton and L. Zhang, “Numerical simulations of the direct shear test,” *Chemical engineering & technology*, vol. 26, no. 2, pp. 153–156, 2003.
- [159] A. S. Suiker and N. A. Fleck, “Frictional collapse of granular assemblies,” *Journal of Applied Mechanics*, vol. 71, no. 3, pp. 350–358, 2004.
- [160] P.-E. Peyneau and J.-N. Roux, “Solidlike behavior and anisotropy in rigid frictionless bead assemblies,” *Physical Review E*, vol. 78, no. 4, p. 041307, 2008.
- [161] Wikipedia, “Confusion matrix — wikipedia, the free encyclopedia,” 2017.
- [162] D. Michie, D. J. Spiegelhalter, and C. C. Taylor, “Machine learning, neural and statistical classification,” 1994.
- [163] I. D. Quintanilla, G. Combe, F. Emeriault, J.-B. Toni, C. Voivret, and J. F. Ferrellec, “Wear of sharp aggregates in a rotating drum,” in *EPJ Web of Conferences*, vol. 140, p. 07009, EDP Sciences, 2017.