Continuous-time Martingale Optimal Transport and Optimal Skorokhod Embedding

Gaoyue Guo

NNT : 2016SACL038

THESE DE DOCTORAT DE L'UNIVERSITE PARIS-SACLAY

préparée à

L'ÉCOLE POLYTECHNIQUE

ÉCOLE DOCTORALE N°574

Mathématiques Hadamard (EDMH)

Spécialité de doctorat : Mathématiques Appliquées

par

Mr. Gaoyue Guo

Transport Optimal Martingale en Temps Continu et Plongement de Skorokhod Optimal.

Thèse présentée et soutenue à Palaiseau, le 27 Octobre 2016

Composition du jury :

Gilles Pagès	(Université Paris VI)	Président du jury
Erhan Bayraktar	(University of Michigan)	Rapporteur
Pierre Vallois	(Université de Lorraine)	Rapporteur
Nizar Touzi	(Ecole Polytechnique)	Directeur de thèse
Guillaume Carlier	(Université Paris Dauphine)	Examinateur
Pierre Henry-Labordère	(Société Générale)	Examinateur
Huyên Pham	(Université Paris Diderot)	Examinateur
Xiaolu Tan	(Université Paris Dauphine)	Examinateur

Contents

Transport Optimal Martingale en Temps Continu et Plongement de Skorokhod Optimal

Résumé : Cette thèse présente trois principaux sujets de recherche, les deux premiers étant indépendants et le dernier indiquant la relation des deux premières problématiques dans un cas concret.

Dans la première partie nous nous intéressons au problème de transport optimal martingale dans l'espace de Skorokhod, dont le premier but est d'étudier systématiquement la tension des plans de transport martingale. On s'intéresse tout d'abord à la semicontinuité supérieure du problème primitif par rapport aux distributions marginales. En utilisant la $S-$topologie introduite par Jakubowski [67], on dérive la semicontinuité supérieure et on montre la première dualité. Nous donnons en outre deux problèmes duaux concernant la surcouverture robuste d'une option exotique, et nous établissons les dualités correspondantes, en adaptant le principe de la programmation dynamique et l'argument de discrétisation initié dans Dolinsky & Soner [37].

La deuxième partie de cette thèse traite le problème du plongement de Skorokhod optimal. On formule tout d'abord ce problème d'optimisation en termes de mesures de probabilité sur un espace élargi et ses problèmes duaux. En utilisant l'approche classique de la dualité convexe et la théorie d'arrêt optimal, nous obtenons les résultats de dualité. Nous rapportons aussi ces résultats au transport optimal martingale dans l'espace des fonctions continues, d'où les dualités correspondantes sont dérivées pour une classe particulière de fonctions de paiement. Ensuite, on fournit une preuve alternative du principe de monotonie établi dans Beiglböck, Cox & Huesmann [5], qui permet de caractériser les optimiseurs par leur support géométrique. Nous montrons à la fin un résultat de stabilité qui contient deux parties : la stabilité du problème d'optimisation par rapport aux marginales cibles et le lien avec un autre problème du plongement optimal.

La dernière partie concerne l'application de contrôle stochastique au transport optimal martingale avec la fonction de paiement dépendant du temps local, et au plongement de Skorokhod. Pour le cas d'une marginale, nous retrouvons les optimiseurs pour les problèmes primitifs et duaux via les solutions de Vallois, et montrons en conséquence l'optimalité des solutions de Vallois, ce qui regroupe le transport optimal martingale et le plongement de Skorokhod optimal. Quand au cas de deux marginales, on obtient une généralisation de la solution de Vallois.

Mots-clés : transport optimal martingale, (sur)couverture robuste, $S-$topologie, principe de la programmation dynamique, plongement de Skorokhod optimal, arrêt optimal, principe de monotonie, métrique de Lévy-Prokhorov/Wasserstein, temps local, contrôle stochastique, solution de Vallois.

Continuous-time Martingale Optimal Transport and Optimal Skorokhod Embedding

Abstract : This PhD dissertation presents three research topics, the first two being independent and the last one relating the first two issues in a concrete case.

In the first part we focus on the martingale optimal transport problem on the Skorokhod space, which aims at studying systematically the tightness of martingale transport plans. Using the $S-$topology introduced by Jakubowski [67], we obtain the desired tightness which yields the upper semicontinuity of the primal problem with respect to the marginal distributions, and further the first duality. Then, we provide also two dual formulations that are related to the robust superhedging in financial mathematics, and we establish the corresponding dualities by adapting the dynamic programming principle and the discretization argument initiated by Dolinsky & Soner [37].

The second part of this dissertation addresses the optimal Skorokhod embedding problem under finitely-many marginal constraints. We formulate first this optimization problem by means of probability measures on an enlarged space as well as its dual problems. Using the classical convex duality approach together with the optimal stopping theory, we obtain the duality results. We also relate these results to the martingale optimal transport on the space of continuous functions, where the corresponding dualities are derived for a special class of reward functions. Next, We provide an alternative proof of the monotonicity principle established in Beiglböck, Cox & Huesmann [5], which characterizes the optimizers by their geometric support. Finally, we show a stability result that is twofold : the stability of the optimization problem with respect to target marginals and the relation with another optimal embedding problem.

The last part concerns the application of stochastic control to the martingale optimal transport with a payoff depending on the local time, and the Skorokhod embedding problem. For the one-marginal case, we recover the optimizers for both primal and dual problems through Vallois' solutions, and show further the optimality of Vallois' solutions, which relates the martingale optimal transport and the optimal Skorokhod embedding. As for the two-marginal case, we obtain a generalization of Vallois' solution.

Keywords : martingale optimal transport, robust (super)hedging, $S-$topology, dynamic programming principle, optimal Skorokhod embedding, optimal stopping, monotonicity principle, Lévy-Prokhorov/Wasserstein metric, local time, stochastic control, Vallois' solution.

Notations [1].

(i) $\mathbf{C} \equiv \mathbf{C}([0,1],\mathbb{R})$ is the space of continuous functions $\omega : [0,1] \to \mathbb{R}$ with $\omega_0 = 0$, $\mathbf{D} \equiv \mathbf{D}([0,1],\mathbb{R}^d)$ is the Skorokhod space of càdlàg (right continuous with left limits) functions $\omega : [0,1] \to \mathbb{R}^d$ for some given $d \in \mathbb{N}^*$, and Ω is the space of continuous functions $\omega : \mathbb{R}_+ \to \mathbb{R}$ with $\omega_0 = 0$. For $\omega \in \mathbf{C}$ (resp. $\omega \in \mathbf{D}$), set $\|\omega\| := \sup_{0\le t\le 1} |\omega_t|$.

(ii) $\mathbf{0} = (0,\cdots,0)$, $\mathbf{1} = (1,\cdots,1) \in \mathbb{R}^d$. $\mathbb{1}_E$ stands for the indicator function relative to E. For $a,b \in \mathbb{R}$, set $a \wedge b := \min(a,b)$ and $a \vee b := \max(a,b)$.

(iii) $\mathbf{P}$ and Λ represent respectively the space of probability measures μ on $\mathbb{R}^d$ with finite first moment and the space of continuous functions $\lambda : \mathbb{R}^d \to \mathbb{R}$ with linear growth. In particular, $\delta_x \in \mathbf{P}$ stands for the Dirac measure on $x \in \mathbb{R}^d$. Denote further

$$\mu(\lambda) \;:=\; \int_{\mathbb{R}^d} \lambda(x)\mu(dx) \text{ for all } \mu \in \mathbf{P} \text{ and } \lambda \in \Lambda.$$

For $i = 1,\cdots,d$, $K \in \mathbb{R}$ and $p \in \mathbb{R}_+$, set

$$\mu(x_i) \;:=\; \int_{\mathbb{R}^d} x_i\mu(dx),\;\; \mu((x_i-K)^+) \;:=\; \int_{\mathbb{R}^d}(x_i-K)^+\mu(dx) \text{ and } \mu(|x|^p) \;:=\; \int_{\mathbb{R}^d}|x|^p\mu(dx).$$

(iv) For a set $\mathbb{T} \subseteq \mathbb{R}_+$, where $\mathbb{T} \subseteq [0,1]$ or $\mathbb{T} = \{1,\cdots,m\}$ with $m \in \mathbb{N}^*$, denote the product spaces

$$\begin{aligned}
\mathbf{P}^{\mathbb{T}} &:= \Big\{\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} :\; \mu_t \in \mathbf{P} \text{ for all } t \in \mathbb{T}\Big\},\\
\Lambda^{\mathbb{T}} &:= \Big\{\boldsymbol{\lambda} = (\lambda_{t_i})_{1\le i\le k} :\; t_i \in \mathbb{T},\; \lambda_{t_i} \in \Lambda \text{ for all } i = 1,\cdots,k \text{ and } k \in \mathbb{N}\Big\}.
\end{aligned}$$

If $\mathbb{T} = \{1,\cdots,m\}$, we write equally $\mathbf{P}^{\mathbb{T}} = \mathbf{P}^m$ and $\Lambda^{\mathbb{T}} = \Lambda^m$.

(v) UI, a.s., a.e., u.s.c., l.s.c., w.l.o.g., w.r.t. l.h.s. r.h.s. are respectively the abbreviations of uniformly integrable, almost surely, almost every, upper semicontinuous, lower semincontinuous, without loss of generality, with respect to, left-hand side, right-hand side.

(vi) Notice that the above notations depend on $d \in \mathbb{N}^*$, but we will not specify this dependence in the sequel. We emphasize that $d = 1$ from Chapter IV to the end.

1. For ease of presentation, we only list the notations that are simultaneously used in the following chapters.

Chapitre I

Introduction (Français)

Cette thèse s'articule en deux thématiques indépandants. Le premier axe traite le transport optimal martingale dans l'espace de Skorokhod et le deuxième axe concerne le plongement de Skorokhod optimal.

La thèse est divisée en trois parties. La première partie considère le problème de transport optimal martingale en temps continu dans l'espace de Skorokhod, voir le Chapitre III ou Guo, Tan & Touzi [47]. La contribution principale est d'étudier systématiquement la tension des plans de transport martingale par la $S-$topologie introduite par Jakubowski et on montre l'existence des plans optimaux. Ensuite, on étudie le problème de transport optimal martingale par ses problèmes duaux et établit les dualités qui ont un lien avec la couverture robuste en finance.

La deuxième partie de cette thèse est consacrée à trois contributions au plongement de Skorokhod optimal. Nous présentons d'abord une formulation faible du problème en termes des mesures de probabilité sur un espace élargi, et puis nous introduisons deux formulations duales et prouvons les dualités correspondantes, voir le Chapitre IV ou Guo, Tan & Touzi [48] . Nous nous attardons ensuite sur le principe de monotonie introduit par Beiglböck, Cox & Huesmann et nous donnons une preuve alternative et simplifiée qui est basée sur les dualités précédentes. A la fin de cette partie, un résultat de stabilité est fourni en utilisant les métriques de Lévy-Prokhorov et Wasserstein, voir le Chapitre V ou Guo, Tan & Touzi [49] et Guo [45].

Dans la dernière partie, on considère l'application de contrôle stochastique au transport optimal martingale dans l'espace des fonctions continues avec la fonction de paiement dépendant uniquement du temps local, et au plongement de Skorokhod, voir le Chapitre VI ou Claisse, Guo & Henry-Labordère [20]. Dans le cas d'une marginale, on produit les optimiseurs explicites pour les problèmes primitifs et duaux à l'aide des solutions de Vallois, ce qui relie le transport optimal martingale et le plongement de Skorokhod optimal. Nous étendons ensuite l'analyse au cas de deux marginales et construisons une généralisation de la solution de Vallois. Enfin, un cas spécial de plusieurs marginales est étudié, où les temps d'arrêt donnés par Vallois sont bien ordonnés.

1 Préliminaires : couverture robuste

Les deux thématiques de la thèse sont étudiés récemment pour leurs nombreuses applications financières, notamment la couverture robuste. Avant de les attaquer, on commence par une introduction brève de la littérature relative qui permet de bien comprendre leur motivation financière.

Supposons qu'il y a d actifs sous-jacents et une option exotique [1] associée dans le marché. L'un des problèmes essentiels dans les mathématiques financières est d'évaluer cette option pour les vendeurs et acheteurs. Soient $X = (X_t)_{t\in I}$ un processus stochastique $d-$dimensionnel qui modélise les actifs sous-jacents, où $I = \{0, 1, \cdots, m\}$ (en temps discret) ou $I = [0, 1]$ (en temps continu) selon la formulation du problème, et $\xi = \xi(X)$ la fonction de paiement de l'option à considérer.

Un critère important pour déterminer le prix vient du souhait de la couverture. Du point de vue d'un vendeur (resp. acheteur), le prix convenable de vente (resp. achat) est une richesse permettant de construire un portefeuille qui sur-couvre/sur-réplique (resp. sous-couvre/sous-réplique) le paiement $\xi(X)$ au temps terminal. Dans la suite, nous allons préciser la surcouverture (resp. souscouverture) en temps discret et nous supposons

$$d \quad = \quad m \quad = \quad 1$$

afin de simplifier la présentation. La formulation en temps continu est similaire et paraît dans la Section 2.2.

1.1 Stratégies dynamique et statique

Pour construire un portefeuille avec une richesse initiale $z \in \mathbb{R}$, le vendeur (resp. acheteur) est autorisé à prendre une stratégie dynamique. C'est-à-dire qu'il détermine le nombre $H = H(X_0)$ d'actifs qu'il souhaite détenir à la date $t = 0$. En supposant que le taux d'intérêt soit nul, ce qui peut toujours être atteint par l'actualisation, le profit au temps terminal $t = 1$ est

$$H(X_0)(X_1 - X_0).$$

1. Une option s'écrivant sur des actifs sous-jacents est un contrat financier qui permet l'acheteur d'être payé par le vendeur à une date fixée selon sa fonction de paiement. Une option est dite exotique si la fonction de paiement dépend des sous-jacents de manière trajectorielle.

Notons que le nombre H ne dépend que de X_0 puisque H ne dépend que de l'information jusqu'au moment où la décision est prise.

En plus de la stratégie dynamique, le vendeur (resp. acheteur) est supposé de pouvoir prendre une position statique des calls et des puts [2] pour les maturités $t = 0, 1$ et tous les strikes $K \in \mathbb{R}$. Alors sous l'hypothèse de non-arbitrage du marché, on a la proposition suivante d'après Breeden & Litzenberger [16].

Proposition I.1.1. *Sous l'hypothèse de non-arbitrage, pour $i = 0, 1$, il existe une mesure de probabilité μ_i sur $\mathbb{R}$ telle que $X_i \sim \mu_i$, ou de manière équivalente,*

$$c_i(K) \;=\; \int_{\mathbb{R}} (x-K)^+ \mu_i(dx) \quad et \quad p_i(K) \;=\; \int_{\mathbb{R}} (K-x)^+ \mu_i(dx) \text{ pour tout } K \in \mathbb{R}.$$

En plus, on a $\mu_i = c_i'' = p_i''$ au sens des distributions.

Par conséquent, toute option avec la fonction de paiement $\lambda(X_i)$ a un prix non ambigu à la date $t = 0$ donné par

$$\int_{\mathbb{R}} \lambda(x) \mu_i(dx) \;\; =: \;\; \mu_i(\lambda) \tag{I.1.1}$$

pour $i = 0, 1$. En effet, supposons d'abord que λ soit suffisamment régulière, alors on obtient pour $x_i := \mu_i(x)$

$$\lambda(x) \;=\; \lambda(x_i) \;+\; (x - x_i)\lambda'(x_i) \;+\; \int_{-\infty}^{x_i} (K-x)^+ \lambda''(K) dK \;+\; \int_{x_i}^{+\infty} (x-K)^+ \lambda''(K) dK,$$

qui est la formule de Carr-Madan et donne une décomposition de la fonction de paiement λ à une combinaison linéaire des calls et des puts. On déduit immédiatement par la linéarité des prix que le prix correspondant à $\lambda(X_i)$ est bien

$$\lambda(x_i) \;+\; \int_{-\infty}^{x_i} p_i(K) \lambda''(K) dK \;+\; \int_{x_i}^{+\infty} c_i(K) \lambda''(K) dK,$$

où l'option de la fonction de paiement $X_i - x_i$ coûte zéro par le fait $x - x_i = (x - x_i)^+ - (x_i - x)^+$. En combinant la relation $\mu_i = c_i'' = p_i''$ et l'intégration par parties, l'expression au-dessus peut s'écrire comme $\mu_i(\lambda)$. Par un argument de densité, on déduit que pour toute fonction mesurable λ, l'option avec la fonction de paiement $\lambda(X_i)$ a un prix donné par (I.1.1).

2. Call (resp. Put) de maturité t et strike K est une option qui permet le détenteur de recevoir $(X_t - K)^+$ (resp. $(K - X_t)^+$) à la date t.

1.2 Couverture robuste

Sous l'hypothèse de non-arbitrage, le vendeur (resp. acheteur) peut construire un portefeuille semi-statique $(\lambda_0, \lambda_1, H) \in \mathbb{L}^1(\mu_0) \times \mathbb{L}^1(\mu_1) \times \mathbb{L}^0(\mathbb{R})$ avec le coût $\mu_0(\lambda_0) + \mu_1(\lambda_1)$ à la date $t = 0$. C'est-à-dire que son portefeuille est composé :

— d'une stratégie statique, qui consiste à acheter les options correspondant à $\lambda_0(X_0)$ et $\lambda_1(X_1)$ en payant $\mu_0(\lambda_0) + \mu_1(\lambda_1)$ à la date $t = 0$;

— d'une stratégie dynamique, qui consiste à acheter $H(X_0)$ d'actifs à la date $t = 0$ et les vendre à la date $t = 1$.

Sa richesse au temps terminal est donc donnée par

$$\lambda_0(X_0) \; + \; \lambda_1(X_1) \; + \; H(X_0)(X_1 - X_0).$$

Le portefeuille est dit une surcouverture (resp. souscouverture) si

$$\lambda_0(x_0) \; + \; \lambda_1(x_1) \; + \; H(x_0)(x_1 - x_0) \; \geq \; (\text{resp. } \leq) \; \xi(x_0, x_1) \text{ pour tout } (x_0, x_1) \in \mathbb{R}^2. \quad \text{(I.1.2)}$$

Notons qu'il n'y a aucun modèle imposé sur la dynamique de (X_0, X_1), donc la couverture s'appelle aussi couverture robuste, *c.à.d.* couverture indépendante du modèle. Soit $\overline{\mathcal{D}}^m$ (resp. $\underline{\mathcal{D}}^m$) l'ensemble de toutes les surcouvertures (resp. souscouvertures) et définissons

$$\overline{\mathsf{D}}^m(\mu_0, \mu_1) \; \Big(\text{resp. } \underline{\mathsf{D}}^m(\mu_0, \mu_1)\Big) \; := \; \inf_{(\lambda_0, \lambda_1, H) \in \overline{\mathcal{D}}^m} \; \Big(\text{resp. } \sup_{(\lambda_0, \lambda_1, H) \in \underline{\mathcal{D}}^m}\Big) \; \mu_0(\lambda_0) \; + \; \mu_1(\lambda_1). \quad \text{(I.1.3)}$$

Remarque I.1.2. (i) *L'hypothese de non-arbitrage est l'une des hypothèses fondamentales dans les mathématiques financières. On va voir dans la suite qu'elle implique que X est une martingale et par conséquent (μ_0, μ_1) est un peacock qui sera précisé ultérieurement.*

(ii) *Sans perte de généralité, on adopte maintenant la formulation de maximisation, et écrit simplement $\overline{\mathcal{D}}^m = \mathcal{D}^m$ et $\overline{\mathsf{D}}^m = \mathsf{D}^m$. Cette formulation est plus adaptée aux applications financières.*

2 Transport optimal martingale : du discret au continu

2.1 Du transport optimal classique au transport optimal martingale

Pour comprendre le lien entre la couverture robuste et le soi-disant transport optimal martingale, nous commençons par une introduction brève du transport optimal classique, et puis nous passons à sa généralisation au transport optimal martingale en temps discret dont le problème dual peut être interprété exactement comme la couverture robuste. En outre, on va voir que la compacité (faible) de l'ensemble des plans de transport joue un rôle essentiel pour résoudre tous les problèmes relatifs.

2.1.1 Problèmes de Monge et Monge-Kantorovich

Monge a considéré le problème suivant en 1781 : Etant donné deux mesures de probabilité μ_0 et μ_1 sur $\mathbb{R}^d$, une application mesurable $T : \mathbb{R}^d \to \mathbb{R}^d$ est appelée un plan de transport si $\mu_0 \circ T^{-1} = \mu_1$, où $\mu_0 \circ T^{-1}$ désigne la mesure image par T. Le problème de Monge consiste à trouver le profit maximal donné par

$$\mathsf{P}^M(\mu_0, \mu_1) \;\; := \;\; \sup_{T:\; \mu_0 \circ T^{-1} = \mu_1} \int_{\mathbb{R}^d} \xi\Big(x, T(x)\Big)\mu_0(dx), \tag{I.2.1}$$

où $\xi : \mathbb{R}^{2d} \to \mathbb{R}$ est une fonction mesurable. Ce problème est difficile à résoudre à cause de contraintes entièrement non linéaires, et restait ouvert pendant de nombreuses années.

Une percée a été faite par Kantorovich en 1940 en relâchant les contraintes, puis en introduisant la formulation duale. C'est-à-dire, soit $\mathcal{P}(\mathbb{R}^{2d})$ l'ensemble des mesures de probabilité sur $\mathbb{R}^{2d}$. Définissons

$$\mathcal{P}(\mu_0, \mu_1) \;\; := \;\; \Big\{\mathbb{P} \in \mathcal{P}(\mathbb{R}^{2d}) :\; \mathbb{P} \circ X_i^{-1} \;=\; \mu_i \text{ pour } i = 0, 1\Big\}, \tag{I.2.2}$$

où (X_0, X_1) indique le processus canonique sur $\mathbb{R}^d \times \mathbb{R}^d$, *c.à.d.* $X_0(x_0, x_1) := x_0$ et $X_1(x_0, x_1) := x_1$ pour tout $(x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d$. On s'intéresse dans la suite au problème de Monge-Kantorovich

$$\mathsf{P}^{MK}(\mu_0, \mu_1) \;\; := \;\; \sup_{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)} \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1)\Big]. \tag{I.2.3}$$

Kantorovich a introduit une formulation duale définie par

$$\mathsf{D}^{MK}(\mu_0, \mu_1) \;\; := \;\; \inf_{(\lambda_0, \lambda_1) \in \mathcal{D}^{MK}} \mu_0(\lambda_0) \; + \; \mu_1(\lambda_1), \tag{I.2.4}$$

où

$$\mathcal{D}^{MK} := \Big\{(\lambda_0, \lambda_1) \in \mathbb{L}^1(\mu_0) \times \mathbb{L}^1(\mu_1) : \lambda_0(x_0) + \lambda_1(x_1) \geq \xi(x_0, x_1) \text{ pour tout } (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d\Big\}.$$

La validité de cette notion vient du souhait d'obtenir la dualité suivante. En effet, l'ensemble

$\mathcal{P}(\mu_0, \mu_1)$ peut être montré compact par un argument simple, et la dualité résulte du théorème Min-Max directement.

Théorème I.2.1. *Supposons que $\xi : \mathbb{R}^{2d} \to \mathbb{R}$ soit semicontinue supérieurement et majorée. Alors il existe des optimiseurs pour les problèmes (I.2.3) et (I.2.4), c.à.d. il existe $\mathbb{P}^* \in \mathcal{P}(\mu_0, \mu_1)$ et $(\lambda_0^*, \lambda_1^*) \in \mathcal{D}^{MK}$ tels que*

$$\mathbb{E}^{\mathbb{P}^*}\Big[\xi(X_0, X_1)\Big] \;=\; \mathsf{P}^{MK}(\mu_0, \mu_1) \quad \textit{et} \quad \mu_0(\lambda_0^*) \;+\; \mu_1(\lambda_1^*) \;=\; \mathsf{D}^{MK}(\mu_0, \mu_1).$$

En plus, on a la dualité souhaitée, c.à.d. $\mathsf{P}^{MK} = \mathsf{D}^{MK}$.

Remarque I.2.2. *Le problème de transport optimal a donné lieu à une littérature prolifique, où la dualité joue un rôle essentiel. Elle permet non seulement d'étudier le problème primitif d'un autre point de vue, i.e. établir le principe de monotonie qui permet de caractériser l'optimiseur $\mathbb{P}^*$, mais aussi de donner une approche numérique pour calculer $\mathsf{P}^{MK} = \mathsf{D}^{MK}$. Nous renvoyons le lecteur à Rachev & Rüschendorf [90] ou Villani [104] pour un compte rendu complet de la littérature, ainsi que les applications relatives.*

2.1.2 Transport optimal martingale en temps discret

Motivés par la couverture robuste en mathématiques financières, Beiglböck, Henry-Labordère & Penkner [7] et Galichon, Henry-Labordère & Touzi [43] ont introduit en même temps un problème de transport optimal avec la contrainte de martingale en temps discret et en temps continu, respectivement. Nous considérrons ici le cas discret pour illustrer l'idée. On définit le sous-ensemble $\mathcal{M}(\mu_0, \mu_1) \subseteq \mathcal{P}(\mu_0, \mu_1)$ par

$$\mathcal{M}(\mu_0, \mu_1) \;:=\; \Big\{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1) : \; \mathbb{E}^{\mathbb{P}}\Big[X_1 \Big| X_0\Big] \;=\; X_0, \;\; \mathbb{P}\text{ - p.s.}\Big\}. \tag{I.2.5}$$

Chaque élément $\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)$ est appelé un plan de transport (martingale). Le problème de transport optimal martingale correspondant vise à considérer

$$\mathsf{P}^m(\mu_0, \mu_1) \;:=\; \sup_{\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)} \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1)\Big]. \tag{I.2.6}$$

Comme dans le transport optimal classique, Beiglböck, Henry-Labordère & Penkner ont montré une relation de dualité entre $\mathsf{P}^m(\mu_0, \mu_1)$ et la couverture robuste $\mathsf{D}^m(\mu_0, \mu_1)$ qui est introduite dans la Section 1.2 :

$$\mathsf{D}^m(\mu_0, \mu_1) \;:=\; \inf_{(\lambda_0, \lambda_1, H) \in \mathcal{D}^m} \mu_0(\lambda_0) \;+\; \mu_1(\lambda_1), \tag{I.2.7}$$

où

$$\mathcal{D}^m \;:=\; \Big\{(\lambda_0,\lambda_1,H)\in\mathbb{L}^1(\mu_0)\times\mathbb{L}^1(\mu_1)\times\mathbb{L}^0(\mathbb{R}^d):\; \lambda_0(x_0)\;+\;\lambda_1(x_1)$$
$$+\;H(x_0)(x_1-x_0)\;\ge\;\xi(x_0,x_1)\text{ pour tout }(x_0,x_1)\in\mathbb{R}^d\times\mathbb{R}^d\Big\}. \tag{I.2.8}$$

Rappelons que $\mathbf{P}$ et Λ représentent respectivement l'espace des mesures de probabilité sur $\mathbb{R}^d$ avec le premier moment fini et l'espace des fonctions continues sur $\mathbb{R}^d$ à croissance linéaire. Une paire $(\mu_0,\mu_1)\in\mathbf{P}\times\mathbf{P}$ est dite croissante pour l'ordre convexe si l'inégalité $\mu_0(\lambda)\le\mu_1(\lambda)$ est vraie pour toutes les fonctions convexes $\lambda\in\Lambda$. Cette relation est notée $\mu_0\preceq\mu_1$ et on a une caractérisation de l'ensemble $\mathcal{M}(\mu_0,\mu_1)$.

Théorème I.2.3 (Strassen)**.** *Soient μ_0 et $\mu_1\in\mathbf{P}$ deux mesures de probabilité sur $\mathbb{R}^d$. Alors*

$$\mathcal{M}(\mu_0,\mu_1)\;\neq\;\emptyset\quad\textit{si et seulement si}\quad\mu_0\;\preceq\;\mu_1.$$

C'est facile de montrer que $\mathcal{M}(\mu_0,\mu_1)$ est un sous-ensemble fermé du compact $\mathcal{P}(\mu_0,\mu_1)$, ce qui implique l'existence d'un optimiseur du problème (I.2.6). En outre,

$$\begin{aligned}
&\inf_{(\lambda_0,\lambda_1,H)\in\mathcal{D}^m}\;\mu_0(\lambda_0)\;+\;\mu_1(\lambda_1),\\
=\;&\inf_{H\in\mathbb{L}^0(\mathbb{R}^d)}\;\inf_{(\lambda_0,\lambda_1)\in\mathcal{D}^{MK}(H)}\;\mu_0(\lambda_0)\;+\;\mu_1(\lambda_1)\\
=\;&\inf_{H\in\mathbb{L}^0(\mathbb{R}^d)}\;\sup_{\mathbb{P}\in\mathcal{P}(\mu_0,\mu_1)}\;\mathbb{E}^{\mathbb{P}}\Big[\xi(X_0,X_1)\;-\;H(X_0)(X_1-X_0)\Big]\\
=\;&\sup_{\mathbb{P}\in\mathcal{P}(\mu_0,\mu_1)}\;\inf_{H\in\mathbb{L}^0(\mathbb{R}^d)}\;\mathbb{E}^{\mathbb{P}}\Big[\xi(X_0,X_1)\;-\;H(X_0)(X_1-X_0)\Big]\\
=\;&\sup_{\mathbb{P}\in\mathcal{M}(\mu_0,\mu_1)}\;\mathbb{E}^{\mathbb{P}}\Big[\xi(X_0,X_1)\;-\;H(X_0)(X_1-X_0)\Big].
\end{aligned}$$

où

$$\mathcal{D}^{MK}(H)\;:=\;\Big\{(\lambda_0,\lambda_1)\in\mathbb{L}^1(\mu_0)\times\mathbb{L}^1(\mu_1):\;\lambda_0(x_0)\;+\;\lambda_1(x_1)$$
$$\ge\;\xi(x_0,x_1)\;-\;H(x_0)(x_1-x_0)\text{ pour tout }(x_0,x_1)\in\mathbb{R}^d\times\mathbb{R}^d\Big\}, \tag{I.2.9}$$

la deuxième égalité est la dualité classique du Théorème I.2.1, et la troisième égalité vient du théorème Min-Max. On obtient donc les mêmes résultats résumés dans le théorème suivant, voir Beiglböck, Henry-Labordère & Penkner [7] pour les détailles.

Théorème I.2.4 (Beiglböck, Henry-Labordère & Penkner). *Supposons que ξ soit semicontinue supérieurement et majorée. Alors pour toute paire $(\mu_0, \mu_1) \in \mathbf{P} \times \mathbf{P}$ telle que $\mu_0 \preceq \mu_1$, on a la dualité $\mathsf{P}^m = \mathsf{D}^m$. De plus il existe un optimiseur du problème* (I.2.6).

Remarque I.2.5. (i) *Sous l'hypothèse de non-arbitrage, le processus X est requis d'être une martingale. Par conséquent, (μ_0, μ_1) doit être un peacock pour que $\mathcal{M}(\mu_0, \mu_1)$ ne soit pas vide. D'après la littérature classique en mathématiques financières, chaque mesure martingale $\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)$ représente un modèle du marché vérifiant l'hypothèse de non-arbitrage. Donc sous l'incertitude du marché, le problème primitif P^m désigne la borne supérieure des prix de non-arbitrage.*

Comme décrit dans la Section 1.2, le problème dual D^m désigne le coût minimal de la surcouverture robuste. Donc la dualité $\mathsf{P}^m = \mathsf{D}^m$ signifie que la borne supérieure des prix de non-arbitrage coïncide avec le coût minimal de la surcouverture robuste.

(ii) *La compacité de $\mathcal{M}(\mu_0, \mu_1)$ et la régularité de ξ sont cruciales pour prouver le Théorème I.2.4.*

(iii) *L'existence d'optimiseurs du problème dual* (I.2.7) *n'est pas assurée en général. Néanmoins, Beiglböck, Nutz & Touzi ont proposé une autre formulation duale dans [11] et prouvé simultanément l'existence d'un optimiseur du problème dual et la dualité sous une condition de régularité plus faible pour ξ.*

2.2 Transport optimal martingale en temps continu

Le problème de transport optimal martingale admet aussi naturellement une version en temps continu, pour lequel on voit que l'approche en temps discret ne fonctionne plus sans la compacité de l'ensemble des plans de transport, voir la Remarque I.2.7 pour la description des difficultés en temps continu. Nous exploitons d'autres méthodes pour étudier ce problème et établissons les résultats principaux de la Section 2. En réalité il n'y a pas unicité de la formulation en temps continu, et nous allons expliquer cela.

2.2.1 Formulation

Nous choisissons ici de ne considérer que des lois de probabilité sur l'espace de Skorokhod. Rappelons que $\mathbf{D}$ est l'espace de Skorokhod des fonctions càdlàg $\omega = (\omega_t)_{0\le t\le 1}$ à valeurs dans $\mathbb{R}^d$. Soient $X = (X_t)_{0\le t\le 1}$ le processus canonique, *c.à.d.* $X_t(\omega) := \omega_t$ et $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ sa filtration naturelle, *c.à.d.* $\mathcal{F}_t := \sigma(X_s, 0 \le s \le t)$. Soit $\mathcal{P}$ l'ensemble des mesures de probabilité sur

$(\mathbf{D}, \mathcal{F}_1)$ et chaque élément $\mathbb{P} \in \mathcal{P}$ est appelé une mesure martingale si X est une $\mathbb{P}-$martingale. On considère la collection de toutes les mesures martingales notée $\mathcal{M}$.

Soit $1 \in \mathbb{T} \subseteq [0,1]$ un sous-ensemble et rappelons que le $\mathbb{T}-$produit de $\mathbf{P}$ est défini par

$$\mathbf{P}^{\mathbb{T}} \;\; := \;\; \Big\{ \boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} : \;\; \mu_t \in \mathbf{P} \text{ pour tout } t \in \mathbb{T} \Big\}.$$

Définition I.2.6. *Une famille de probabilités $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}^{\mathbb{T}}$ est applée un peacock si on a $\mu_s \preceq \mu_t$ pour tout $s, t \in \mathbb{T}$ tels que $s \le t$. Un peacock $\boldsymbol{\mu}$ est dit càdlàg si l'application $t \mapsto \mu_t$ est càdlàg sur $\mathbb{T}$ par rapport à la convergence faible. Soit $\mathbf{P}^{\preceq}$ l'ensemble des peacocks càdlàg.*

Pour un peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, définissons par $\mathcal{M}(\boldsymbol{\mu}) \subseteq \mathcal{M}$ le sous-ensemble des plans de transport martingale $\mathbb{P}$, *c.à.d.* $\mathbb{P} \circ X_t^{-1} = \mu_t$ pour tout $t \in \mathbb{T}$. Etant donné une fonction $\mathcal{F}_1-$mesurable $\xi : \mathbf{D} \to \mathbb{R}$, le problème de transport optimal martingale est défini par

$$\mathsf{P}(\boldsymbol{\mu}) \;\; := \;\; \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big], \tag{I.2.10}$$

où ξ est supposée de satisfaire $\mathbb{E}^{\mathbb{P}}[|\xi|] < +\infty$ pour toute $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$. Ensuite, toutes les questions naturelles qui se posent dans le cadre en temps discret peuvent apparaître ici :

— Quel est le problème dual correspondant de (I.2.10) ?

— Y a-t-il existence des optimiseurs pour les problèmes primitif et dual ?

— Est-ce qu'on a la dualité ?

Remarque I.2.7. (i) *Lorsque $\mathbb{T}$ est fini ou $\mathbb{T} = [0,1]$, il résulte respectivement de Kellerer [72] et Hirsch & Roynette [54, 55] que, pour chaque $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, l'ensemble $\mathcal{M}(\boldsymbol{\mu})$ est non vide, voir Hirsch, Profeta, Roynette & Yor [53] pour une littérature complète du peacock.*

(ii) *La première difficulté en temps continu vient de la définition du problème dual. En effet, quand on passe du temps discret au temps continu, la partie dynamique devient donc*

$$H(X_0)(X_1 - X_0) \;\; \Longrightarrow \;\; \int_0^1 H_t dX_t$$

pour certain processus $H = (H_t)_{0\le t\le 1}$ convenable. Mais comme le processus X n'est pas a priori à variation finie, et on ne peut donc pas définir $\int_0^1 H_t dX_t$ de manière trajectorielle.

Dans la littérature existante, il y a deux formulations duales pour le problème (I.2.10). Galichon, Henry-Labordère & Touzi [43] ont étudié une classe de plans de transport martingale définis par des équations différentielles stochastiques et introduit un problème dual au sens quasi-sûr et cette formulation sera précisée dans les Sections 3.2.2 et 4.1. Ils ont appliqué l'approche de *contrôle stochastique et déduit la dualité. Une autre contribution importante vient de Dolinsky & Soner [36, 37] et Hou & Obłój [63], où le problème dual a une définition trajectorielle.*

(iii) *Par rapport au cas en temps discret, l'ensemble $\mathcal{M}(\boldsymbol{\mu})$ n'est généralement pas tendu*[3] *par rapport aux topologies habituelles, ce qui est montré par l'exemple suivant. Sans la compacité cruciale, les arguments dans le cadre classique ne sont plus adaptés pour traiter les questions relatives.*

Exemple I.2.8. *Soit $M = (M_0, M_1, M_2)$ une martingale telle que $\mathbb{P}[M_0 \neq M_1$ et $M_1 \neq M_2] > 0$. Définissons $\mathbb{P}_n := \mathbb{P} \circ (M^n)^{-1}$ pour tout $n \geq 3$, où $M^n = (M_t^n)_{0 \leq t \leq 1}$ est donnée par*

$$M_t^n \;:=\; M_0 \mathbb{1}_{[0, \frac{1}{2} - \frac{1}{n})}(t) \;+\; M_1 \mathbb{1}_{[\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n})}(t) \;+\; M_2 \mathbb{1}_{[\frac{1}{2} + \frac{1}{n}, 1]}(t).$$

On a immédiatement $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu})$ pour tout $n \geq 3$ avec $\mathbb{T} = \{0, 1\}$ et $\boldsymbol{\mu} = (\mathbb{P} \circ M_0^{-1}, \mathbb{P} \circ M_2^{-1})$. Néanmoins, il résulte du Théorème VI.3.21 dans Jacod & Shiryaev [66] que, la suite $(\mathbb{P}_n)_{n \geq 3}$ n'est pas tendue par rapport à la topologie de Skorokhod, et donc pas compacte.

2.2.2 Résultats principaux

On introduit trois formulations duales du problème (I.2.10). La contribution principale est d'étudier systématiquement la tension de l'ensemble $\mathcal{M}(\boldsymbol{\mu})$ par la $S-$topologie introduite dans Jakubowski [67]. On munit l'espace $\mathbf{P}^{\mathbb{T}}$ avec une topologie de type Wasserstein. La tension de $\mathcal{M}(\boldsymbol{\mu})$ implique la semicontinuité supérieure de l'application $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$, et par conséquent la première dualité, obtenue en pénalisant les contraintes marginales. En utilisant le principe de la programmation dynamique et l'argument de discrétisation introduit dans Dolinsky & Soner [37], on obtient les deux dualités suivantes.

Premier problème dual Rappelons que Λ est l'ensemble des fonctions continues sur $\mathbb{R}^d$ à croissance linéaire et $\Lambda^{\mathbb{T}}$ est le $\mathbb{T}-$produit de Λ, *c.à.d*

$$\Lambda^{\mathbb{T}} \;:=\; \Big\{ \boldsymbol{\lambda} = (\lambda_{t_i})_{1 \leq i \leq k} : \; t_i \in \mathbb{T}, \; \lambda_{t_i} \in \Lambda \text{ pour tout } i = 1, \cdots, k \text{ et } k \in \mathbb{N} \Big\}.$$

Pour tout $\boldsymbol{\lambda} = (\lambda_{t_i})_{1 \leq i \leq k} \in \Lambda^{\mathbb{T}}$ et $\omega = (\omega_t)_{0 \leq t \leq 1} \in \mathbf{D}$, on définit

$$\boldsymbol{\lambda}(\omega) \;:=\; \sum_{i=1}^k \lambda_{t_i}(\omega_{t_i}) \quad \text{et} \quad \boldsymbol{\mu}(\boldsymbol{\lambda}) \;:=\; \sum_{i=1}^k \mu_{t_i}(\lambda_{t_i}).$$

3. Un ensemble $\mathcal{E}$ de mesures de probabilité est dit tendu (par rapport à une convergence selon le contexte) si pour tout $\varepsilon > 0$ il existe un ensemble compact K tel que $\mathbb{P}[K] \geq 1 - \varepsilon$ pour toute $\mathbb{P} \in \mathcal{E}$.

Le premier problème dual est défini par

$$\mathsf{D}_1(\boldsymbol{\mu}) \;:=\; \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\big[\xi(X) \;-\; \boldsymbol{\lambda}(X)\big]\Big\}. \tag{I.2.11}$$

Le problème dual D_1 semble le Lagrangien des problèmes d'optimisation sous des contraintes, où les contraintes marginales $\boldsymbol{\mu}$ sont pénalisées par les multiplicateurs Lagrangiens $\boldsymbol{\lambda}$.

Deuxième problème dual Le deuxième problème dual pénalise en outre la contrainte martingale. Pour des raisons techniques, on se restreint au cas où le processus X est positif et démarre d'un point donné et normalisé à $\mathbf{1} = (1,\cdots,1) \in \mathbb{R}^d$. Définissons $\mathbf{D}_+ := \{\omega \in \mathbf{D} : \omega_0 = \mathbf{1}$ et $\omega_t \in \mathbb{R}^d_+$ pour tout $t \in [0,1]\}$ et $\mathcal{M}_+ := \{\mathbb{P} \in \mathcal{M} : \mathrm{supp}(\mathbb{P}) \subseteq \mathbf{D}_+\}$. Par conséquent, pour faire en sorte que $\mathcal{M}_+ \cap \mathcal{M}(\boldsymbol{\mu}) \neq \emptyset$, $\boldsymbol{\mu}$ doit satisfaire

$$\mu_0(dx) \;=\; \delta_{\mathbf{1}}(dx) \quad \text{et} \quad \mathrm{supp}(\mu_1) \;\subseteq\; \mathbb{R}^d_+. \tag{I.2.12}$$

Soit $\mathbb{F}^U = (\mathcal{F}^U_t)_{0\le t\le 1}$ la filtration universellement completée, *c.à.d.* $\mathcal{F}^U_t := \cap_{\mathbb{P}\in\mathcal{P}}\mathcal{F}^{\mathbb{P}}_t$, où $\mathcal{F}^{\mathbb{P}}_t$ est la tribu completée de $\mathcal{F}_t$ sous $\mathbb{P}$.

Définition I.2.9. *Un processus $\mathbb{F}^U-$adapté $S : [0,1]\times\mathbf{D}_+ \to \mathbb{R}$ est appelé une $\mathcal{M}_+-$surmartingale si il est une $\mathbb{P}-$surmartingale pour toute $\mathbb{P} \in \mathcal{M}_+$. Soient $\mathcal{S}$ la collection de tous ces processus et $\mathcal{S}_0 \subseteq \mathcal{S}$ le sous-ensemble des processus partant de zéro. Définissons*

$$\mathcal{D}_2(\xi) \;:=\; \Big\{(\boldsymbol{\lambda}, S) \in \Lambda^{\mathbb{T}} \times \mathcal{S}_0 : \; \boldsymbol{\lambda}(\omega) \;+\; S_1(\omega) \;\geq\; \xi(\omega) \text{ pour tout } \omega \in \mathbf{D}_+\Big\}.$$

Pour un peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ vérifiant la condition (I.2.12) , le deuxième problème dual est défini par

$$\mathsf{D}_2(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda},S)\in\mathcal{D}_2(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{I.2.13}$$

Remarque I.2.10. *Notons que la surmartingale $S \in \mathcal{S}$ n'est pas requise d'avoir une certaine régularité. Si elle était càdlàg alors d'après Théorème 2.1 dans Kramkov [74], pour chaque $\mathbb{P} \in \mathcal{M}_+$ il existerait un processus prévisible $H^{\mathbb{P}} = (H^{\mathbb{P}}_t)_{0\le t\le 1}$ et un processus optionnel croissant $A^{\mathbb{P}} = (A^{\mathbb{P}}_t)_{0\le t\le 1}$ de telle sorte que*

$$S_t \;=\; S_0 \;+\; \int_0^t H^{\mathbb{P}}_s dX_s \;-\; A^{\mathbb{P}}_t \text{ pour tout } t \in [0,1], \; \mathbb{P}\text{ - p.s.}$$

Cependant, on ne sait pas si l'on peut agréger la dernière représentation, c.à.d. trouver des *processus H et A tels que $(H, A) = (H^{\mathbb{P}}, A^{\mathbb{P}})$, $\mathbb{P}$ - presque sûrement pour toute $\mathbb{P} \in \mathcal{M}_+$. D'après Nutz [80], sous l'hypothèse du continu, ce résultat d'agrégation est vrai lorsque X est à valeurs dans l'espace des fonctions continues puisqu'on peut alors identifier l'intégrant dans l'intégrale stochastique par $H_t = H_t^{\mathbb{P}} = d\langle S, X\rangle_t / d\langle X\rangle_t$, où les processus $\langle S, X\rangle$ et $\langle X\rangle$ désignent respectivement la co-variation quadratique de S et X et la variation quadratique de X, et ils peuvent être définis de manière universelle.*

Troisième problème dual Le troisième problème dual est à peu près similaire au deuxième, mais plus fort. Il peut être considéré comme une représentation intégrale des surmartingales sous une famille non-dominée de mesures.

Definition I.2.11. *Un processus $\mathbb{F}-$adapté $H : [0, 1] \times \mathbf{D}_+ \to \mathbb{R}^d$ est appelé une stratégie dynamique si $t \mapsto H_t(\omega)$ est continue à gauche et de variation bornée pour tout $\omega \in \mathbf{D}_+$, et $(H \cdot X)$ est une $\mathbb{P}-$surmartingale pour toute $\mathbb{P} \in \mathcal{M}_+$, où*

$$(H \cdot X)_t \;:=\; H_t \cdot X_t \;-\; H_0 \cdot X_0 \;-\; \int_0^t X_s \cdot dH_s \text{ pour tout } t \in [0, 1],$$

avec $\int_0^{\cdot} X_s \cdot dH_s$ représentant l'intégration scalaire de Lebesgue-Stieltjes. Soit $\mathcal{H}$ l'ensemble de toutes les stratégies dynamiques.

Pour un peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ vérifiant la condition (I.2.12), le troisième problème dual est défini par

$$\mathsf{D}_3(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda}, H) \in \mathcal{D}_3(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}), \tag{I.2.14}$$

où

$$\mathcal{D}_3(\xi) \;:=\; \Big\{(\boldsymbol{\lambda}, H) \in \Lambda^{\mathbb{T}} \times \mathcal{H} :\; \boldsymbol{\lambda}(\omega) \;+\; (H \cdot X)_1(\omega) \;\geq\; \xi(\omega) \text{ pour tout } \omega \in \mathbf{D}_+\Big\}.$$

Enfin, le théorème suivant résume les résultats de dualité.

Théorème I.2.12. *Sous des conditions appropriées, il existe une mesure de probabilité $\mathbb{P}^* \in \mathcal{M}(\boldsymbol{\mu})$ telle que*

$$\mathbb{E}^{\mathbb{P}^*}\Big[\xi(X)\Big] \;=\; \mathsf{P}(\boldsymbol{\mu}) \;=\; \mathsf{D}_1(\boldsymbol{\mu}) \;=\; \mathsf{D}_2(\boldsymbol{\mu}) \;=\; \mathsf{D}_3(\boldsymbol{\mu}).$$

Remarque I.2.13. (i) *Comme dans le cadre en temps discret, la formulation duale (I.2.14) désigne la fortune minimale permettant de construire une surcouverture robuste de ξ. La seule* *différence est qu'on se restreint à une classe plus petite de stratégies dynamiques.*

(ii) *On peut également formuler le problème en supposant X est à valeurs dans l'espace des fonctions continues, voir Dolinsky & Soner [36] et Hou & Obłój [63]. On considère la formulation dans l'espace des fonctions continues dans les Sections 3.2.2 et 4.1, où le problème est formulé différemment et étudié par d'autre approches.*

3 Plongement de Skorokhod optimal : du fort au faible

3.1 Motivation et formulation

Le problème du plongement de Skorokhod vise à représenter une mesure de probabilité donnée sur $\mathbb{R}$ par la loi du mouvement brownien arrêté à un temps d'arrêt bien choisi. Initié par Skorokhod, ce problème a été résolu de très nombreuse manières différentes. En particulier, un grand nombre de solutions du plongement de Skorokhod satisfont certaine notamment optimalité. Par exemple, la solution de Root minimise la moyenne du plongement, la solution d'Azéma-Yor maximise la loi du maximum du mouvement brownien, et Vallois construit les solutions qui optimisent l'espérance d'une fonction convexe du temps local, *etc.*

Le but de la Section 3 est d'unifier le problème d'optimisation parmi les temps d'arrêt du plongement, nommé le plongement de Skorokhod optimal, et de l'étudier systématiquement, *c.à.d.* l'existence et la caractérisation d'optimiseurs, voir Sections 3.2, 3.3 et 3.4 pour plus de résultats.

3.1.1 Lien avec la couverture robuste

Sachant que chaque martingale continue partant de zéro est un mouvement brownien changeant du temps, on voit que ce problème d'optimisation est lié au problème de transport optimal martingale dans l'espace des fonctions continues. Dans cette partie on illustre formellement cette idée et indique que ce sujet est motivé par l'étude de la couverture robuste en finance.

Soient $B = (B_t)_{t\geq 0}$ un mouvement brownien (standard) et μ une mesure de probabilité centrée sur $\mathbb{R}$. Pour un plongement arbitraire τ, considérons le processus $X = (X_t)_{0\leq t\leq 1}$ défini par

$$X_t \quad := \quad B_{\frac{t}{1-t}\wedge\tau} \text{ pour tout } t \in [0,1].$$

D'après la construction, X est une martingale continue avec $X_1 \sim \mu$. Inversement, pour chaque martingale continue, il résulte du théorème de Dambis-Dubins-Schwarz qu'il existe un mouvement brownien $W = (W_t)_{t\geq 0}$ tel que presque sûrement

$$X_t \;=\; W_{\langle X\rangle_t},$$

d'où vient le lien entre ces deux problèmes.

En outre, il est bien connu que dans de nombreux cas, le plongement optimal permet d'expliciter les optimiseurs du problème de transport optimal martingale et de son problème dual, autrement dit, le modèle associé au prix extrémal et la stratégie optimale de la couverture robuste. Pour plus de résultats relatives, nous renvoyons le lecteur à Hobson [56], Hobson & Obłój [23], Cox & Obłój [24, 25], Hobson & Klimmek [58], Obłój & Spoida [84], Galichon, Henry-Labordère & Touzi [43], Henry-Labordère, Obłój, Spoida & Touzi [50], Cox, Obłój & Touzi [26], *etc.*

3.1.2 Formulation

Ce problème d'optimisation n'a été étudié de manière systématique que dans l'article [5] récent de Beiglböck, Cox & Huesmann, où ils ont formulé ce problème dans le cas d'une marginale et pu retrouver tous les résultats connus précédents par cette formulation unificatrice. Nous considérons une extension de cette formulation dans [5] au cas d'un nombre fini de contraintes marginales.

Rappelons que Ω est l'espace des fonctions continues $\omega = (\omega_t)_{t\geq 0}$ sur $\mathbb{R}_+$ telles que $\omega_0 = 0$. Soient $B = (B_t)_{t\geq 0}$ le processus canonique, *c.à.d.* $B_t(\omega) := \omega_t$, $\mathbb{P}_0$ la mesure de Wiener, $\mathbb{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ la filtration naturelle de B, *c.à.d.* $\mathcal{F}^0_t := \sigma(B_s, s\leq t)$ et $\mathbb{F}^a = (\mathcal{F}^a_t)_{t\geq 0}$ la filtration augmentée sous $\mathbb{P}_0$.

Soit $m \geq 1$ un entier fixé et définissons $\Theta := \{\theta = (\theta_1,\cdots,\theta_m) : 0 \leq \theta_1 \leq \cdots \leq \theta_m\}$. On définit ensuite l'espace canonique élargi par $\overline{\Omega} := \Omega \times \Theta$, où tous les éléments de $\overline{\Omega}$ sont notés par $\bar{\omega} = (\omega, \theta)$. Désignons par (B, T) avec $T = (T_1, \cdots, T_m)$ l'élément canonique sur $\overline{\Omega}$, *c.à.d.* $B_t(\bar{\omega}) := \omega_t$ et $T(\bar{\omega}) := \theta$ pour tout $\bar{\omega} = (\omega,\theta) \in \overline{\Omega}$. La filtration canonique est notée par $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$, où $\overline{\mathcal{F}}_t$ est générée par $(B_s)_{0\leq s\leq t}$ et tous les ensembles $\{T_k \leq s\}$ pour tout $s \in [0,t]$ et $k = 1,\cdots,m$. En particulier, toutes les variables aléatoires $T_1,\cdots,T_m$ sont les $\overline{\mathbb{F}}-$temps d'arrêt.

Soit $\overline{\mathcal{P}}(\overline{\Omega})$ l'espace de toutes les mesures de probabilité sur $\overline{\Omega}$, et on définit

$$\begin{aligned}\overline{\mathcal{P}} \;:=\; \Big\{ & \overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) : \; B \text{ est un } \overline{\mathbb{F}} - \text{mouvement brownien et} \\ & B_{T_m\wedge\cdot} := (B_{T_m\wedge t})_{t\geq 0} \text{ est uniformément intégrable sous } \overline{\mathbb{P}} \Big\}. \end{aligned} \tag{I.3.1}$$

Rappelons que $\mathbf{P}$ est l'espace des probabilités sur $\mathbb{R}$ avec le premier moment fini et $\mathbf{P}^m$ est son $m-$produit. Rappelons aussi qu'un vecteur $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m} \in \mathbf{P}^m$ est appelé un peacock si $\mu_k \preceq \mu_{k+1}$ est valable pour tout $k = 1, \cdots, m-1$. Un peacock est dit centré si chaque élément μ_k est de moyenne nulle. Pour un peacock centré $\boldsymbol{\mu}$, définissons l'ensemble des plongements

$$\overline{\mathcal{P}}(\boldsymbol{\mu}) \quad := \quad \Big\{ \overline{\mathbb{P}} \in \overline{\mathcal{P}} : \ \overline{\mathbb{P}} \circ B_{T_k}^{-1} \ = \ \mu_k \text{ pour tout } k = 1, \cdots, m \Big\}. \tag{I.3.2}$$

Soit $\Phi : \overline{\Omega} \to \mathbb{R}$ une fonction mesurable, alors Φ est dite non-anticipative si $\Phi(\omega, \theta) = \Phi(\omega_{\theta_m \wedge \cdot}, \theta)$ pour tout $(\omega, \theta) \in \overline{\Omega}$. Le plongement de Skorokhod optimal est défini par

$$\overline{P}(\boldsymbol{\mu}) \quad := \quad \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})} \ \mathbb{E}^{\overline{\mathbb{P}}} \Big[\Phi(B, T) \Big]. \tag{I.3.3}$$

Ici on suppose que $\mathbb{E}^{\overline{\mathbb{P}}}[|\Phi|] < +\infty$ pour toute $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$.

Remarque I.3.1. (i) *Sous chaque $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$, le vecteur des temps d'arrêt $(T_1, \cdots, T_m)$ est un plongement de Skorokhod. Notons que les temps d'arrêt ne sont pas requis d'être par rapport à la filtration brownienne $\mathbb{F}^0$, mais à la filtration $\overline{\mathbb{F}}$ qui est plus large. C'est pour cela on les appelle aussi plongements faibles.*

(ii) *D'après le lien avec le transport optimal martingale dans la Section 3.1.1, chaque mesure $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ peut aussi être vue comme un modèle du marché vérifiant l'hypothèse de non-arbitrage, où m désigne le nombre de maturités. Voir Section 3.2.2 pour plus de détailles.*

(iii) *Récement, Källblad, Tan & Touzi [69] ont généralisé les résultats de dualité au cas d'un nombre infini de marginales dans le cadre des options lookback.*

3.2 Dualités du plongement de Skorokhod optimal

Inspiré par l'approche qu'on utilise pour le problème de transport optimal martingale, on introduit ici deux problèmes duaux et prouve les dualités correspondantes. En utilisant les résultats de dualité, on dérive aussi les dualités pour le transport optimal martingale dans l'espace des fonctions continues. En plus, avec les dualités obtenues, on prouve dans la Section 3.3 le principe de monotonie qui caractérise les optimiseurs du plongement de Skorokhod optimal, et dans la Section 3.4 un résultat de stabilité.

3.2.1 Résultats de dualité

Nous nous intéressons à démontrer deux dualités au cas d'un nombre fini de marginales. Tout d'abord, suivant l'approche de la dualité convexe, nous convertissons le plongement de Skorokhod optimal à un infimum de problèmes d'arrêt optimal classiques, où la démarche ici s'inspire grandement du transport optimal martingale, et nous obtenons la première dualité. Ensuite, on applique l'approche de la programmation dynamique pour obtenir la deuxième dualité requise.

Premier problème dual Soit $\mathcal{T}^a$ la collection de toutes les familles croissantes de $\mathbb{F}^a-$temps d'arrêt $\tau = (\tau_k)_{1\le k\le m}$ telles que le processus arrêté $B_{\tau_m\wedge\cdot}$ est uniformément intégrable. Rappelons que Λ est l'espace des fonctions continues sur $\mathbb{R}$ à croissance linéaire et Λ^m est son $m-$produit.

Pour $\boldsymbol{\lambda} = (\lambda_k)_{1\le k\le m} \in \Lambda^m$ et $(\omega,\theta) \in \overline{\Omega}$, on définit

$$\boldsymbol{\lambda}(\omega,\theta) \ := \ \sum_{k=1}^m \lambda_k(\omega_{\theta_k}) \quad \text{et} \quad \boldsymbol{\mu}(\boldsymbol{\lambda}) \ := \ \sum_{k=1}^m \mu_k(\lambda_k).$$

Alors le premier problème dual de (I.3.3) est donné par

$$\overline{D}_0(\boldsymbol{\mu}) \ := \ \inf_{\boldsymbol{\lambda}\in\Lambda^m} \left\{ \boldsymbol{\mu}(\boldsymbol{\lambda}) \ + \ \sup_{\tau\in\mathcal{T}^a} \ \mathbb{E}^{\mathbb{P}_0}\Big[\Phi(B,\tau) \ - \ \boldsymbol{\lambda}(B,\tau)\Big]\right\}. \tag{I.3.4}$$

Deuxième problème dual Quant au deuxième problème dual, on revient à l'espace $\overline{\Omega}$. Etant donné $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$, un processus $\overline{\mathbb{F}}-$optionnel $S : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ est appelé une $\overline{\mathbb{P}}-$surmartingale forte si l'inégalité

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[S_{\tau_2}\middle|\overline{\mathcal{F}}_{\tau_1}\right] \ \le \ S_{\tau_1}, \ \overline{\mathbb{P}}\text{ - p.s.}$$

est vraie pour tous les $\overline{\mathbb{F}}-$temps d'arrêt $\tau_1 \le \tau_2$. Soit $\mathbb{L}^2_{loc}$ l'espace des processus $\overline{\mathbb{F}}-$progressivement mesurables $\overline{H} = (\overline{H}_t)_{t\ge 0}$ tels que

$$\int_0^t \overline{H}_s^2 ds \ < \ +\infty \text{ pour tout } t \in \mathbb{R}_+, \ \overline{\mathbb{P}}\text{ - p.s. pour toute } \overline{\mathbb{P}} \in \overline{\mathcal{P}}.$$

On introduit ensuite un sous-ensemble de $\mathbb{L}^2_{loc}$ défini par

$$\overline{\mathcal{H}} \;:=\; \Big\{\overline{H} \in \mathbb{L}^2_{loc}:\; (\overline{H}\cdot B)_{T_m\wedge\cdot} \;:=\; \int_0^{T_m\wedge\cdot} \overline{H}_t dB_t \text{ est une } \overline{\mathbb{P}}\text{ - surmartingale forte pour toute } \overline{\mathbb{P}} \in \overline{\mathcal{P}}\Big\}.$$

Définissons

$$\overline{\mathcal{D}} \;:=\; \Big\{(\boldsymbol{\lambda}, \overline{H}) \in \Lambda^m \times \overline{\mathcal{H}}:\; \boldsymbol{\lambda}(B,T) \;+\; (\overline{H}\cdot B)_{T_m} \;\geq\; \Phi(B,T),\; \overline{\mathbb{P}}\text{ - p.s. pour toute } \overline{\mathbb{P}} \in \overline{\mathcal{P}}\Big\},$$

et le deuxième problème dual est donné par

$$\overline{D}(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda},\overline{H})\in\overline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{I.3.5}$$

Le résultat principal est le suivant :

Théorème I.3.2. *Sous des conditions appropriées, il existe $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ telle que*

$$\mathbb{E}^{\overline{\mathbb{P}}^*}\Big[\Phi(B,T)\Big] \;=\; \overline{P}(\boldsymbol{\mu}) \;=\; \overline{D}_0(\boldsymbol{\mu}) \;=\; \overline{D}(\boldsymbol{\mu}).$$

Remarque I.3.3. *Les résultats de dualité sont obtenus sous les conditions plus générales que celles dans [5], ce qui contient un exemple important, où Φ dépend du temps local de B en zéro, voir aussi Cox, Hobson & Obłój [23] et Claisse, Guo & Henry-Labordère [20].*

L'exemple suivant montre que notre formulation faible n'est généralement pas équivalente à la formulation forte, où tous les temps d'arrêt du plongement sont par rapport à une filtration brownienne.

Exemple I.3.4. *Pour $m = 1$, prenons $\mu = \delta_{\{0\}}/3 + \delta_{\{1\}}/3 + \delta_{\{-1\}}/3$ et $\Phi(\omega,\theta) = \mathbb{1}_{\{0\}}(\theta)$. Définissons*

$$\tau_0 \;:=\; \inf\Big\{t \in \mathbb{R}_+:\; |B_t| \geq 1\Big\} \quad et \quad \overline{\mathbb{P}}_0 \;:=\; \frac{1}{3}\mathbb{P}_0 \circ (B,0)^{-1} \;+\; \frac{2}{3}\mathbb{P}_0 \circ (B,\tau_0)^{-1},$$

alors $\overline{\mathbb{P}}_0 \in \mathcal{P}(\mu)$ et $\mathbb{E}^{\overline{\mathbb{P}}_0}[\Phi(B,T)] = 1/3$. En outre, pour tout $\tau \in \mathcal{T}^a$ tel que $\mathbb{P}_0 \circ B_\tau^{-1} = \mu$, on a $\mathbb{P}_0[\tau > 0] > 0$. Vu que la filtration brownienne augmentée vérifie la condition zéro-un de Blumenthal, on obtient $\mathbb{P}_0[\tau > 0] = 1$, ce qui implique que

$$\sup_{\tau\in\mathcal{T}^a:\; B_\tau\sim\mu} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi(B,\tau)\Big] \;=\; 0 \;<\; \frac{1}{3} \;\leq\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big].$$

Ensuite, on donne un exemple où la dualité n'est plus valable lorsque l'hypothèse de semicontinuité supérieure n'est pas vérifiée.

Exemple I.3.5. *Pour $m = 1$, prenons $\mu = \delta_{\{1\}}/2 + \delta_{\{-1\}}/2$ et $\Phi(\omega,\theta) = \mathbb{1}_{\mathbb{Q}}(\theta)$. On note tout d'abord que $\overline{\mathcal{P}}(\mu)$ ne contient qu'un élément qui est la probabilité induite par (B,τ_0), où $\tau_0 := \inf\{t \in \mathbb{R}_+ : |B_t| \geq 1\}$. En effet, pour toute $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)$, on a $\mathbb{E}^{\overline{\mathbb{P}}}[T] = \mathbb{E}^{\overline{\mathbb{P}}}[B_T^2] = \mathbb{E}^{\overline{\mathbb{P}}}[\tau_0]$ et $T \geq \tau_0$, $\overline{\mathbb{P}}$ - p.s. En plus, notons que le temps d'attaint τ_0 a une distribution sans atome sur $\mathbb{R}_+$, alors*

$$\overline{P}(\mu) \;=\; \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big] \;=\; \mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\mathbb{1}_{\mathbb{Q}}(\tau_0)\Big] \;=\; 0.$$

Quant aux problèmes duaux, comme $\lambda \in \Lambda$ est continue, on a par le fait que chaque temps d'arrêt peut être approximé par des temps d'arrêt à valeurs dans $\mathbb{Q}$

$$\sup_{\tau \in \mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\mathbb{1}_{\mathbb{Q}}(\tau) - \lambda(B_\tau)\Big] \;=\; \sup_{\tau \in \mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[1 - \lambda(B_\tau)\Big] \text{ pour toute } \lambda \in \Lambda.$$

Donc, il résulte de la Définition (I.3.4) que $\overline{D}_0(\mu) = \inf_{\lambda\in\Lambda}\{\mu(\lambda) + \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}[1-\lambda(B_\tau)]\} = 1$. Suivant le même raisonnement, on peut déduire $\overline{D}(\mu) = 1$, ce qui donne l'écart de dualité :

$$\overline{P}(\mu) \;=\; 0 \;<\; 1 \;=\; \overline{D}_0(\mu) \;=\; \overline{D}(\mu).$$

3.2.2 Application au transport optimal martingale

Compte tenu du fait que toute martingale continue partant de zéro est un mouvement brownien changeant du temps, on montre que le Théorème I.3.2 implique les dualités pour une classe de problèmes de transport optimal martingale dans l'espace des fonctions continues.

Rappelons que $\mathbf{C}$ est l'espace de fonctions continues $\omega = (\omega_t)_{0\leq t\leq 1}$ telles que $\omega_0 = 0$. On définit le processus canonique $X = (X_t)_{0\leq t\leq 1}$, *c.à.d.* $X_t(\omega) := \omega_t$, et sa filtration naturelle $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq 1}$, *c.à.d.* $\mathcal{F}_t := \sigma(X_s, s \leq t)$. Désignons en plus par $\mathcal{M}$ la collection de toutes les mesures martingales $\mathbb{P}$, *c.à.d.* mesures de probabilité sur $(\mathbf{C}, \mathcal{F}_1)$ sous lesquelles X est une martingale. Soient $\mathbb{T} = \{t_1 < \cdots < t_m = 1\} \subseteq (0,1]$ un ensemble fini et $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$ un peacock centré, on définit l'ensemble des plans de transport martingale

$$\mathcal{M}(\boldsymbol{\mu}) \;:=\; \Big\{\mathbb{P} \in \mathcal{M} :\; \mathbb{P}\circ X_{t_k}^{-1} \;=\; \mu_k \text{ pour tout } k = 1,\cdots,m\Big\}.$$

Pour une fonction $\mathcal{F}_1-$mesurable $\xi : \mathbf{C} \to \mathbb{R}$, le problème de transport optimal martingale est défini par

$$\mathsf{P}(\boldsymbol{\mu}) \;:=\; \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big], \tag{I.3.6}$$

où ξ est supposée de vérifier $\mathbb{E}^{\mathbb{P}}[|\xi|] < +\infty$ pour toute $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$. On donne ensuite les problèmes duaux correspondants. Soit $\mathcal{H}$ la collection de tous les processus $\mathbb{F}-$progressivement mesurables $H : [0,1]\times\mathbf{C} \to \mathbb{R}$ tels que

$$\int_0^1 H_t^2 d\langle X\rangle_t \;<\; +\infty,\ \mathbb{P}\text{ - p.s. et l'intégrale stochastique } (H\cdot X) \;:=\; \int_0^{\cdot} H_t dX_t$$
$$\text{est une } \mathbb{P}-\text{surmartingale pour toute } \mathbb{P}\in\mathcal{M}. \tag{I.3.7}$$

Ici $\langle X\rangle$ est un processus croissant et $\mathbb{F}-$progressivement mesurable, qui prend la valeur dans $[0,+\infty]$, tel que $\langle X\rangle$ coïncide avec la variation quadratique de X, $\mathbb{P}$ - presque-sûre pour toute $\mathbb{P}\in\mathcal{M}$. En fait, l'existence de cette version universelle de la variation quadratique découle du résultat d'agrégation des intégrales stochastiques de Karandikar [70]. Alors les deux problèmes duaux sont donnés par

$$\mathsf{D}_0(\boldsymbol{\mu}) \;:=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m}\left\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}}\mathbb{E}^{\mathbb{P}}\;\Big[\xi\Big(X\Big)-\boldsymbol{\lambda}(X)\Big]\right\} \;\text{ et }\; \mathsf{D}(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda},H)\in\mathcal{D}} \boldsymbol{\mu}(\boldsymbol{\lambda}), \tag{I.3.8}$$

où

$$\boldsymbol{\lambda}(X) \;:=\; \sum_{i=1}^m \lambda_i(X_{t_i})$$

et

$$\mathcal{D} \;:=\; \Big\{(\boldsymbol{\lambda},H)\in\Lambda^m\times\mathcal{H}:\; \boldsymbol{\lambda}(X) \;+\; (H\cdot X)_1 \;\geq\; \xi(X),\ \mathbb{P}\text{ - p.s. pour toute } \mathbb{P}\in\mathcal{M}\Big\}.$$

Soient $\langle X\rangle_t^{-1} := \inf\{s\in\mathbb{R}_+ : \langle X\rangle_s > t\}$ et

$$W_t \;:=\; X_{\langle X\rangle_t^{-1}}\mathbb{1}_{\{t<\langle X\rangle_1\}} \;+\; \Big(X_1 \;+\; \widehat{W}_{t-\langle X\rangle_1}\Big)\mathbb{1}_{\{t\geq\langle X\rangle_1\}} \text{ pour tout } t\in\mathbb{R}_+,$$

où $\widehat{W}$ est un mouvement brownien indépendant de $\mathcal{F}_1$. Il résulte du théorème de Dambis-Dubins-Schwarz que, le processus $W=(W_t)_{t\geq 0}$ est un mouvement brownien, voir Revuz & Yor [92, Théorème 1.7, Chapitre V].

Le théorème suivant montre que, à l'aide du plongement de Skorokhod optimal, on peut établir les dualités de transport optimal martingale avec les trajectoires continues.

Théorème I.3.6. *Supposons que* $\xi(X) = \Phi(W, \langle X\rangle_{t_1}, \cdots, \langle X\rangle_{t_m})$ *pour une fonction mesurable* $\Phi : \overline{\Omega} \to \mathbb{R}$*. Alors sous des conditions appropriées on a*

$$\mathsf{P}(\boldsymbol{\mu}) \quad = \quad \mathsf{D}_0(\boldsymbol{\mu}) \quad = \quad \mathsf{D}(\boldsymbol{\mu}).$$

3.3 Principe de monotonie du plongement de Skorokhod optimal

3.3.1 Motivation

Comme introduit au début de la Section 3.1, certaines solutions du plongement de Skorokhod dans la littérature existante jouissent d'une propriété d'optimalité par rapport à un critère. Beiglböck, Cox & Huesmann [5] ont introduit le principe de monotonie dans le cas d'une marginale, ce qui permet de caractériser les plongements optimaux. Grace à ce principe, tous les plongements optimaux connus peuvent être interprétés uniformément.

La contribution ici est de fournir une preuve alternative et simplifiée du principe de monotonie, qui est basée sur les dualités obtenues dans la Section 3.2.1. Notre preuve suit une application délicate du théorème de section optionnelle ainsi que l'argument remarquable de conditionnement introduit dans [5]. Nous soulignons que nous nous concentrons sur le cas d'une marginale dans la Section 3.3.

3.3.2 Résultats principaux

Etant donné un ensemble $\overline{\Gamma} \subseteq \overline{\Omega}$, on définit $\overline{\Gamma}^{<}$ par

$$\overline{\Gamma}^{<} \quad := \quad \Big\{\bar{\omega} = (\omega, \theta) \in \overline{\Omega} : \ \bar{\omega}_{\theta\wedge\cdot} \ = \ \bar{\omega}'_{\theta\wedge\cdot} \text{ pour certain } \bar{\omega}' \in \overline{\Gamma} \text{ avec } \theta' > \theta\Big\}.$$

Pour tout $\bar{\omega} = (\omega, \theta)$ et $\bar{\omega}' = (\omega', \theta') \in \overline{\Omega}$, la concaténation $\bar{\omega} \otimes \bar{\omega}' \in \overline{\Omega}$ est définie par

$$\bar{\omega} \otimes \bar{\omega}' \quad := \quad \Big(\omega \otimes_\theta \omega', \ \theta + \theta'\Big),$$

où

$$(\omega \otimes_\theta \omega')_t \quad := \quad \omega_t \mathbb{1}_{[0,\theta)}(t) \ + \ \Big(\omega_\theta \ + \ \omega'_{t-\theta}\Big)\, \mathbb{1}_{[\theta,+\infty)}(t) \text{ pour tout } t \in \mathbb{R}_+.$$

Définition I.3.7. *Une paire* $(\bar{\omega}, \bar{\omega}') \in \overline{\Omega} \times \overline{\Omega}$ *est appelée une paire stop-go si* $\omega_\theta = \omega'_{\theta'}$ *et*

$$\Phi(\bar{\omega}) \;+\; \Phi(\bar{\omega}' \otimes \bar{\omega}'') \;>\; \Phi(\bar{\omega} \otimes \bar{\omega}'') \;+\; \Phi(\bar{\omega}') \textit{ pour tout } \bar{\omega}'' \in \overline{\Omega}^{+},$$

où $\overline{\Omega}^{+} := \{\bar{\omega} = (\omega, \theta) \in \overline{\Omega} : \theta > 0\}$*. Soit* SG *l'ensemble de toutes les paires stop-go.*

Théorème I.3.8. *Soit* $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\mu)$ *un optimiseur du problème* (I.3.3)*. Alors, sous des conditions appropriées, il existe un ensemble borélien* $\overline{\Gamma}^* \subseteq \overline{\Omega}$ *tel que*

$$\overline{\mathbb{P}}^* \left[\overline{\Gamma}^*\right] \;=\; 1 \quad \textit{et} \quad \mathrm{SG} \;\cap\; \left(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\right) \;=\; \emptyset.$$

Exemple I.3.9. *Il est connu que la solution de Root* $\overline{\mathbb{P}}_{Root}$ *résout le problème d'optimisation, c.à.d., étant donné une fonction strictement concave* $h : \mathbb{R}_+ \to \mathbb{R}$*, il existe une barrière* $\mathcal{R} \subseteq \mathbb{R}_+ \times \mathbb{R}$ *telle que*

$$\overline{P}(\mu) \;=\; \mathbb{E}^{\overline{\mathbb{P}}_{Root}}[h(B, T)] \quad \textit{avec} \quad T \;=\; \inf\left\{t \in \mathbb{R}_+ :\; (t, B_t) \in \mathcal{R}\right\}, \; \overline{\mathbb{P}}_{Root}\textit{ - p.s.}$$

Ici un ensemble borélien $\mathcal{R} \subseteq \mathbb{R}_+ \times \mathbb{R}$ *est dit une barrière si* $(s, x) \in \mathcal{R}$ *et* $s < t$ *impliquent* $(t, x) \in \mathcal{R}$*. Nous montrons ici que le principe de monotonie permet de retrouver la solution de Root. D'après le Théorème I.3.8, prenons l'ensemble* $\overline{\Gamma}^*$ *tel que*

$$(B, T) \;\in\; \overline{\Gamma}^*, \; \overline{\mathbb{P}}_{Root}\textit{ - p.s.} \quad \textit{et} \quad \mathrm{SG} \;\cap\; \left(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\right) \;=\; \emptyset.$$

En outre, la concavité de h *implique que*

$$\mathrm{SG} \;=\; \left\{\left((\omega, \theta), (\omega', \theta')\right) \in \overline{\Omega} \times \overline{\Omega} :\; \omega_\theta = \omega'_{\theta'} \textit{ et } \theta < \theta'\right\}.$$

Définissons la barrière par

$$\mathcal{R} \;:=\; \left\{(t, x) \in \mathbb{R}_+ \times \mathbb{R} :\; \textit{il existe } (\omega, \theta) \in \overline{\Gamma}^* \textit{ tel que } \omega_\theta = x \textit{ et } t < \theta\right\},$$

alors il résulte de $\mathrm{SG} \cap (\overline{\Gamma}^{*<} \times \overline{\Gamma}^*) = \emptyset$ *que* $\mathcal{R}$ *est bien la barrière requise.*

En fait le résultat démontré est un peu plus fort dans le sens suivant. Pour l'optimiseur $\overline{\mathbb{P}}^*$, on peut trouver un ensemble SG^* dépendant de $\overline{\mathbb{P}}^*$ tel que

$$\mathrm{SG} \;\cap\; \left(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\right) \;\subseteq\; \mathrm{SG}^* \;\cap\; \left(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\right) \;=\; \emptyset.$$

D'après Stroock & Varadhan [98], $\overline{\mathbb{P}}^*$ admet une famille de distributions de probabilité conditionnelles régulières $(\overline{\mathbb{P}}^*_{\bar\omega})_{\bar\omega\in\overline{\Omega}}$ par rapport à $\overline{\mathcal{F}}^B := \sigma(B_t, t \geq 0)$ qui est la filtration engendrée par B on $\overline{\Omega}$. Notons que pour $\bar\omega = (\omega,\theta)$, la mesure $\overline{\mathbb{P}}^*_{\bar\omega}$ est indépendante de θ, alors nous pouvons désigner cette famille par $(\overline{\mathbb{P}}^*_{\omega})_{\omega\in\Omega}$. Ensuite, pour chaque $\bar\omega\in\overline{\Omega}$, on définit une probabilité $\overline{\mathbb{Q}}^1_{\bar\omega}$ sur $\overline{\Omega}$ par

$$\overline{\mathbb{Q}}^1_{\bar\omega}[\overline{A}] \;:=\; \int_\Omega \overline{\mathbb{P}}^*_{\omega\otimes_\theta\omega'}[\overline{A}]\;\mathbb{P}_0(d\omega') \text{ pour tout } \overline{A}\in\overline{\mathcal{F}} \;:=\; \bigvee_{t\geq 0}\overline{\mathcal{F}}_t. \tag{I.3.9}$$

On définit en outre, pour tout $\bar\omega\in\overline{\Omega}$, une probabilité $\overline{\mathbb{Q}}^2_{\bar\omega}$ par

$$\overline{\mathbb{Q}}^2_{\bar\omega}[\overline{A}] \;:=\; \overline{\mathbb{Q}}^1_{\bar\omega}\left[\overline{A}\middle|T>\theta\right]\mathbb{1}_{\{\overline{\mathbb{Q}}^1_{\bar\omega}[T>\theta]>0\}} \;+\; \mathbb{P}^{\theta,\omega}_0\otimes\delta_{\{\theta\}}[\overline{A}]\mathbb{1}_{\{\overline{\mathbb{Q}}^1_{\bar\omega}[T>\theta]=0\}} \tag{I.3.10}$$

pour tout $\overline{A}\in\overline{\mathcal{F}}$, où $\mathbb{P}^{\theta,\omega}_0$ est la mesure de Wiener décalée sur Ω donnée par

$$\mathbb{P}^{\theta,\omega}_0[A] \;:=\; \mathbb{P}_0\left[\omega\otimes_\theta B\in A\right] \text{ pour tout } A\in\mathcal{F}^0 \;:=\; \bigvee_{t\geq 0}\mathcal{F}^0_t.$$

On introduit finalement une probabilité décalée $\overline{\mathbb{Q}}^*_{\bar\omega}$ par

$$\overline{\mathbb{Q}}^*_{\bar\omega}[\overline{A}] \;:=\; \overline{\mathbb{Q}}^2_{\bar\omega}\left[\bar\omega\otimes(B,T)\in\overline{A}\right] \text{ pour tout } \overline{A}\in\overline{\mathcal{F}},$$

et puis définit un nouvel ensemble SG^* par

$$\mathrm{SG}^* \;:=\; \left\{(\bar\omega,\bar\omega') :\; \omega_\theta=\omega'_{\theta'} \text{ et } \Phi(\bar\omega)+\mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\left[\Phi(\bar\omega'\otimes\cdot)\right] \;>\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\left[\Phi(\bar\omega\otimes\cdot)\right]+\Phi(\bar\omega')\right\}. \tag{I.3.11}$$

Enfin, on obtient une caractérisation plus précise du support de $\overline{\mathbb{P}}^*$.

Théorème I.3.10. *Sous les mêmes conditions de Théorème I.3.8, il existe un ensemble borélien $\overline{\Gamma}^*\subseteq\overline{\Omega}$ tel que*

$$\overline{\mathbb{P}}^*\left[\overline{\Gamma}^*\right] \;=\; 1 \quad \textit{et} \quad \mathrm{SG}^* \;\cap\; \left(\overline{\Gamma}^{*<}\times\overline{\Gamma}^*\right) \;=\; \emptyset.$$

Exemple I.3.11. *Grâce au Théorème I.3.10, la solution de Vallois peut être reconstruit, c.à.d. si $\Phi(B,T)=F(L^B_T)$, où $F:\mathbb{R}_+\to\mathbb{R}$ est une fonction convexe et $L^B=(L^B_t)_{t\geq0}$ est le temps local du mouvement brownien en zéro, alors il existe une fonction croissante $\phi_+:\mathbb{R}_+\to\mathbb{R}_+$ et une fonction décroissante $\phi_-:\mathbb{R}_+\to\mathbb{R}_-$ telles que*

$$T \;=\; \inf\left\{t\in\mathbb{R}_+ :\; B_t\notin\left(\phi_-(L^B_t),\phi_+(L^B_t)\right)\right\},\; \overline{\mathbb{P}}_{Val}\text{ - p.s.}$$

où $\overline{\mathbb{P}}_{Val}$ est l'optimiseur du problème (I.3.3).

3.4 Stabilité du plongement de Skorokhod optimal

3.4.1 Motivation

On se rappelle que, sous l'hypothèse de non-arbitrage, les contraintes marginales dans le problème du plongement de Skorokhod optimal viennent de l'hypothèse que les prix des calls sont connus pour tous les strikes dans le marché. La situation plus générale est qu'il n'y qu'un nombre fini de calls dans le marché. Par conséquent, on considère un autre problème d'optimisation similaire au plongement de Skorokhod optimal, où le mouvement brownien arrêté n'a besoin de reproduire qu'un nombre fini de prix des calls.

Soient $\boldsymbol{K} = (K_i)_{1\le i\le n}$ le vecteur de strikes avec $K_1 < \cdots < K_n$ qui sont disponibles dans le marché et $\boldsymbol{C} = (C_{i,j})_{1\le i\le m, 1\le j\le n}$ la matrice de prix des calls correspondante, où m et n désignent respectivement le nombre de maturités et le nombre de strikes. Sous l'hypothèse de non-arbitrage, il existe un peacock centré $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m}$ tel que $C_{i,j} = \mu_i\left((x-K_j)^+\right)$ pour tout $i = 1,\cdots,m$ et $j = 1,\cdots,n$. Rappelons que

$$\overline{\mathcal{P}} \;:=\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) :\; B \text{ est un } \overline{\mathbb{F}} - \text{mouvement brownien et} \\ B_{T_m\wedge\cdot} \text{ est uniformément intégrable sous } \overline{\mathbb{P}}\Big\}.$$

Une probabilité $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$ est appelée un $(\boldsymbol{K},\boldsymbol{C})-$plongement si

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[(B_{T_i} - K_j)^+\right] \;=\; C_{i,j} \text{ pour tout } i = 1,\cdots,m \text{ et } j = 1,\cdots,n. \tag{I.3.12}$$

Soit $\overline{\mathcal{P}}(\boldsymbol{K},\boldsymbol{C})$ la collection des $(\boldsymbol{K},\boldsymbol{C})-$plongements. De même, nous pouvons définir le problème du plongement optimal par

$$\overline{P}(\boldsymbol{K},\boldsymbol{C}) \;:=\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\boldsymbol{K},\boldsymbol{C})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big], \tag{I.3.13}$$

où on suppose que $\mathbb{E}^{\overline{\mathbb{P}}}[|\Phi|] < +\infty$ pour toute $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$.

3.4.2 Résultats principaux

Notons que l'existence de $\boldsymbol{\mu}$ est garantie par l'hypothèse de non-arbitrage, néanmoins, $\boldsymbol{\mu}$ est inconnu puisqu'on ne peut pas le déterminer par un nombre fini de prix. L'objective ici est de capturer $\overline{P}(\boldsymbol{\mu})$ par $\overline{P}(\boldsymbol{K},\boldsymbol{C})$ et d'étudier le comportement asymptotique de $\overline{P}(\boldsymbol{K},\boldsymbol{C})$.

La première question appraît immédiatement : Etant donné de plus en plus de calls, est-ce que le problème d'optimisation $\overline{P}(\boldsymbol{K},\boldsymbol{C})$ converge vers $\overline{P}(\boldsymbol{\mu})$? Alors le théorème suivant nous garantit que notre problème d'optimisation est cohérent à notre intuition.

Théorème I.3.12. *Soit $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m}$ un peacock centré fixé. Sous des conditions appropriées on a*

$$\lim_{n\to\infty} \overline{P}(\boldsymbol{K}^n,\boldsymbol{C}^n) \;\; = \;\; \overline{P}(\boldsymbol{\mu}),$$

pour toute suite $(\boldsymbol{K}^n,\boldsymbol{C}^n)_{n\ge 1}$ telle que $\cup_{n\ge 1}\boldsymbol{K}^n$ soit dense dans $\mathbb{R}$, où $C^n_{i,j} = \mu_i((x-K^n_j)^+)$ pour tout $i=1,\cdots,m$ et $j=1,\cdots,n$.

En suite, on s'intéresse à estimer la vitesse de convergence au moyen de métriques différentes sur l'espace des mesures de probabilité. On suppose désormais que μ_m admet le moment d'ordre p fini avec $p>1$, *c.à.d.* $\mu_m(|x|^p) \equiv V < +\infty$. Alors la proposition suivante implique que l'application $\boldsymbol{\mu} \mapsto \overline{P}(\boldsymbol{\mu})$ est höldérienne par rapport aux métriques de Lévy-Prokhorov et Wasserstein.

Soit $q>1$ le nombre conjugué de p, *c.à.d.* $1/p+1/q=1$. On note respectivement par $\rho(\cdot,\cdot)$ la métrique de Lévy-Prokhorov et par $\mathcal{W}_1(\cdot,\cdot)$ la métrique de Wasserstein, *c.à.d.* pour toutes les deux mesures de probabilité μ et $\nu \in \mathbf{P}$, on a

$$\begin{aligned}
\rho(\mu,\nu) &:= \inf\Big\{\varepsilon>0:\; F_\mu(x-\varepsilon)-\varepsilon \le F_\nu(x) \le F_\mu(x+\varepsilon)+\varepsilon \text{ pour tout } x\in\mathbb{R}\Big\},\\
\mathcal{W}_1(\mu,\nu) &:= \inf_{\mathbb{P}\in\mathcal{P}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}\Big[|X-Y|\Big],
\end{aligned}$$

où F_μ (resp. F_ν) représente la fonction de répartition de μ (resp. ν) et $\mathcal{P}(\mu,\nu)$ est défini par (I.2.2).

Proposition I.3.13. *Sous des conditions appropriées, il existe une constante $C>0$ telle que*
(i) *si $m=1$*

$$\Big|\overline{P}(\mu) \;-\; \overline{P}(\nu)\Big| \;\;\le\;\; C\rho(\mu,\nu)^{1/2q};$$

(ii) *si $m\ge 1$*

$$\Big|\overline{P}(\boldsymbol{\mu}) \;-\; \overline{P}(\boldsymbol{\nu})\Big| \;\;\le\;\; C\sum_{i=1}^m \mathcal{W}_1(\mu_i,\nu_i)^{\frac{p-2}{p-1}}.$$

Pour analyser la vitesse de convergence, on considère un sous-ensemble $\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C})$ défini par

$$\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C}) \;:=\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{K},\boldsymbol{C}) :\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[|B_{T_m}|^p\Big] \;=\; V\Big\}.$$

Cette restriction provient d'une information complémentaire : une option de la fonction de paiement $|B_{T_m}|^p$ est observée dans le marché. Définissons de manière similaire

$$\overline{P}^V(\boldsymbol{K},\boldsymbol{C}) \;:=\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big].$$

On obtient l'estimation de la vitesse de convergence suivante.

Théorème I.3.14. *Sous des conditions appropriées, il existe une constante $C > 0$ telle que*
(i) *si $m = 1$*

$$0 \;\le\; \overline{P}^V(\boldsymbol{K}^n,\boldsymbol{C}^n) \;-\; \overline{P}(\mu) \;\le\; C\left((\Delta\boldsymbol{K}^n)^{1/4q} \;+\; |\boldsymbol{K}^n|^{-p/4q^2}\right);$$

(ii) *si $m \ge 1$*

$$0 \;\le\; \overline{P}^V(\boldsymbol{K}^n,\boldsymbol{C}^n) \;-\; \overline{P}(\boldsymbol{\mu}) \;\le\; C|\boldsymbol{K}^n|^{\frac{p-2}{p-1}}\left(\sqrt{\Delta\boldsymbol{K}^n} \;+\; |\boldsymbol{K}^n|^{-p/2q}\right)^{\frac{p-2}{p-1}},$$

où

$$|\boldsymbol{K}^n| \;:=\; (K_1^n)^- \wedge (K_n^n)^+ \quad et \quad \Delta\boldsymbol{K}^n \;:=\; \max_{1<i\le n}\left(K_i^n - K_{i-1}^n\right).$$

4 Approche contrôle stochastique : options sur temps local

4.1 Motivation et formulation

On considère l'application de contrôle stochastique au transport optimal martingale dans l'espace des fonctions continues avec la fonction de paiement dépendant uniquement du temps local, et au plongement de Skorokhod. Rappelons que $\mathbf{C}$ est l'espace des functions continues $\omega = (\omega_t)_{0\le t\le 1}$ avec $\omega_0 = 0$. Comme les définitions dans la Section 3.2.2, $X = (X_t)_{0\le t\le 1}$ est le processus canonique, $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ est sa filtration naturelle et $\mathcal{M}$ désigne l'ensemble de toutes les mesures martingales $\mathbb{P}$ sur $(\mathbf{C},\mathcal{F}_1)$. On fixe un peacock centré $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m}$ et un ensemble de maturités $\mathbb{T} = \{t_1 < \cdots < t_m = 1\} \subseteq (0,1]$, et considère la collection des plans de transport martingale

$$\mathcal{M}(\boldsymbol{\mu}) \;\; := \;\; \Big\{ \mathbb{P} \in \mathcal{M} : \;\; \mathbb{P} \circ X_{t_k}^{-1} \;\; = \;\; \mu_k \text{ pour tout } k = 1, \cdots, m \Big\}$$

Soit $L = (L_t)_{0 \le t \le 1}$ le temps local du processus X en zéro, qui est défini du point de vue des semimartingales. Pour une fonction convexe $F : \mathbb{R}_+ \to \mathbb{R}$, on considère les problèmes de transport optimal martingale

$$\overline{\mathsf{P}}(\boldsymbol{\mu}) \;\; := \;\; \sup_{\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[F(L_1)\Big], \tag{I.4.1}$$

$$\underline{\mathsf{P}}(\boldsymbol{\mu}) \;\; := \;\; \inf_{\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[F(L_1)\Big]. \tag{I.4.2}$$

Soit $\overline{\mathcal{H}}$ (resp. $\underline{\mathcal{H}}$) l'ensemble des processus $\mathbb{F}-$progressivement mesurables $H : [0, 1] \times \mathbb{C} \to \mathbb{R}$ tels que,

$$\int_0^1 H_t^2 d\langle X \rangle_t \;\; < \;\; +\infty, \;\; \mathbb{P}\text{ - p.s. et l'intégrale stochastique } (H \cdot X) \;\; := \;\; \int_0^{\cdot} H_t dX_t$$
$$\text{est une } \mathbb{P} - \text{surmartingale (resp. sousmartingale) pour toute } \mathbb{P} \in \mathcal{M}. \tag{I.4.3}$$

Définissons en outre

$$\overline{\mathcal{D}} \text{ (resp. } \underline{\mathcal{D}}) \;\; := \;\; \Big\{ (\boldsymbol{\lambda}, H) \in \Lambda^m \times \overline{\mathcal{H}} \text{ (resp. } \underline{\mathcal{H}}) : \;\; \boldsymbol{\lambda}(X) \;\; + \;\; (H \cdot X)_1$$
$$\ge \;\; \text{(resp. } \;\le) \;\; F(L_1), \;\; \mathbb{P}\text{ - p.s. pour toute } \mathbb{P} \in \mathcal{M} \Big\} \tag{I.4.4}$$

Alors les problèmus duaux sont définis par

$$\overline{\mathsf{D}}(\boldsymbol{\mu}) \text{ (resp. } \underline{\mathsf{D}}(\boldsymbol{\mu})) \;\; := \;\; \inf_{(\boldsymbol{\lambda}, H) \in \overline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda}) \text{ (resp. } \sup_{(\boldsymbol{\lambda}, H) \in \underline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda})). \tag{I.4.5}$$

Remarque I.4.1. *Comme dans la Remarque I.2.5, chaque mesure martingale $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$ représente un modèle du marché vérifiant l'hypothèse de non-arbitrage. Les problèmes primitifs $\overline{\mathsf{P}}(\boldsymbol{\mu})$ et $\underline{\mathsf{P}}(\boldsymbol{\mu})$ sont les bornes supérieure et inférieure des prix déterminés sous tous les modèles. Les problèmes duaux $\overline{\mathsf{D}}(\boldsymbol{\mu})$ et $\underline{\mathsf{D}}(\boldsymbol{\mu})$ peuvent être interprétés comme les prix minimal de la surcouverture et maximal de la souscouverture.*

4.2 Résultats principaux

4.2.1 Cas d'une marginale : solutions de Vallois et dualités

Dans le cas d'une marginale, on produit les optimiseurs explicites pour les problèmes primitifs et duaux à l'aide des solutions de Vallois.

Soit $B = (B_t)_{t\geq 0}$ un mouvement brownien. D'après Vallois [102], il existe une fonction croissante (resp. décroissante) continue à droite $\overline{\phi}_+ : \mathbb{R}_+ \to \mathbb{R}_+$ (resp. $\underline{\phi}_+ : \mathbb{R}_+^* \to \mathbb{R}_+$) et une fonction décroissante (croissante) continue à droite $\overline{\phi}_- : \mathbb{R}_+ \to \mathbb{R}_-$ (resp. $\underline{\phi}_- : \mathbb{R}_+^* \to \mathbb{R}_-$) telles que, $B_{\overline{\tau}\wedge\cdot}$ (resp. $B_{\underline{\tau}\wedge\cdot}$) est uniformément intégrable et $B_{\overline{\tau}} \sim \mu$ (resp. $B_{\underline{\tau}} \sim \mu$), où

$$\overline{\tau} \;:=\; \inf\Big\{t>0 : B_t \notin \Big(\overline{\phi}_-(L_t^B), \overline{\phi}_+(L_t^B)\Big)\Big\} \;\Big(\text{resp. } \underline{\tau} \;:=\; \inf\Big\{t>0 : B_t \notin \Big(\underline{\phi}_-(L_t^B), \underline{\phi}_+(L_t^B)\Big)\Big\}\Big).$$

Ici $L^B = (L_t^B)_{t\geq 0}$ est le temps local de B en zéro. Définissons $\overline{\gamma} : \mathbb{R}_+ \to \mathbb{R}_+$ (resp. $\underline{\gamma} : \mathbb{R}_+ \to \mathbb{R}_+$) par $\overline{\gamma}(0) = 0$ (resp. $\underline{\gamma}(0) = 0$) et pour tout $l > 0$

$$\overline{\gamma}(l) \;:=\; \frac{1}{2}\int_0^l \left(\frac{1}{\overline{\phi}_+(s)} - \frac{1}{\overline{\phi}_-(s)}\right) ds \;\left(\text{resp. } \underline{\gamma}(l) \;:=\; \frac{1}{2}\int_0^l \left(\frac{1}{\underline{\phi}_+(s)} - \frac{1}{\underline{\phi}_-(s)}\right) ds\right)$$

On définit en outre

$$\overline{\lambda}'(\pm x) \;:=\; \overline{\lambda}'(0\pm) \;+\; \int_0^{\overline{\psi}_\pm(\pm x)} \frac{dy}{\overline{\phi}_\pm(y)} e^{\overline{\gamma}(y)} \int_y^{+\infty} e^{-\overline{\gamma}(z)} F''(dz) \text{ pour tout } x > 0, \quad \text{(I.4.6)}$$

$$\overline{\lambda}'(0\pm) \;:=\; \pm F'(0) \;\pm\; \int_0^{+\infty} e^{-\overline{\gamma}(z)} F''(dz), \quad \text{(I.4.7)}$$

$$\overline{\lambda}(0) \;:=\; F(0), \quad \text{(I.4.8)}$$

et

$$\underline{\lambda}'(\pm x) \;:=\; \underline{\lambda}'(0\pm) \;-\; \int_{\underline{\psi}_\pm(\pm x)}^{+\infty} \frac{dy}{\underline{\phi}_\pm(y)} e^{\underline{\gamma}(y)} \int_y^{+\infty} e^{-\underline{\gamma}(z)} F''(dz) \text{ pour tout } x > 0, \quad \text{(I.4.9)}$$

$$\underline{\lambda}'(0\pm) \;:=\; \pm F'(+\infty), \quad \text{(I.4.10)}$$

$$\underline{\lambda}(0) \;:=\; F(0) \;-\; \int_0^{+\infty} e^{\underline{\gamma}(y)} dy \int_y^{+\infty} e^{-\underline{\gamma}(z)} F''(dz), \quad \text{(I.4.11)}$$

où $\psi_\pm$ (resp. $\underline{\psi}_\pm$) représente l'inverse continue à droite de $\overline{\phi}_\pm$ (resp. $\underline{\phi}_\pm$), F' est la dérivée à droite et $F''(dz)$ est la dérivée seconde au sens de distribution.

Le théorème suivant nous donne les optimiseurs explicites and les dualités souhaitées.

Théorème I.4.2. *Sous des conditions appropriées,* $\overline{\mathsf{P}}(\mu)$ *(resp.* $\underline{\mathsf{P}}(\mu)$*) est atteint par la loi de la* *martingale* $(B_{\overline{\tau}\wedge t/(1-t)})_{0\le t\le 1}$ *(resp.* $(B_{\underline{\tau}\wedge t/(1-t)})_{0\le t\le 1}$*), et* $\overline{\mathsf{D}}(\mu)$ *(resp.* $\underline{\mathsf{D}}(\mu)$*) est atteint par* $(\overline{\lambda}, \overline{H})$ *(resp.* $(\underline{\lambda}, \underline{H})$*), où* $\overline{\lambda}$ *(resp.* $\underline{\lambda}$*) est donnée par* (I.4.6), (I.4.7), (I.4.8) *(resp.* (I.4.9), (I.4.10), (I.4.11)*), et* $\overline{H} = (\overline{H}_t)_{0\le t\le 1}$ *(resp.* $\underline{H} = (\underline{H}_t)_{0\le t\le 1}$*) est défini par*

$$\overline{H}_t \;:=\; -\overline{A}_+(L_t)\mathbb{1}_{\{X_t>0\}} - \overline{A}_-(L_t)\mathbb{1}_{\{X_t\le 0\}} \;\Big(\text{resp. } \underline{H}_t \;:=\; -\underline{A}_+(L_t)\mathbb{1}_{\{X_t>0\}} - \underline{A}_-(L_t)\mathbb{1}_{\{X_t\le 0\}}\Big),$$

où $\overline{A}_\pm : \mathbb{R}_+ \to \mathbb{R}$ *(resp.* $\underline{A}_\pm : \mathbb{R}_+ \to \mathbb{R}$*) est donnée par*

$$\overline{A}_\pm(l) \;:=\; \overline{\lambda}'(0\pm) \;+\; \int_0^l \frac{dy}{\overline{\phi}_\pm(y)} e^{\overline{\gamma}(y)} \int_y^{+\infty} e^{-\overline{\gamma}(z)} F''(dz)$$
$$\left(\text{resp. } \underline{A}_\pm(l) \;:=\; \underline{\lambda}'(0\pm) \;-\; \int_l^{+\infty} \frac{dy}{\underline{\phi}_\pm(y)} e^{\underline{\gamma}(y)} \int_y^{+\infty} e^{-\underline{\gamma}(z)} F''(dz)\right).$$

En outre on a la dualité $\overline{\mathsf{P}}(\mu) = \overline{\mathsf{D}}(\mu)$ *(resp.* $\underline{\mathsf{P}}(\mu) = \underline{\mathsf{D}}(\mu)$*).*

4.2.2 Cas de deux marginales : généralisation de la solution de Vallois

Nous étendons ensuite l'analyse au cas de deux marginales. Pour des raisons techniques, on suppose dans cette section μ_i sont symétriques et admettent une densité notée $\mu_i(x)dx$ pour $i = 1, 2$.

Pour le premier temps d'arrêt, on utilise la solution de Vallois qui reproduit μ_1, *c.à.d.*

$$\tau_1 \;:=\; \inf\left\{t > 0 : \;|B_t| \;\ge\; \phi_1(L_t^B)\right\},$$

où $\phi_1 : \mathbb{R}_+ \to \mathbb{R}_+$ est l'inverse de ψ_1 définie par

$$\psi_1(x) \;:=\; \int_0^x \frac{y\mu_1(y)}{\mu_1([y, +\infty))} dy \text{ pour tout } x \in \mathbb{R}_+. \tag{I.4.12}$$

Pour le deuxième temps d'arrêt, on cherche une fonction croissante $\phi_2 : \mathbb{R}_+ \to \mathbb{R}_+$ telle que $B_{\tau_2\wedge\cdot}$ est uniformément intégrable et $B_{\tau_2} \sim \mu_2$, où

$$\tau_2 \;:=\; \inf\left\{t \ge \tau_1 : \;|B_t| \;\ge\; \phi_2(L_t^B)\right\}.$$

De manière similaire, ϕ_2 est définie via son inverse ψ_2. Tout d'abord, on fixe

$$\psi_2(x) \;:=\; \int_0^x \frac{y\,\mu_2(y)}{\mu_2([y, +\infty))} dy \text{ pour tout } x \in [0, x_1], \tag{I.4.13}$$

où x_1 est le plus petit réel tel que la fonction à droite de (I.4.13) dépasse ψ_1. Afin de garantir que $x_1 > 0$, on impose $\delta\mu := \mu_2 - \mu_1 \leq 0$ sur un voisinage de zéro. Si $x_1 = +\infty$, la construction est finie, ce qui correspond au cas où les solutions de Vallois sont ordonnées. Autrement, on procède par récurence :

(i) si $x_{2i-1} < +\infty$, on écrit pour tout $x \in (x_{2i-1}, x_{2i}]$,

$$\psi_2(x) \quad := \quad \psi_2(x_{2i-1}) \ + \ \int_{x_{2i-1}}^{x} \frac{y\delta\mu(y)}{\delta\mu([y,+\infty))}dy, \tag{I.4.14}$$

où x_{2i} est le plus petit nombre tel que ψ_1 dépasse la fonction à droite de (I.4.14) au-dessous ;

(ii) si $x_{2i} < +\infty$, on écrit pour tout $x \in (x_{2i}, x_{2i+1}]$,

$$\psi_2(x) \quad := \quad \psi_2(x_{2i}) \ + \ \int_{x_{2i}}^{x} \frac{y\mu_2(y)}{\mu_2([y,+\infty))}dy, \tag{I.4.15}$$

où x_{2i+1} est le plus petit nombre tel que la fonction à droite de (I.4.15) au-dessus dépasse ψ_1.

Alors on obtient les temps d'arrêt du plongement souhaités par le théorème suivant.

Théorème I.4.3. *Sous des conditions appropriées, si ψ_2 est donnée par (I.4.14) et (I.4.15), alors $B_{\tau_2\wedge\cdot}$ est uniformément intégrable et $B_{\tau_i} \sim \mu_i$ pour $i = 1, 2$.*

4.2.3 Un cas particulier

Enfin, un cas spécial de plusieurs marginales est étudié, où les temps d'arrêt donnés par Vallois sont bien ordonnés. Dans le cas d'un nombre fini de marginales $(\mu_k)_{1\leq k\leq m}$, le problème réduit essentiellement au cas d'une marginale. Dans le cadre d'un nombre infini de marginales $(\mu_t)_{0\leq t\leq 1}$, on construit une martingale de Markov et calcule explicitement son générateur. En particulier, cela induit un nouvel exemple de faux mouvement brownien.

Hypothèse I.4.4. (i) *Pour tous les $0 \leq s \leq t \leq 1$, $\tau_s \leq \tau_t$, ou autrement dit, $\phi_s \leq \phi_t$, où ϕ_s et ϕ_t sont les fonctions de Vallois associées à μ_s et μ_t ;*

(ii) *L'application $(t, x) \mapsto \psi_t(x)$ admet les dérivées partielles du premier ordre, où ψ_t désigne l'inverse à droite de ϕ_t.*

Le résultat suivant donne le générateur du processus de Markov $(B_{\tau_t})_{0\leq t\leq 1}$.

Théorème I.4.5. *Sous Hypothèse I.4.4, $(B_{\tau_t})_{0\leq t\leq 1}$ est une martingale de Markov inhomogène dont le générateur est donné par*

$$\mathcal{L}_t f(x) \quad = \quad \frac{\partial_t\psi_t(|x|)}{\partial_x\psi_t(|x|)} \left(\frac{e^{\gamma_t\circ\psi_t(|x|)}}{2|x|} \int_{|x|}^{+\infty} \Big(f(y) + f(-y) - 2f(x)\Big) de^{-\gamma_t\circ\psi_t(y)} \ - \ f'(x) \right),$$

pour toute fonction bornée et suffisamment régulière $f : \mathbb{R} \to \mathbb{R}$.

En particulier, le processus $(B_{\tau_t})_{0\leq t\leq 1}$ est un processus de saut pur, qui correspond au modèle introduit dans Carr, Geman, Madan & Yor [19].

Chapitre II

Introduction (English)

The Phd dissertation presents two independent research topics. The first one deals with the martingale optimal transport on the Skorokhod space, and the second one for the optimal Skorokhod embedding.

The thesis is divided into three parts. The first part considers the continuous-time martingale optimal transport problem on the Skorokhod space, see Chapter III or Guo, Tan & Touzi [47]. Our main contribution is to study systematically the tightness of martingale transport plans by means of the $S-$topology introduced by Jakubowski and obtain the existence of optimal plans. Then, we study the martingale optimal transport problem by its dual problems and establish the dualities that have a link to robust hedging in finance.

The second part of the thesis is devoted to three contributions for the optimal Skorokhod embedding problem. We first introduce a weak formulation of the problem where the embedding stopping times are identified in terms of probability measures on an enlarged space, then we derive two dual formulations and prove the corresponding dualities, see Chapter IV or Guo, Tan & Touzi [48]. Next, we revisit the monotonicity principle established by Beiglböck, Cox & Huesmann [5] and provide an alternative and simplified proof based on the obtained dualities. Finally, a stability result is provided using Lévy-Prokhov and Wasserstein metrics, see Chapter V or Guo, Tan & Touzi [49] and Guo [45].

In the last part, we consider the application of stochastic control to the martingale optimal transport on the space of continuous functions with a payoff depending on the local time, and the Skorokhod embedding problem, see Chapter VI or Claisse, Guo & Henry-Labordère [20]. For the one-marginal case, we recover the explicit optimizers for both primal and dual problems using Vallois' solutions, which relate the martingale optimal transport and the optimal Skorokhod embedding. We then extend the analysis to the two-marginal case and construct a generalization of Vallois' solution. Finally, a special multi-marginal case is studied, where the stopping times given by Vallois are well ordered.

1 Preliminaries : robust hedging

These two subjects are recently studied for the purpose of their applications in finance, especially the robust hedging. Before showing that they are related to robust hedging, we begin by a brief introduction of robust hedging that allows to illustrate their financial motivation.

Assume that there are d underlying assets and a relatied exotic option [1] in the market. One of the fundamental problems in financial mathematics is to evaluate this option for both sellers and buyers. Let $X = (X_t)_{t\in I}$ be a $d-$dimensional stochastic process which models the underlying assets, where $I = \{0, 1, \cdots, m\}$ (discrete-time case) or $I = [0, 1]$ (continuous-time case) according to the problem formulation, and $\xi = \xi(X)$ be the payment function of the option to consider.

An important criterion to determine the price comes from the desire of robust hedging. From the viewpoint of a seller (resp. buyer), the ideal selling (resp. buying) price is some wealth allowing to build a portfolio to superhedge/super-replicate (resp. subhedge/sub-replicate) the payment $\xi(X)$ at terminal time. Next, let us specify the superhedging (resp. subhedging) for the discrete-time case. For ease of presentation, we assume

$$d \quad = \quad m \quad = \quad 1.$$

The continuous-time formulation is quite similar and will appear in Section 2.2.

1.1 Dynamic and static strategies

To construct a portfolio associated to initial wealth $z \in \mathbb{R}$, the seller (resp. buyer) is authorized to take a dynamic strategy. Namely, he determines the number $H = H(X_0)$ of assets that he wants to hold at $t = 0$, and the gain at $t = 1$ is thus

$$H(X_0)(X_1 - X_0).$$

Notice that the number H depends only on X_0 since H relies merely on the information available at the time where the investment decision is taken.

In addition to the dynamic strategy, the seller (resp. buyer) is assumed to be able to take a static position of calls and puts [2] for the maturities $t = 0, 1$ and all the strikes $K \in \mathbb{R}$. Then

1. An option writing on underlying assets is a financial contract that allows the buyer to be paid by the seller at a fixed date according to the payment function. An option is said to be exotic if the payment function depends pathwisely on the assets.

2. A call (resp. put) of maturity t and strike K is an option that allows the holder to receive the cash equal

under the hypothesis of no-arbitrage, the proposition below follows by Breeden & Litzenberger [16].

Proposition II.1.1. *Under the hypothesis of no-arbitrage, there exists a probability measure μ_i on $\mathbb{R}$ for $i = 0, 1$ such that $X_i \sim \mu_i$, or equivalently,*

$$c_i(K) \;=\; \int_{\mathbb{R}} (x-K)^+ \mu_i(dx) \quad and \quad p_i(K) \;=\; \int_{\mathbb{R}} (K-x)^+ \mu_i(dx) \ \text{pour tout } K \in \mathbb{R}.$$

Moreover, one has $\mu_i = c_i'' = p_i''$ in the sense of distributions.

Therefore, any option with payment function $\lambda(X_i)$ has a non ambiguous at $t = 0$ given by

$$\int_{\mathbb{R}} \lambda(x)\mu_i(dx) \;=:\; \mu_i(\lambda) \tag{II.1.1}$$

for $i = 0, 1$. Indeed, assume first that λ is smooth, then we obtain for $x_i := \mu_i(x)$ that

$$\lambda(x) \;=\; \lambda(x_i) \;+\; (x - x_i)\lambda'(x_i) \;+\; \int_{-\infty}^{x_i} (K-x)^+ \lambda''(K)dK \;+\; \int_{x_i}^{+\infty} (x-K)^+ \lambda''(K)dK.$$

This is Carr-Madan's formula and gives a decomposition of the payment function λ to a linear combination of calls and puts. We deduce immediately by the linearity of prices that the corresponding price to $\lambda(X_i)$ is

$$\lambda(x_i) \;+\; \int_{-\infty}^{x_i} p_i(K)\lambda''(K)dK \;+\; \int_{x_i}^{+\infty} c_i(K)\lambda''(K)dK,$$

where the option with payment function $X_i - x_i$ costs zero by observing that $x - x_i = (x-x_i)^+ - (x_i - x)^+$. By combining the relation $\mu_i = c_i'' = p_i''$ and integration by parts, the expression above can be written as $\mu_i(\lambda)$. By a density argument, we deduce that for each measurable function λ, the option with the payment function $\lambda(X_i)$ has a price given by (II.1.1).

1.2 Robust hedging

Under the hypothesis of no-arbitrage, the seller (resp. buyer) may construct a semi-static portfolio $(\lambda_0, \lambda_1, H) \in \mathbb{L}^1(\mu_0) \times \mathbb{L}^1(\mu_1) \times \mathbb{L}^0(\mathbb{R})$ that costs $\mu_0(\lambda_0) + \mu_1(\lambda_1)$ at $t = 0$. Namely, its portfolio consists of

— a static strategy, which aims at buying the options corresponding to $\lambda_0(X_0)$ and $\lambda_1(X_1)$ by paying $\mu_0(\lambda_0) + \mu_1(\lambda_1)$ at $t = 0$;

to $(X_t - K)^+$ (resp. $(K - X_t)^+$) at time t.

— a dynamic strategy, which aims at buying $H(X_0)$ assets at $t = 0$ and selling them at $t = 1$.

Its final wealth at $t = 1$ is given by

$$\lambda_0(X_0) \ + \ \lambda_1(X_1) \ + \ H(X_0)(X_1 - X_0).$$

The portfolio is said to be a superhedging / super-replication (resp. subhedging / sub-replication) if

$$\lambda_0(x_0) \ + \ \lambda_1(x_1) \ + \ H(x_0)(x_1 - x_0) \ \geq \ (\text{resp. } \leq) \ \xi(x_0, x_1) \text{ for all } (x_0, x_1) \in \mathbb{R}^2. \quad \text{(II.1.2)}$$

Notice that no special model is imposer on the dynamic of (X_0, X_1), thus the hedging is called robust, *i.e.* model-independent hedging. Let $\overline{\mathcal{D}}^m$ (resp. $\underline{\mathcal{D}}^m$) be the set of all superhedgings (resp. subhedgings) and define

$$\overline{\mathsf{D}}^m(\mu_0, \mu_1) \ \Big(\text{resp. } \underline{\mathsf{D}}^m(\mu_0, \mu_1)\Big) \ := \ \inf_{(\lambda_0, \lambda_1, H) \in \overline{\mathcal{D}}^m} \ \Big(\text{resp. } \sup_{(\lambda_0, \lambda_1, H) \in \underline{\mathcal{D}}^m}\Big) \ \mu_0(\lambda_0) \ + \ \mu_1(\lambda_1) \quad \text{(II.1.3)}$$

Remark II.1.2. (i) *The hypothesis of no-arbitrage is one of the fundamental hypotheses in financial mathematics. It is shown later that this hypothesis implies that X is a martingale and consequently, (μ_0, μ_1) is a peacock that is to specify afterwards.*

(ii) *Without loss of generality, we adopt in the following the formulation of maximization, and write without confusion that $\overline{\mathcal{D}}^m = \mathcal{D}^m$ and $\overline{\mathsf{D}}^m = \mathsf{D}^m$. This formulation is more adapted to financial applications.*

2 Martingale optimal transport : from discrete to continuous

2.1 From classical optimal transport to martingale optimal transport

To understand the relation between robust hedging and martingale optimal transport, we begin by a brief introduction of classic optima transport, and then we pass to its generalization, *i.e.* discrete-time martingale optimal transport, where its dual problem can be exactly interpreted as robust hedging. Moreover, it is shown that the (weak) compactness of the set of transport plans plays an essential role to solve the relative problems.

2.1.1 Monge and Monge-Kantorovich problems

In 1781, Monge considered the problem as follows : Given two probability measures μ_0 and μ_1 on $\mathbb{R}^d$, a measurable map $T : \mathbb{R}^d \to \mathbb{R}^d$ is called a transport plan if $\mu_0 \circ T^{-1} = \mu_1$, where $\mu_0 \circ T^{-1}$ stands for the image measure by T. Then the problem of Monge consists in searching the maximal gain given by

$$\mathsf{P}^M \quad := \sup_{T:\ \mu_0 \circ T^{-1} = \mu_1} \int_{\mathbb{R}^d} \xi\Big(x, T(x)\Big)\mu_0(dx), \tag{II.2.1}$$

where $\xi : \mathbb{R}^{2d} \to \mathbb{R}$ is a measurable function. This problem is difficult to solve because of the full nonlinearity of the constraints, and remained open for many years.

A breakthrough was made by Kantorovich in 1940 by relaxing the constraints and then introducing the dual formulation. Namely, let $\mathcal{P}(\mathbb{R}^{2d})$ be the set of probability measures on $\mathbb{R}^{2d}$, and

$$\mathcal{P}(\mu_0, \mu_1) \quad := \quad \Big\{\mathbb{P} \in \mathcal{P}(\mathbb{R}^{2d}) : \ \mathbb{P} \circ X_i^{-1} \ = \ \mu_i \text{ for } i = 0, 1\Big\}, \tag{II.2.2}$$

where (X_0, X_1) denotes the canonical process on $\mathbb{R}^d \times \mathbb{R}^d$, *i.e.* $X_0(x_0, x_1) := x_0$ and $X_1(x_0, x_1) := x_1$ for all $(x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d$. We are interested in the following Monge-Kantorovich problem

$$\mathsf{P}^{MK}(\mu_0, \mu_1) \quad := \sup_{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)} \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1)\Big]. \tag{II.2.3}$$

Kantorovich introduced a dual formulation defined by

$$\mathsf{D}^{MK}(\mu_0, \mu_1) \quad := \inf_{(\lambda_0, \lambda_1) \in \mathcal{D}^{MK}} \mu_0(\lambda_0) \ + \ \mu_1(\lambda_1), \tag{II.2.4}$$

where

$$\mathcal{D}^{MK} := \Big\{(\lambda_0, \lambda_1) \in \mathbb{L}^1(\mu_0) \times \mathbb{L}^1(\mu_1) : \lambda_0(x_0) + \lambda_1(x_1) \ \geq \ \xi(x_0, x_1) \text{ for all } (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d\Big\}.$$

The set $\mathcal{P}(\mu_0, \mu_1)$ is shown to be compact by a simple argument, and the desired duality follows directly from the Min-Max theorem.

Theorem II.2.1. *Assume that $\xi : \mathbb{R}^{2d} \to \mathbb{R}$ is upper semicontinuous and bounded from above. Then there exist optimizers for the primal and dual problems* (II.2.3) *and* (II.2.4)*, i.e. there* *exist* $\mathbb{P}^* \in \mathcal{P}(\mu_0, \mu_1)$ *and* $(\lambda_0^*, \lambda_1^*) \in \mathcal{D}^{MK}$ *such that.*

$$\mathbb{E}^{\mathbb{P}^*}\Big[\xi(X_0, X_1)\Big] \;=\; \mathsf{P}^{MK}(\mu_0, \mu_1) \quad \textit{and} \quad \mu_0(\lambda_0^*) \;+\; \mu_1(\lambda_1^*) \;=\; \mathsf{D}^{MK}(\mu_0, \mu_1).$$

Moreover, the desired duality holds, i.e. $\mathsf{P}^{MK} = \mathsf{D}^{MK}$.

Remark II.2.2. *The classical optimal transport problem has given rise to a prolific literature, where the duality is one of the essential results. The duality allows not only to study the primal problem from another viewpoint, i.e. deriving the monotonicity principle that characterizes the optimizer* $\mathbb{P}^*$*, but also to provide a numerical scheme to compute* $\mathsf{P}^{MK} = \mathsf{D}^{MK}$*. We refer the reader to Rachev & Rüschendorf [90], or Villani [104] for a comprehensive account of the literature, as well as the relative applications.*

2.1.2 Discrete-time martingale optimal transport

Motivated by the robust hedging in financial mathematics, Beiglböck, Henry-Labordère & Penkner [7] and Galichon, Henry-Labordère & Touzi [43] have studied at the same time an optimal transport problem under martingale constraint, respectively for discrete-time and for continuous-time. Define the subset $\mathcal{M}(\mu_0, \mu_1) \subseteq \mathcal{P}(\mu_0, \mu_1)$ by

$$\mathcal{M}(\mu_0, \mu_1) \;:=\; \Big\{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1) : \; \mathbb{E}^{\mathbb{P}}\Big[X_1 \Big| X_0\Big] \;=\; X_0, \; \mathbb{P}\text{ - a.s.}\Big\}. \tag{II.2.5}$$

Each element $\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)$ is called a (martingale) transport plan. The corresponding martingale optimal transport problem is defined by

$$\mathsf{P}^m(\mu_0, \mu_1) \;:=\; \sup_{\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)} \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1)\Big]. \tag{II.2.6}$$

Similar to the classical optimal transport, Beiglböck, Henry-Labordère & Penkner proved a duality result between $\mathsf{P}^m(\mu_0, \mu_1)$ and the robust hedging $\mathsf{D}^m(\mu_0, \mu_1)$ that is introduced in Section 1.2 :

$$\mathsf{D}^m(\mu_0, \mu_1) \;:=\; \inf_{(\lambda_0, \lambda_1, H) \in \mathcal{D}^m} \mu_0(\lambda_0) \;+\; \mu_1(\lambda_1), \tag{II.2.7}$$

where

$$
\begin{aligned}
\mathcal{D}^m \quad := \quad & \Big\{(\lambda_0, \lambda_1, H) \in \mathbb{L}^1(\mu_0) \times \mathbb{L}^1(\mu_1) \times \mathbb{L}^0(\mathbb{R}^d) : \ \lambda_0(x_0) \ + \ \lambda_1(x_1) \\
& + H(x_0)(x_1 - x_0) \ \geq \ \xi(x_0, x_1) \text{ for all } (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d \Big\}.
\end{aligned} \tag{II.2.8}
$$

Recall that $\mathbf{P}$ and Λ represent respectively the space of all probability measures on $\mathbb{R}^d$ with finite first moment and the space of continuous functions on $\mathbb{R}^d$ with linear growth. A pair $(\mu_0, \mu_1) \in \mathbf{P} \times \mathbf{P}$ is said to be increasing in convex ordering if the inequality $\mu_0(\lambda) \leq \mu_1(\lambda)$ holds for all convex functions $\lambda \in \Lambda$. This relation is denoted by $\mu_0 \preceq \mu_1$ and one has a characterization of the set $\mathcal{M}(\mu_0, \mu_1)$.

Theorem II.2.3 (Strassen). *Let μ_0 and $\mu_1 \in \mathbf{P}$ be two probability measures on $\mathbb{R}^d$. Then*

$$
\mathcal{M}(\mu_0, \mu_1) \ \neq \ \emptyset \quad \textit{if and only if} \quad \mu_0 \ \preceq \ \mu_1.
$$

It is easy to check that, $\mathcal{M}(\mu_0, \mu_1)$ is a closed subset of the compact set $\mathcal{P}(\mu_0, \mu_1)$, which yields the existence of optimizers of the problem (II.2.6). Moreover,

$$
\begin{aligned}
& \inf_{(\lambda_0, \lambda_1, H) \in \mathcal{D}^m} \ \mu_0(\lambda_0) \ + \ \mu_1(\lambda_1), \\
= \ & \inf_{H \in \mathbb{L}^0(\mathbb{R}^d)} \ \inf_{(\lambda_0, \lambda_1) \in \mathcal{D}^{MK}(H)} \ \mu_0(\lambda_0) \ + \ \mu_1(\lambda_1) \\
= \ & \inf_{H \in \mathbb{L}^0(\mathbb{R}^d)} \ \sup_{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)} \ \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1) \ - \ H(X_0)(X_1 - X_0)\Big] \\
= \ & \sup_{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1)} \ \inf_{H \in \mathbb{L}^0(\mathbb{R}^d)} \ \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1) \ - \ H(X_0)(X_1 - X_0)\Big] \\
= \ & \sup_{\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)} \ \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1) \ - \ H(X_0)(X_1 - X_0)\Big].
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{D}^{MK}(H) \quad := \quad & \Big\{(\lambda_0, \lambda_1) \in \mathbb{L}^1(\mu_0) \times \mathbb{L}^1(\mu_1) : \ \lambda_0(x_0) \ + \ \lambda_1(x_1) \\
& \geq \ \xi(x_0, x_1) \ - \ H(x_0)(x_1 - x_0) \text{ for all } (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d \Big\},
\end{aligned} \tag{II.2.9}
$$

the second equality is the classical duality in Theorem II.2.1, and the third equality follows by the Min-Max theorem. Then we obtain the same results summarized in the theorem below, see Beiglböck, Henry-Labordère & Penkner [7] for more details.

Theorem II.2.4 (Beiglböck, Henry-Labordère & Penkner)**.** *Assume that ξ is upper semicontinuous and bounded from above. Then for any pair $(\mu_0, \mu_1) \in \mathbf{P} \times \mathbf{P}$ satisfying $\mu_0 \preceq \mu_1$, the duality $\mathsf{P}^m = \mathsf{D}^m$ holds. Moreover, there exists an optimizer of the problem (II.2.6).*

Remark II.2.5. (i) *Under the hypothesis of no-arbitrage, the process X is required to be a martingale. Therefore, (μ_0, μ_1) should be a peacock in order to ensure that $\mathcal{M}(\mu_0, \mu_1)$ is non empty. Il follows from the classical literature that, each martingale measure $\mathbb{P} \in \mathcal{M}(\mu_0, \mu_1)$ stands for a market model satisfying the hypothesis of no-arbitrage. Then under market uncertainty, the primal problem P^m denotes the upper bound of no-arbitrage prices.*

As described in Section 1.2, the dual problem D^m represents the minimal cost of robust superhedging. Then the duality $\mathsf{P}^m = \mathsf{D}^m$ shows that the upper bound of no-arbitrage prices is equal to the minimal cost of robust superhedging.

(ii) *The compactness of $\mathcal{M}(\mu_0, \mu_1)$ and the regularity of ξ are crucial to prove Theorem II.2.4.*

(iii) *The existence of optimizers for the dual problem* (II.2.7) *is not ensured in general. Nevertheless, Beiglböck, Nutz & Touzi proposed another dual formulation in [11] and proved simultaneously the existence of optimizers for the dual problem and the duality under a weaker regularity condition on ξ.*

2.2 Continuous-time martingale optimal transport

The martingale optimal transport problem admits naturally a continuous-time version, for which we find that the discrete-time arguments no longer work without the compactness of the set of transport plans, see Remark II.2.7 for the summary of difficulties in the continuous-time case. We apply other methods to study this problem and establish the main results in Section 2.2. In the existing literature, the formulation is not unique. We will explain this in the following.

2.2.1 Formulation

We consider only the probability measures on the Skorokhod space. Recall that $\mathbf{D}$ is the Skorokhod space of càdlàg functions $\omega = (\omega_t)_{0 \le t \le 1}$ taking values in $\mathbb{R}^d$. Denote by $X = (X_t)_{0 \le t \le 1}$ the canonical process, *i.e.* $X_t(\omega) := \omega_t$ and by $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le 1}$ its natural filtration, *i.e.* $\mathcal{F}_t := \sigma(X_s, 0 \le s \le t)$. Let $\mathcal{P}$ the set of all probability measures on $(\mathbf{D}, \mathcal{F}_1)$ and each element $\mathbb{P} \in \mathcal{P}$ is called a martingale measure if X is a $\mathbb{P}-$martingale. Denote by $\mathcal{M}$ the collection of all martingale measures.

Let $1 \in \mathbb{T} \subseteq [0,1]$ be a subset and recall that the $\mathbb{T}-$product of $\mathbf{P}$ is defined by

$$\mathbf{P}^{\mathbb{T}} \ := \ \Big\{ \boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} : \ \mu_t \in \mathbf{P} \text{ for all } t \in \mathbb{T} \Big\}.$$

Definition II.2.6. *A family of probability measures $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}^{\mathbb{T}}$ is called a peacock if $\mu_s \preceq \mu_t$ holds for all $s,t \in \mathbb{T}$ such that $s \leq t$. A peacock $\boldsymbol{\mu}$ is said to be càdlàg if the map $t \mapsto \mu_t$ is càdlàg on $\mathbb{T}$ with respect to the weak convergence. Denote by $\mathbf{P}^{\preceq}$ the set of all càdlàg peacocks.*

For a peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, define by $\mathcal{M}(\boldsymbol{\mu}) \subseteq \mathcal{M}$ the subset of martingale transport plans $\mathbb{P}$, *i.e.* $\mathbb{P} \circ X_t^{-1} = \mu_t$ for all $t \in \mathbb{T}$. Then for an $\mathcal{F}_1-$measurable function $\xi : \mathbf{D} \to \mathbb{R}$, the martingale optimal transport problem is defined by

$$\mathsf{P}(\boldsymbol{\mu}) \ := \ \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big]. \tag{II.2.10}$$

Then, all the questions that we ask for the discrete-time case appear as following :

— What is the corresponding dual problem of (II.2.10) ?
— Do the optimizers exist for the primal and dual problems ?
— Does the duality holds ?

Remark II.2.7. (i) *For $\mathbb{T}$ is finite or $\mathbb{T} = [0,1]$, it follows respectively by Kellerer [72] and Hirsch & Roynette [54, 55] that, for any $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, the set $\mathcal{M}(\boldsymbol{\mu})$ is non empty, see Hirsch, Profeta, Roynette & Yor [53] for a complete literature about peacock.*

(ii) *The first problem in continuous-time case comes from the definition of the dual problem. Indeed, when passing from discrete-time case to continuous-time case, the dynamic part becomes*

$$H(X_0)(X_1 - X_0) \ \Longrightarrow \ \int_0^1 H_t dX_t$$

pour some suitable process $H = (H_t)_{0\leq t\leq 1}$. However, since the process X is not of bounded variation, then it is not trivial to define $\int_0^1 H_t dX_t$ pathwisely.

In the existing literature, there are two dual formulations for the problem (II.2.10), Galichon, Henry-Labordère & Touzi [43] studied a class of transport plans defined by stochastic differential equations and introduced a quasi-sure dual problem. They applied a stochastic control approach and deduced the duality. Another important contribution is due to Dolinsky & Soner [36, 37] and Hou & Obłój [63], where the dual problem has a pathwise definition.

(iii) *In contrast with the discrete-time case, the set $\mathcal{M}(\boldsymbol{\mu})$ is generally not tight*[3] *with respect to the usual topologies, which is shown by the example below. Without the crucial compactness, the arguments in the classical setting fail to be adapted to handle the related issues.*

Example II.2.8. *Let $M = (M_0, M_1, M_2)$ be a martingale on some probability space such that $\mathbb{P}[M_0 \neq M_1$ and $M_1 \neq M_2] > 0$. Define $\mathbb{P}_n := \mathbb{P} \circ (M^n)^{-1}$ for all $n \geq 3$, where $M^n = (M_t^n)_{0 \leq t \leq 1}$ is defined by*

$$M_t^n \;:=\; M_0 \mathbb{1}_{[0, \frac{1}{2} - \frac{1}{n})}(t) \;+\; M_1 \mathbb{1}_{[\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n})}(t) \;+\; M_2 \mathbb{1}_{[\frac{1}{2} + \frac{1}{n}, 1]}(t).$$

Clearly, $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu})$ for all $n \geq 3$ with $\mathbb{T} = \{0, 1\}$ and $\boldsymbol{\mu} = (\mathbb{P} \circ M_0^{-1}, \mathbb{P} \circ M_2^{-1})$. However, it follows from Theorem VI.3.21 in Jacod & Shiryaev [66] that, the sequence $(\mathbb{P}_n)_{n \geq 3}$ is not tight with respect to the Skorokhod topology, and thus not compact.

2.2.2 Main results

We introduce three dual formulations of the primal problem (II.2.10). The main contribution is to study systematically the tightness of the set $\mathcal{M}(\boldsymbol{\mu})$ by means of the $S-$topology introduced in Jakubowski [67]. Endowing the space $\mathbf{P}^{\mathbb{T}}$ with a topology of Wasserstein kind, the tightness of $\mathcal{M}(\boldsymbol{\mu})$ yields the upper semicontinuity of the map $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ and further the first duality, obtained by penalizing the marginal constraints. Using the dynamic programming principle and the discretization argument introduced in Dolinsky & Soner [37], we obtain the second and third dualities from the first duality.

First dual problem Recall that Λ is the set of continuous functions on $\mathbb{R}^d$ with linear growth and $\Lambda^{\mathbb{T}}$ is its $\mathbb{T}-$product, *i.e.*

$$\Lambda^{\mathbb{T}} \;:=\; \Big\{ \boldsymbol{\lambda} = (\lambda_{t_i})_{1 \leq i \leq k} : \; t_i \in \mathbb{T}, \; \lambda_{t_i} \in \Lambda \text{ for all } i = 1, \cdots, k \text{ and } k \in \mathbb{N} \Big\}.$$

For every $\boldsymbol{\lambda} = (\lambda_{t_i})_{1 \leq i \leq k} \in \Lambda^{\mathbb{T}}$ and $\omega = (\omega_t)_{0 \leq t \leq 1} \in \mathbf{D}$, denote

$$\boldsymbol{\lambda}(\omega) \;:=\; \sum_{i=1}^k \lambda_{t_i}(\omega_{t_i}) \quad \text{and} \quad \boldsymbol{\mu}(\boldsymbol{\lambda}) \;:=\; \sum_{i=1}^k \mu_{t_i}(\lambda_{t_i}).$$

3. A set $\mathcal{E}$ of probability measures is said to be tight (with respect to some convergence) if for any $\varepsilon > 0$, there exists a compact set K such that $\mathbb{P}[K] \geq 1 - \varepsilon$ for all $\mathbb{P} \in \mathcal{E}$.

The first dual problem is defined by

$$\mathsf{D}_1(\boldsymbol{\mu}) \quad := \quad \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \left\{ \boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) \;-\; \boldsymbol{\lambda}(X)\Big] \right\}. \tag{II.2.11}$$

The dual problem D_1 seems like the Lagrangian of optimization problems under constraints, where the marginal constraints $\boldsymbol{\mu}$ are penalized by the Lagrange multipliers $\boldsymbol{\lambda}$.

Second dual problem The second dual problem penalizes further the martingale constraint. For technical reasons, we focus on the case where the process X is non-negative and starts at some given point that may be normalized to $\mathbf{1} := (1, \cdots, 1) \in \mathbb{R}^d$. Define $\mathbf{D}_+ := \{\omega \in \mathbf{D} : \omega_0 = \mathbf{1}$ and $\omega_t \in \mathbb{R}^d_+$ for all $t \in [0,1]\}$ and $\mathcal{M}_+ := \{\mathbb{P} \in \mathcal{M} : \mathrm{supp}(\mathbb{P}) \subseteq \mathbf{D}_+\}$. In order to ensure that $\mathcal{M}_+ \cap \mathcal{M}(\boldsymbol{\mu}) \neq \emptyset$, $\boldsymbol{\mu}$ must satisfy

$$\mu_0(dx) \;=\; \delta_{\mathbf{1}}(dx) \quad \text{and} \quad \mathrm{supp}(\mu_1) \;\subseteq\; \mathbb{R}^d_+. \tag{II.2.12}$$

Denote by $\mathbb{F}^U = (\mathcal{F}^U_t)_{0\le t\le 1}$ the universally completed filtration, *i.e.* $\mathcal{F}^U_t := \cap_{\mathbb{P}\in\mathcal{P}} \mathcal{F}^{\mathbb{P}}_t$, where $\mathcal{F}^{\mathbb{P}}_t$ is the completed σ-field of $\mathcal{F}_t$ under $\mathbb{P}$.

Definition II.2.9. *An $\mathbb{F}^U-$adapted process $S : [0,1]\times\mathbf{D} \to \mathbb{R}$ is called an $\mathcal{M}_+-$supermartingale if it is a $\mathbb{P}-$supermartingale for all $\mathbb{P} \in \mathcal{M}_+$. Denote by $\mathcal{S}$ the collection of all such processes and by $\mathcal{S}_0 \subseteq \mathcal{S}$ the subset of processes starting at zero. Denote further*

$$\mathcal{D}_2(\xi) \quad := \quad \Big\{(\boldsymbol{\lambda}, S) \in \Lambda^{\mathbb{T}} \times \mathcal{S}_0 : \; \boldsymbol{\lambda}(\omega) \;+\; S_1(\omega) \;\ge\; \xi(\omega) \text{ for all } \omega \in \mathbf{D}_+\Big\}.$$

For a peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ satisfying the condition (II.2.12) , the second dual problem is defined by

$$\mathsf{D}_2(\boldsymbol{\mu}) \quad := \quad \inf_{(\boldsymbol{\lambda},S)\in\mathcal{D}_2(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{II.2.13}$$

Remark II.2.10. *Notice that the supermartingale $S \in \mathcal{S}$ is not required to have any regularity. If it were càdlàg then it would follow from Theorem 2.1 in Kramkov [74] that, for every $\mathbb{P} \in \mathcal{M}_+$ there exist a predictable process $H^{\mathbb{P}} = (H^{\mathbb{P}}_t)_{0\le t\le 1}$ and an optional non-decreasing process $A^{\mathbb{P}} = (A^{\mathbb{P}}_t)_{0\le t\le 1}$ such that*

$$S_t \;=\; S_0 \;+\; \int_0^t H^{\mathbb{P}}_s dX_s \;-\; A^{\mathbb{P}}_t \text{ for all } t \in [0,1], \; \mathbb{P} \text{ - a.s.}$$

However, it is not clear whether one can aggregate the last representation, i.e. find predictable processes H and A such that $(H, A) = (H^{\mathbb{P}}, A^{\mathbb{P}})$, $\mathbb{P}$ - almost surely. It follows by Nutz [80], *under the continuum hypothesis, the aggregation result holds when X takes values in the space of continuous functions, since we may identify the integrand in the stochastic integral by $H_t = H_t^{\mathbb{P}} = d\langle S, X\rangle_t / d\langle X\rangle_t$, where the processes $\langle S, X\rangle$ and $\langle X\rangle$ denote the quadratic co-variation of S and X and the quadratic variation of X, and they can be defined universally.*

Third dual problem The third dual problem is roughly parallel to the second one, but a bit stronger. It can be seen as an integral representation of supermartingales under a non-dominated family of measures.

Definition II.2.11. *An $\mathbb{F}-$adapted process $H : [0,1] \times \mathbf{D}_+ \to \mathbb{R}^d$ is called a dynamic strategy if $t \mapsto H_t(\omega)$ is left-continuous and of bounded variation for every $\omega \in \mathbf{D}_+$, and $(H \cdot X)$ is a $\mathbb{P}-$supermartingale for all $\mathbb{P} \in \mathcal{M}_+$, where*

$$(H \cdot X)_t \;\; := \;\; H_t \cdot X_t \;-\; H_0 \cdot X_0 \;-\; \int_0^t X_s \cdot dH_s \text{ for all } t \in [0,1],$$

with $\int_0^\cdot X_s \cdot dH_s$ refering to the scalar Lebesgue-Stieltjes integration. Denote by $\mathcal{H}$ the set of all dynamic strategies.

For a peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ satisfying the condition (II.2.12), the third dual problem is defined by

$$\mathsf{D}_3(\boldsymbol{\mu}) \;\; := \inf_{(\boldsymbol{\lambda}, H) \in \mathcal{D}_3(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}), \tag{II.2.14}$$

where

$$\mathcal{D}_3(\xi) \;\; := \;\; \Big\{ (\boldsymbol{\lambda}, H) \in \Lambda^{\mathbb{T}} \times \mathcal{H} : \; \boldsymbol{\lambda}(\omega) \;+\; (H \cdot X)_1(\omega) \;\geq\; \xi(\omega) \text{ for all } \omega \in \mathbf{D}_+ \Big\}.$$

The following theorem summarizes the duality results obtained in this thesis.

Theorem II.2.12. *Under suitable conditions, there exists a probability measure $\mathbb{P}^* \in \mathcal{M}(\boldsymbol{\mu})$ such that*

$$\mathbb{E}^{\mathbb{P}^*}\Big[\xi(X)\Big] \;\; = \;\; \mathsf{P}(\boldsymbol{\mu}) \;\; = \;\; \mathsf{D}_1(\boldsymbol{\mu}) \;\; = \;\; \mathsf{D}_2(\boldsymbol{\mu}) \;\; = \;\; \mathsf{D}_3(\boldsymbol{\mu}).$$

Remark II.2.13. (i) *Similar to the discrete-time framework, the dual formulation (II.2.14) represents the minimum fortune to build a robust superhedging of ξ. The only difference here is that we are restricted to a smaller class of dynamic strategies.*

(ii) *We can also formulate the problem assuming that X takes values in the space of continuous functions, see Dolinsky & Soner [36] and Hou & Obłój [63]. We consider the formulation on* *the space of continuous functions in Sections 3.2.2 and 4.1, where the problem is formulated differently and approached in another way.*

3 Optimal Skorokhod embedding : from strong to weak

3.1 Motivation and formulation

The Skorokhod embedding problem (SEP) aims at representing a probability measure on real line by the law of Brownian motion stopped at some stopping time. Initiated by Skorokhod, various solutions are constructed by different methods. In particular, a lot of existing solutions satisfy some particular optimality. For example, Root's solution minimizes the average embedding, Azéma-Yor's solution maximizes the running maximum law of Brownian motion, and Vallois constructed the solutions that optimize the expectation of a convex function of local time, *etc.*

The objective is to unify the optimization problem among the embedding stopping times, named optimal Skorokhod embedding problem, and study it systematically, *i.e.* the existence and characterization of optimizers, see Sections 3.2, 3.3 and 3.4.

3.1.1 Link to robust hedging

Knowing that each continuous martingale starting at zero is a time-changed Brownian motion, we find that this optimization problem is related to the martingale optimal transport problem on the space of continuous functions. In this section we explain roughly this idea and shows that this problem is also motivated by the study of robust hedging in finance.

Let $B = (B_t)_{t\geq 0}$ be a (standard) Brownian motion and μ be a centered probability measure on $\mathbb{R}$. For any embedding τ, consider the process $X = (X_t)_{0\leq t\leq 1}$ defined by

$$X_t \quad := \quad B_{\frac{t}{1-t}\wedge\tau} \text{ for all } t \in [0, 1].$$

According to the construction, X is a martingale with $X_1 \sim \mu$. Conversely, for each continuous martingale, it follows from the Dambis-Dubins-Schwarz theorem that there is a Brownian motion $W = (W_t)_{t\geq 0}$ such that almost surely

$$X_t \quad = \quad W_{\langle X\rangle_t} \text{ for all } t \in [0, 1].$$

It is known that, in several cases, solving the optimal Skorokhod embedding problem may provide the explicit optimizers for both the martingale optimal transport problem and its dual problem, *i.e.* the market model linked to the extremal no-arbitrage price and the optimal robust hedging strategy. For more results, we refer to Hobson [56], Hobson & Obłój [23], Cox & Obłój [24, 25], Hobson & Klimmek [58], Obłój & Spoida [84], Galichon, Henry-Labordère & Touzi [43], Henry-Labordère, Obłój, Spoida & Touzi [50], Cox, Obłój & Touzi [26], *etc.*

3.1.2 Formulation

The optimization problem was first systematically studied by Beiglböck, Cox & Huesmann [5], where they formulated the problem for the one-marginal case and recovered all previous known results by a unified formulation. They introduced a weak formulation, where they considered a family of embedding stopping times with respect to a larger filtration. We consider an extension of the weak formulation of the optimal SEP in [5] to the case of finitely-many marginal constraints.

Recall that Ω is the space of all continuous paths $\omega = (\omega_t)_{t\geq 0}$ on $\mathbb{R}_+$ such that $\omega_0 = 0$. Let $B = (B_t)_{t\geq 0}$ be the canonical process, *i.e.* $B_t(\omega) := \omega_t$, $\mathbb{P}_0$ be the Wiener measure, $\mathbb{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ be the natural filtration, *i.e.* $\mathcal{F}^0_t := \sigma(B_s, s \leq t)$, and $\mathbb{F}^a = (\mathcal{F}^a_t)_{t\geq 0}$ be the augmented filtration under $\mathbb{P}_0$.

Let $m \geq 1$ be a fixed integer and set $\Theta = \{\theta = (\theta_1, \cdots, \theta_m) : 0 \leq \theta_1 \leq \cdots \leq \theta_m\}$. Define the enlarged canonical space by $\overline{\Omega} := \Omega \times \Theta$, and all the elements of $\overline{\Omega}$ are denoted by $\bar{\omega} = (\omega, \theta)$. Denote further by (B, T) with $T = (T_1, \cdots, T_m)$ the canonical element on $\overline{\Omega}$, *i.e.* $B_t(\bar{\omega}) := \omega_t$ and $T(\bar{\omega}) := \theta$ for all $\bar{\omega} = (\omega, \theta) \in \overline{\Omega}$. The canonical filtration is denoted by $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$, where $\overline{\mathcal{F}}_t$ is generated by $(B_s)_{0\leq s\leq t}$ and all the sets $\{T_k \leq s\}$ for all $s \in [0, t]$ and $k = 1, \cdots, m$. In particular, all random variables $T_1, \cdots, T_m$ are $\overline{\mathbb{F}}-$stopping times.

Let $\overline{\mathcal{P}}(\overline{\Omega})$ be the space of all probability measures on $\overline{\Omega}$, and define

$$\overline{\mathcal{P}} \ := \ \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) : \ B \text{ is an } \overline{\mathbb{F}} - \text{Brownian motion and} \\ B_{T_m \wedge \cdot} := (B_{T_m \wedge t})_{t\geq 0} \text{ is UI under } \overline{\mathbb{P}}\Big\}. \tag{II.3.1}$$

Recall that $\mathbf{P}$ is the space of all probability measures on $\mathbb{R}$ with finite first moment and $\mathbf{P}^m$ is its $m-$product. Recall also that a vector $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m} \in \mathbf{P}^m$ is called a peacock if $\mu_k \preceq \mu_{k+1}$ holds for all $k = 1, \cdots, m-1$. This peacock is centered if each element μ_k has zero mean. Given a centered peacock $\boldsymbol{\mu}$, define the set of embeddings

$$\overline{\mathcal{P}}(\boldsymbol{\mu}) \ := \ \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}} : \ \overline{\mathbb{P}} \circ B^{-1}_{T_k} \ = \ \mu_k \text{ for all } k = 1, \cdots, m\Big\}. \tag{II.3.2}$$

Let $\Phi : \overline{\Omega} \to \mathbb{R}$ be a measurable function, then Φ is called non-anticipative if $\Phi(\omega, \theta) = \Phi\big(\omega_{\theta_m \wedge \cdot}, \theta\big)$ for all $(\omega, \theta) \in \overline{\Omega}$. Then the optimal SEP is defined by

$$\overline{P}(\boldsymbol{\mu}) \;\; := \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big]. \tag{II.3.3}$$

Remark II.3.1. (i) *Under each $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$, the vector of stopping times $(T_1, \cdots, T_m)$ is an embedding. Notice that the stopping times are with respect to the filtration $\overline{\mathbb{F}}$ which is larger than the Brownian filtration $\mathbb{F}^0$, which explains why our formulation is said to be weak.*

(ii) *It follows from the relation to martingale optimal transport in Section 3.1.1, each measure $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ can also be considered as a market model satisfying the hypothesis of no-arbitrage, where m denotes the number of maturities. See Section 3.2.2 for more details.*

(iii) *Recently, Källblad, Tan & Touzi [69] have generalized the results to the case of infinitely-many marginals in the context of lookback options.*

3.2 Dualities of optimal Skorokhod embedding

Inspired by the approach used for the problem of martingale optimal transport, we introduce here two dual problems and prove the corresponding dualities. Using the duality results, we may establish the dualities for martingale optimal transport in the space of continuous functions. In addition, with the obtained dualities obtained, we prove the monotony principle that characterizes optimizers of optimal Skorokhod embedding in Section 3.3, and a stability result in Section 3.4.

3.2.1 Duality results

We revisit the duality result of Beiglböck, Cox & Huesmann [5] and extend the duality to the multi-marginal case under more general conditions. Our approach uses tools of a completely different nature. First, by following the convex duality approach, we convert the optimal Skorokhod embedding problem into an infimum of classical optimal stopping problems, which yields the first duality. Next, we use the standard dynamic programming approach to obtain the second duality.

First dual problem Denote by $\mathcal{T}^a$ the collection of all increasing families of $\mathbb{F}^a$−stopping times $\tau = (\tau_k)_{1 \le k \le m}$ such that the stopped process $B_{\tau_m \wedge \cdot}$ is uniformly integrable. Recall that Λ is the space of continuous functions on $\mathbb{R}$ with linear growth and Λ^m is its m−product.

For $\boldsymbol{\lambda} = (\lambda_k)_{1\le k\le m} \in \Lambda^m$ and $(\omega,\theta) \in \overline{\Omega}$, we denote

$$\boldsymbol{\lambda}(\omega,\theta) \;:=\; \sum_{k=1}^m \lambda_k(\omega_{\theta_k}) \quad \text{and} \quad \boldsymbol{\mu}(\boldsymbol{\lambda}) \;:=\; \sum_{k=1}^m \mu_k(\lambda_k).$$

Then the first dual problem for the optimal Skorokhod embedding problem (II.3.3) is given by

$$\overline{D}_0(\boldsymbol{\mu}) \;:=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m} \Big\{ \boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi(B,\tau) \;-\; \boldsymbol{\lambda}(B,\tau)\Big]\Big\}. \tag{II.3.4}$$

Second dual problem The second dual problem is formulated on the enlarged space $\overline{\Omega}$. Given $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$, an $\overline{\mathbb{F}}-$optional process $S : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ is called a strong $\overline{\mathbb{P}}-$supermartingale if

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[S_{\tau_2}\middle|\overline{\mathcal{F}}_{\tau_1}\right] \;\le\; S_{\tau_1},\; \overline{\mathbb{P}}\text{ - a.s.}$$

holds for all $\overline{\mathbb{F}}-$stopping times $\tau_1 \le \tau_2$. Let $\mathbb{L}^2_{loc}$ be the space of all $\overline{\mathbb{F}}-$progressively measurable processes $\overline{H} = (\overline{H}_t)_{t\ge 0}$ such that

$$\int_0^t \overline{H}_t^2 dt \;<\; +\infty \text{ for all } t \ge 0,\; \overline{\mathbb{P}}\text{ - a.s. for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}.$$

For $\overline{H} \in \mathbb{L}^2_{loc}$, the stochastic integral $(\overline{H}\cdot B) := \int_0^\cdot \overline{H}_t dB_t$ is well defined $\overline{\mathbb{P}}$ - a.s. for all $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$. We introduce a subset of processes :

$$\overline{\mathcal{H}} \;:=\; \left\{\overline{H} \in \mathbb{L}^2_{loc} :\; (\overline{H}\cdot B)_{T_m\wedge\cdot} \text{ is a } \mathbb{P}\text{ - strong supermartingale for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}\right\}.$$

Denote further

$$\overline{\mathcal{D}} \;:=\; \left\{(\boldsymbol{\lambda},\overline{H}) \in \Lambda^m \times \overline{\mathcal{H}} :\; \boldsymbol{\lambda}(B,T) \;+\; (\overline{H}\cdot B)_{T_m} \;\ge\; \Phi(B,T),\; \overline{\mathbb{P}}\text{ - a.s. for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}\right\},$$

and the second dual problem is given by

$$\overline{D}(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda},\overline{H})\in\overline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{II.3.5}$$

Theorem II.3.2. *Under suitable conditions, there is some $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ such that*

$$\mathbb{E}^{\overline{\mathbb{P}}^*}\Big[\Phi(B,T)\Big] \;=\; \overline{P}(\boldsymbol{\mu}) \;=\; \overline{D}_0(\boldsymbol{\mu}) \;=\; \overline{D}(\boldsymbol{\mu}).$$

Remark II.3.3. *We prove the dualities under weaker conditions on Φ of [5]. In particular, this* *allows to recover the important example, where Φ depends on the local time of B at zero, see e.g Cox, Hobson & Obłój [23] and Claisse, Guo & Henry-Labordère [20].*

The following example shows that our formulation is generally not equivalent to the strong formulation, where the embedding stopping times are with respect to a Brownian filtration.

Example II.3.4. *For $m = 1$, take $\mu = \delta_{\{0\}}/3 + \delta_{\{1\}}/3 + \delta_{\{-1\}}/3$ and $\Phi(\omega, \theta) = \mathbb{1}_{\{0\}}(\theta)$. Define*

$$\tau_0 \;:=\; \inf\Big\{t \in \mathbb{R}_+ :\; |B_t| \geq 1\Big\} \quad \textit{and} \quad \overline{\mathbb{P}}_0 \;:=\; \frac{1}{3}\mathbb{P}_0 \circ (B, 0)^{-1} \;+\; \frac{2}{3}\mathbb{P}_0 \circ (B, \tau_0)^{-1},$$

then $\overline{\mathbb{P}}_0 \in \mathcal{P}(\mu)$ and $\mathbb{E}^{\overline{\mathbb{P}}_0}[\Phi(B, T)] = \frac{1}{3}$. Further, let $\tau \in \mathcal{T}^a$ such that $\mathbb{P}_0 \circ B_\tau^{-1} = \mu$, then $\mathbb{P}_0[\tau > 0] > 0$. Since the augmented Brownian filtration satisfies Blumenthal's zero-one law, then $\mathbb{P}_0[\tau > 0] = 1$, which implies that

$$\sup_{\tau \in \mathcal{T}^a:\; B_\tau \sim \mu} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi(B, \tau)\Big] \;=\; 0 \;<\; \frac{1}{3} \;\leq\; \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big].$$

Next, we provide an example where the duality fails when Φ has no regularity in θ.

Example II.3.5. *For $m = 1$, take $\mu = \delta_{\{1\}}/2 + \delta_{\{-1\}}/2$ and $\Phi(\omega, \theta) = \mathbb{1}_{\mathbb{Q}}(\theta)$. We first notice that $\overline{\mathcal{P}}(\mu)$ has only one element, which is the probability measure induced by (B, τ_0), where $\tau_0 := \inf\{t \in \mathbb{R}_+ : |B_t| \geq 1\}$. Indeed, for any $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)$, one has $\mathbb{E}^{\overline{\mathbb{P}}}[T] = \mathbb{E}^{\overline{\mathbb{P}}}[B_T^2] = \mathbb{E}^{\overline{\mathbb{P}}}[\tau_0]$ and $T \geq \tau_0$, $\overline{\mathbb{P}}$ - a.s. Moreover, since the hitting time τ_0 has an atom-less distribution on $\mathbb{R}_+$, then*

$$\overline{P}(\mu) \;=\; \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big] \;=\; \mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\mathbb{1}_{\mathbb{Q}}(\tau_0)\Big] \;=\; 0.$$

As for the dual problem, we notice that $\lambda \in \Lambda$ is a continuous function, and one can approximate any stopping time by stopping times taking value in $\mathbb{Q}$, then

$$\sup_{\tau \in \mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\mathbb{1}_{\mathbb{Q}}(\tau) - \lambda(B_\tau)\Big] \;=\; \sup_{\tau \in \mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[1 - \lambda(B_\tau)\Big] \textit{ for all } \lambda \in \Lambda.$$

Then by the definition (II.3.4), $\overline{D}_0(\mu) = \inf_{\lambda \in \Lambda}\{\mu(\lambda) + \sup_{\tau \in \mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}[1 - \lambda(B_\tau)]\} = 1$. Similarly, we can easily deduce that $\overline{D}(\mu) = 1$, which yields

$$\overline{P}(\mu) \;=\; 0 \;<\; 1 \;=\; \overline{D}_0(\mu) \;=\; \overline{D}(\mu).$$

3.2.2 Application to martingale optimal transport

In view of the fact that each continuous martingale starting at zero is a time-changed Brownian motion, we show that our duality results induce the dualities for a class of martingale optimal transport problems on the space of continuous functions.

Recall that $\mathbf{C}$ is the space of continuous funcitons $\omega = (\omega_t)_{0\le t\le 1}$ such that $\omega_0 = 0$. Define the canonical process $X = (X_t)_{0\le t\le 1}$, *i.e* $X_t(\omega) := \omega_t$ and its natural filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$, *i.e.* $\mathcal{F}_t := \sigma(X_s, s \le t)$. Denote further by $\mathcal{M}$ the collection of all martingale measures $\mathbb{P}$, *i.e.* probability measures on $(\mathbf{C}, \mathcal{F}_1)$ under which X is a martingale. Let $\mathbb{T} := \{t_1 < \cdots < t_m = 1\} \subseteq (0, 1]$ be a subset and define the set of martingale transport plans

$$\mathcal{M}(\boldsymbol{\mu}) \;\; := \;\; \Big\{\mathbb{P} \in \mathcal{M} : \; \mathbb{P} \circ X_{t_k}^{-1} \;=\; \mu_k \text{ for all } k = 1, \cdots, m\Big\}.$$

Given a measurable function $\xi : \mathbf{C} \to \mathbb{R}$, the corresponding martingale transport problem is defined by

$$\mathsf{P}(\boldsymbol{\mu}) \;\; := \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big]. \tag{II.3.6}$$

We give next the corresponding dual problems. Denote by $\mathcal{H}$ the collection of all $\mathbb{F}-$progressively measurable processes $H : [0, 1] \times \mathbf{C} \to \mathbb{R}$ such that

$$\int_0^1 H_t^2 d\langle X\rangle_t \;<\; +\infty, \; \mathbb{P}\text{ - a.s. and the stochastic integral } (H \cdot X) \;:=\; \int_0^\cdot H_t dX_t$$
$$\text{is a } \mathbb{P} - \text{supermartingale for all } \mathbb{P} \in \mathcal{M}. \tag{II.3.7}$$

Here, it follows by Karandikar [70] that, there is a non-decreasing $\mathbb{F}-$progressively measurable process $\langle X\rangle$ taking value in $[0, +\infty]$, such that $\langle X\rangle$ coincides with the quadratic variation of X, $\mathbb{P}$ - a.s. for all $\mathbb{P} \in \mathcal{M}$. Then the two dual problems are given by

$$\mathsf{D}_0(\boldsymbol{\mu}) \;:=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m} \left\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) - \boldsymbol{\lambda}(X)\Big]\right\} \;\text{ and }\; \mathsf{D}(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda},H)\in\mathcal{D}} \boldsymbol{\mu}(\boldsymbol{\lambda}), \tag{II.3.8}$$

where

$$\boldsymbol{\lambda}(X) \;\; := \;\; \sum_{i=1}^m \lambda_i(X_{t_i})$$

and

$$\mathcal{D} \;:=\; \Big\{(\boldsymbol{\lambda}, H) \in \Lambda^m \times \mathcal{H}:\; \boldsymbol{\lambda}(X) \;+\; (H \cdot X)_1 \;\geq\; \xi(X),\; \mathbb{P}\text{ - a.s. for all } \mathbb{P} \in \mathcal{M}\Big\}.$$

Denote $\langle X\rangle_t^{-1} := \inf\{s \in \mathbb{R}_+ : \langle X\rangle_s > t\}$ and

$$W_t \;:=\; X_{\langle X\rangle_t^{-1}} \mathbb{1}_{\{t<\langle X\rangle_1\}} \;+\; \Big(X_1 \;+\; \widehat{W}_{t-\langle X\rangle_1}\Big)\mathbb{1}_{\{t\geq\langle X\rangle_1\}} \text{ for all } t \in \mathbb{R}_+,$$

where $\widehat{W}$ is a Brownian motion independent of $\mathcal{F}_1$. It follows from the Dambis-Dubins-Schwarz theorem that, the process $W = (W_t)_{t\geq 0}$ is a Brownian motion, see *e.g.* Revuz & Yor [92, Theorem 1.7, Chapter V].

The following theorem shows that, with the help of the optimal Skorokhod embedding problem, we may establish the dualities for the martingale optimal transport problem with continuous paths.

Theorem II.3.6. *Assume that $\xi(X) = \Phi(W, \langle X\rangle_{t_1}, \cdots, \langle X\rangle_{t_n})$ holds for some measurable function $\Phi : \overline{\Omega} \to \mathbb{R}$. Then under suitable conditions one has*

$$\mathsf{P}(\boldsymbol{\mu}) \;=\; \mathsf{D}_0(\boldsymbol{\mu}) \;=\; \mathsf{D}(\boldsymbol{\mu}).$$

3.3 Monotonicity principle of optimal Skorokhod embedding

3.3.1 Motivation

As introduced at the beginning of Section 3.1, it is well known that some existing solutions to the Skorokhod embedding problem enjoy an optimality property with respect to some criterion. Beiglböck, Cox & Huesmann [5] introduced the monotonicity principle for the one-marginal case, which allows to characterize optimal embeddings. To the best of our knowledge, all well known solutions to the Skorokhod embedding problem with optimality property can be interpreted through this unifying principle.

The contribution here is to provide an alternative and simplified proof of the monotonicity principle, based on the dualities obtained in Section 3.2.1. Our proof follows a delicate application of the optional cross-section theorem, and a clever conditioning argument introduced in [5]. We emphasize that we focus on the one-marginal case throughout Section 3.3.

3.3.2 Main results

For a subset $\overline{\Gamma} \subseteq \overline{\Omega}$, denote $\overline{\Gamma}^<$ by

$$\overline{\Gamma}^< \;:=\; \Big\{\bar{\omega} = (\omega, \theta) \in \overline{\Omega} : \; \bar{\omega}_{\theta \wedge \cdot} \;=\; \bar{\omega}'_{\theta \wedge \cdot} \text{ for some } \bar{\omega}' \in \overline{\Gamma} \text{ with } \theta' > \theta\Big\}.$$

For every $\bar{\omega} = (\omega, \theta)$ and $\bar{\omega}' = (\omega', \theta') \in \overline{\Omega}$, the concatenation $\bar{\omega} \otimes \bar{\omega}' \in \overline{\Omega}$ is defined by

$$\bar{\omega} \otimes \bar{\omega}' \;:=\; \Big(\omega \otimes_\theta \omega', \; \theta + \theta'\Big),$$

where

$$\left(\omega \otimes_\theta \omega'\right)_t \;:=\; \omega_t \mathbb{1}_{[0,\theta)}(t) \;+\; \Big(\omega_\theta \;+\; \omega'_{t-\theta}\Big)\,\mathbb{1}_{[\theta,+\infty)}(t) \text{ for all } t \in \mathbb{R}_+.$$

Definition II.3.7. *A pair $(\bar{\omega}, \bar{\omega}') \in \overline{\Omega} \times \overline{\Omega}$ is said to be a stop-go pair if $\omega_\theta = \omega'_{\theta'}$ and*

$$\Phi(\bar{\omega}) \;+\; \Phi(\bar{\omega}' \otimes \bar{\omega}'') \;\;>\;\; \Phi(\bar{\omega} \otimes \bar{\omega}'') \;+\; \Phi(\bar{\omega}') \textit{ for all } \bar{\omega}'' \in \overline{\Omega}^+,$$

where $\overline{\Omega}^+ := \{\bar{\omega} = (\omega, \theta) \in \overline{\Omega} : \theta > 0\}$. Denote by SG *the set of all stop-go pairs.*

Theorem II.3.8. *Let $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\mu)$ be an optimizer of the optimal Skorokhod embedding problem (II.3.3). Then under suitable conditions, there exists a Borel subset $\overline{\Gamma}^* \subseteq \overline{\Omega}$ such that*

$$\overline{\mathbb{P}}^*\left[\overline{\Gamma}^*\right] \;=\; 1 \quad \textit{and} \quad \mathrm{SG} \,\cap\, \Big(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\Big) \;=\; \emptyset.$$

Example II.3.9. *It is known that Root's solution $\overline{\mathbb{P}}_{Root}$ solves the optimization problem, i.e. given a strictly concave function $h : \mathbb{R}_+ \to \mathbb{R}$, there exists a barrier $\mathcal{R} \subseteq \mathbb{R}_+ \times \mathbb{R}$ such that*

$$\overline{P}(\mu) \;=\; \mathbb{E}^{\overline{\mathbb{P}}_{Root}}[h(B,T)] \quad \textit{with} \quad T \;=\; \inf\Big\{t \in \mathbb{R}_+ : \; (t, B_t) \in \mathcal{R}\Big\}, \; \overline{\mathbb{P}}_{Root} \text{ - } a.s.$$

Here a Borel set $\mathcal{R} \subseteq \mathbb{R}_+ \times \mathbb{R}$ is called a barrier if $(s,x) \in \mathcal{R}$ and $s < t$ imply $(t,x) \in \mathcal{R}$. We show here that the monotonicity principle allows to recover Root's solution. It follows by Theorem II.3.8, pick the set $\overline{\Gamma}^*$ such that

$$(B,T) \;\in\; \overline{\Gamma}^*, \; \overline{\mathbb{P}}_{Root} \text{ - } a.s. \quad \textit{and} \quad \mathrm{SG} \,\cap\, \Big(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\Big) \;=\; \emptyset.$$

Moreover, the concavity of h yields that

$$\mathrm{SG} \;=\; \Big\{\Big((\omega,\theta),(\omega',\theta')\Big) \in \overline{\Omega} \times \overline{\Omega} : \; \omega_\theta = \omega'_{\theta'} \textit{ and } \theta < \theta'\Big\}.$$

Define the barrier by

$$\mathcal{R} \;:=\; \Big\{(t,x)\in\mathbb{R}_+\times\mathbb{R}:\ \textit{there exists } (\omega,\theta)\in\overline{\Gamma}^*\ \textit{such that } \omega_\theta = x\ \textit{and } t<\theta\Big\},$$

then it follows from $\mathrm{SG}\cap(\overline{\Gamma}^{*<}\times\overline{\Gamma}^*)=\emptyset$ *that* $\mathcal{R}$ *is the required barrier.*

In fact, we prove the following more general result which requires more technical notations. For the optimizer $\overline{\mathbb{P}}^*$, we have a set SG^* depending on $\overline{\mathbb{P}}^*$ such that

$$\mathrm{SG}\;\cap\;\Big(\overline{\Gamma}^{*<}\times\overline{\Gamma}^*\Big)\;\subseteq\;\mathrm{SG}^*\;\cap\;\Big(\overline{\Gamma}^{*<}\times\overline{\Gamma}^*\Big)\;=\;\emptyset.$$

Next let us explain how to construct this set SG^*. It follows by Stroock & Varadhan [98] that $\overline{\mathbb{P}}^*$ admits a family of regular conditional probability distributions $(\overline{\mathbb{P}}^*_{\bar\omega})_{\bar\omega\in\overline{\Omega}}$ with respect to $\overline{\mathcal{F}}^B := \sigma(B_t, t\geq 0)$ that is the filtration generated by B on $\overline{\Omega}$. Notice that for $\bar\omega=(\omega,\theta)$, the measure $\overline{\mathbb{P}}^*_{\bar\omega}$ is independent of θ, then we may denote this family by $(\overline{\mathbb{P}}^*_\omega)_{\omega\in\Omega}$. Next, for every $\bar\omega\in\overline{\Omega}$, define a probability $\overline{\mathbb{Q}}^1_{\bar\omega}$ on $\overline{\Omega}$ by

$$\overline{\mathbb{Q}}^1_{\bar\omega}[\overline{A}] \;:=\; \int_\Omega \overline{\mathbb{P}}^*_{\omega\otimes_\theta\omega'}[\overline{A}]\;\mathbb{P}_0(d\omega')\ \text{for all } \overline{A}\in\overline{\mathcal{F}}\;:=\;\bigvee_{t\geq 0}\overline{\mathcal{F}}_t. \tag{II.3.9}$$

We define further, for every $\bar\omega\in\overline{\Omega}$, a probability $\overline{\mathbb{Q}}^2_{\bar\omega}$ by

$$\overline{\mathbb{Q}}^2_{\bar\omega}[\overline{A}] \;:=\; \overline{\mathbb{Q}}^1_{\bar\omega}\Big[\overline{A}\Big|T>\theta\Big]\,\mathbb{1}_{\{\overline{\mathbb{Q}}^1_{\bar\omega}[T>\theta]>0\}} \;+\; \mathbb{P}_0^{\theta,\omega}\otimes\delta_{\{\theta\}}[\overline{A}]\mathbb{1}_{\{\overline{\mathbb{Q}}^1_{\bar\omega}[T>\theta]=0\}} \tag{II.3.10}$$

for all $\overline{A}\in\overline{\mathcal{F}}$, where $\mathbb{P}_0^{\theta,\omega}$ is the shifted Wiener measure on Ω defined by

$$\mathbb{P}_0^{\theta,\omega}[A] \;:=\; \mathbb{P}_0\left[\omega\otimes_\theta B\in A\right]\ \text{ for all } A\in\mathcal{F}^0\;:=\;\bigvee_{t\geq 0}\mathcal{F}^0_t.$$

We finally introduce a shifted probability $\overline{\mathbb{Q}}^*_{\bar\omega}$ by

$$\overline{\mathbb{Q}}^*_{\bar\omega}[\overline{A}] \;:=\; \overline{\mathbb{Q}}^2_{\bar\omega}\Big[\bar\omega\otimes(B,T)\in\overline{A}\Big]\ \text{ for all } \overline{A}\in\overline{\mathcal{F}},$$

and then define a new set SG^* by

$$\mathrm{SG}^* \;:=\; \Big\{(\bar\omega,\bar\omega'):\omega_\theta=\omega'_{\theta'}\ \text{and}\ \Phi(\bar\omega)+\mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega'\otimes\cdot)\Big]\;>\;\mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega\otimes\cdot)\Big]+\Phi(\bar\omega')\Big\} \tag{II.3.11}$$

Finally, we obtain a more accurate characterization of the support of $\overline{\mathbb{P}}^*$.

Theorem II.3.10. *Under the same conditions of Theorem II.3.8, there exists a Borel subset* $\overline{\Gamma}^* \subseteq \overline{\Omega}$ *such that*

$$\overline{\mathbb{P}}^*\left[\overline{\Gamma}^*\right] \;=\; 1 \quad and \quad \mathrm{SG}^* \;\cap\; \left(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\right) \;=\; \emptyset.$$

Example II.3.11. *It follows by Theorem II.3.10 that, Vallois' solution can be recovered, i.e. if* $\Phi(B,T) = F(L_T^B)$*, where* $F : \mathbb{R}_+ \to \mathbb{R}$ *is a convex function and* $L^B = (L_t^B)_{t\geq 0}$ *is the local time of* B *at zero, then there exists a non-decreasing function* $\phi_+ : \mathbb{R}_+ \to \mathbb{R}_+$ *and a non-increasing function* $\phi_- : \mathbb{R}_+ \to \mathbb{R}_-$ *such that*

$$T \;=\; \inf\left\{t \in \mathbb{R}_+ :\; B_t \notin \left(\phi_-(L_t^B), \phi_+(L_t^B)\right)\right\},\; \overline{\mathbb{P}}_{Val} \text{ - } a.s.$$

where $\overline{\mathbb{P}}_{Val}$ *is the optimizer of problem* (II.3.3).

3.4 Stability of optimal Skorokhod embedding

3.4.1 Motivation

Recall that, under the hypothesis of no arbitrage, the marginal constraints in the Skorokhod embedding problem follow from the fact that the prices of calls are known for all strikes in the market. A more general situation is that there is a finite number of calls in the market. Therefore, we consider another optimization problem similar to optimal Skorokhod embedding problem, where the stopped Brownian motion only needs to reproduce a finite number of prices of calls.

Let $\boldsymbol{K} = (K_i)_{1\leq i\leq n}$ be the vector of strikes with $K_1 < \cdots < K_n$ that are available in the market and $\boldsymbol{C} = (C_{i,j})_{1\leq i\leq m, 1\leq j\leq n}$ be the corresponding price matrix of calls, where m and n denote respectively the number of maturities and the number of strikes. Under the hypothesis of no-arbitrage, *i.e.* there exists a centered peacock $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$ such that $C_{i,j}^n = \mu_i((x - K_j^n)^+)$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$. Recall that

$$\overline{\mathcal{P}} \;:=\; \left\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) : B \text{ is an } \overline{\mathbb{F}} - \text{Brownian motion and } B_{T_m\wedge\cdot} \text{ is UI under } \overline{\mathbb{P}}\right\}.$$

A probability $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$ is called a $(\boldsymbol{K}, \boldsymbol{C})-$embedding if

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[(B_{T_i} - K_j)^+\right] \;=\; C_{i,j}, \;\text{ for all } i = 1, \cdots, m \text{ and } j = 1, \cdots, n. \tag{II.3.12}$$

Each $(\boldsymbol{K}, \boldsymbol{C})-$embedding can be seen as a model calibrated to the market. Denote by $\overline{\mathcal{P}}(\boldsymbol{K}, \boldsymbol{C})$

the collection of all $(\boldsymbol{K}, \boldsymbol{C})-$embeddings. Similarly, we may define the optimal embedding problem by

$$\overline{P}(\boldsymbol{K}, \boldsymbol{C}) \quad := \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{K}, \boldsymbol{C})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big]. \tag{II.3.13}$$

3.4.2 Main results

Notice that the existence of μ is ensured by the hypothesis of no-arbitrage. However, μ is unknown since it can not be determined by a finite number of prices of calls. The objective here is to capture $\overline{P}(\boldsymbol{\mu})$ by $\overline{P}(\boldsymbol{K}, \boldsymbol{C})$ and to study the asymptotic behavior of $\overline{P}(\boldsymbol{K}, \boldsymbol{C})$.

The first question appearing immediately is the following : Given more and more call options, does the above optimization problem $\overline{P}(\boldsymbol{K}, \boldsymbol{C})$ converge to $\overline{P}(\boldsymbol{\mu})$? Then the following theorem ensures that our optimization problem is consistent with our intuition.

Theorem II.3.12. *Under suitable conditions one has*

$$\lim_{n \to \infty} \overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n) \quad = \quad \overline{P}(\boldsymbol{\mu}),$$

for any sequence $(\boldsymbol{K}^n, \boldsymbol{C}^n)_{n \geq 1}$ such that $\cup_{n \geq 1} \boldsymbol{K}^n$ is dense in $\mathbb{R}$, where $C^n_{i,j} = \mu_i((x - K^n_j)^+)$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$.

Next, we are interested in analyzing the convergence rate by means of different metrics on the space of probability measures. Assume that μ_m has a finite moment of order $p > 1$, *i.e.* $\mu_m(|x|^p) \equiv V < +\infty$. Then the following proposition shows that the map $\boldsymbol{\mu} \mapsto \overline{P}(\boldsymbol{\mu})$ is Hölder with respect to Lévy-Prokhorov and Wasserstein distances.

Let q be the conjugate number of p, *i.e.* $1/p + 1/q = 1$. We denote respectively by $\rho(\cdot, \cdot)$ the Lévy-Prokhorov metric and by $\mathcal{W}_1(\cdot, \cdot)$ the Wasserstein metric, *i.e.* for any two probability measures μ and $\nu \in \mathbf{P}$, one has

$$\begin{aligned} \rho(\mu, \nu) \quad &:= \quad \inf \Big\{\varepsilon > 0 : \ F_\mu(x - \varepsilon) - \varepsilon \leq F_\nu(x) \leq F_\mu(x + \varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\Big\}, \\ \mathcal{W}_1(\mu, \nu) \quad &:= \quad \inf_{\mathbb{P} \in \mathcal{P}(\mu, \nu)} \mathbb{E}^{\mathbb{P}}\Big[|X - Y|\Big], \end{aligned}$$

where F_μ (resp. F_ν) denotes the cumulative distribution function of μ (resp. ν) and $\mathcal{P}(\mu, \nu)$ is defined by (II.2.2).

Proposition II.3.13. *Under suitable conditions, there is a constant $C > 0$ such that*

(i) *if* $m = 1$

$$\left|\overline{P}(\mu) \ - \ \overline{P}(\nu)\right| \ \leq \ C\rho(\mu,\nu)^{1/2q};$$

(ii) *if* $m \geq 1$

$$\left|\overline{P}(\boldsymbol{\mu}) \ - \ \overline{P}(\boldsymbol{\nu})\right| \ \leq \ C\sum_{i=1}^{m}\mathcal{W}_1(\mu_i,\nu_i)^{\frac{p-2}{p-1}}.$$

For the purpose of estimating the convergence rate, we need to consider a subset $\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C})$ defined by

$$\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C}) \ := \ \left\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{K},\boldsymbol{C}) : \ \mathbb{E}^{\overline{\mathbb{P}}}\Big[|B_{T_m}|^p\Big] \ = \ V\right\}.$$

The restriction of the embeddings comes from an additional information : a Power option of the last maturity is quoted in the market. Put similarly

$$\overline{P}^V(\boldsymbol{K},\boldsymbol{C}) \ := \ \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big].$$

Theorem II.3.14. *Under suitable conditions, there is a constant $C > 0$ such that*

(i) *if* $m = 1$

$$0 \ \leq \ \overline{P}^V(\boldsymbol{K}^n,\boldsymbol{C}^n) \ - \ \overline{P}(\mu) \ \leq \ C\left((\Delta\boldsymbol{K}^n)^{1/4q} \ + \ |\boldsymbol{K}^n|^{-p/4q^2}\right);$$

(ii) *if* $m \geq 1$

$$0 \ \leq \ \overline{P}^V(\boldsymbol{K}^n,\boldsymbol{C}^n) \ - \ \overline{P}(\boldsymbol{\mu}) \ \leq \ C|\boldsymbol{K}^n|^{\frac{p-2}{p-1}}\left(\sqrt{\Delta\boldsymbol{K}^n} \ + \ |\boldsymbol{K}^n|^{-p/2q}\right)^{\frac{p-2}{p-1}},$$

where

$$|\boldsymbol{K}^n| \ := \ (K_1^n)^- \wedge (K_n^n)^+ \quad \textit{and} \quad \Delta\boldsymbol{K}^n \ := \ \max_{1<i\leq n}\left(K_i^n - K_{i-1}^n\right).$$

4 Stochastic control approach for options on local time

4.1 Motivation and formulation

We consider the application of stochastic control to martingale optimal transport on the space of continuous functions with payoff depending on the local time, and to the Skorokhod embedding problem. Recall that $\mathbf{C}$ is the space of continuous functions $\omega = (\omega_t)_{0\le t\le 1}$ with $\omega_0 = 0$. With the same notations of Section 3.2.2, let $X = (X_t)_{0\le t\le 1}$ be the canonical process, $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ be its natural filtration and $\mathcal{M}$ be the set of martingale measures $\mathbb{P}$ on $(\mathbf{C}, \mathcal{F}_1)$. We fix a centered peacock $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m}$ and a set of maturities $\mathbb{T} = \{t_1 < \cdots < t_m = 1\} \subseteq (0,1]$, and consider the the collection of martingale transport plans

$$\mathcal{M}(\boldsymbol{\mu}) \;:=\; \Big\{\mathbb{P}\in\mathcal{M} :\; \mathbb{P}\circ X_{t_k}^{-1} \;=\; \mu_k \text{ for all } k = 1,\dots,m\Big\}.$$

Let $L = (L_t)_{0\le t\le 1}$ be the local time of X at zero. For a convex function $F:\mathbb{R}_+\to\mathbb{R}$, define the martingale optimal transport problems

$$\overline{\mathsf{P}}(\boldsymbol{\mu}) \;:=\; \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[F(L_1)\Big], \tag{II.4.1}$$

$$\underline{\mathsf{P}}(\boldsymbol{\mu}) \;:=\; \inf_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[F(L_1)\Big]. \tag{II.4.2}$$

Denote by $\overline{\mathcal{H}}$ (resp. $\underline{\mathcal{H}}$) the collection of all $\mathbb{F}-$progressively measurable processes $H:[0,1]\times\mathbf{C}\to\mathbb{R}$ such that,

$$\int_0^1 H_t^2 d\langle X\rangle_t \;<\; +\infty,\; \mathbb{P}\text{ - a.s. and the stochastic integral } (H\cdot X) \;:=\; \int_0^{\cdot} H_t dX_t$$
$$\text{is a } \mathbb{P}-\text{supermartingale (resp. submartingale) for all } \mathbb{P}\in\mathcal{M}. \tag{II.4.3}$$

Define further

$$\overline{\mathcal{D}} \text{ (resp. } \underline{\mathcal{D}}) \;:=\; \Big\{(\boldsymbol{\lambda}, H)\in\Lambda^m\times\overline{\mathcal{H}} \text{ (resp. } \underline{\mathcal{H}}) :\; \boldsymbol{\lambda}(X) \;+\; (H\cdot X)_1$$
$$\ge\; \text{(resp. } \le) \; F(L_1),\; \mathbb{P}\text{ - a.s. for all } \mathbb{P}\in\mathcal{M}\Big\} \tag{II.4.4}$$

Then the dual problems are defined by

$$\overline{\mathsf{D}}(\boldsymbol{\mu}) \text{ (resp. } \underline{\mathsf{D}}(\boldsymbol{\mu})) \;:=\; \inf_{(\boldsymbol{\lambda},H)\in\overline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda}) \text{ (resp. } \sup_{(\boldsymbol{\lambda},H)\in\underline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda})). \tag{II.4.5}$$

4.2 Main results

4.2.1 One-marginal case : Vallois' solutions and dualities

For the one-marginal case, we provide the explicit optimizers for the primal and dual problems using Vallois' solutions.

Let $B = (B_t)_{t\geq 0}$ be a Brownian motion, then it follows by Vallois [102] that, there exist a non-decreasing (resp. non-increasing) right-continuous $\overline{\phi}_+ : \mathbb{R}_+ \to \mathbb{R}_+$ (resp. $\underline{\phi}_+ : \mathbb{R}_+^* \to \mathbb{R}_+$) and a non-increasing (resp. non-decreasing) right-continuous function $\overline{\phi}_- : \mathbb{R}_+ \to \mathbb{R}_-$ (resp. $\underline{\phi}_- : \mathbb{R}_+^* \to \mathbb{R}_-$) such that $B_{\overline{\tau}\wedge\cdot}$ (resp. $B_{\underline{\tau}\wedge\cdot}$) is uniformly integrable and $B_{\overline{\tau}} \sim \mu$ (resp. $B_{\underline{\tau}} \sim \mu$), where

$$\overline{\tau} \;:=\; \inf\Big\{t>0 : B_t \notin \Big(\overline{\phi}_-(L_t^B), \overline{\phi}_+(L_t^B)\Big)\Big\} \;\Big(\text{resp. } \underline{\tau} \;:=\; \inf\Big\{t>0 : B_t \notin \Big(\underline{\phi}_-(L_t^B), \underline{\phi}_+(L_t^B)\Big)\Big\}\Big).$$

Here $L^B = (L_t^B)_{t\geq 0}$ denotes the local time of B at zero. Define $\overline{\gamma} : \mathbb{R}_+ \to \mathbb{R}_+$ (resp. $\underline{\gamma} : \mathbb{R}_+ \to \mathbb{R}_+$) by $\overline{\gamma}(0) = 0$ (resp. $\underline{\gamma}(0) = 0$) and for all $l > 0$

$$\overline{\gamma}(l) \;:=\; \frac{1}{2}\int_0^l \left(\frac{1}{\overline{\phi}_+(s)} - \frac{1}{\overline{\phi}_-(s)}\right) ds \;\left(\text{resp. } \underline{\gamma}(l) \;:=\; \frac{1}{2}\int_0^l \left(\frac{1}{\underline{\phi}_+(s)} - \frac{1}{\underline{\phi}_-(s)}\right) ds\right)$$

Define further

$$\overline{\lambda}'(\pm x) \;:=\; \overline{\lambda}'(0\pm) \;+\; \int_0^{\overline{\psi}_\pm(\pm x)} \frac{dy}{\overline{\phi}_\pm(y)} e^{\overline{\gamma}(y)} \int_y^{+\infty} e^{-\overline{\gamma}(z)} F''(dz) \text{ for all } x>0, \tag{II.4.6}$$

$$\overline{\lambda}'(0\pm) \;:=\; \pm F'(0) \;\pm\; \int_0^{+\infty} e^{-\overline{\gamma}(z)} F''(dz), \tag{II.4.7}$$

$$\overline{\lambda}(0) \;:=\; F(0), \tag{II.4.8}$$

and

$$\underline{\lambda}'(\pm x) \;:=\; \underline{\lambda}'(0\pm) \;-\; \int_{\underline{\psi}_\pm(\pm x)}^{+\infty} \frac{dy}{\underline{\phi}_\pm(y)} e^{\underline{\gamma}(y)} \int_y^{+\infty} e^{-\underline{\gamma}(z)} F''(dz) \text{ for all } x>0, \tag{II.4.9}$$

$$\underline{\lambda}'(0\pm) \;:=\; \pm F'(+\infty), \tag{II.4.10}$$

$$\underline{\lambda}(0) \;:=\; F(0) \;-\; \int_0^{+\infty} e^{\underline{\gamma}(y)} dy \int_y^{+\infty} e^{-\underline{\gamma}(z)} F''(dz), \tag{II.4.11}$$

where $\overline{\psi}_\pm$ (resp. $\underline{\psi}_\pm$) denote the right-continuous inverses of $\overline{\phi}_\pm$ (resp. $\underline{\phi}_\pm$), F' is the right derivative, $F''(dz)$ is the second derivative of F in the sense of distribution.

Theorem II.4.1. *Under suitable conditions,* $\overline{\mathsf{P}}(\mu)$ *(resp.* $\underline{\mathsf{P}}(\mu)$*) is attained by the law of* $(B_{\overline{\tau}\wedge t/(1-t)})_{0\leq t\leq 1}$

(resp. $(B_{\underline{\tau}\wedge t/(1-t)})_{0\le t\le 1}$*) and* $\overline{\mathsf{D}}(\mu)$ *(resp.* $\underline{\mathsf{D}}(\mu)$*) is attained by* $(\overline{\lambda},\overline{H})$ *(resp.* $(\underline{\lambda},\underline{H})$*), where* $\overline{\lambda}$ *(resp.* $\underline{\lambda}$*) is given by* (II.4.6), (II.4.7), (II.4.8)*(resp.* (II.4.9), (II.4.10), (II.4.11)*) and* $\overline{H} = (\overline{H}_t)_{0\le t\le 1}$ *(resp.* $\underline{H} = (\underline{H}_t)_{0\le t\le 1}$*) is defined as follows :*

$$\overline{H}_t \;:=\; -\overline{A}_+(L_t)\mathbb{1}_{\{X_t>0\}} - \overline{A}_-(L_t)\mathbb{1}_{\{X_t\le 0\}} \;\Big(\textit{resp. } \underline{H}_t \;:=\; -\underline{A}_+(L_t)\mathbb{1}_{\{X_t>0\}} - \underline{A}_-(L_t)\mathbb{1}_{\{X_t\le 0\}}\Big),$$

where $\overline{A}_\pm : \mathbb{R}_+ \to \mathbb{R}$ *(resp.* $\underline{A}_\pm : \mathbb{R}_+ \to \mathbb{R}$*) is given by*

$$\overline{A}_\pm(l) \;:=\; \overline{\lambda}'(0\pm) \;+\; \int_0^l \frac{dy}{\overline{\phi}_\pm(y)} e^{\overline{\gamma}(y)} \int_y^{+\infty} e^{-\overline{\gamma}(z)} F''(dz)$$
$$\left(\textit{resp. } \underline{A}_\pm(l) \;:=\; \underline{\lambda}'(0\pm) \;-\; \int_l^{+\infty} \frac{dy}{\underline{\phi}_\pm(y)} e^{\underline{\gamma}(y)} \int_y^{+\infty} e^{-\underline{\gamma}(z)} F''(dz)\right).$$

Moreover, the duality $\overline{\mathsf{P}}(\mu) = \overline{\mathsf{D}}(\mu)$ *(resp.* $\underline{\mathsf{P}}(\mu) = \underline{\mathsf{D}}(\mu)$*) holds.*

4.2.2 Two-marginal case : generalization of Vallois' solution

Next, we extend the analysis to the two-marginal case, where we construct a new solution to the two-marginal Skorokhod embedding problem as a generalization of Vallois' solution. For technical reasons, we assume in this section that μ_i are symmetric and have a density denoted by $\mu_i(x)dx$ for $i = 1, 2$.

For the first stopping time, we take the solution given by Vallois [102] that embeds μ_1, *i.e.*,

$$\tau_1 \;:=\; \inf\left\{t > 0 : \; |B_t| \;\ge\; \phi_1(L_t^B)\right\},$$

where $\phi_1 : \mathbb{R}_+ \to \mathbb{R}_+$ is the inverse of ψ_1 given by

$$\psi_1(x) \;:=\; \int_0^x \frac{y\mu_1(y)}{\mu_1([y,+\infty))} dy \text{ for all } x \in \mathbb{R}_+. \tag{II.4.12}$$

For the second stopping time, we look for an increasing function $\phi_2 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $B_{\tau_2\wedge\cdot}$ is uniformly integrable and $B_{\tau_2} \sim \mu_2$, where

$$\tau_2 \;:=\; \inf\left\{t \ge \tau_1 : \; |B_t| \;\ge\; \phi_2(L_t^B)\right\}.$$

As before, ϕ_2 is defined through its inverse ψ_2. First, we set

$$\psi_2(x) \;:=\; \int_0^x \frac{y\mu_2(y)}{\mu_2([y,+\infty))} dy \text{ for all } x \in [0, x_1], \tag{II.4.13}$$

where x_1 is the smallest element such that the r.h.s. above exceeds ψ_1. To ensure that $x_1 > 0$, we need to assume that $\delta\mu := \mu_2 - \mu_1 \leq 0$ on a neighborhood of zero. If $x_1 = +\infty$, the construction is over. This corresponds to the case where Vallois' embeddings are well ordered. Otherwise, we proceed by induction as follows :
(i) if $x_{2i-1} < +\infty$, we set for all $x \in (x_{2i-1}, x_{2i}]$,

$$\psi_2(x) \quad := \quad \psi_2(x_{2i-1}) \ + \ \int_{x_{2i-1}}^{x} \frac{y\delta\mu(y)}{\delta\mu([y, +\infty))} dy, \tag{II.4.14}$$

where x_{2i} is the smallest element such that ψ_1 exceeds the r.h.s. above ;
(ii) if $x_{2i} < +\infty$, we set for all $x \in (x_{2i}, x_{2i+1}]$,

$$\psi_2(x) \quad := \quad \psi_2(x_{2i}) \ + \ \int_{x_{2i}}^{x} \frac{y\mu_2(y)}{\mu_2([y, +\infty))} dy, \tag{II.4.15}$$

where x_{2i+1} is the smallest element such that the r.h.s. above exceeds $\psi_1^{\boldsymbol{\mu}}$.

Then we obtain the required embeddings by the following theorem.

Theorem II.4.2. *Under suitable conditions, if ψ_2 is given by (II.4.14) and (II.4.15), then $B_{\tau_2\wedge\cdot}$ is uniformly integrable and $B_{\tau_i} \sim \mu_i$ for $i = 1, 2$.*

4.2.3 A special case

Finally, a special multi-marginal case is studied, where the embedding stopping times given by Vallois's rule are well ordered. For the m-marginal case, *i.e.* $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$, the problem essentially reduces to the one-marginal case. In the full marginal setting, *i.e.* $\boldsymbol{\mu} = (\mu_t)_{0\leq t\leq 1}$, we construct a remarkable Markov martingale and compute its generator explicitly. In particular, this provides a new example of fake Brownian motion.

Assumption II.4.3. (i) *For all $0 \leq s \leq t \leq 1$, $\tau_s \leq \tau_t$, or equivalently, $\phi_s \leq \phi_t$, where ϕ_s and ϕ_t are the Vallois' functions corresponding to μ_s and μ_t ;*
(ii) *The map $(t, x) \mapsto \psi_t(x)$ admits first-order partial derivatives, where ψ_t denotes the right-continuous inverse of ϕ_t.*

The next result gives the generator of the Markov process $(B_{\tau_t})_{0\leq t\leq 1}$.

Theorem II.4.4. *Under Assumption II.4.3, $(B_{\tau_t})_{0\leq t\leq 1}$ is an inhomogeneous Markov martingale whose generator is given by*

$$\mathcal{L}_t f(x) \quad = \quad \frac{\partial_t\psi_t(|x|)}{\partial_x\psi_t(|x|)} \left(\frac{e^{\gamma_t\circ\psi_t(|x|)}}{2|x|} \int_{|x|}^{+\infty} \Big(f(y) + f(-y) - 2f(x)\Big) de^{-\gamma_t\circ\psi_t(y)} \ - \ f'(x) \right),$$

for all smooth bounded functions $f : \mathbb{R} \to \mathbb{R}$.

In particular, the process $(B_{\tau_t})_{0 \leq t \leq 1}$ is a pure jump process, which corresponds to an example of the local Lévy model introduced in Carr, Geman, Madan & Yor [19].

Chapitre III

Tightness and dualities of martingale transport on the Skorokhod space

1 Introduction

Initialed by the famous work of Monge and Kantorovich, the optimal transport problem concerns the optimal transfer of mass from one location to another. Namely, let $\mathcal{P}(\mathbb{R}^d)$ be the space of probability measures on the Euclidean space $\mathbb{R}^d$. For any given probabilities $\mu_0,\ \mu_1 \in \mathcal{P}(\mathbb{R}^d)$, put

$$\mathcal{P}(\mu_0,\mu_1) \;\; := \;\; \Big\{\mathbb{P} \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d):\ \mathbb{P} \circ X_i^{-1} \;\; = \;\; \mu_i \text{ for } i = 0, 1\Big\}, \qquad \text{(III.1.1)}$$

where (X_0, X_1) denotes the canonical process on $\mathbb{R}^d \times \mathbb{R}^d$, *i.e.* $X_0(x_0, x_1) = x_0$ and $X_1(x_0, x_1) = x_1$ for all $(x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d$. Then the optimal transport problem consists in optimizing the expectation of some measurable function $\xi : \mathbb{R}^{2d} \to \mathbb{R}$ among all probability measures in $\mathcal{P}(\mu_0, \mu_1)$. Various related issues are studied, such as the duality theory and optimality results. We refer to Rachev & Rüschendorf [90] and Villani [104] for a comprehensive account of the literature.

Recently, a martingale optimal transport problem was introduced in Beiglböck, Henry-Labordère & Penkner [7] in discrete-time, see also Galichon, Henry-Labordère & Touzi [43] for the continuous-time case, where a maximization problem is considered over a subset $\mathcal{M}(\mu_0, \mu_1) := \{\mathbb{P} \in \mathcal{P}(\mu_0, \mu_1) : \mathbb{E}^{\mathbb{P}}[X_1|X_0] = X_0,\ \mathbb{P}\text{ - a.s.}\}$, *i.e.*

$$\mathsf{P}^m(\mu_0, \mu_1) \;\; := \sup_{\mathbb{P}\in\mathcal{M}(\mu_0,\mu_1)} \mathbb{E}^{\mathbb{P}}\Big[\xi(X_0, X_1)\Big].$$

Each element of $\mathcal{M}(\mu_0, \mu_1)$ is called a martingale transport plan. Similarly to the classical setting, the corresponding dual problem is defined by

$$\mathsf{D}^m(\mu_0, \mu_1) \;\; := \inf_{(\lambda_0,\lambda_1,H)\in\mathcal{D}^m} \mu_0(\lambda_0) \; + \; \mu_1(\lambda_1),$$

with $\mathcal{D}^m$ being the collection of triplets $(\lambda_0, \lambda_1, H)$, where $\lambda_0 \in \mathbb{L}^1(\mu_0)$, $\lambda_1 \in \mathbb{L}^1(\mu_1)$ and $H \in \mathbb{L}^0(\mathbb{R}^d)$ satisfy

$$\lambda_0(x_0) \; + \; \lambda_1(x_1) \; + \; H(x_0)(x_1 - x_0) \;\; \geq \;\; \xi(x_0, x_1) \text{ for all } (x_0, x_1) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{III.1.2}$$

From the financial viewpoint, the process (X_0, X_1) may model d underlying assets in the market. When $d = 1$, as observed by Breeden & Litzenberger [16], the marginal distributions of X_0 and X_1 are recovered by the market prices of calls for all strikes, and any Vanilla option has a non-ambiguous price as the integral of its payoff function with respect to the marginal. Therefore, the inequality (III.1.2) represents a super-replication of ξ, which consists of a dynamic trading of the underlying and a static position of Vanilla options at different maturities. Since there is no specific model imposed on (X_0, X_1), the dual problem may be interpreted as the robust superhedging of ξ. Similar to the classical setting, the duality $\mathsf{P}^m(\mu_0, \mu_1) = \mathsf{D}^m(\mu_0, \mu_1)$ holds under quite general conditions.

It is natural to consider its generalization to the continuous-time case. Let $\mathcal{Z} := \{\omega = (\omega_t)_{0\leq t\leq 1} : \omega_t \in \mathbb{R}^d \text{ for all } t \in [0, 1]\}$, where $\mathcal{Z}$ is either the space of continuous functions or the Skorokhod space of càdlàg functions. Denote by $X = (X_t)_{0\leq t\leq 1}$ the canonical process and by $\mathcal{M}$ the set of all martingale measures $\mathbb{P}$, *i.e.* X is a $\mathbb{P}-$martingale. For a given family of probability measures $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}}$, where $\mathbb{T} \subseteq [0, 1]$ is a subset, define by $\mathcal{M}(\boldsymbol{\mu}) \subseteq \mathcal{M}$ the subset of martingale transport plans $\mathbb{P}$, *i.e.* $\mathbb{P} \circ X_t^{-1} = \mu_t$ for all $t \in \mathbb{T}$. Then for a measurable function $\xi : \mathcal{Z} \to \mathbb{R}$, the martingale optimal transport problem is defined by

$$\mathsf{P}(\boldsymbol{\mu}) \;\; := \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big]. \tag{III.1.3}$$

In contrast with the discrete-time case, the set $\mathcal{M}(\boldsymbol{\mu})$ is generally not tight with respect to the usual topologies. Without the crucial compactness, the arguments in the classical setting fail to be adapted to handle the related issues.

In the existing literature, there are two dual formulations for the problem (III.1.3). Galichon, Henry-Labordère & Touzi studied a class of transport plans defined by stochastic differential equations in [43] and introduced a quasi-sure dual problem. They applied a stochastic control approach and deduced the corresponding duality. Another important contribution is due to

Dolinsky & Soner [36, 37] and Hou & Obłój [63], where the dual problem is still pathwisely formulated as in (III.1.2). By discretizing the paths and a technical construction of approximated martingale measures, they avoid the compactness issue and derive the duality.

In particular, when $\mathcal{Z}$ denotes the space of continuous functions, the martingale optimal transport problem can be studied by means of the optimal Skorokhod embedding problem in view of the Dambis-Dubins-Schwarz theorem. Following the seminal paper of Hobson [56], this methodology generated developments in many directions, see *e.g.* Brown, Hobson & Rogers [17], Cox & Obłój [24, 25], Cox, Hobson & Obłój [23], Cox, Obłój & Touzi [26], Cox & Wang [28], Davis, Obłój & Raval [30], Gassiat, Oberhauser & dos Reis [44], Hobson & Klimmek [58–60], Hobson & Neuberger [61], Madan & Yor [76], *etc.* A thorough literature is provided in the survey papers Hobson [57] and Obłój [83].

The main contribution of Chapter III is to study systematically the tightness of the set $\mathcal{M}(\boldsymbol{\mu})$ by means of the $S-$topology introduced in Jakubowski [67]. Endowing properly the space of marginal laws with a Wasserstein kind topology, the tightness yields the upper semicontinuity of the map $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ and further the first duality, obtained by penalizing the marginal constraints. Using respectively the dynamic programming principle and the discretization argument introduced in Dolinsky & Soner [37], we obtain the second and third dualities from the first one.

In addition, the above analysis immediately gives rise to a stability consequence. Denote $\overline{\mathsf{P}} := \mathsf{P}$ and $\underline{\mathsf{P}}(\boldsymbol{\mu}) := \inf_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}[\xi(X)]$, then it is shown that the map $\boldsymbol{\mu} \mapsto \overline{\mathsf{P}}(\boldsymbol{\mu})$ (resp. $\boldsymbol{\mu} \mapsto \underline{\mathsf{P}}(\boldsymbol{\mu})$) is upper (resp. lower) semicontinuous, which yields the stability, *i.e.* for any sequence $(\boldsymbol{\mu}^n)_{n\geq 1}$ converging to $\boldsymbol{\mu}$, there exists a sequence $(\varepsilon_n)_{n\geq 1} \subseteq \mathbb{R}_+$ converging to zero such that

$$\Big[\underline{\mathsf{P}}(\boldsymbol{\mu}^n),\ \overline{\mathsf{P}}(\boldsymbol{\mu}^n)\Big] \ \subseteq\ \Big[\underline{\mathsf{P}}(\boldsymbol{\mu}) - \varepsilon_n,\ \overline{\mathsf{P}}(\boldsymbol{\mu}) + \varepsilon_n\Big] \text{ for all } n \geq 1,$$

which implies that the interval of no-arbitrage prices is stable with respect to the market.

Chapter III is organized as follows. We formulate the martingale optimal transport problem and introduce its dual problems in Section 2. In Section 3, the duality results are provided and we reduce the infinitely-many marginal constraints to the finitely-many marginal constraints. In Sections 4 and 5 we focus on the case of finitely-many marginals and provide all related proofs.

2 Martingale optimal transport

For all $0 \le s < t$, denote by $\mathbf{D}([s,t],\mathbb{R}^d)$ the space of càdlàg functions defined on $[s,t]$ taking values in $\mathbb{R}^d$. Set in particular $\mathbf{D} := \mathbf{D}([0,1],\mathbb{R}^d)$ with generic element denoted by ω. Denote further by $X = (X_t)_{0\le t\le 1}$ the canonical process, *i.e.* $X_t(\omega) := \omega_t$ and by $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ its natural filtration, *i.e.* $\mathcal{F}_t := \sigma(X_s, s \le t)$. Let $\mathcal{P}$ be the set of probability measures on $(\mathbf{D}, \mathcal{F}_1)$. A probability measure $\mathbb{P} \in \mathcal{P}$ is called a martingale measure if the canonical process X is a $\mathbb{P}-$martingale. Denote by $\mathcal{M}$ the collection of all martingale measures.

2.1 Peacock and martingale optimal transport

Recall that $\mathbf{P}$ is the space of all probability measures μ on $\mathbb{R}^d$ with finite first moment and Λ is the set of continuous functions $\lambda : \mathbb{R}^d \to \mathbb{R}$ with linear growth. A pair $(\mu,\nu) \in \mathbf{P}\times\mathbf{P}$ is said to be increasing in convex ordering if $\mu(\lambda) \le \nu(\lambda)$ holds for all convex functions $\lambda \in \Lambda$. This relation is denoted by $\mu \preceq \nu$. For a subset $1 \in \mathbb{T} \subseteq [0,1]$, recall that the $\mathbb{T}-$product of $\mathbf{P}$ is given by

$$\mathbf{P}^{\mathbb{T}} \;:=\; \Big\{\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}}:\; \mu_t \in \mathbf{P} \text{ for all } t \in \mathbb{T}\Big\}.$$

Definition III.2.1. *A family of probability measures $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}^{\mathbb{T}}$ is called a peacock ($\mathbb{T}-$peacock) if $\mu_s \preceq \mu_t$ holds for all $s,t \in \mathbb{T}$ such that $s \le t$. A peacock $\boldsymbol{\mu}$ is said to be càdlàg if the map $t \mapsto \mu_t$ is càdlàg on $\mathbb{T}$ with respect to the weak convergence. Denote by $\mathbf{P}^{\preceq}$ the set of all càdlàg peacocks.*

For each peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, define the set of martingale transport plans

$$\mathcal{M}(\boldsymbol{\mu}) \;:=\; \Big\{\mathbb{P} \in \mathcal{M}:\; \mathbb{P}\circ X_t^{-1} \;=\; \mu_t \text{ for all } t \in \mathbb{T}\Big\}. \tag{III.2.1}$$

We may assume without loss of generality that $\mathbb{T}$ is closed under the lower limit topology, *i.e.* the topology generated by all half-open intervals $[s,t) \subseteq [0,1]$, see *e.g.* Steen & Seebach [97]. Indeed, denote by $\overline{\mathbb{T}}$ the closure of $\mathbb{T}$ under the lower limit topology, then it follows that the law of X_t for $t \in \overline{\mathbb{T}}$ is uniquely determined by the right continuity of X. This implies that $\mathcal{M}(\overline{\boldsymbol{\mu}}) = \mathcal{M}(\boldsymbol{\mu})$, where $\overline{\boldsymbol{\mu}} := (\overline{\mu}_t)_{t\in\overline{\mathbb{T}}}$ is defined by

$$\overline{\mu}_t \;:=\; \lim_{n\to\infty}\mu_{t_n} \text{ for any sequence } (t_n)_{n\ge1} \subseteq \mathbb{T} \text{ decreasing to } t. \tag{III.2.2}$$

Remark III.2.2. (i) *Since $\mu_{t_n} \preceq \mu_1$ for all n, we have*

$$\mu_{t_n}\Big((x_i - K)^+\Big) \;\; \le \;\; \mu_1\Big((x_i - K)^+\Big) \text{ for all } i = 1, \cdots, d \text{ and } K \in \mathbb{R},$$

thus showing that the sequence $(\mu_{t_n})_{n\ge1}$ is uniformly integrable. In particular, $(\mu_{t_n})_{n\ge1}$ is tight, and we may verify immediately by a direct density argument that any two possible accumulation points $\overline{\mu}_t$ and $\overline{\mu}'_t$ coincides, i.e. $\overline{\mu}_t = \overline{\mu}'_t$. Hence the sequence $(\mu_{t_n})_{n\ge1}$ converges weakly, justifying the convergence in (III.2.2) is well defined.

(ii) *When $\mathbb{T} = [0,1]$, $\mathcal{M}(\boldsymbol{\mu})$ is nonempty by Strassen's theorem, see e.g. Hirsch & Roynette [54, 55] and Kellerer [72]. For a general closed $\mathbb{T}$, we may extend $\boldsymbol{\mu}$ to some $\overline{\boldsymbol{\mu}} = (\overline{\mu}_t)_{0\le t\le1}$ by $\overline{\mu}_t := \mu_{\overline{t}}$ with $\overline{t} := \inf\{s \ge t : s \in \mathbb{T}\}$. Clearly, $\overline{\boldsymbol{\mu}} \in \mathbf{P}^{[0,1]}$ is a càdlàg peacock and $\overline{\mu}_t = \mu_t$ for all $t \in \mathbb{T}$. Hence $\mathcal{M}(\boldsymbol{\mu}) \supseteq \mathcal{M}(\overline{\boldsymbol{\mu}})$ is again nonempty.*

Let $\xi : \mathbf{D} \to \mathbb{R}$ be a measurable function. For every peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, define the martingale optimal transport problem by

$$\mathsf{P}(\boldsymbol{\mu}) \;\; := \;\; \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big], \tag{III.2.3}$$

where $\mathbb{E}^{\mathbb{P}}[\xi] := \mathbb{E}^{\mathbb{P}}[\xi^+] - \mathbb{E}^{\mathbb{P}}[\xi^-]$ with the convention $+\infty - \infty = -\infty$.

2.2 Dual problems

First dual problem Recall that

$$\Lambda^{\mathbb{T}} \;\; := \;\; \Big\{\boldsymbol{\lambda} = (\lambda_{t_i})_{1\le i\le k} : \; t_i \in \mathbb{T}, \; \lambda_{t_i} \in \Lambda \text{ for all } i = 1, \cdots, k, \; k \in \mathbb{N}\Big\}.$$

For every $\boldsymbol{\lambda} = (\lambda_{t_i})_{1\le i\le k} \in \Lambda^{\mathbb{T}}$ and $\omega \in \mathbf{D}$, denote

$$\boldsymbol{\lambda}(\omega) \;\; := \;\; \sum_{i=1}^{k} \lambda_{t_i}(\omega_{t_i}) \quad \text{and} \quad \boldsymbol{\mu}(\boldsymbol{\lambda}) \;\; := \;\; \sum_{i=1}^{k} \mu_{t_i}(\lambda_{t_i}).$$

Next, we introduce three dual formulations. Roughly speaking, as X is required to be a martingale and to have the given marginal laws in problem (III.2.3), then we penalize respectively these two constraints. The first dual problem is defined by

$$\mathsf{D}_1(\boldsymbol{\mu}) \;\; := \;\; \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \; + \; \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) \; - \; \boldsymbol{\lambda}(X)\Big]\Big\}. \tag{III.2.4}$$

The dual problem D_1 seems like that Lagrangian of constrained optimization problems, where the marginal constraints $\boldsymbol{\mu}$ are penalized by the Lagrange multipliers $\boldsymbol{\lambda}$.

Second dual problem The second dual problem penalizes further the martingale constraint and has close analogues in the mathematical finance literature. For technical reasons, we focus on the case where the underlying process X is non-negative and starts at some fixed price that may be normalized to $\mathbf{1} := (1, \cdots, 1) \in \mathbb{R}^d$. Namely, define the set of market scenarios

$$\mathbf{D}_+ \quad := \quad \left\{ \omega \in \Omega : \; \omega_0 = \mathbf{1} \text{ and } \omega_t \in \mathbb{R}^d_+ \text{ for all } t \in [0,1] \right\}$$

and the set of all possible models $\mathcal{M}_+ := \{\mathbb{P} \in \mathcal{M} : \operatorname{supp}(\mathbb{P}) \subseteq \mathbf{D}_+\}$. Consequently, the market calibration $\boldsymbol{\mu}$ should satisfy

$$\mu_0(dx) \;=\; \delta_{\mathbf{1}}(dx) \quad \text{and} \quad \operatorname{supp}(\mu_1) \;\subseteq\; \mathbb{R}^d_+. \tag{III.2.5}$$

Moreover, let us denote by $\mathbb{F}^U = (\mathcal{F}^U_t)_{0\le t\le 1}$ the universally completed filtration, *i.e.* $\mathcal{F}^U_t := \cap_{\mathbb{P}\in\mathcal{P}} \mathcal{F}^{\mathbb{P}}_t$, where $\mathcal{F}^{\mathbb{P}}_t$ is the completed σ-field of $\mathcal{F}_t$ under $\mathbb{P}$.

Definition III.2.3. *An $\mathbb{F}^U-$adapted process $S : [0,1]\times\mathbf{D}_+ \to \mathbb{R}$ is called an $\mathcal{M}_+-$supermartingale if it is a $\mathbb{P}-$supermartingale for all $\mathbb{P} \in \mathcal{M}_+$. Denote by $\mathcal{S}$ the collection of such processes and by $\mathcal{S}_0 \subseteq \mathcal{S}$ the subset of processes starting at zero. Denote further*

$$\mathcal{D}_2(\xi) \quad := \quad \left\{ (\boldsymbol{\lambda}, S) \in \Lambda^{\mathbb{T}} \times \mathcal{S}_0 : \; \boldsymbol{\lambda}(\omega) \; + \; S_1(\omega) \; \ge \; \xi(\omega) \textit{ for all } \omega \in \mathbf{D}_+ \right\}.$$

For a peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ satisfying (III.2.5) , the second dual problem is defined by

$$\mathsf{D}_2(\boldsymbol{\mu}) \quad := \inf_{(\boldsymbol{\lambda},S)\in\mathcal{D}_2(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{III.2.6}$$

Remark III.2.4. *Notice that the supermartingale $S \in \mathcal{S}$ is not required to have any regularity. If it were càdlàg then it would follow from Theorem 2.1 in Kramkov [74] that, for every $\mathbb{P} \in \mathcal{M}_+$ there exist a predictable process $H^{\mathbb{P}} = (H^{\mathbb{P}}_t)_{0\le t\le 1}$ and an optional non-decreasing process $A^{\mathbb{P}} = (A^{\mathbb{P}}_t)_{0\le t\le 1}$ such that*

$$S_t \;=\; S_0 \;+\; \int_0^t H^{\mathbb{P}}_s dX_s \;-\; A^{\mathbb{P}}_t \textit{ for all } t \in [0,1], \;\; \mathbb{P} \textit{ - a.s.}$$

However, it is not clear whether one can aggregate the last representation, i.e. find predictable processes H and A such that $(H, A) = (H^{\mathbb{P}}, A^{\mathbb{P}})$, $\mathbb{P}$ - almost surely. It follows by Nutz [80], *under the continuum hypothesis, the aggregation result holds when X takes values in the space of continuous functions, since we may identify the integrand in the stochastic integral by $H_t = H_t^{\mathbb{P}} = d\langle S, X\rangle_t / d\langle X\rangle_t$, where the processes $\langle S, X\rangle$ and $\langle X\rangle$ denote the quadratic co-variation of S and X and the quadratic variation of X, and they can be defined universally.*

Third dual problem The third dual problem is roughly parallel to the second one, but a bit stronger. It can be seen as an integral representation of supermartingales under a non-dominated family of measures.

The return from a zero-initial cost dynamic trading, defined by a suitable process $H = (H_t)_{0\leq t\leq 1}$, is given by the stochastic integral $(H \cdot X)$ which we define similarly to Dolinsky & Soner [37]. We restrict $H : [0,1] \to \mathbb{R}^d$ to be left-continuous with bounded variation. Then, we may define the stochastic integral by integration by parts :

$$(H \cdot X)_t \;:=\; H_t \cdot X_t \;-\; H_0 \cdot X_0 \;-\; \int_0^t X_s \cdot dH_s \text{ for all } t \in [0,1], \tag{III.2.7}$$

where $\int_0^t X_s \cdot dH_s$ refers to the scalar Lebesgue-Stieltjes integration.

Definition III.2.5. *An $\mathbb{F}-$adapted process $H : [0,1] \times \mathbf{D}_+ \to \mathbb{R}^d$ is called a dynamic strategy if $t \mapsto H_t(\omega)$ is left-continuous and of bounded variation for every $\omega \in \mathbf{D}_+$ and $(H \cdot X)$ is a $\mathbb{P}-$supermartingale for all $\mathbb{P} \in \mathcal{M}_+$. Let $\mathcal{H}$ be the set of all dynamic strategies and define the set of robust super-replications*

$$\mathcal{D}_3(\xi) \;:=\; \Big\{(\boldsymbol{\lambda}, H) \in \Lambda^{\mathbb{T}} \times \mathcal{H} :\; \boldsymbol{\lambda}(\omega) \;+\; (H \cdot X)_1(\omega) \;\geq\; \xi(\omega) \textit{ for all } \omega \in \mathbf{D}_+\Big\}.$$

For a peacock $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ satisfying (III.2.5), the third dual problem is defined by

$$\mathsf{D}_3(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda}, H)\in\mathcal{D}_3(\xi)} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{III.2.8}$$

Remark III.2.6. (i) *As discussed in the introduction, the process X stands for d risky assets, and the formulation* (III.2.8) *denotes the minimal cost allowing to construct a robust superhedging of ξ using the underlying assets and Vanilla options.*

(ii) *It is clear by definition that the weak duality $\mathsf{P}(\boldsymbol{\mu}) \leq \mathsf{D}_1(\boldsymbol{\mu})$ holds. Moreover, if the peacock $\boldsymbol{\mu}$ satisfies* (III.2.5)*, then*

$$\mathsf{P}(\boldsymbol{\mu}) \;\leq\; \mathsf{D}_1(\boldsymbol{\mu}) \;\leq\; \mathsf{D}_2(\boldsymbol{\mu}) \;\leq\; \mathsf{D}_3(\boldsymbol{\mu}).$$

3 Main results

We aim to study the existence of optimal martingale transport plans and establish the dualities in a systematic way. Before providing these results in Sections 3.2 and 3.3, we first introduce some topology on $\mathbf{D}$ and the associated space of probability measures in Section 3.1.

3.1 Preliminaries

In the classical optimal transport problem, the relevant results, *i.e.* existence of optimizers, duality, rely essentially on the compactness of $\mathcal{M}(\mu_0, \mu_1)$. However, when passing to the continuous-time case, as shown by Example III.3.1 below, the set $\mathcal{M}(\boldsymbol{\mu})$ is in general not tight with respect to the topologies L^∞ (uniform topology) and J_1 (Skorokhod topology). For our purpose, we endow $\mathbf{D}$ with the $S-$topology introduced by Jakubowski [67] such that the Borel σ-field agrees with the projection σ-field $\mathcal{F}_1$, and more importantly, the $S-$topology facilitates the tightness issue and both Skorokhod's representation theorem and Prohorov's theorem hold true. Before introducing the $S-$topology, we give an example which shows that the topologies L^∞ and J_1 are not convenient to handle the tightness of $\mathcal{M}(\boldsymbol{\mu})$.

Example III.3.1. *Let $M = (M_0, M_1, M_2)$ be a discrete-time martingale on some probability space such that $\mathbb{P}[M_0 \neq M_1$ and $M_1 \neq M_2] > 0$. Define $\mathbb{P}_n := \mathbb{P} \circ (M^n)^{-1}$ for $n \geq 3$, where $M^n = (M^n_t)_{0 \leq t \leq 1}$ is defined by*

$$M^n_t \;:=\; M_0 \mathbb{1}_{[0, \frac{1}{2} - \frac{1}{n})}(t) \;+\; M_1 \mathbb{1}_{[\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n})}(t) \;+\; M_2 \mathbb{1}_{[\frac{1}{2} + \frac{1}{n}, 1]}(t).$$

Clearly, $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu})$ for all $n \geq 3$ with $\mathbb{T} = \{0, 1\}$ and $\boldsymbol{\mu} = (\mathbb{P} \circ M_0^{-1}, \mathbb{P} \circ M_2^{-1})$. However, it follows from Theorem VI.3.21 in Jacod & Shiryaev [66] that, the sequence $(\mathbb{P}_n)_{n \geq 3}$ is not J_1-tight and thus not $L^\infty-$tight.

Definition III.3.2 ($S-$topology)**.** *The $S-$topology on $\mathbf{D}$ is the sequential topology induced by the following $S-$convergence, i.e. a set $F \subseteq \mathbf{D}$ is closed under $S-$topology if it contains all limits of its $S-$convergent subsequences, where the $S-$convergence (denoted by $\xrightarrow{S}$) is defined as follows. Let $(\omega^n)_{n \geq 0} \subseteq \mathbf{D}$, we say that $\omega^n \xrightarrow{S} \omega^0$ as $n \to \infty$ if for each $\varepsilon > 0$, we may find a sequence $(v^n_\varepsilon)_{n \geq 0} \subseteq \mathbf{D}$ such that*

$$v^n_\varepsilon \text{ has bounded variation, } \|\omega^n - v^n_\varepsilon\| \;\leq\; \varepsilon \text{ for all } n \geq 0$$

and

$$\lim_{n\to\infty}\int_{[0,1]} f(t)\cdot dv_\varepsilon^n(t) \;=\; \int_{[0,1]} f(t)\cdot dv_\varepsilon^0(t) \text{ for all continuous function } f,$$

where $\|\cdot\|$ *denotes the uniform norm. We denote by* $\xrightarrow{S^*}$ *the convergence induced by the* $S-$*topology.*

Remark III.3.3. (i) *It is shown in [67] that the* $S-$*topology is not metrizable. However, its associated Borel* σ*-field coincides with* $\mathcal{F}_1$*. In a metric space, a subset is sequentially closed if and only if it is closed; but in a non-metrizable space, a sequentially closed set may not be closed. In particular, a sequentially closed set under* $\xrightarrow{S^*}$ *may not be closed under* $S-$*topology (which is equivalent to be sequentially closed under* $\xrightarrow{S}$*). More precisely, it is shown in Remark 2.6 of [67] that the convergence* $\xrightarrow{S^*}$ *is weaker than the original one* $\xrightarrow{S}$*. However, this is not a real problem for our case, since we know, from [67],*

$$\begin{aligned}&\omega^n \xrightarrow{S^*} \omega, \text{ if and only if, in every subsequence } (n_k)_{k\geq 1},\\ &\text{one may find a further subsequence } (n_{k_l})_{l\geq 1} \text{ such that } \omega^{n_{k_l}} \xrightarrow{S} \omega.\end{aligned}$$

Hence, a function ξ *is* $S-$*continuous (resp. semicontinuous) if and only if* ξ *is* S^*-*continuous (resp. semicontinuous).*

(ii) *The functions* $\omega\mapsto\omega_{i,1}$*,* $\omega\mapsto\int_0^1\omega_{i,t}dt$ *and* $\omega\mapsto\int_0^1|\omega_t|dt$ *for* $i=1,\cdots,d$ *are* $S-$*continuous. The functions* $\omega\mapsto\|\omega\|$ *and* $\omega\mapsto\sup_{0\leq t\leq 1}\omega_{i,t}$ *for* $i=1,\cdots,d$ *are* $S-$*lower semicontinuous.*

Since the $S-$topology is not metrizable, then instead of the usual weak convergence, we use another convergence of probability measures introduced in [67], which induces a simple criteria for $S-$tightness and preserves Prohorov's theorem, *i.e.* tightness yields relative compactness.

Definition III.3.4. *Let* $(\mathbb{P}_n)_{n\geq 1}$ *be a sequence of probability measures on the space* $(\mathbf{D},\mathcal{F}_1)$*. We say* $\mathbb{P}_n \overset{*}{\Longrightarrow}_D \mathbb{P}$ *if for each subsequence* $(\mathbb{P}_{n_k})_{k\geq 1}$*, one can find a further subsequence* $(\mathbb{P}_{n_{k_l}})_{l\geq 1}$ *and stochastic processes* $(Y^l)_{l\geq 1}$ *and* Y *defined on the probability space* $([0,1],\mathcal{B}_{[0,1]},\ell)$ *endowed with the Lebesgue measure* ℓ*, such that* $\mathcal{L}(Y^l)=\mathbb{P}_{n_{k_l}}$ *for all* $l\geq 1$*,* $\mathcal{L}(Y)=\mathbb{P}$*,*

$$Y^l(e) \;\xrightarrow{S^*}\; Y(e) \quad \text{as} \quad l \;\longrightarrow\; \infty \text{ for all } e\in[0,1],$$

and for each $\varepsilon>0$*, there exists an* S^*-*compact subset* $K_\varepsilon\subseteq\Omega$ *such that*

$$\inf_{l\geq 1}\,\mathbb{P}_{n_{k_l}}\Big[X\in K_\varepsilon\Big] \;>\; 1-\varepsilon.$$

It follows from [67], see also Theorem III.6.1, that the convergence $\stackrel{*}{\Longrightarrow}_D$ implies in some sense the convergence of finite dimensional distributions that is specified later, and more importantly, the limit of every convergent sequence of martingale measures is still a martingale measure.

Remark III.3.5. *Meyer & Zheng [78] have also introduced a topology on $\mathbf{D}$ with $d = 1$, called pseudo-path topology, by considering the occupation measure induced by every path $\omega \in \mathbf{D}$ on $[0, 1] \times \mathbb{R}$. We notice that the convergence under the $S-$topology induces that under the pseudo-path topology, and hence it is easier to obtain the relative compactness of a sequence of martingale measures, but one has less continuous functionals defined on $\mathbf{D}$. In particular, the simple maps $\omega \mapsto \|\omega\|$, $\omega \mapsto \omega_1$ are not upper semicontinuous under the pseudo-path topology, which makes it unsuitable to study the current martingale optimal transport problem.*

We next introduce the Wasserstein distance for the purpose of deriving the duality $\mathsf{P} = \mathsf{D}_1$. Recall the set $\mathcal{P}(\mu, \nu)$ introduced in (III.1.1).

Definition III.3.6. *The Wasserstein distance of order 1 is defined by*

$$\mathcal{W}_1(\mu, \nu) \;\; := \inf_{\mathbb{P} \in \mathcal{P}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}\Big[\big|X_0 - X_1\big|\Big] \text{ for all } \mu, \nu \in \mathbf{P}.$$

A sequence $(\mu^n)_{n \geq 1} \subseteq \mathbf{P}$ converges to $\mu \in \mathbf{P}$ under $\mathcal{W}_1$ if $\mathcal{W}_1(\mu^n, \mu) \to 0$ as $n \to \infty$ or, equivalently, $\lim_{n\to\infty} \mu^n(\lambda) = \mu(\lambda)$ for all $\lambda \in \Lambda$, see e.g. Theorem 6.9 in Villani [104].

For $(\boldsymbol{\mu}^n = (\mu_t^n)_{t\in\mathbb{T}})_{n\geq 1} \subseteq \mathbf{P}^{\preceq}$ and $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}} \in \mathbf{P}^{\preceq}$, we say that $\boldsymbol{\mu}^n$ converges to $\boldsymbol{\mu}$ if μ_t^n converges to μ_t under $\mathcal{W}_1$ for all $t \in \mathbb{T}$ and this convergence is denoted by $\xrightarrow{\mathcal{W}_1^{\mathbb{T}}}$. We now provide a crucial tightness result which is a consequence from [67].

Let $\mathbb{T}_0 \subseteq \mathbb{T}$ be the collection of all condensation points under the lower limit topology, *i.e.* $t = 1$ or $[t, t + \varepsilon) \cap \mathbb{T}$ is uncountable for any $\varepsilon > 0$.

Definition III.3.7. *A sequence of probability measures $(\mathbb{P}_n)_{n\geq 1}$ on $\mathbf{D}$ is said to be $S-$tight, if for any $\varepsilon > 0$ one has a compact K under $S-$topology, such that*

$$\mathbb{P}_n[K] \;\; \geq \;\; 1 - \varepsilon \text{ for all } n \geq 1.$$

Lemma III.3.8. *Let $(\mathbb{P}_n)_{n\geq 1}$ be a sequence of probability measures such that $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu}^n)$ with $\boldsymbol{\mu}^n \in \mathbf{P}^{\mathbb{T}}_{\preceq}$, satisfying*

$$\boldsymbol{\mu}^n \;\; \xrightarrow{\mathcal{W}_1^{\mathbb{T}}} \;\; \boldsymbol{\mu} \;\in\; \mathbf{P}^{\preceq} \quad \text{as} \quad n \;\longrightarrow\; \infty. \tag{III.3.1}$$

(i) *Then,* $(\mathbb{P}_n)_{n\ge1}$ *is* $S-$*tight, i.e. any subsequence* $(\mathbb{P}_{n_k})_{k\ge1}$ *admits a further convergent subsequence under* $\overset{*}{\Longrightarrow}_D$. *Moreover, any limit point* $\mathbb{P}$ *of* $(\mathbb{P}_n)_{n\ge1}$ *is again a martingale measure.*

(ii) *Assume in addition that* $\mathbb{T}_0 = \mathbb{T}$, *then* $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$.

Proof. (i) By Theorem III.6.1 in Jakubowski [67], it is clear that $(\mathbb{P}_n)_{n\ge1}$ is $S-$tight and there exist a convergent subsequence $(\mathbb{P}_{n_k})_{k\ge1}$ with limit $\mathbb{P} \in \mathcal{P}$. Moreover, one has a countable subset $\mathcal{T} \subseteq [0,1)$ such that for any finite set $\{u_1,\cdots,u_r\} \subseteq [0,1]\setminus\mathcal{T}$,

$$\mathbb{P}_{n_k}\circ(X_{u_1},\cdots,X_{u_r})^{-1} \xrightarrow{\mathcal{L}} \mathbb{P}\circ(X_{u_1},\cdots,X_{u_r})^{-1} \quad \text{as} \quad k \longrightarrow \infty. \tag{III.3.2}$$

Let $s,t\in[0,1]\setminus\mathcal{T}$ such that $s<t$, and take a finite subset $\{u_1,\cdots,u_r\}\subseteq[0,s]\setminus\mathcal{T}$ and a sequence of bounded continuous functions $\{f_i\}_{1\le i\le r}$. Notice that for every $u\in[0,1]$, X_u is uniformly integrable with respect to $(\mathbb{P}_n)_{n\ge1}$. Indeed,

$$\begin{aligned}\lim_{R\to\infty}\sup_{n\ge1}\mathbb{E}^{\mathbb{P}_n}\Big[|X_u|\mathbb{1}_{\{|X_u|\ge R\}}\Big] &\le \lim_{R\to\infty}\sup_{n\ge1}\mathbb{E}^{\mathbb{P}_n}\Big[\Big(|X_u|-R/2\Big)^+\Big]\\ &\le \lim_{R\to\infty}\sup_{n\ge1}\mathbb{E}^{\mathbb{P}_n}\Big[\Big(|X_1|-R/2\Big)^+\Big]\\ &= \lim_{R\to\infty}\sup_{n\ge1}\ \mu_1^n\Big((|x|-R/2)_+\Big) \;=\; 0.\end{aligned} \tag{III.3.3}$$

Combining (III.3.2) and (III.3.3), one has

$$\mathbb{E}^{\mathbb{P}}\Big[f_1(X_{u_1})\cdots f_r(X_{u_r})\,(X_t-X_s)\Big] \;=\; \lim_{k\to\infty}\mathbb{E}^{\mathbb{P}_{n_k}}\Big[f_1(X_{u_1})\cdots f_r(X_{u_r})\,(X_t-X_s)\Big] \;=\; 0.$$

Since $\mathcal{T}$ is at most countable, it follows that $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}_s]=X_s$ for any $s,t\in[0,1]\setminus\mathcal{T}$ such that $s<t$. It follows by the right continuity of X that $\mathbb{P}\in\mathcal{M}$.

(ii) To prove that $\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})$, it remains to show that $\mathbb{P}\circ X_t^{-1}=\mu_t$ for all $t\in\mathbb{T}$. When $t\in\mathbb{T}\setminus\mathcal{T}$, by the convergence (III.3.2) and the fact that $\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^{\mathbb{T}}} \boldsymbol{\mu}$, it follows that $\mathbb{P}\circ X_t^{-1}=\mu_t$. Further, notice that $\mathbb{T}_0=\mathbb{T}$, then for every $t\in\mathbb{T}$, there exists a sequence $(t_i)_{i\ge1}\subseteq\mathbb{T}\setminus\mathcal{T}$ decreasing to t. Using again the right continuity of X, we conclude $\mathbb{P}\circ X_t^{-1}=\lim_{i\to\infty}\mathbb{P}\circ X_{t_i}^{-1}=\mu_t$. $\square$

As a consequence, the set $\mathcal{M}(\boldsymbol{\mu})$ is $S-$tight and it is closed if $\mathbb{T}_0=\mathbb{T}$. The following example shows that the closeness may fail when $\mathbb{T}_0\ne\mathbb{T}$.

Example III.3.9. *Let* $\mathbb{T}=\{0,1\}$ *and consider a random variable* Y *such that* $\mathbb{P}[Y=1]=\mathbb{P}[Y=-1]=1/2$. *Define* $\mathbb{P}_n:=\mathbb{P}\circ(M^n)^{-1}$ *for* $n\ge1$, *where* $M^n=(M_t^n)_{0\le t\le1}$ *is defined by*

$$M_t^n \;:=\; Y\mathbb{1}_{[\frac{1}{n},1]}(t).$$

Define a peacock $\boldsymbol{\mu} = (\mu_0, \mu_1)$ by $\mu_0 = \delta_{\{0\}}$ and $\mu_1 = \mathbb{P} \circ Y^{-1}$. Obviously, $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu})$ for all $n \geq 1$. However, the limit of $(\mathbb{P}_n)_{n\geq 1}$ is some measure $\mathbb{P}_0$ such that $X_t = X_0$, $\mathbb{P}_0$ - a.s. and $\mathbb{P}_0 \circ X_0^{-1} = \mu_1$, which does not lie in $\mathcal{M}(\boldsymbol{\mu})$.

3.2 Finitely-many marginal constraints

We start by studying the finitely-many marginal case and assume throughout this subsection that $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$. Denote $\Delta t_i := t_i - t_{i-1}$ for all $i = 1, \cdots, m$ and $\Delta\mathbb{T} := \min_{1\leq i\leq m} \Delta t_i$. Let us formulate some conditions on the function ξ. We shall see later that the usual examples satisfy our conditions.

Assumption III.3.10. $\limsup_{n\to\infty} \xi(\omega^n) \leq \xi(\omega)$ *holds for all $(\omega^n)_{n\geq 1} \subseteq \mathbf{D}$ and $\omega \in \mathbf{D}$ such that*

$$\omega^n \xrightarrow{S^*} \omega \quad \textit{and} \quad \omega^n_{t_i} \longrightarrow \omega_{t_i} \textit{ for all } i = 0, \cdots, m-1, \ \textit{as } n \longrightarrow \infty.$$

For $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_m) \in \mathbb{R}^m_+$ such that $|\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$, let $f_{\boldsymbol{\varepsilon}}$ (forward function) and $b_{\boldsymbol{\varepsilon}}$ (backward function) be two non-decreasing functions on $[0,1]$ defined by

$$f_{\boldsymbol{\varepsilon}}(t) := \sum_{i=1}^m \mathbb{1}_{(t_{i-1},t_i]}(t) \left(t_{i-1} + \frac{\Delta t_i}{\Delta t_i - \varepsilon_i} (t - t_{i-1} - \varepsilon_i)^+ \right), \tag{III.3.4}$$

$$b_{\boldsymbol{\varepsilon}}(t) := \sum_{i=1}^m \mathbb{1}_{(t_{i-1},t_i]}(t) \left(t_i - \left(\Delta t_i - \frac{\Delta t_i}{\Delta t_i - \varepsilon_i} (t - t_{i-1}) \right)^+ \right). \tag{III.3.5}$$

Assumption III.3.11. *There is a continuous function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ with $\alpha(0) = 0$ such that the following inequality holds for any $\boldsymbol{\varepsilon} \in \mathbb{R}^m_+$ satisfying $|\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$*

$$|\xi(\omega) - \xi(\omega_{f_{\boldsymbol{\varepsilon}}})| \leq \alpha(|\boldsymbol{\varepsilon}|) \left(1 + \sum_{i=0}^m |\omega_{t_i}| + \int_0^1 |\omega_t| dt \right), \tag{1}$$

$$|\xi(\omega) - \xi(\omega_{b_{\boldsymbol{\varepsilon}}})| \leq \alpha(|\boldsymbol{\varepsilon}|) \left(1 + \sum_{i=0}^m |\omega_{t_i}| + \int_0^1 |\omega_t| dt \right), \tag{2}$$

where $\omega_{f_{\boldsymbol{\varepsilon}}}$ (resp. $\omega_{b_{\boldsymbol{\varepsilon}}}$) denotes the composition of ω and $f_{\boldsymbol{\varepsilon}}$ (resp. $b_{\boldsymbol{\varepsilon}}$).

Theorem III.3.12. *Le ξ be bounded from above and satisfies Assumptions III.3.10 and III.3.11 (1). Then for all $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$:*

(i) *The duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu})$ holds.*

(ii) *Assuming further that ξ is bounded, the duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_2(\boldsymbol{\mu})$ holds for all $\boldsymbol{\mu}$ satisfying* (III.2.5).

To establish the duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_3(\boldsymbol{\mu})$, we need more regularity conditions on ξ. Let $\mathbf{D}([s,t],\mathbb{R}^d)$ denote the Skorokhod space of functions $\omega : [s,t] \to \mathbb{R}^d$ et $\rho_{[s,t]}$ denote the Skorokhod metric, *i.e.*

$$\rho_{[s,t]}(\omega,\omega') \;\; := \inf_{\gamma\in\Gamma([s,t])} \Big\{ \|\omega_\gamma - \omega'\| \bigvee \|\gamma - Id\| \Big\},$$

where, $\Gamma([s,t])$ denotes the set of continuous and strictly increasing functions $\gamma : [s,t] \to [s,t]$ such that $\gamma(s) = s$ and $\gamma(t) = t$ and Id denotes the identity function. Define a distance $\rho_{\mathbb{T}}$ on $\mathbf{D}$ by

$$\rho_{\mathbb{T}}(\omega,\omega') \;\; := \;\; \sum_{i=1}^m \rho_{[t_{i-1},t_i]}(\omega,\omega') \; + \; \Big|\int_0^1 (\omega_u - \omega'_u)\, du\Big| \;\text{ for all } \omega,\omega' \in \mathbf{D}, \qquad \text{(III.3.6)}$$

Clearly, $|\omega_{t_i} - \omega'_{t_i}| \le \rho_{\mathbb{T}}(\omega,\omega')$ for all $\omega,\omega' \in \Omega$ and $i = 0,\cdots,m$.

Assumption III.3.13. *ξ is locally $\rho_{\mathbb{T}}-$uniformly continuous, i.e. for every $R>0$, there exists a continuous increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\beta(0) = 0$, such that*

$$|\xi(\omega) \;-\; \xi(\omega')| \;\le\; \beta\left(\rho_{\mathbb{T}}(\omega,\omega')\right) \;\textit{for all } \omega,\omega' \in \mathbf{D} \textit{ with } \|\omega\|,\|\omega'\| \le R.$$

Theorem III.3.14. *Let ξ be bounded and $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ satisfying* (III.2.5)*. Then under Assumptions III.3.10, III.3.11 and III.3.13, the duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_3(\boldsymbol{\mu})$ holds.*

Remark III.3.15. (i) *Using the pathwise Doob's inequality in Acciaio, Beiglböck, Penkner, Schachermayer & Temme [1], the boundeness condition in Theorem III.3.14 may be removed when $\mu_1(|x|^p) < +\infty$ for some $p > 1$, see also Dolinsky & Soner [37].*

(ii) *It is easy to check that the payoff functions of usual exotic options satisfy the above assumptions. Indeed, let $F : \mathbb{R}^k \to \mathbb{R}$ be a bounded and Lipschitz function. Then Assumptions III.3.10, III.3.11 and III.3.13 are satisfied for*

1. *$\xi(X) = F(X_{t_1},\cdots,X_{t_m})$ (Vanilla option) when for $k = md$;*
2. *$\xi(X) = F(\sup_{0\le t\le 1} X_{1,t},\cdots,\sup_{0\le t\le 1} X_{d,t})$ (lookback option) when $k = d$, where $X_t = (X_{1,t},\cdots,X_{d,t})$ for all $t \in [0,1]$ and F is non-increasing with respect to each component.*
3. *$\xi(X) = F(\int_0^1 X_{1,t}dt,\cdots,\int_0^1 X_{d,t}dt)$ (Asian option) when $k = d$.*

(iii) *The proofs of Theorem III.3.12 and III.3.14 are respectively provided in Sections 4.2, 4.3 and 5.*

3.3 Infinitely-many marginal constraints

Using an approximation argument, we then obtain the duality results for the martingale transport problem under infinitely-many marginal constraints.

Proposition III.3.16. *Let ξ be S^*-upper semicontinuous and bounded from above. For all $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$:*

(i) *Assume that there exists an increasing sequence of finite sets $\{\mathbb{T}_n\}_{n\geq 1}$ such that $1 \in \mathbb{T}_n \subseteq \mathbb{T}$ for all $n \geq 1$ and $\cup_{n\geq 1}\mathbb{T}_n$ is dense in $\mathbb{T}$ under the lower limit topology. Then*

$$\lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) \;=\; \mathsf{P}(\boldsymbol{\mu}) \quad \textit{with} \quad \boldsymbol{\mu}^n \;:=\; (\mu_t)_{t\in\mathbb{T}_n}.$$

(ii) *Assume $\mathbb{T}_0 = \mathbb{T}$, then there exists an optimizer $\mathbb{P}^* \in \mathcal{M}(\boldsymbol{\mu})$, i.e.*

$$\mathsf{P}(\boldsymbol{\mu}) \;=\; \mathbb{E}^{\mathbb{P}^*}\Big[\xi(X)\Big]. \tag{III.3.7}$$

Proof. (i) It follows by the definition of $\boldsymbol{\mu}^n$ that $\mathsf{P}(\boldsymbol{\mu}^n)$ is non-increasing with respect to n. Take a sequence $(\mathbb{P}_n)_{n\geq 1}$ such that $\mathbb{P}_n \in \mathcal{M}(\boldsymbol{\mu}^n)$ and

$$\mathsf{P}(\boldsymbol{\mu}) \;\leq\; \lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) \;=\; \lim_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}[\xi].$$

By Lemma III.3.8 (i), there is a convergent subsequence $(\mathbb{P}_{n_k})_{k\geq 1}$ with some limit $\mathbb{P} \in \mathcal{M}$. It follows by the same arguments in the proof of Lemma III.3.8 that $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$ and, the upper semicontinuity of ξ yields

$$\lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) \;=\; \lim_{k\to\infty} \mathbb{E}^{\mathbb{P}_{n_k}}[\xi] \;\leq\; \mathbb{E}^{\mathbb{P}}[\xi] \;\leq\; \mathsf{P}(\boldsymbol{\mu}).$$

(ii) Take a maximizing sequence $(\mathbb{P}_n)_{n\geq 1} \subseteq \mathcal{M}(\boldsymbol{\mu})$, then we may get a limit point $\mathbb{P}^*$ and by Lemma III.3.8 (ii), $\mathbb{P}^*$ is the required optimizer. □

Consequently, we obtain immediately the dualities for a general $\mathbb{T}$ through Proposition III.3.16.

Theorem III.3.17. *Let ξ be S^*-upper semicontinuous and bounded from above and $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$, consider an increasing sequence of finite sets $\{\mathbb{T}_n\}_{n\geq 1}$ such that $1 \in \mathbb{T}_n \subseteq \mathbb{T}$ and $\cup_{n\geq 1}\mathbb{T}_n$ is dense in $\mathbb{T}$. Set $\boldsymbol{\mu}^n := (\mu_t)_{t\in\mathbb{T}_n}$:*

(i) *Assume that $\mathsf{P}(\boldsymbol{\mu}^n) = \mathsf{D}_1(\boldsymbol{\mu}^n)$ for all $n \geq 1$. Then $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu})$.*

(ii) *Assume further that $\boldsymbol{\mu}$ satisfies (III.2.5) and $\mathsf{P}(\boldsymbol{\mu}^n) = \mathsf{D}_2(\boldsymbol{\mu}^n) = \mathsf{D}_3(\boldsymbol{\mu}^n)$ for all $n \geq 1$. Then $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_2(\boldsymbol{\mu}) = \mathsf{D}_3(\boldsymbol{\mu})$.*

Proof. It is enough to show (i). Notice by definition that $\mathsf{D}_1(\boldsymbol{\mu}^n) \geq \mathsf{D}_1(\boldsymbol{\mu})$ for all $n \geq 1$, then it follows by Proposition III.3.16 (i) that

$$\mathsf{P}(\boldsymbol{\mu}) \quad = \quad \lim_{n\to\infty} \mathsf{P}(\boldsymbol{\mu}^n) \quad \geq \quad \mathsf{D}_1(\boldsymbol{\mu}).$$

Then the proof is fulfilled by the weak duality $\mathsf{P}(\boldsymbol{\mu}) \leq \mathsf{D}_1(\boldsymbol{\mu})$. □

Remark III.3.18. *In the present setting, the marginal constraint $\boldsymbol{\mu} = (\mu_t)_{t\in\mathbb{T}}$ is given by a family of joint distributions on $\mathbb{R}^d$. If we replace the marginal distribution μ_t by, either d marginal distributions $(\mu_t^1, \cdots, \mu_t^d)$ on $\mathbb{R}$, or a joint distribution $\vec{\mu}_{\vec{t}}$ on $\mathbb{R}^{m\times d}$ for some $\vec{t} := (t_1, \cdots, t_m)$ with $0 \leq t_1 < \cdots < t_m \leq 1$, then all the arguments still hold true and we can obtain similar duality results as in Theorems III.3.12, III.3.14 and III.3.17.*

4 Proof of $\mathsf{P} = \mathsf{D}_1 = \mathsf{D}_2$

In the following, we focus on the finite-marginal case, *i.e.* $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$ and start by proving the first duality. To prove the equality $\mathsf{P} = \mathsf{D}_1$, we shall apply the following well-known result from convex analysis.

Theorem III.4.1 (Fenchel-Moreau)**.** *Let (E, Σ) be a Hausdorff locally convex space and $F : E \to \mathbb{R}$ be a concave and upper semicontinuous function. Then F is equal to its biconjugate F^{**} which is defined by*

$$F^{**}(e) \quad := \quad \inf_{e^*\in E^*} \left\{ \langle e, e^*\rangle \ + \ \sup_{e'\in E} \Big(F(e') \ - \ \langle e', e^*\rangle \Big) \right\},$$

where E^* denotes the dual space of E.

Next we show that the map $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ is $\mathcal{W}_1^{\mathbb{T}}-$upper semicontinuous and concave and then identify its dual space to be $\Lambda^{\mathbb{T}}$ by $\langle \boldsymbol{\mu}, \boldsymbol{\lambda}\rangle = \boldsymbol{\mu}(\boldsymbol{\lambda})$.

4.1 Space of signed Borel measures on $\mathbb{R}^d$ and its dual space

Let $\mathbf{M}$ denote the space of all finite signed Borel measures μ on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d} \Big(1 + |x|\Big) |\mu|(dx) \quad < \quad +\infty.$$

It is clear that $\mathbf{M}$ is a linear vector space. We endow $\mathbf{M}$ with a topology (of Wasserstein kind) induced by the following convergence : Let $(\mu^n)_{n\geq 0} \subseteq \mathbf{M}$ be a sequence of bounded signed measures, we say μ^n converges to μ^0 if

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \lambda(x)\mu^n(dx) \;=\; \int_{\mathbb{R}^d} \lambda(x)\mu^0(dx) \text{ for all } \lambda \in \Lambda.$$

Notice that the above topology restricted on the subspace $\mathbf{P} \subseteq \mathbf{M}$ of probability measures is exactly that induced by the Wasserstein distance. As for the space $\mathbf{M}_0$ of all finite signed Borel measures on $\mathbb{R}^d$ equipped with the weak convergence topology, it is well known that its dual space $\mathbf{M}_0^*$ can be identified as the space of all bounded continuous functions Λ_0, see *e.g.* Deuschel & Stroock [35]. The following lemma identifies the dual space of $\mathbf{M}$.

Lemma III.4.2. *The space* $\mathbf{M}$ *is a Hausdorff locally convex space, and the duality relation*

$$(\lambda, \mu) \;\in\; \Lambda \times \mathbf{M} \;\;\longmapsto\;\; \mu(\lambda)$$

determines a representation of $\mathbf{M}^*$ *as* Λ.

The proof is almost the same as that of $\mathbf{M}_0^* = \Lambda_0$. For completeness, we provide a short proof in Appendix. For the finite set $\mathbb{T}$, let us endow $\mathbf{M}^{\mathbb{T}}$ with the product topology and obviously, the dual space of $\mathbf{M}^{\mathbb{T}}$ is given by $\Lambda^{\mathbb{T}}$.

4.2 Proof of $\mathsf{P} = \mathsf{D}_1$

In preparation for the first duality, we show first the upper semicontinuity of $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ in the context of Theorem III.3.12 (i). For $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_m) \in \mathbb{R}_+^m$ such that $|\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$, we introduce

$$\mathcal{M}^{\varepsilon}(\boldsymbol{\mu}) \;\;:=\;\; \Big\{\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu}) :\; X_t \;=\; X_{t_k} \text{ on } [t_k, t_k + \varepsilon_k) \text{ for all } k = 0, \cdots, m-1,\; \mathbb{P}\text{ - a.s.}\Big\}$$

and

$$\mathsf{P}^{\varepsilon}(\boldsymbol{\mu}) \;\;:=\;\; \sup_{\mathbb{P}\in\mathcal{M}^{\varepsilon}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big].$$

Proposition III.4.3. *Let* ξ *be bounded from above and satisfying Assumptions III.3.10 and III.3.11 (1), then* $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ *is* $\mathcal{W}_1^{\mathbb{T}}$*−upper semicontinuous on* $\mathbf{P}^{\preceq}$.

Proof. (i) First notice $\mathsf{P}^{\varepsilon}(\boldsymbol{\mu}) \le \mathsf{P}(\boldsymbol{\mu})$ since $\mathcal{M}^{\varepsilon}(\boldsymbol{\mu}) \subseteq \mathcal{M}(\boldsymbol{\mu})$. Next, for each $\mathbb{P} \in \mathcal{M}(\boldsymbol{\mu})$, define $\mathbb{P}^{\varepsilon} := \mathbb{P} \circ X_{f_\varepsilon}^{-1}$, where f_ε is defined in (III.3.4). It is clear that $\mathbb{P}^{\varepsilon} \in \mathcal{M}^{\varepsilon}(\boldsymbol{\mu})$ and $\mathbb{E}^{\mathbb{P}^{\varepsilon}}[\xi(X)] =$

$\mathbb{E}^{\mathbb{P}}[\xi(X_{f_\varepsilon})]$. It follows by Assumption III.3.11 (1),

$$\begin{aligned}\mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big] &\le \mathbb{E}^{\mathbb{P}}\Big[\xi(X_{f_\varepsilon})\Big] + \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + (m+2)\mathbb{E}^{\mathbb{P}}\Big[|X_1|\Big]\Big)\\ &= \mathbb{E}^{\mathbb{P}^\varepsilon}\Big[\xi(X)\Big] + \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + (m+2)\mu_1(|x|)\Big)\\ &\le \mathsf{P}^\varepsilon(\boldsymbol{\mu}) + \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + (m+2)\mu_1(|x|)\Big),\end{aligned}$$

which implies that

$$\mathsf{P}(\boldsymbol{\mu}) = \inf_{0<|\boldsymbol{\varepsilon}|<\Delta\mathbb{T}} \mathsf{P}^\varepsilon(\boldsymbol{\mu}) + \alpha(|\boldsymbol{\varepsilon}|)\Big(1 + (m+2)\mu_1(|x|)\Big).$$

(ii) In order to prove the upper semicontinuity of $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$, it suffices to verify that $\boldsymbol{\mu} \mapsto \mathsf{P}^\varepsilon(\boldsymbol{\mu})$ is upper semicontinuous. To see this, let $(\boldsymbol{\mu}^n)_{n\ge1} \subseteq \mathbf{P}^{\preceq}$ be a sequence such that $\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^{\mathbb{T}}} \boldsymbol{\mu} \in \mathbf{P}^{\preceq}$. By definition, we have a sequence $(\mathbb{P}_n)_{n\ge1}$ such that $\mathbb{P}_n \in \mathcal{M}^\varepsilon(\boldsymbol{\mu}^n)$ and

$$\limsup_{n\to\infty} \mathsf{P}^\varepsilon(\boldsymbol{\mu}^n) = \limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}[\xi].$$

Then one may find a convergent subsequence $(\mathbb{P}_{n_k})_{k\ge1}$ with limit $\mathbb{P} \in \mathcal{M}$. It follows by exactly the same arguments as in Lemma III.3.8 (ii) that $\mathbb{P} \in \mathcal{M}^\varepsilon(\boldsymbol{\mu})$. Since ξ is bounded from above, then it follows from Fatou's lemma that

$$\limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}[\xi] = \lim_{k\to\infty} \mathbb{E}^{\mathbb{P}_{n_k}}[\xi] \le \mathbb{E}^{\mathbb{P}}[\xi] \le \mathsf{P}^\varepsilon(\boldsymbol{\mu}),$$

which concludes the proof. $\square$

Now we are ready to provide the first duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu})$. To apply the Fenchel-Moreau theorem, we need to embed $\mathbf{P}^{\preceq}$ to a locally convex space. Recall that $\mathbf{M}$ is the space of all finite signed measures μ such that

$$\int_{\mathbb{R}^d} \Big(1 + |x|\Big)|\mu|(dx) < +\infty,$$

and $\mathbf{M}^{\mathbb{T}}$ is its $\mathbb{T}-$product. We then extend the map P from $\mathbf{P}^{\preceq}$ to $\mathbf{M}^{\mathbb{T}}$ by

$$\widetilde{\mathsf{P}}(\boldsymbol{\mu}) := \begin{cases}\mathsf{P}(\boldsymbol{\mu}), & \text{if } \boldsymbol{\mu} \in \mathbf{P}^{\mathbb{T}}_{\preceq},\\ -\infty, & \text{otherwise.}\end{cases}$$

Proof of Theorem III.3.12 (i). The concavity of the map $\boldsymbol{\mu} \mapsto \mathsf{P}(\boldsymbol{\mu})$ is immediate from its definition. Together with the upper semicontinuity of Proposition III.4.3, we may directly verify that the extended map $\widetilde{\mathsf{P}}$ is also $\mathcal{W}_1^{\mathbb{T}}-$upper semicontinuous and concave. Then, combining the Fenchel-Moreau theorem and Lemma III.4.2, it follows that for all $\boldsymbol{\mu} \in \mathbf{M}^{\mathbb{T}}$,

$$\widetilde{\mathsf{P}}(\boldsymbol{\mu}) \;=\; \widetilde{\mathsf{P}}^{**}(\boldsymbol{\mu}),$$

where $\widetilde{\mathsf{P}}^{**}$ denotes the biconjugae of $\widetilde{\mathsf{P}}$. In particular, for $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ one has

$$\begin{aligned}
&\mathsf{P}(\boldsymbol{\mu}) \;=\; \widetilde{\mathsf{P}}(\boldsymbol{\mu}) \;=\; \widetilde{\mathsf{P}}^{**}(\boldsymbol{\mu})\\
=\;& \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;-\; \widetilde{\mathsf{P}}^{*}(\boldsymbol{\lambda})\Big\} \;=\; \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;-\; \inf_{\boldsymbol{\nu}\in\mathbf{M}^{\mathbb{T}}} \Big\{\boldsymbol{\nu}(\boldsymbol{\lambda}) \;-\; \widetilde{\mathsf{P}}(\boldsymbol{\nu})\Big\}\Big\}\\
\geq\;& \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\boldsymbol{\nu}\in\mathbf{P}^{\preceq}} \Big\{\sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\nu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) \;-\; \boldsymbol{\lambda}(X)\Big]\Big\}\Big\}\\
=\;& \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) \;-\; \boldsymbol{\lambda}(X)\Big]\Big\} \;=\; \mathsf{D}_1(\boldsymbol{\mu}) \;\geq\; \mathsf{P}(\boldsymbol{\mu}),
\end{aligned}$$

which yields $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_1(\boldsymbol{\mu})$. □

4.3 Proof of $\mathsf{D}_1 = \mathsf{D}_2$

For technical reasons, we need to restrict the static strategy $\boldsymbol{\lambda}$ to a smaller class of functions $\Lambda_{lip}^{\mathbb{T}} \subseteq \Lambda^{\mathbb{T}}$ defined by

$$\Lambda_{lip}^{\mathbb{T}} \;:=\; \Big\{\boldsymbol{\lambda} = (\lambda_{t_i})_{1\leq i\leq m} : \;\lambda_{t_i} \text{ is boundedly supported and Lipschitz for all } i = 1, \cdots, m\Big\}.$$

Proposition III.4.4. *Under the conditions of Theorem III.3.12* (ii) *one has*

$$\mathsf{D}_1(\boldsymbol{\mu}) \;=\; \inf_{\boldsymbol{\lambda}\in\Lambda_{lip}^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) \;-\; \boldsymbol{\lambda}(X)\Big]\Big\}. \tag{III.4.1}$$

Proof. Clearly, by the definition of D_1 and the fact that $\mu_0 = \delta_{\mathbf{1}}(dx)$ and $\operatorname{supp}(\mu_1) \subseteq \mathbb{R}_+^d$, one obtains

$$\begin{aligned}
\mathsf{D}_1(\boldsymbol{\mu}) \;&\geq\; \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \sup_{\mathbb{P}\in\mathcal{M}_+} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}(X)\Big]\Big\}\\
&\geq\; \sup_{\mathbb{P}\in\mathcal{M}_+} \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}(X)\Big]\Big\}\\
&=\; \mathsf{P}(\boldsymbol{\mu}) \;=\; \mathsf{D}_1(\boldsymbol{\mu}),
\end{aligned}$$

by Theorem III.3.12 (i). Hence

$$\mathsf{D}_1(\boldsymbol{\mu}) \;=\; \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}} \left\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}(X)\Big]\right\}.$$

Next for every $\boldsymbol{\lambda} = (\lambda_{t_i})_{1\le i\le m} \in \Lambda^{\mathbb{T}}$, there exists some constant $L > 0$ such that for all $i = 1, \cdots, m$,

$$\lambda^L_{t_i}(x) \;:=\; \lambda_{t_i}(x) \;-\; L(1+\mathbf{1}\cdot x) \;\le\; 0 \text{ for all } x\in\mathbb{R}^d_+.$$

Denote $\boldsymbol{\lambda}^L := (\lambda^L_{t_i})_{1\le i\le m}$, then for every martingale measure $\mathbb{P}\in\mathcal{M}_+$, we have

$$\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}(X)\Big] \;=\; \boldsymbol{\mu}(\boldsymbol{\lambda}^L) \;+\; \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^L(X)\Big].$$

Further, for each $R > 0$, let $\psi_R : \mathbb{R}^d \to [0,1]$ be some continuous function such that

$$\psi_R(x) \;=\; 1 \text{ whenever } |x| \;\le\; R \text{ and } \psi_R(x) \;=\; 0 \text{ whenever } |x| \;>\; R+1.$$

Let $\boldsymbol{\lambda}^{L,R} := \big(\lambda^{L,R}_{t_i}\big)_{1\le i\le m}$ with $\lambda^{L,R}_{t_i}(x) := \lambda^L_{t_i}(x)\psi_R(x) \ge \lambda^L_{t_i}(x)$, then

$$\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L,R}(X)\Big] \;\le\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L}(X)\Big].$$

On the other hand, for all $\mathbb{P}\in\mathcal{M}_+$ we have by the monotone convergence theorem

$$\lim_{R\to\infty} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L,R}(X)\Big] \;=\; \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L}(X)\Big].$$

Hence

$$\lim_{R\to\infty}\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L,R}(X)\Big] \;=\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L}(X)\Big].$$

It follows that

$$\begin{aligned}\lim_{R\to\infty}\Big(\boldsymbol{\mu}(\boldsymbol{\lambda}^{L,R}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L,R}(X)\Big]\Big) \;&=\; \boldsymbol{\mu}(\boldsymbol{\lambda}^L) \;+\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}^{L}(X)\Big] \\ &=\; \boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi \;-\; \boldsymbol{\lambda}(X)\Big].\end{aligned}$$

Finally, by a convolution argument each $\lambda^{L,R}_{t_i}$ can be approximated uniformly by some Lipschitz function that is also boundedly supported, which yields the required result. □

For all $(\omega, t) \in \mathbf{D}_+ \times [0,1]$, denote by $\mathcal{B}^{sem}_{\omega,t} \subseteq \mathcal{P}$ the set of probability measures $\mathbb{P}$ such that $\mathbb{P}[X_s = \omega_s \text{ for all } 0 \le s \le t] = 1$ and $(X_s)_{s\ge t}$ is a non-negative semimartingale under $\mathbb{P}$. Denote further

$$\mathcal{M}^{loc}_{\omega,t} \;:=\; \Big\{\mathbb{P} \in \mathcal{B}^{sem}_{\omega,t} :\; (X_s)_{s\ge t} \text{ is a local martingale under } \mathbb{P}\Big\}.$$

Write in particular $\mathcal{B}^{sem} = \mathcal{B}^{sem}_{\omega,0}$ and $\mathcal{M}^{loc} = \mathcal{M}^{loc}_{\omega,0}$. Let $\zeta : \mathbf{D} \to \mathbb{R}$ be a measurable function and put

$$V_t(\omega) \;:=\; \sup_{\mathbb{P}\in\mathcal{M}^{loc}_{\omega,t}} \mathbb{E}^{\mathbb{P}}\Big[\zeta(X)\Big]. \tag{III.4.2}$$

Our objective now is to show that the process $(V_t)_{0\le t\le 1}$ is $\mathbb{F}^U-$adapted and that the dynamic programming principle holds. To achieve this, we use the related results in Neufeld & Nutz [81, 82]. Let $\mathbb{P} \in \mathcal{B}^{sem}$ be a semimartingale measure with the triplet $(B^{\mathbb{P}}, C^{\mathbb{P}}, \nu^{\mathbb{P}})$ of predictable semimartingale characteristics, see *e.g.* Chapter II of Jacod & Shiryaev [66]. Notice that

$$\mathcal{M}^{loc} \;=\; \Big\{\mathbb{P} \in \mathcal{B}^{sem} :\; B^{\mathbb{P}}_t = 0 \text{ for all } t \in [0,1]\Big\}.$$

By Theorem 2.5 in [81], the map $\mathbb{P} \mapsto (B^{\mathbb{P}}, C^{\mathbb{P}}, \nu^{\mathbb{P}})$ is measurable, then it follows that $\mathcal{M}^{loc}$ is Borel. Moreover, by the same arguments we have the following lemma.

Lemma III.4.5. *The set* $\{(\omega, t, \mathbb{P}) \in \mathbf{D} \times [0,1] \times \mathcal{P} :\; \mathbb{P} \in \mathcal{M}^{loc}_{\omega,t}\}$ *is Borel.*

By Theorem 2.1 in [82], we have the following lemma.

Lemma III.4.6. *Let* $\mathbb{P} \in \mathcal{M}^{loc}_{\omega,t}$ *and* τ *be an* $\mathbb{F}-$*stopping time taking values in* $[t,1]$.
(i) *There is a family of conditional probability* $(\mathbb{P}_\omega)_{\omega\in\mathbf{D}}$ *of* $\mathbb{P}$ *with respect to* $\mathcal{F}_\tau$ *such that* $\mathbb{P}_\omega \in \mathcal{M}^{loc}_{\omega,\tau(\omega)}$ *for* $\mathbb{P}$ *- almost every* $\omega \in \mathbf{D}$.
(ii) *Assume that there exists a family of probability measures* $(\mathbb{Q}_\omega)_{\omega\in\mathbf{D}_+}$ *such that*

$$\mathbb{Q}_\omega \in \mathcal{M}^{loc}_{\omega,\tau(\omega)} \text{ for } \mathbb{P} \text{ - a.e. } \omega \in \mathbf{D}, \text{ and the map } \omega \mapsto \mathbb{Q}_\omega \text{ is } \mathcal{F}_\tau - \text{measurable},$$

then $\mathbb{P} \otimes \mathbb{Q} \in \mathcal{M}^{loc}_{\omega,t}$, *where*

$$\mathbb{P} \otimes \mathbb{Q}[\cdot] \;:=\; \int_\Omega \mathbb{Q}_\omega[\cdot]\mathbb{P}(d\omega).$$

The dynamic programming principle follows by Lemmas III.4.5 and III.4.6, and as a consequence we have the following proposition.

Proposition III.4.7. *Assume that ζ is bounded, then the process $V = (V_t)_{0\le t\le 1}$ defined in* (III.4.2) *is an $\mathcal{M}_+-$supermartingale, i.e. $V \in \mathcal{S}$.*

Proposition III.4.8. *Let ζ be a measurable and bounded function, then one has*

$$\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\zeta(X)\Big] \;=\; \inf\Big\{V_0:\; (V_t)_{0\le t\le 1}\in\mathcal{S} \text{ such that } V_1(\omega) \;\ge\; \zeta(\omega) \text{ for all } \omega\in\Omega_+\Big\}.$$

Proof. By Proposition III.4.7 with the process V defined in (III.4.2), it remains to show that

$$\sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\zeta(X)\Big] \;=\; \sup_{\mathbb{P}\in\mathcal{M}^{loc}} \mathbb{E}^{\mathbb{P}}\Big[\zeta(X)\Big].$$

It is clear that $\sup_{\mathbb{P}\in\mathcal{M}_+}\mathbb{E}^{\mathbb{P}}[\zeta(X)] \le \sup_{\mathbb{P}\in\mathcal{M}^{loc}}\mathbb{E}^{\mathbb{P}}[\zeta(X)]$ since $\mathcal{M}_+\subseteq\mathcal{M}^{loc}$, then it suffices to prove the converse inequality. For each $\mathbb{P}\in\mathcal{M}^{loc}$, there exists an increasing sequence of stopping times $(\sigma_n)_{n\ge 1}$ such that $\sigma_n\to+\infty$, $\mathbb{P}$ - almost surely and $X_{\sigma_n\wedge\cdot}$ is a $\mathbb{P}-$martingale, where $X_{\sigma_n\wedge\cdot} := (X_{\sigma_n\wedge t})_{0\le t\le 1}$. Hence

$$\mathbb{E}^{\mathbb{P}}\Big[\zeta(X_{\sigma_n\wedge\cdot})\Big] \;\le\; \sup_{\mathbb{Q}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{Q}}\Big[\zeta(X)\Big].$$

The required result follows from the dominated convergence theorem. □

Proof of Theorem III.3.12 (ii). It remains to show $\mathsf{D}_1(\boldsymbol{\mu}) \ge \mathsf{D}_2(\boldsymbol{\mu})$. Indeed, one has by Proposition III.4.4,

$$\mathsf{D}_1(\boldsymbol{\mu}) \;=\; \inf_{\boldsymbol{\lambda}\in\Lambda^{\mathbb{T}}_{lip}} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\xi(X) \;-\; \boldsymbol{\lambda}(X)\Big]\Big\}.$$

For each $\varepsilon > 0$, by Proposition III.4.8 there exist a vector $\boldsymbol{\lambda}^{\varepsilon}\in\Lambda^{\mathbb{T}}_{lip}$ and a process $V^{\varepsilon} = (V^{\varepsilon}_t)_{0\le t\le 1}\in\mathcal{S}$ such that

$$\mathsf{D}_1(\boldsymbol{\mu}) \;+\; \varepsilon \;\ge\; \boldsymbol{\mu}(\boldsymbol{\lambda}^{\varepsilon}) \;+\; V^{\varepsilon}_0 \quad\text{and}\quad V^{\varepsilon}_1(\omega) \;\ge\; \xi(\omega) \;-\; \boldsymbol{\lambda}^{\varepsilon}(\omega).$$

This implies that $\mathsf{D}_1(\boldsymbol{\mu})+\varepsilon \ge \mathsf{D}_2(\boldsymbol{\mu})$, and the required result by the arbitrariness of ε. □

5 Proof of $\mathsf{D}_1 = \mathsf{D}_3$

Now let us turn to prove the third duality $\mathsf{D}_1 = \mathsf{D}_3$ in Theorem III.3.14. We will follow the idea in Dolinsky & Soner [37] to discretize the underlying paths and then use the classical constrained duality result of Föllmer & Kramkov [42]. The proof in [37] relies on the min-max theorem and the explicit approximation of martingale measures. We emphasize that the present proof is less technically involved than [37] as the marginals constraints have already been removed by the first duality.

5.1 Reduction of ξ to be boundedly supported

In this section we denote $\mathsf{P}(\boldsymbol{\mu},\xi)$ and $\mathsf{D}_3(\boldsymbol{\mu},\xi)$ in place of $\mathsf{P}(\boldsymbol{\mu})$ and $\mathsf{D}_3(\boldsymbol{\mu})$ to emphasize the dependence on ξ, then clearly for any ξ, $\xi' : \mathbf{D} \to \mathbb{R}$ and $c \in \mathbb{R}$, one has

$$\mathsf{D}_3(\boldsymbol{\mu},\xi+\xi') \;\le\; \mathsf{D}_3(\boldsymbol{\mu},\xi) \;+\; \mathsf{D}_3(\boldsymbol{\mu},\xi') \quad \text{and} \quad \mathsf{D}_3(\boldsymbol{\mu},\xi+c) \;=\; \mathsf{D}_3(\boldsymbol{\mu},\xi) \;+\; c.$$

In particular for $c \in \mathbb{R}_+$ one has

$$\mathsf{D}_3(\boldsymbol{\mu},c\xi) \;=\; c\mathsf{D}_3(\boldsymbol{\mu},\xi).$$

Hence, under the conditions of Theorem III.3.14, we may assume that $0 \le \xi \le 1$. Indeed, we show next that it suffices to establish the duality $\mathsf{P}(\boldsymbol{\mu},\xi) = \mathsf{D}_3(\boldsymbol{\mu},\xi)$ for ξ that is boundedly supported. For all $R > 0$, define the continuous function $\chi_R : \mathbb{R}_+ \to [0,1]$ by

$$\chi_R(x) \;:=\; \mathbb{1}_{[0,R]}(x) \;+\; (R+1-x)\mathbb{1}_{(R,R+1]}(x) \text{ for all } x \in \mathbb{R}_+.$$

Denote further for $R > 0$

$$\xi_R(\omega) \;:=\; \xi(\omega)\chi_R(\|\omega\|) \text{ for all } \omega \in \mathbf{D}.$$

Notice that $0 \le \xi \le 1$ yields $\xi_R(\omega) \le \xi(\omega) \le \xi_R(\omega) + \mathbb{1}_{\{\|\omega\| \ge R\}}$, then it follows that

$$\mathsf{D}_3(\boldsymbol{\mu},\xi_R) \;\le\; \mathsf{D}_3(\boldsymbol{\mu},\xi) \;\le\; \mathsf{D}_3(\boldsymbol{\mu},\xi_R) \;+\; \mathsf{D}_3(\boldsymbol{\mu},\mathbb{1}_{\{\|X\| \ge R\}}). \tag{III.5.1}$$

Lemma III.5.1. *Let ξ be bounded and $\boldsymbol{\mu} \in \mathbf{P}^{\preceq}$ satisfying (III.2.5). Then*

$$\mathsf{D}_3(\boldsymbol{\mu},\xi) \;=\; \lim_{R\to\infty} \mathsf{D}_3(\boldsymbol{\mu},\xi_R)$$

Proof. It is enough to prove by (III.5.1) that

$$\lim_{R\to+\infty} \mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X\|\geq R\}}) \;=\; 0.$$

This is indeed a direct consequence of the pathwise inequality, see *e.g.* Lemma 2.3 of Brown, Hobson & Rogers [17]

$$\mathbb{1}_{\{\|X_i\|\geq R\}} \;\leq\; \frac{(|X_{i,1}|-K)^+}{R-K} \;+\; \mathbb{1}_{\{\|X_i\|\geq R\}}\frac{R-X_{i,1}}{R-K} \text{ for all } i=1,\cdots,d$$

holds for every $0<K<R$. It follows by taking $K=R/2d$ that

$$\mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X\|\geq R\}}) \;\leq\; \sum_{i=1}^d \mathsf{D}_3(\boldsymbol{\mu}, \mathbb{1}_{\{\|X_i\|\geq R/d\}}) \;\leq\; \frac{2d}{R}\sum_{i=1}^d \mu_1\Big((x_i-\frac{R}{2d})^+\Big).$$

The proof is fulfilled by letting $R\to+\infty$. □

Next we show that ξ_R inherits almost the same properties as ξ.

Lemma III.5.2. *For each $R>0$:*

(i) *If ξ satisfies Assumptions III.3.10 and III.3.11, then so does ξ_R.*

(ii) *If ξ satifies Assumption III.3.13, then*

$$\begin{aligned} &\xi_R \;\; \textit{is } L^\infty\textit{-uniformly continuous, and} \\ \xi_R(\omega) \;-\; \xi_R(\omega') \;&\leq\; \beta\Big(\rho_{\mathbb{T}}(\omega,\omega')\Big) \textit{ for all } \omega,\omega'\in\Omega \textit{ such that } \|\omega'\|\leq\|\omega\| \end{aligned} \tag{III.5.2}$$

for some continuous increasing function $\beta:\mathbb{R}_+\to\mathbb{R}_+$ with $\beta(0)=0$.

Proof. (i) follows by the fact that $\omega\mapsto\|\omega\|$ is S^*-lower semicontinuous and $\|\omega_{f_\varepsilon}\|=\|\omega_{b_\varepsilon}\|=\|\omega\|$. Let us turn to show (ii) . Notice that ξ is $\rho_{\mathbb{T}}-$uniformly continuous on $\{\omega:\|\omega\|\leq R\}$, *i.e.* there exists a continuous increasing function $\beta:\mathbb{R}_+\to\mathbb{R}_+$ with $\beta(0)=0$ such that for all $\|\omega\|,\|\omega'\|\leq R$

$$|\xi(\omega) \;-\; \xi(\omega')| \;\leq\; \beta\Big(\rho_{\mathbb{T}}(\omega,\omega')\Big).$$

Hence, for any $\omega,\omega'\in\Omega$ such that $\|\omega'\|\leq\|\omega\|$, one has

$$\xi_R(\omega) \;-\; \xi_R(\omega') \;\leq\; \mathbb{1}_{\{\|\omega\|\leq R\}}\Big(\xi(\omega) \;-\; \xi(\omega')\Big) \;\leq\; \beta\Big(\rho_{\mathbb{T}}(\omega,\omega')\Big).$$

Moreover,

$$\begin{aligned} |\xi_R(\omega') - \xi_R(\omega)| &\leq |\xi(\omega') - \xi(\omega)|\chi_R(\|\omega\|) + |\chi_R(\|\omega\|) - \chi_R(\|\omega'\|)| \\ &\leq \beta\Big(\rho_{\mathbb{T}}(\omega, \omega')\Big) + \|\omega - \omega'\| \\ &\leq \beta\Big(2\|\omega - \omega'\|\Big) + \|\omega - \omega'\|, \end{aligned}$$

which yields the $L^\infty-$uniform continuity of ξ_R. □

Therefore, in the following it suffices to consider the function ξ that is boundedly supported such that the Assumptions III.3.10, III.3.11 and Condition (III.5.2) hold. Similar to the proof of the duality $\mathsf{P}(\boldsymbol{\mu}) = \mathsf{D}_2(\boldsymbol{\mu})$, it remains to prove a duality without marginal constraints.

5.2 Duality without marginal constraints

We consider in this section the optimization problem without marginal constraints. Let $\zeta : \Omega \to \mathbb{R}$ be bounded and define

$$\mathsf{P}(\zeta) := \sup_{\mathbb{P}\in\mathcal{M}_+} \mathbb{E}^{\mathbb{P}}\Big[\zeta(X)\Big] \quad \text{and} \quad \mathsf{D}(\zeta) := \inf_{(z,H)\in\mathcal{D}(\zeta)} z. \tag{III.5.3}$$

where, with the same definition of integral in (III.2.7),

$$\mathcal{D}(\zeta) := \Big\{(z, H) \in \mathbb{R}\times\mathcal{H} : z + (H\cdot X)_1(\omega) \geq \zeta(\omega) \text{ for all } \omega \in \mathbf{D}_+\Big\}.$$

We provide immediately a duality result for the above optimization problems, and leave its proof in Section 5.3.

Theorem III.5.3. *For the given $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$, suppose further that ζ satisfies the Assumptions III.3.10, III.3.11 and Condition* (III.5.2)*, then*

$$\mathsf{P}(\zeta) = \mathsf{D}(\zeta). \tag{III.5.4}$$

By exactly the same arguments as in the proof of Theorem III.3.12 (ii), the duality $\mathsf{P}(\boldsymbol{\mu}, \xi_R) = \mathsf{D}_3(\boldsymbol{\mu}, \xi_R)$ follows immediately by taking $\zeta = \xi_R - \boldsymbol{\lambda}$ in Theorem III.5.3.

Proof of Theorem III.3.14. Using Lemma III.5.1 as well as the first duality $\mathsf{P} = \mathsf{D}_1$ for ξ_R, one has

$$\mathsf{P}(\boldsymbol{\mu}, \xi) \geq \lim_{R\to\infty} \mathsf{P}(\boldsymbol{\mu}, \xi_R) = \lim_{R\to\infty} \mathsf{D}_3(\boldsymbol{\mu}, \xi_R) = \mathsf{D}_3(\boldsymbol{\mu}, \xi).$$

Hence we conclude the proof by the weak duality $\mathsf{P}(\boldsymbol{\mu},\xi) \le \mathsf{D}_3(\boldsymbol{\mu},\xi)$. $\square$

5.3 Proof of Theorem III.5.3

Recall that $\mathbb{T} = \{0 = t_0 < \cdots < t_m = 1\}$, $\Delta t_i = t_i - t_{i-1}$ for $i = 1, \cdots, m$ and $\Delta\mathbb{T} = \min_{1\le i\le m} \Delta t_i$. Let $\zeta : \mathbf{D} \to \mathbb{R}$ be measurable and boundedly supported. Then for each $0 \le \delta < \Delta\mathbb{T}$, denote $\mathbb{D}^\delta := \mathbb{D}([0, 1+\delta], \mathbb{R}^d)$ and all its elements by ω^δ. Put $\mathbb{T}^\delta := \{0 = t_0^\delta < \cdots < t_m^\delta = 1+\delta\}$, where $t_i^\delta \;:=\; k_\delta t_i$ for all $i = 0, \cdots, m$ with $k_\delta \;:=\; 1+\delta$. Define $\zeta^\delta : \mathbf{D}^\delta \to \mathbb{R}$ by

$$\zeta^\delta(\omega^\delta) \;:=\; \zeta(\bar\omega^\delta), \text{ where } \bar\omega^\delta \in \mathbf{D} \text{ is defined by } \bar\omega_t^\delta \;:=\; \omega_{k_\delta t}^\delta \text{ for all } t \in [0,1]. \tag{III.5.5}$$

Proposition III.5.4. *Assume that ζ satisfies Assumptions III.3.10, III.3.11 and Condition (III.5.2). Then :*

(i) *For all $0 \le \delta < \Delta\mathbb{T}$, the ζ^δ defined by (III.5.5) satisfies Assumptions III.3.10, III.3.11 and Condition (III.5.2).*

(ii) *There is a continuous function $\eta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\eta(0) = 0$ such that for all $0 \le \delta < \delta' < \Delta\mathbb{T}$ the following inequality holds*

$$\left|\zeta^\delta(\omega^\delta) \;-\; \zeta^{\delta'}(\omega^{\delta',\delta})\right| \;\le\; \eta\left(\frac{\delta'-\delta}{1+\delta'}\right)\left(1 \;+\; \sum_{i=0}^m |\omega_{t_i^\delta}^\delta| \;+\; \int_0^{1+\delta} |\omega_t^\delta| dt\right) \text{ for all } \omega^\delta \in \Omega^\delta,$$

where $\omega^{\delta',\delta} \in \mathbf{D}^{\delta'}$ is defined by

$$\omega_t^{\delta',\delta} \;:=\; \omega^\delta_{(t - t_i^{\delta'} + t_i^\delta)\wedge t_{i+1}^\delta} \text{ for all } t \in [t_i^{\delta'}, t_{i+1}^{\delta'}] \text{ and } i = 0, \cdots, m-1.$$

Proof. (i) will be proved in Lemmas III.5.10, III.5.11 and III.5.12 in Section 5.5.

(ii) Clearly, $\zeta^{\delta'}(\omega^{\delta',\delta}) \;=\; \zeta(\bar\omega^{\delta',\delta})$, where

$$\bar\omega_t^{\delta',\delta} \;:=\; \omega_{k_{\delta'}t}^{\delta',\delta} \text{ for all } t \in [0,1].$$

Direct computation reveals that $\bar\omega^{\delta,'\delta} = \bar\omega^\delta \circ b_\varepsilon$ with

$$\boldsymbol{\varepsilon} \;:=\; \frac{\delta'-\delta}{1+\delta'}\big(\Delta t_1, \cdots, \Delta t_m\big).$$

Hence by Assumption III.3.11 one obtains

$$\begin{aligned}
\left|\zeta^{\delta}(\omega^{\delta}) \;-\; \zeta^{\delta'}(\omega^{\delta',\delta})\right| \;&=\; \left|\zeta(\bar{\omega}^{\delta}) \;-\; \zeta(\bar{\omega}^{\delta',\delta})\right| \\
&\leq\; \alpha(|\boldsymbol{\varepsilon}|)\left(1 \;+\; \sum_{i=0}^{m}|\bar{\omega}^{\delta}_{t_i}| \;+\; \int_0^1 |\bar{\omega}^{\delta}_t|dt\right) \\
&\leq\; \alpha\left(\Delta\mathbb{T}\frac{\delta'-\delta}{1+\delta'}\right)\left(1 \;+\; \sum_{i=0}^{m}|\omega^{\delta}_{t_i^{\delta}}| \;+\; \int_0^{1+\delta} |\omega^{\delta}_t|dt\right).
\end{aligned}$$

The proof is completed by taking $\eta(\cdot) = \alpha(\Delta\mathbb{T}\times\cdot)$. □

We are now ready to prove the required duality. Define

$$\mathbf{D}^{\delta}_{+} \;:=\; \left\{\omega^{\delta}\in\mathbb{D}^{\delta} :\; \omega^{\delta}_0 = \mathbf{1} \text{ and } \omega^{\delta}_t\in\mathbb{R}^d_+ \text{ for all } t\in[0,1+\delta]\right\}$$

and the corresponding martingale optimal transport problem

$$\mathsf{P}_{\delta} \;:=\; \sup_{\mathbb{P}\in\mathcal{M}^{\delta}_{+}} \mathbb{E}^{\mathbb{P}}\left[\zeta^{\delta}(X^{\delta})\right],$$

where similarly, $X^{\delta} = (X^{\delta}_t)_{0\leq t\leq k_{\delta}}$ denotes the canonical process and $\mathcal{M}^{\delta}_{+}$ denotes the set of martingale measures supported on $\mathbf{D}^{\delta}_{+}$. The dual problem is slightly different. Denote further

$$\mathbf{D}^{c,\delta}_{+} \;:=\; \left\{\omega^{\delta}\in\mathbf{D}^{\delta}_{+} :\; \omega^{\delta}_{t_i^{\delta}-} \;=\; \omega^{\delta}_{t_i^{\delta}} \text{ for all } i=1,\cdots,m\right\}$$

and define the dual problem by

$$\mathsf{D}^{c}_{\delta} \;:=\; \inf_{(z,H)\in\mathcal{D}^{c}_{\delta}} z,$$

with $\mathcal{D}^{c}_{\delta}$ given by

$$\mathcal{D}^{c}_{\delta} \;:=\; \left\{(z^{\delta},H^{\delta})\in\mathbb{R}\times\mathcal{H}^{\delta} :\; z^{\delta} \;+\; (H^{\delta}\cdot\omega^{\delta})_1 \;\geq\; \zeta^{\delta}(\omega^{\delta}) \text{ for all } \omega^{\delta}\in\mathbf{D}^{c,\delta}_{+}\right\},$$

where, similarly to Definition III.2.5, $\mathcal{H}^{\delta}$ denotes the collection of all left-continuous adapted processes with bounded variation such that the stochastic integral $(H^{\delta}\cdot X^{\delta})$ is a supermartingale under all probability measures in $\mathcal{M}^{\delta}_{+}$.

The main technical step for our result is the following.

Lemma III.5.5. *Suppose that ζ satisfies Assumptions III.3.10, III.3.11 and Condition* (III.5.2).

Then

$$\mathsf{D}^c_\delta \;\le\; \mathsf{P}_\delta \text{ for all } \delta \;\ge\; 0. \tag{III.5.6}$$

The proof of Lemma III.5.5 is adapted from Dolinsky & Soner [37] and is reported in Section 5.4.

Lemma III.5.6. *Suppose that ζ satisfies Assumptions III.3.10, III.3.11 and Condition* (III.5.2)*. Then*

$$\liminf_{\delta\searrow 0} \mathsf{D}^c_\delta \;\ge\; \mathsf{D}(\zeta) \quad \textit{and} \quad \limsup_{\delta\searrow 0} \mathsf{P}_\delta \;\le\; \mathsf{P}(\zeta). \tag{III.5.7}$$

Proof. (i) For each $(z^\delta, H^\delta) \in \mathcal{D}^c_\delta$ with $\delta > 0$ let us construct a robust super-replication on $\mathbf{D}_+$. For any $\omega \in \mathbf{D}_+$ define $H_0(\omega) = H^\delta_0(\omega^{\delta,0})$ and

$$H_t(\omega) \;:=\; H^\delta_{t-t_i+t_i^\delta}(\omega^{\delta,0}) \text{ for all } t \in (t_i, t_{i+1}] \text{ and } i = 0, \cdots, m-1,$$

where $\omega^{\delta,0} \in \mathbf{D}^{c,\delta}_+$ is defined as before by

$$\omega^{\delta,0}_t \;=\; \omega_{(t-t_i^\delta+t_i)\wedge t_{i+1}} \text{ for all } t \in [t^\delta_i, t^\delta_{i+1}] \text{ and } i = 0, \cdots, m-1.$$

It is clear that H is $\mathbb{F}-$adapted, left-continuous, with bounded variation, and $(H \cdot X)$ is a supermartingale under every $\mathbb{P} \in \mathcal{M}_+$, hence $H \in \mathcal{H}$. Moreover,

$$z^\delta \;+\; \left(H^\delta \cdot \omega^{\delta,0}\right)_{1+\delta} \;\ge\; \zeta^\delta(\omega^{\delta,0}) \text{ for all } \omega \in \mathbf{D}_+. \tag{III.5.8}$$

Notice that $(H^\delta \cdot \omega^{\delta,0})_{1+\delta} = (H \cdot \omega)_1$, thus we obtain by Assumption III.3.11 and Condition (III.5.8)

$$z^\delta \;+\; (H\cdot\omega)_1 \;\ge\; \zeta(\omega) \;-\; \eta\left(\frac{\delta}{1+\delta}\right)\left(1 \;+\; \sum_{i=0}^m |\omega_{t_i}| \;+\; \int_0^1 |\omega_t| dt\right) \text{ for all } \omega \in \mathbf{D}_+,$$

which yields $\mathsf{D}^c_\delta + (1+(m+2)d)\eta(\frac{\delta}{1+\delta}) \;\ge\; \mathsf{D}(\zeta)$ and therefore

$$\liminf_{\delta\searrow 0} \mathsf{D}^c_\delta \;\ge\; \mathsf{D}(\zeta).$$

(ii) Let $(\delta_n)_{n\geq 1}$ be such that $\delta_n > 0$ and $\delta_n \searrow 0$. Then there is a sequence $(\mathbb{P}_n)_{n\geq 1}$ such that

$$\limsup_{n\to\infty} \mathsf{P}_{\delta_n} \;=\; \limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}\left[\zeta^{\delta_n}(X^{\delta_n})\right].$$

For any fixed $\delta_0 > 0$, we assume without loss of generality that $\delta_n \leq \delta_0$ for all $n \geq 1$. Then for each $n \geq 1$, let us define $\bar{\mathbb{P}}_n := \mathbb{P}_n \circ (\tilde{X}^{\delta_n})^{-1}$ where $\tilde{X}^{\delta_n}(\omega^{\delta_n}) := X^{\delta_0}(\omega^{\delta_0,\delta_n})$ is the extended process from $\mathbf{D}^{\delta_n}$ to $\mathbf{D}^{\delta_0}$. It follows by Proposition III.5.4 (ii) that

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}_n}\left[\zeta^{\delta_n}(X^{\delta_n})\right] \;&\leq\; (1 \;+\; (m+2)d)\,\eta\left(\frac{\delta_0-\delta_n}{1+\delta_0}\right) \;+\; \mathbb{E}^{\mathbb{P}_n}\left[\zeta^{\delta_0}(\tilde{X}^{\delta_n})\right]\\
&=\; \Big(1 \;+\; (m+2)d\Big)\eta\Big(\frac{\delta_0-\delta_n}{1+\delta_0}\Big) \;+\; \mathbb{E}^{\bar{\mathbb{P}}_n}\Big[\zeta^{\delta_0}(X^{\delta_0})\Big].
\end{aligned}$$

Again by the same argument in Proposition III.4.3 we obtain

$$\limsup_{n\to\infty} \mathsf{P}_{\delta_n} \;\leq\; 2\,(1 \;+\; (m+2)d)\,\eta\left(\frac{\delta_0}{1+\delta_0}\right) \;+\; \mathsf{P}(\zeta)$$

which yields the required result since $\delta_0 > 0$ is arbitrary. □

Proof of Theorem III.5.3. Let $(z,H) \in \mathcal{D}(\zeta)$, we know by definition $z + (H\cdot\omega)_1 \geq \zeta(\omega)$, $\forall \omega \in \mathbf{D}_+$. Taking expectation over each sides, it follows that

$$z \;\geq\; \mathbb{E}^{\mathbb{P}}\Big[\zeta(X)\Big] \text{ for all } \mathbb{P} \in \mathcal{M}_+.$$

Then we get the weak duality $\mathsf{P}(\zeta) \leq \mathsf{D}(\zeta)$. The reverse inequality follows by Lemmas III.5.5 and III.5.6. □

5.4 Proof of Lemma III.5.5

The arguments are mainly adapted from Dolinsky & Soner [37] and the main idea is to discretize the paths on the Skorokhod space. By Proposition III.5.4 (i), the proof of $\mathsf{D}^c_\delta \leq \mathsf{P}_\delta$ is not altered by the value of δ. We therefore consider $\delta = 0$ in this subsection.

5.4.1 A probabilistic superhedging problem

For all $n \in \mathbb{N}$, put

$$A^{(n)} \;:=\; \Big\{2^{-n}q:\; q\in\mathbb{N}^d\Big\} \;\text{ and }\; B^{(n)} \;:=\; \Big\{i\sqrt{d}2^{-n}:\; i\in\mathbb{N}\Big\}\cup\Big\{\sqrt{d}2^{-n}/j:\; j\in\mathbb{N}^*\Big\}.$$

We then define a subspace $\hat{\Omega} := \hat{\Omega}^{(n)} \subseteq \mathbf{D}_+$ as follows.

Definition III.5.7. *A path $\omega \in \mathbf{D}_+$ belongs to $\hat{\Omega}$ if there exist non-negative integers $0 = K_0 < K_1 + 1 < \cdots < K_m + m$ and a partition $\Big\{0 = \hat{\tau}_0 < \hat{\tau}_1 < \cdots < \hat{\tau}_{K_m+m} = 1\Big\}$ such that $\hat{\tau}_{K_i+i} = t_i$ for $1 \leq i \leq m$ and*

$$\omega_t \;\; = \;\; \sum_{i=0}^{m-1} \Bigg(\sum_{k=K_i+i}^{K_{i+1}+i-1} \omega_{\hat{\tau}_k} \mathbb{1}_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t) \; + \; \omega_{t_{i+1}} \mathbb{1}_{[\hat{\tau}_{K_{i+1}+i}, t_{i+1})}(t) \Bigg) \; + \; \omega_1 \mathbb{1}_{\{t=1\}},$$

where $\omega_{t_i} \in A^{(n)}$ for $1 \leq i \leq m$ and for $0 \leq i < m$

$$\begin{aligned} \omega_{\hat{\tau}_k} \;\; &\in \;\; A^{(n+k-K_i-i)}, \;\; K_i + i < k < K_{i+1} + i + 1, \\ \hat{\tau}_k - \hat{\tau}_{k-1} \;\; &\in \;\; B^{(n+k-K_i-i)}, \;\; K_i + i < k < K_{i+1} + i + 1. \end{aligned}$$

Notice that $\hat{\Omega}$ is countable, then there exists a probability measure $\hat{\mathbb{P}} := \hat{\mathbb{P}}^{(n)}$ on $\mathbf{D}_+$ supported on $\hat{\Omega}$ which gives positive weight to every element of $\hat{\Omega}$. In particular, the canonical process X has finitely many jumps $\hat{\mathbb{P}}$ - almost surely. Denote by $\hat{\mathbb{F}}$ the completed filtration of $\mathbb{F}$ under $\hat{\mathbb{P}}$. Put

$$\hat{\mathcal{H}}^{(n)} \;\; := \;\; \Big\{ \hat{H} : \;\; [0,1] \times \mathbf{D}_+ \to \mathbb{R}^d \text{ is } \hat{\mathbb{F}} - \text{predictable such that } \|\hat{H}\| \leq n \Big\}$$

and

$$\hat{\mathcal{A}}^{(n)} \;\; := \;\; \Big\{ \hat{H} \in \hat{\mathcal{H}}^{(n)} : \;\; (\hat{H} \cdot X)_t \;\; \geq \;\; K \text{ for all } t \in [0,1], \;\; \hat{\mathbb{P}} \text{ - a.s. for some } K \in \mathbb{R} \Big\}.$$

Let

$$\hat{\mathcal{D}}^{(n)}(\zeta) \;\; := \;\; \Big\{ (z, \hat{H}) \in \mathbb{R} \times \hat{\mathcal{A}}^{(n)} : \;\; z \; + \; (\hat{H} \cdot X)_1 \;\; \geq \;\; \zeta(X), \;\; \hat{\mathbb{P}} \text{ - a.s.} \Big\}$$

and define the robust superhedging problem under the dominating measure $\hat{\mathbb{P}}$

$$\mathsf{D}^{(n)}(\zeta) \;\; := \inf_{(z, \hat{H}) \in \hat{\mathcal{D}}^{(n)}(\zeta)} z.$$

Let $\hat{\mathcal{P}} \subseteq \mathcal{P}$ be the subset of probability measures supported on $\hat{\Omega}$, and $\hat{\mathcal{M}}_n \subseteq \hat{\mathcal{P}}$ be the subset of probability measures $\mathbb{Q}$ that have the following properties :

$$\mathbb{E}^{\mathbb{Q}}\Big[\sum_{k=1}^{K_m+m}\Big|\mathbb{E}^{\mathbb{Q}}\Big[X_{\hat{\tau}_k}\Big|\mathcal{F}_{\hat{\tau}_k-}\Big] \; - \; X_{\hat{\tau}_{k-1}}\Big|\Big] \;\; \leq \;\; \frac{1}{n},$$

where $0 < \hat{\tau}_1(\omega) < \cdots < \hat{\tau}_{K_m+m-1}(\omega) < 1$ are the jumps times of the piecewise constant process $X(\omega)$ with $\hat{\tau}_0(\omega) = 0$ and $\hat{\tau}_{K_m+m}(\omega) = 1$. Then the required result $\mathsf{D}^c(\zeta) \leq \mathsf{P}(\zeta)$ follows from the following Propositions III.5.8 and III.5.9.

Proposition III.5.8. *Assume that ζ satisfies Assumptions III.3.10, III.3.11 and is $\mathbb{L}^\infty-$uniformly continuous, then*

$$\limsup_{n\to\infty} \mathsf{D}^{(n)}(\zeta) \;\; \leq \;\; \mathsf{P}(\zeta).$$

Proof. (i) From Example 2.3 and Proposition 4.1 in Föllmer & Kramkov [42], it follows that

$$\mathsf{D}^{(n)}(\zeta) \;\; = \;\; \sup_{\mathbb{Q}\in\hat{\mathcal{P}}} \; \mathbb{E}^{\mathbb{Q}}\Big[\zeta \; - \; n\sum_{k=1}^{K_m+m}\Big|X_{\hat{\tau}_{k-1}} \; - \; \mathbb{E}^{\mathbb{Q}}\Big[X_{\hat{\tau}_k}\Big|\mathcal{F}_{\hat{\tau}_k-}\Big]\Big|\Big].$$

Since $0 \leq \zeta \leq 1$, we determine that $\mathsf{D}^{(n)}(\zeta) \geq 0$ and we have for every $\mathbb{Q} \in \hat{\mathcal{P}}\backslash\hat{\mathcal{M}}_n$,

$$\mathbb{E}^{\mathbb{Q}}\Big[\zeta \; - \; n\sum_{k=1}^{K_m+m}\Big|X_{\hat{\tau}_{k-1}} \; - \; \mathbb{E}^{\mathbb{Q}}\Big[X_{\hat{\tau}_k}\Big|\mathcal{F}_{\hat{\tau}_k-}\Big]\Big|\Big] \;\; \leq \;\; 0,$$

which yields

$$\mathsf{D}^{(n)}(\zeta) \;\; \leq \;\; \sup_{\mathbb{Q}\in\hat{\mathcal{M}}_n} \; \mathbb{E}^{\mathbb{Q}}\Big[\zeta(X)\Big].$$

(ii) Let us take a sequence $(\mathbb{Q}_n)_{n\geq 1}$ with $\mathbb{Q}_n \in \hat{\mathcal{M}}_n$ such that

$$\limsup_{n\to\infty} \sup_{\mathbb{Q}\in\hat{\mathcal{M}}_n} \; \mathbb{E}^{\mathbb{Q}}\Big[\zeta(X)\Big] \;\; = \;\; \limsup_{n\to\infty} \mathbb{E}^{\mathbb{Q}_n}\Big[\zeta(X)\Big].$$

Since under each $\mathbb{Q}_n$ the canonical process X is piecewise constant with jump times $0 < \hat{\tau}_1 < \cdots < \hat{\tau}_{K_m+m-1} < 1$, X is a $\mathbb{Q}_n-$semimartingale. Then we have the decomposition $X = M^{\mathbb{Q}_n} - A^{\mathbb{Q}_n}$, where $A^{\mathbb{Q}_n}$ is a predictable process of bounded variation and $M^{\mathbb{Q}_n}$ is a martingale under $\mathbb{Q}_n$. Moreover, $A^{\mathbb{Q}_n}$ is identified by

$$A_t^{\mathbb{Q}_n} \;=\; \sum_{k=1}^{K_m+m-1} \mathbb{1}_{[\hat\tau_k,\hat\tau_{k+1})}(t) \sum_{j=1}^{k} \Big[X_{\hat\tau_{j-1}} \;-\; \mathbb{E}^{\mathbb{Q}_n}\Big[X_{\hat\tau_j} \Big| \mathcal{F}_{\hat\tau_j -} \Big] \Big] \text{ for all } t \in [0,1),$$

and $A_1^{\mathbb{Q}_n} = \lim_{t\to 1} A_t^{\mathbb{Q}_n}$. It follows then $\mathbb{E}^{\mathbb{Q}_n}[|X_1 - M_1^{\mathbb{Q}_n}|] \le \mathbb{E}^{\mathbb{Q}_n}[|A_1^{\mathbb{Q}_n}|] \le 1/n$ and

$$\mathbb{Q}_n\Big[\|A^{\mathbb{Q}_n}\| \ge n^{-1/2} \Big] \;\le\; n^{1/2}\mathbb{E}^{\mathbb{Q}_n}\Bigg[\sum_{k=1}^{K_m+m-1} \Big| X_{\hat\tau_{k-1}} \;-\; \mathbb{E}^{\mathbb{Q}_n}\Big[X_{\hat\tau_k} \Big| \mathcal{F}_{\hat\tau_k -} \Big] \Big| \Bigg] \;\le\; n^{-1/2}.$$

Since ζ is $\mathbb{L}^\infty-$uniformly continuous, one obtains

$$\limsup_{n\to\infty} \mathbb{E}^{\mathbb{Q}_n}\Big[\zeta(X) \Big] \;\le\; \limsup_{n\to\infty} \mathbb{E}^{\mathbb{Q}_n}\Big[\zeta(M^{\mathbb{Q}_n}) \Big].$$

Let $\mathbb{P}_n = \mathbb{Q}_n \circ (M^{\mathbb{Q}_n})^{-1}$, then

$$\begin{aligned}
\sup_{n\ge 1} \; \mathbb{E}^{\mathbb{P}_n}\Big[|X_1| \Big] \;&=\; \sup_{n\ge 1} \; \mathbb{E}^{\mathbb{Q}_n}\Big[|M_1^{\mathbb{Q}_n}| \Big] \\
&\le\; \sup_{n\ge 1} \; \mathbb{E}^{\mathbb{Q}_n}\Big[|M_1^{\mathbb{Q}_n} \;-\; X_1| \Big] \;+\; \sup_{n\ge 1} \; \mathbb{E}^{\mathbb{Q}_n}\Big[X_1 \Big] \\
&\le\; \sup_{n\ge 1} \; \mathbb{E}^{\mathbb{Q}_n}\Big[|M_1^{\mathbb{Q}_n} \;-\; X_1| \Big] \;+\; \sup_{n\ge 1} \; \mathbb{E}^{\mathbb{Q}_n}\Big[X_1 \;-\; M_1^{\mathbb{Q}_n} \Big] \;+\; \sup_{n\ge 1} \; \mathbb{E}^{\mathbb{Q}_n}\Big[M_1^{\mathbb{Q}_n} \Big] \\
&\le\; 1 \;+\; \frac{2}{n} \;\le\; 3.
\end{aligned}$$

By Assumptions III.3.10 and III.3.11, it follows that for any $\boldsymbol{\varepsilon} \in \mathbb{R}_+^m$ such that $0 < |\boldsymbol{\varepsilon}| < \Delta\mathbb{T}$

$$\limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}\Big[\zeta(X) \Big] \;\le\; \limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}\Big[\zeta(X_{f_{\boldsymbol{\varepsilon}}}) \Big] \;+\; \Big(1 \;+\; (m+2)d \Big) \alpha(|\boldsymbol{\varepsilon}|).$$

Again with the same reasoning, we may prove

$$\limsup_{n\to\infty} \mathbb{E}^{\mathbb{P}_n}\Big[\zeta(X_{f_{\boldsymbol{\varepsilon}}}) \Big] \;\le\; \mathsf{P}(\zeta).$$

Since $\boldsymbol{\varepsilon}$ is arbitrary we get

$$\limsup_{n\to\infty} \Bigg[\sup_{\mathbb{Q}\in\hat{\mathcal{M}}_n} \; \mathbb{E}^{\mathbb{Q}}\Big[\zeta(X) \Big] \Bigg] \;\le\; \mathsf{P}(\zeta),$$

and hence the required result. $\square$

5.4.2 Time-space discretization

<u>*Discretization*</u> : For each $\omega \in \mathbf{D}_+^c$ let us define $\tau_k := \tau_k^{(n)}(\omega)$ and $K_i := K_i^{(n)}(\omega)$ by

$$
\begin{aligned}
\tau_0 &:= 0, \qquad K_0 := 0, \\
\tau_1 &:= t_1 \wedge \sqrt{d}2^{-n} \wedge \inf\Big\{t > 0:\ |\omega_t - \omega_0| \geq 2^{-n}\Big\}, \\
\tau_{k+1} &:= t_1 \wedge \Big(\tau_k + \Delta\tau_k\Big) \wedge \inf\Big\{t > \tau_k:\ |\omega_t - \omega_{\tau_k}| \geq 2^{-n}\Big\},\ \Delta\tau_k = \tau_k - \tau_{k-1} \text{ for } k \geq 1.
\end{aligned}
$$

Set further

$$K_1 \ := \ \min\{k \in \mathbb{N}:\ \tau_k = t_1\}.$$

Recursively, we define for $1 \leq i \leq m-1$ and $k \geq K_i$,

$$
\begin{aligned}
\tau_{K_i+1} &:= t_{i+1} \wedge \Big(t_i + \sqrt{d}2^{-n}\Big) \wedge \inf\Big\{t > t_i:\ |\omega_t - \omega_{t_i}| \geq 2^{-n}\Big\}, \\
\tau_{k+1} &:= t_{i+1} \wedge \Big(\tau_k + \Delta\tau_k\Big) \wedge \inf\Big\{t > \tau_k:\ |\omega_t - \omega_{\tau_k}| \geq 2^{-n}\Big\} \text{ for } k \geq K_i + 1
\end{aligned}
$$

and

$$K_{i+1} \ := \ \min\Big\{k \in \mathbb{N}:\ \tau_k = t_{i+1}\Big\}.$$

Notice that the above τ_k, $k \geq 0$ are all stopping times with respect to the right-continuous filtration $\mathbb{F}^+ = (\mathcal{F}_{t+})_{t\geq 0}$, and

$$0 \ = \ \tau_0 \ < \ \tau_1 \ \cdots \ < \ \tau_{K_m} \ = \ 1 \quad \text{and} \quad \tau_{K_i} \ = \ t_i \text{ for all } i = 1, \cdots, m.$$

Moreover, for $0 \leq i \leq m-1$, $K_i < k \leq K_{i+1}$ and $t \in [\tau_{k-1}, \tau_k)$,

$$|\omega_t - \omega_{\tau_{k-1}}| \ \leq \ 2^{-n} \quad \text{and} \quad \Delta\tau_{k+1} \ \leq \ \Delta\tau_k \ \leq \ 2^{-n}.$$

Also by the continuity of ω at $\tau_{K_i} = t_i$ for all $i = 1, \cdots, m$

$$|\omega_t - \omega_{\tau_{K_i-1}}| \ \leq \ 2^{-n} \text{ for all } t \in [\tau_{K_i-1}, t_i] \text{ and } i = 1, \cdots, m.$$

Lifting : Set $\hat\tau_0 := 0$ and for $0 \le i \le m-1$

$$\begin{aligned}
\hat\tau_{K_i+1} &:= \hat\tau_{K_i} + \sqrt{d}2^{-n},\\
\hat\tau_k &:= \hat\tau_{k-1} + \Big(1-\sqrt{d}2^{-n}/\Delta t_{i+1}\Big)\sup\Big\{\Delta t>0:\ \Delta t\in B^{(n+k-K_i-i)},\ \Delta t<\Delta\tau_{k-1}\Big\},\\
&\quad\ \text{for all } K_i+i+2\le k\le K_{i+1}+i,\\
\hat\tau_{K_{i+1}+i+1} &:= t_{i+1}.
\end{aligned}$$

Denote $\hat\Pi(\omega) = \big(\hat\Pi_t(\omega)\big)_{0\le t\le 1}$ by

$$\begin{aligned}
\hat\Pi_t(\omega) \ :=\ & \sum_{i=0}^{m-1}\Big\{\sum_{k=K_i+i}^{K_{i+1}+i-1}\pi^{(n+k-K_i-i)}(\omega_{\tau_k})\mathbb{1}_{[\hat\tau_k,\hat\tau_{k+1})}(t)\ +\ \pi^{(n)}(\omega_{t_i})\mathbb{1}_{[\hat\tau_{K_r},t_{i+1})}(t)\Big\}\\
&+\ \pi^{(n)}(\omega_1)\mathbb{1}_{\{t=1\}},
\end{aligned}$$

then $\hat\Pi(\omega)\in\hat\Omega$. For each $\hat H\in\hat{\mathcal{A}}$ we may define

$$H_t(\omega)\ :=\ \sum_{k=0}^{K_m-1}\hat H_{\hat\tau_{k+1}(\omega)}\Big(\hat\Pi(\omega)\Big)\mathbb{1}_{(\tau_k(\omega),\tau_{k+1}(\omega)]}(t)\ \text{for all } (\omega,t)\in\mathbf{D}_+\times[0,1]. \tag{III.5.9}$$

We observe that $(\omega,t)\mapsto H_t(\omega)$ is Borel measurable on $\mathbf{D}_+\times[0,1]$, $t\mapsto H_t(\omega)$ is left-continuous, $\omega\mapsto H_t(\omega)$ is $\mathcal{F}_{t+}-$measurable. Hence H is $\mathbb{F}^+-$predictable, which is equivalent to be $\mathbb{F}-$predictable. Further, following the argument of Lemmas 3.5 and 3.6 of Dolinsky & Soner [37], we see that the process H defined by (III.5.9) belongs to $\mathcal{H}$, and more importantly, there exists some constant $C>0$ independent of n such that for all $\omega\in\mathbf{D}_+$,

$$\rho_{\mathbb{T}}(\omega,\hat\Pi(\omega))\ \le\ C2^{-n}\Big(1\ +\ \|\omega\|\Big)\ \text{and}\ \Big|(H\cdot\omega)_1\ -\ (\hat H(\hat\Pi(\omega))\cdot\hat\Pi(\omega))_1\Big|\ \le\ Cn2^{-n}. \tag{III.5.10}$$

Proposition III.5.9. *Assume that ζ satisfies Condition (III.5.2), then one has*

$$\liminf_{n\to\infty}\mathsf{D}^{(n)}(\zeta)\ \ge\ \mathsf{D}^c(\zeta).$$

Proof. Take an arbitrary $(\hat z,\hat H)\in\hat{\mathcal{D}}$. Then for any $\omega\in\mathbf{D}_+$ one has $\hat\Pi(\omega)\in\hat\Omega$ and thus

$$\hat z\ +\ \Big(\hat H\Big(\hat\Pi(\omega)\Big)\cdot\hat\Pi(\omega)\Big)_1\ \ge\ \zeta(\hat\Pi(\omega))\ \text{for all } \omega\in\mathbf{D}_+.$$

Take H constructed as (III.5.9), then by (III.5.10), we have $H \in \mathcal{H}$ and

$$\hat{z} \;+\; (H \cdot \omega)_1 \;\geq\; \zeta(\hat{\Pi}(\omega)) \;-\; Cn2^{-n} \text{ for all } \omega \in \mathbf{D}_+.$$

Moreover, by the construction of $\hat{\Pi}(\omega)$ one has $\|\hat{\Pi}(\omega)\| \leq \|\omega\|$. Notice that ζ is boundedly supported, saying by $\{\omega \in \Omega : \|\omega\| \leq R\}$. Then by (III.5.2) one has a continuous increasing function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\beta(0) = 0$ such that for all $\omega \in \mathbf{D}_+$,

$$\zeta(\hat{\Pi}(\omega)) \;\;\geq\;\; \zeta(\omega) \;-\; \mathbb{1}_{\{\|\omega\| \leq R\}} \beta\Big(\rho_{\mathbb{T}}\Big(\omega, \hat{\Pi}(\omega)\Big)\Big) \;\;\geq\;\; \zeta(\omega) \;-\; \beta\Big(C(1+R)2^{-n}\Big),$$

which implies that $(\hat{z} + \beta(C(1+R)2^{-n}) + Cn2^{-n}, H) \in \mathcal{D}^c(\zeta)$. Hence

$$\mathsf{D}^c(\zeta) \;\leq\; \mathsf{D}^{(n)}(\zeta) \;+\; \beta\Big(C(1+R)2^{-n}\Big) \;+\; Cn2^{-n},$$

which yields the required result. □

5.5 Proof of Proposition III.5.4 (i)

Recal that ξ satisfies Assumptions III.3.10, III.3.11 and Condition (III.5.2). The required statement follows from the three lemmas below.

Lemma III.5.10. $\limsup_{n\to\infty} \xi^\delta(\omega^{\delta,n}) \leq \xi(\omega^{\delta,0})$ *holds for any sequence* $(\omega^{\delta,n})_{n\geq 0} \subseteq \mathbf{D}^\delta$ *such that*

$$\omega^{\delta,n} \;\xrightarrow{S^*}\; \omega^{\delta,0} \quad \textit{and} \quad \omega^{\delta,n}_{t_i^\delta} \;\longrightarrow \omega^{\delta,0}_{t_i^\delta} \textit{ for all } i = 0, \cdots, m-1, \;\; \textit{as } n \longrightarrow \infty.$$

Proof. For the sake of simplicity we may assume that

$$\limsup_{n\to\infty} \xi^\delta(\omega^{\delta,n}) \;\;=\;\; \lim_{n\to\infty} \xi^\delta(\omega^{\delta,n}).$$

Since $(\omega^{\delta,n})_{n\geq 1}$ is $S-$tight, then by the $S-$tightness criteria and the construction in (III.5.5) we determine that $(\bar{\omega}^{\delta,n})_{n\geq 1}$ is again $S-$tight, which yields a convergent subsequence $(\bar{\omega}^{\delta,n_k})_{k\geq 1}$ and a limit $\omega^0 \in \mathbf{D}$, *i.e.* $\bar{\omega}^{\delta,n_k} \xrightarrow{S^*} \omega^0$. Clearly, $\omega^{\delta,n}_{t_i^\delta} \longrightarrow \omega^{\delta,0}_{t_i^\delta}$ implies in particular that $\bar{\omega}^{\delta,n_k}_{t_i} \longrightarrow \bar{\omega}^{\delta,0}_{t_i}$ for all $i = 0, \cdots, m-1$. Next, $\omega^{\delta,n_k} \xrightarrow{S^*} \omega^{\delta,0}$ yields a countable set $\mathcal{T} \subseteq [0, 1+\delta)$ such that

$$\omega^{\delta,n_k}_t \;\longrightarrow\; \omega^{\delta,0}_t \text{ for all } t \in [0, 1+\delta] \backslash \mathcal{T},$$

which yields another countable set $\mathcal{T}' \subseteq [0,1)$ such that

$$\bar{\omega}_t^{\delta,n_k} \longrightarrow \bar{\omega}_t^{\delta,0} \text{ for all } t \in [0,1] \setminus \mathcal{T}'.$$

Hence one has $\bar{\omega}^{\delta,0} = \omega^0$ and thus

$$\bar{\omega}^{\delta,n_k} \xrightarrow{S^*} \bar{\omega}^{\delta,0} \quad \text{and} \quad \bar{\omega}_{t_i}^{\delta,n_k} \longrightarrow \bar{\omega}_{t_i}^{\delta,0} \text{ for all } i = 0, \cdots, m-1,$$

which implies that

$$\lim_{k\to\infty} \xi^\delta(\omega^{\delta,n_k}) \;=\; \lim_{k\to\infty} \xi(\bar{\omega}^{\delta,n_k}) \;\leq\; \xi(\bar{\omega}^{\delta,0}) \;=\; \xi^\delta(\omega^{\delta,0}).$$

□

Lemma III.5.11. *There exists a continuous function $\alpha_\delta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\alpha_\delta(0) = 0$ such that for all $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_m) \in \mathbb{R}_+^m$ sufficiently small one has*

$$\left|\xi^\delta(\omega^\delta) - \xi^\delta\left(\omega^\delta_{f^\delta_{\boldsymbol{\varepsilon}}}\right)\right|, \; \left|\xi^\delta(\omega^\delta) - \xi^\delta\left(\omega^\delta_{b^\delta_{\boldsymbol{\varepsilon}}}\right)\right| \;\leq\; \alpha_\delta(|\boldsymbol{\varepsilon}|)\left(1 + \sum_{i=0}^m |\omega^\delta_{t_i^\delta}| + \int_0^{1+\delta} |\omega_t^\delta| dt\right),$$

where $f^\delta_{\boldsymbol{\varepsilon}}$, $b^\delta_{\boldsymbol{\varepsilon}} : [0, 1+\delta] \to [0, 1+\delta]$ are two non-decreasing functions defined as in (III.3.4) *and* (III.3.5)*.*

Proof. We only prove the inequality on $f^\delta_{\boldsymbol{\varepsilon}}$, while the inequality on $b^\delta_{\boldsymbol{\varepsilon}}$ follows by the same arguments. Define $\bar{f}^\delta_{\boldsymbol{\varepsilon}} : [0, 1+\delta] \to [0, 1+\delta]$ by

$$\bar{f}^\delta_{\boldsymbol{\varepsilon}}(t) \;:=\; \frac{1}{k_\delta} f^\delta_{\boldsymbol{\varepsilon}}(k_\delta t) \text{ for all } t \in [0,1].$$

Thus by the construction of $f^\delta_{\boldsymbol{\varepsilon}}$ we get

$$\begin{aligned}
\bar{f}^\delta_{\boldsymbol{\varepsilon}}(t) \;&=\; \frac{1}{k_\delta} \sum_{i=1}^m \mathbb{1}_{(t^\delta_{i-1}, t^\delta_i]}(k_\delta t)\left(t^\delta_{i-1} + \frac{\Delta t_i^\delta}{\Delta t_i^\delta - \varepsilon_i}\left(k_\delta t - t^\delta_{i-1} - \varepsilon_i\right)^+\right) \\
&=\; \sum_{i=1}^m \mathbb{1}_{(t_{i-1}, t_i]}(t)\left(t_{i-1} + \frac{\Delta t_i}{\Delta t_i - \varepsilon_i / k_\delta}\left(t - t_{i-1} - \frac{\varepsilon_i}{k_\delta}\right)^+\right),
\end{aligned}$$

which implies that

$$\bar{\omega}^\delta_{\bar{f}^\delta_{\boldsymbol{\varepsilon}}(t)} \;=\; \omega^\delta_{k_\delta \bar{f}^\delta_{\boldsymbol{\varepsilon}}(t)} \text{ for all } t \in [0,1]$$

and thus $\bar{\omega}^\delta_{f^\delta_\varepsilon} = \bar{\omega}^\delta \circ \bar{f}^\delta_\varepsilon$. Hence

$$\begin{aligned}
\left|\xi^\delta(\omega^\delta) \ - \ \xi^\delta\Big(\bar{\omega}^\delta_{f^\delta_\varepsilon}\Big)\right| \ &= \ \left|\xi(\bar{\omega}^\delta) \ - \ \xi\Big(\bar{\omega}^\delta \circ \bar{f}^\delta_\varepsilon\Big)\right| \\
&\leq \ \alpha(|\boldsymbol{\varepsilon}|/k_\delta)\Big(1 \ + \ \sum_{i=0}^m |\bar{\omega}^\delta_{t_i}| \ + \ \int_0^1 |\bar{\omega}^\delta_t| dt\Big) \\
&= \ \alpha(|\boldsymbol{\varepsilon}|/k_\delta)\Big(1 \ + \ \sum_{i=0}^m |\omega^\delta_{t^\delta_i}| \ + \ \frac{1}{k_\delta}\int_0^{1+\delta} |\omega^\delta_t| dt\Big) \\
&\leq \ \alpha(|\boldsymbol{\varepsilon}|/k_\delta)\Big(1 \ + \ \sum_{i=0}^m |\omega^\delta_{t^\delta_i}| \ + \ \int_0^{1+\delta} |\omega^\delta_t| dt\Big).
\end{aligned}$$

The proof is completed by taking $\alpha_\delta(\cdot) = \alpha(\cdot/k_\delta)$. □

Lemma III.5.12. *ξ^δ is $L^\infty-$uniformly continuous and satisfies Condition (III.5.2) for $\rho_{\mathbb{T}^\delta}$.*

Proof. For any $\omega^\delta, v^\delta \in \Omega^\delta$ such that $\|v^\delta\| \leq \|\omega^\delta\|$, one has

$$\begin{aligned}
\xi^\delta(\omega^\delta) \ - \ \xi^\delta(v^\delta) \ &= \ \xi(\bar{\omega}^\delta) \ - \ \xi(\bar{v}^\delta) \ \leq \ \beta\Big(\rho_{\mathbf{T}}(\bar{\omega}^\delta, \bar{v}^\delta)\Big), \\
\left|\xi^\delta(\omega^\delta) \ - \ \xi^\delta(v^\delta)\right| \ &= \ \left|\xi(\bar{\omega}^\delta) \ - \ \xi(\bar{v}^\delta)\right|.
\end{aligned}$$

It is thus enough to show that

$$\rho_{[t_{i-1},t_i]}(\bar{\omega}^\delta, \bar{v}^\delta) \ \leq \ \rho_{[t^\delta_{i-1},t^\delta_i]}(\omega^\delta, v^\delta) \text{ for all } i = 1, \cdots, m$$

and

$$\left|\int_0^1 \Big(\bar{\omega}^\delta_t - \bar{v}^\delta_t\Big) dt\right| \ \leq \ \left|\int_0^{1+\delta} \Big(\omega^\delta_t - v^\delta_t\Big) dt\right|.$$

Let $\Gamma_{[s,t]}$ denotes the collection of strictly increasing continuous functions γ defined on $[s,t]$ such that $\gamma(s) = s$ and $\gamma(t) = t$. For any $\gamma^\delta \in \Gamma_{[t^\delta_{i-1},t^\delta_i]}$, define $\gamma \in \Gamma_{[t_{i-1},t_i]}$ by

$$\gamma(t) \ := \ \frac{1}{k_\delta}\gamma^\delta\Big(k_\delta t\Big) \text{ for all } t \in [t_{i-1}, t_i].$$

Hence

$$\begin{aligned}
\sup_{t_{i-1}\leq t\leq t_i}\left|\bar{\omega}^\delta_{\gamma(t)} \ - \ \bar{v}^\delta_t\right| \ &= \ \sup_{t_{i-1}\leq t\leq t_i}\left|\omega^\delta_{k_\delta\gamma(t)} \ - \ v^\delta_{k_\delta t}\right| \ = \ \sup_{t_{i-1}\leq t\leq t_i}\left|\omega^\delta_{\gamma^\delta(k_\delta t)} \ - \ v^\delta_{k_\delta t}\right| \\
= \ \sup_{t^\delta_{i-1}\leq t\leq t^\delta_i}\left|\omega^\delta_{\gamma^\delta(t)} \ - \ v^\delta_t\right|
\end{aligned}$$

and

$$\sup_{t_{i-1}\le t\le t_i}\Big|\gamma(t)\;-\;t\Big| \;=\; \sup_{t_{i-1}\le t\le t_i}\Big|\frac{1}{k_\delta}\gamma^\delta\Big(k_\delta t\Big)\;-\;t\Big| \;=\; \frac{1}{k_\delta}\sup_{t_{i-1}\le t\le t_i}\Big|\gamma^\delta\Big(k_\delta t\Big)\;-\;k_\delta t\Big|$$
$$=\;\frac{1}{k_\delta}\sup_{t^\delta_{i-1}\le t\le t^\delta_i}\Big|\gamma^\delta(t)\;-\;t\Big|\;\le\;\sup_{t^\delta_{i-1}\le t\le t^\delta_i}\Big|\gamma^\delta(t)\;-\;t\Big|,$$

which implies that

$$\rho_{[t_{i-1},t_i]}(\bar\omega^\delta,\bar\upsilon^\delta)\;\;\le\;\;\rho_{[t^\delta_{i-1},t^\delta_i]}(\omega^\delta,\upsilon^\delta).$$

We may thus conclude by

$$\Big|\int_0^1\Big(\bar\omega^\delta_t-\bar\upsilon^\delta_t\Big)dt\Big|\;\;=\;\;\frac{1}{k_\delta}\Big|\int_0^{1+\delta}\Big(\omega^\delta_t-\upsilon^\delta_t\Big)dt\Big|\;\;\le\;\;\Big|\int_0^{1+\delta}\Big(\omega^\delta_t-\upsilon^\delta_t\Big)dt\Big|.$$

□

6 Appendix

6.1 Tightness under $S-$topology

Recall that $\mathbf{D}$ is the Skorokhod space of càdlàg functions on $[0,1]$, with the canonical process $X=(X_t)_{0\le t\le 1}$ and the canonical filtration $\mathbb{F}=(\mathcal{F}_t)_{0\le t\le 1}$, and $\mathcal{P}$ denotes the set of all probability measures on $(\mathbf{D},\mathcal{F}_1)$. A sequence of probability measures $(\mathbb{P}_n)_{n\ge1}\subseteq\mathcal{P}$ is said to be $S-$tight if for any $\varepsilon>0$, there exists a $S-$compact set $K_\varepsilon\subseteq\mathbf{D}$ such that

$$\inf_{n\ge1}\;\mathbb{P}_n\Big[X\in K_\varepsilon\Big]\;\;\ge\;\;1-\varepsilon.$$

The following result is recalled from Jakubowski [67], see Theorem 3.11 and the discussion at the beginning of Section 4 in [67].

Theorem III.6.1 (Jakubowski)**.** (i) *Let $(\mathbb{P}_n)_{n\ge1}\subseteq\mathcal{P}$ be a sequence of probability measures such that X is a $\mathbb{P}_n-$supermartingale for all $n\ge1$, then*

$$\sup_{n\ge1}\sup_{0\le t\le1}\mathbb{E}^{\mathbb{P}_n}\Big[|X_t|\Big]\;<\;+\infty\qquad\Longrightarrow\qquad(\mathbb{P}_n)_{n\ge1}\ \textit{is } S-\textit{tight}.$$

(ii) *Let $(\mathbb{P}_n)_{n\ge1}\subseteq\mathcal{P}$ be a $S-$tight sequence of probability measures. Then there exist a subsequence $(\mathbb{P}_{n_k})_{k\ge1}$, a probability measure $\mathbb{P}\in\mathcal{P}$ and a countable subset $\mathcal{T}\subseteq[0,1)$ such that for* *all finite sets* $\{u_1 < u_2 < \cdots < u_r\} \subseteq [0,1]\backslash\mathcal{T}$,

$$\mathbb{P}_{n_k} \circ (X_{u_1}, \cdots, X_{u_r})^{-1} \;\longrightarrow\; \mathbb{P} \circ (X_{u_1}, \cdots, X_{u_r})^{-1} \quad as \quad k \;\longrightarrow\; \infty. \tag{III.6.1}$$

In particular, $X^{n_k} \overset{*}{\Longrightarrow}_D X^0$ *as* $k \to \infty$.

6.2 Dual space of M

Recall that $\mathbf{M}$ denotes the space of all finite signed measures μ on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d} \Big(1 + |x|\Big)|\mu|(dx) \;\;<\;\; +\infty,$$

and it is equipped with the topology induced by the convergence $\xrightarrow{\mathcal{W}_1}$. We would like identify its dual space as Λ, where the arguments are mainly adapted from Lemma 3.2.3 of Deuschel & Stroock [35]. Notice that the topology on $\mathbf{M}$ is generated by all the following open balls

$$U_{\lambda^1,\cdots,\lambda^m,c}(\mu) \;\;:=\;\; \Big\{\nu \in \mathbf{M}: \;\Big|\mu(\lambda^i) - \nu(\lambda^i)\Big| \;<\; c \text{ for all } 1 \le i \le m\Big\},$$

where $\lambda^i \in \Lambda$ for $1 \le i \le m$ and $c > 0$. Let $\mathcal{O}$ be the collection of open sets generated by the open balls above, then clearly, every open set $U \in \mathcal{O}$ could be expressed as

$$U \;=\; \bigcup_{\alpha} U_{\lambda^1_\alpha,\cdots,\lambda^{n_\alpha}_\alpha,c^\alpha}(\mu^\alpha) \text{ with } \lambda^i_\alpha \in \Lambda \text{ for } 1 \le i \le n_\alpha, n_\alpha \in \mathbb{N} \text{ and } c^\alpha > 0.$$

Theorem III.6.2. *The space* $(\mathbf{M}, \mathcal{O})$ *is a Hausdorff locally convex space, whose dual space can be identified by* $\mathbf{M}^* = \Lambda$.

Proof. (i) First, $(\mathbf{M}, \mathcal{O})$ is clearly a topological vector space. For every $\mu \in \mathbf{M}$, let

$$\mathcal{U}(\mu) \;\;:=\;\; \Big\{U_{\lambda^1_\alpha,\cdots,\lambda^{n_\alpha}_\alpha,c^\alpha}(\mu): \;\lambda^i_\alpha \in \Lambda \text{ for } 1 \le i \le n_\alpha, n_\alpha \in \mathbb{N} \text{ and } c^\alpha > 0\Big\}.$$

By definition, one can check that $\mathcal{U}(\mu)$ is a local basis of μ for every $\mu \in \mathbf{M}$. Moreover, by denoting $\mathbf{0} \in \mathbf{M}$ the null measure, $\mathcal{U}(\mathbf{0})$ is a local basis of absolutely convex absorbent sets and thus $\mathbf{M}$ is a locally convex space.

(ii) Now, let us identify the dual space of $\mathbf{M}$. First, for every $\lambda \in \Lambda$, the map $F_\lambda : \mathbf{M} \to \mathbb{R}$ defined by $F_\lambda(\mu) := \mu(\lambda)$ gives a unique element in $\mathbf{M}^*$, and hence $\Lambda \subseteq \mathbf{M}^*$. On the other hand, for any $F \in \mathbf{M}^*$, we define a function λ^F by

$$\lambda^F(x) \ := \ F(\delta_{\{x\}}) \text{ for all } x \in \mathbb{R}^d.$$

Clearly one has the following implication

$$x_n \ \longrightarrow \ x_0 \quad \Longrightarrow \quad \delta_{\{x_n\}} \xrightarrow{\mathcal{O}} \delta_{\{x_0\}} \quad \Longrightarrow \quad \lambda^F(x_n) \ \longrightarrow \ \lambda^F(x_0),$$

which implies that λ^F is continuous. It follows that the set $F^{-1}\Big((-1,1)\Big)$ is open and thus there exists some $U_{\lambda^1,\cdots,\lambda^m,c}(\mathbf{0})$ such that

$$U_{\lambda^1,\cdots,\lambda^m,c}(\mathbf{0}) \ \subseteq \ F^{-1}\Big((-1,1)\Big),$$

where $\lambda^i \in \Lambda$ for all $i = 1, \cdots, m$ and $c > 0$. Now for any $\mu \in \mathbf{M}$ such that $\sum_{i=1}^m \Big|\mu(\lambda^i)\Big| > 0$, we define

$$\bar{\mu} \ := \ \frac{c\mu}{\sum_{i=1}^m \Big|\mu(\lambda^i)\Big|}.$$

Then $\bar{\mu} \in U_{\lambda^1,\cdots,\lambda^m,c}(\mathbf{0})$ and thus $|F(\bar{\mu})| < 1$. It follows that

$$|F(\mu)| \ \le \ c\sum_{i=1}^m \Big|\mu(\lambda^i)\Big| \text{ for all } \mu \in \mathbf{M},$$

and hence $\lambda^F \in \Lambda$. When μ is a linear combination of Dirac measures, it is obvious that $F(\mu) = \mu(\lambda^F)$. Moreover, since such μ are dense in $\mathbf{M}$, it follows that $F(\mu) = \mu(\lambda^F)$ holds for all $\mu \in \mathbf{M}$. $\square$

Chapitre IV

Optimal Skorokhod embedding under finitely-many marginal constraints

1 Introduction

Given a probability measure μ on $\mathbb{R}$, with finite first moment and centered, the Skorokhod embedding problem (SEP) consists in finding a stopping time τ on a Brownian motion W such that $W_\tau \sim \mu$ and the stopped process $W_{\tau\wedge\cdot} := (W_{\tau\wedge t})_{t\geq 0}$ is uniformly integrable. We refer the readers to the survey paper [83] of Obłój for a comprehensive account of the field.

We consider here its extension to the case of multiple marginal constraints. Namely, let $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$ be some vector of probability distributions. The extended SEP aims to find an increasing family of stopping times $\tau = (\tau_k)_{1\leq k\leq m}$ such that $W_{\tau_k} \sim \mu_k$ for all $k = 1, \cdots, m$ and the stopped process $W_{\tau_m\wedge\cdot}$ is uniformly integrable. Then, we study an associated optimization problem, which consists in maximizing the expected value of some function among all such embeddings.

One of the motivations is to study the robust hedging problem of derivatives consistent with the market prices of call options. Mathematically, the underlying asset is required to be a martingale according to the no-arbitrage condition and the market calibration allows to recover the marginal laws of the underlying at certain maturities. Then by considering all martingales fitting the given marginal distributions, one can obtain the interval of no-arbitrage prices. Based on the fact that every continuous martingale can be considered as a time-changed Brownian motion, Hobson studied the robust superhedging of lookback options in his seminal paper [56] by means of the optimal SEP. The main idea of his pioneering work is to exploit some solution of the SEP satisfying some optimality criteria, which yields the optimal superhedging strategies. Since then, the optimal SEP has received substantial attention from the mathematical finance community and various extensions were achieved in the literature, such as Brown, Hobson &

Rogers [17], Cox & Obłój [24, 25], Cox, Hobson & Obłój [23], Cox, Obłój & Touzi [26], Cox & Wang [28], Davis, Obłój & Raval [30], Gassiat, Oberhauser & dos Reis [44], Hobson & Klimmek [58–60], Hobson & Neuberger [61], Madan & Yor [76], *etc.* A thorough literature is provided in Hobson's survey paper [57].

Beiglböck, Cox & Huesmann generalized this heuristic idea and formulated the optimal SEP in [5], which recovered many previous known results by a unifying formulation. Namely, their main results are twofold. First, they establish the expected identity between the optimal SEP and the corresponding robust superhedging problem. Second, they derive the characterization of optimal embeddings by a geometric property which allows to recover all previous known embeddings in the literature.

The objective of Chapter IV is to revisit the duality result of [5] and to extend the duality to the multi-marginal case under more general conditions. Our approach uses tools of a completely different nature. First, by following the convex duality approach, we convert the optimal SEP into an infimum of classical optimal stopping problems. Next, we use the standard dynamic programming approach to relate such optimal stopping problems to the required duality. We observe that the derived duality allows to reproduce the geometric characterization of the optimal embedding introduced in [5], which is discussed in Chapter V. Finally, we show that our result induces the duality for a class of martingale optimal transport problems on the space of continuous functions.

Chapter IV is organized as follows. In Section 2, we formulate our optimal SEP under finitely-many marginal constraints and provide two duality results. In Section 3, the obtained duality together with the time-change argument yields the duality for the martingale optimal transport problem under multiple marginal constraints. We finally provide the related proofs in Section 4.

2 Optimal Skorokhod embedding and dualities

In this section, we formulate the optimal SEP under finitely-many marginal constraints, as well as its dual problems. We then provide two duality results.

Recall that Ω is the space of all continuous functions $\omega = (\omega_t)_{t\geq 0}$ such that $\omega_0 = 0$, $B = (B_t)_{t\geq 0}$ is the canonical process, *i.e.* $B_t(\omega) := \omega_t$, $\mathbb{P}_0$ is the Wiener measure, $\mathbb{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ is its natural filtration, *i.e.* $\mathcal{F}_t := \sigma(B_s, s \leq t)$, and $\mathbb{F}^a = (\mathcal{F}^a_t)_{t\geq 0}$ is the augmented filtration under $\mathbb{P}_0$. Recall also that, the enlarged canonical space is defined by $\overline{\Omega} := \Omega \times \Theta$, where $\Theta := \{\theta = (\theta_1, \cdots, \theta_m) : 0 \leq \theta_1 \leq \cdots \leq \theta_m\}$, see El Karoui & Tan [40, 41]. All the elements of

$\overline{\Omega}$ are denoted by $\bar{\omega} = (\omega, \theta)$.

Denote further by (B, T) with $T = (T_1, \cdots, T_m)$, the canonical element on $\overline{\Omega}$, *i.e.* $B_t(\bar{\omega}) := \omega_t$ and $T(\bar{\omega}) := \theta$ for all $\bar{\omega} = (\omega, \theta) \in \overline{\Omega}$ and $t \in \mathbb{R}_+$. The enlarged canonical filtration is denoted by $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$, where $\overline{\mathcal{F}}_t$ is generated by $(B_s)_{0\leq s\leq t}$ and all the sets $\{T_k \leq s\}$ for all $s \in [0, t]$ and $k = 1, \cdots, m$. In particular, all random variables $T_1, \cdots, T_m$ are $\overline{\mathbb{F}}-$stopping times.

We endow Ω with the compact convergence topology, and Θ with the classical Euclidean topology, then Ω and $\overline{\Omega}$ are both Polish spaces. In particular, $\mathcal{F}^0 := \bigvee_{t\geq 0} \mathcal{F}^0_t$ and $\overline{\mathcal{F}} := \bigvee_{t\geq 0} \overline{\mathcal{F}}_t$ are respectively the Borel σ-field of the Polish spaces Ω and $\overline{\Omega}$, see Lemma IV.5.1.

Remark IV.2.1. *Notice that B stands for both the canonical processes on Ω and $\overline{\Omega}$, however, this will not lead to any ambiguity in the following as the process B will represent a Brownian motion for our interest, defined on different probability spaces Ω and $\overline{\Omega}$. In contrast with the natural filtration $\mathbb{F}^0$ on Ω, we denote by $\overline{\mathbb{F}}^B = (\overline{\mathcal{F}}^B_t)_{t\geq 0}$ the filtration generated by B on $\overline{\Omega}$, i.e. $\overline{\mathcal{F}}^B_t := \sigma(B_s, s \leq t)$ and by $\overline{\mathcal{F}}^B := \bigvee_{t\geq 0} \overline{\mathcal{F}}^B_t$.*

2.1 Optimal Skorokhod embedding problem

We fix a centered peacock on $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$ on $\mathbb{R}$ throughout this chapter, *i.e.* $\mu_k(x) = 0$ for all $k = 1, \cdots, m$, and $\mu_k \preceq \mu_{k+1}$ for all $k = 1, \cdots, m-1$.

We consider here the problem in a weak setting, *i.e.* the stopping times may be identified by probability measures on the enlarged space $\overline{\Omega}$. Let $\overline{\mathcal{P}}(\overline{\Omega})$ be the space of all probability measures on $(\overline{\Omega}, \overline{\mathcal{F}})$, and define

$$\overline{\mathcal{P}} \;:=\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) : \; B \text{ is an } \overline{\mathbb{F}} - \text{Brownian motion and } B_{T_m\wedge\cdot} \text{ is UI under } \overline{\mathbb{P}}\Big\}. \qquad \text{(IV.2.1)}$$

Set for the peacock $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$

$$\overline{\mathcal{P}}(\boldsymbol{\mu}) \;:=\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}} : \; \overline{\mathbb{P}} \circ B^{-1}_{T_k} \;=\; \mu_k \text{ for all } k = 1, \cdots, m\Big\}. \qquad \text{(IV.2.2)}$$

As a consequence of Strassen's theorem, $\overline{\mathcal{P}}(\boldsymbol{\mu})$ is nonempty.

Let $\Phi : \overline{\Omega} \to \mathbb{R}$ be a Borel measurable function, then Φ is called non-anticipative if $\Phi(\omega, \theta) = \Phi(\omega_{\theta_m\wedge\cdot}, \theta)$ for all $(\omega, \theta) \in \overline{\Omega}$. Define the optimal SEP for a non-anticipative function Φ by

$$\overline{P}(\boldsymbol{\mu}) \;:=\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\boldsymbol{\mu})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big]. \qquad \text{(IV.2.3)}$$

Remark IV.2.2. (i) *A $\boldsymbol{\mu}-$embedding is a collection*

$$\alpha = \Big(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{P}^\alpha, \mathbb{F}^\alpha = (\mathcal{F}_t^\alpha)_{t\geq 0}, W^\alpha, \tau^\alpha = (\tau_1^\alpha, \cdots, \tau_m^\alpha)\Big),$$

where W^α is an $\mathbb{F}^\alpha-$Brownian motion, $\tau_1^\alpha, \cdots, \tau_m^\alpha$ are increasing $\mathbb{F}^\alpha-$stopping times such that $W^\alpha_{\tau_m^\alpha \wedge \cdot}$ is uniformly integrable, and $W^\alpha_{\tau_k^\alpha} \sim \mu_k$ for all $k = 1, \cdots m$. We observe that for every centered peacock μ, every $\mu-$embedding α induces a probability measure $\overline{\mathbb{P}} := \mathbb{P}^\alpha \circ (W^\alpha, \tau^\alpha)^{-1} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$. Conversely, every probability measure $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ together with the canonical space $(\overline{\Omega}, \overline{\mathcal{F}})$, canonical filtration $\overline{\mathbb{F}}$, and canonical element (B, T) is a $\boldsymbol{\mu}-$embedding.

(ii) *The problem (IV.2.3) can be considered as a weak formulation of the optimal SEP. A strong formulation consists in considering all stopping times with respect to the Brownian filtration $\mathbb{F}^0$, and it may not be equivalent to the weak formulation, see Example IV.2.11. Although most of the well known embeddings are "strong" solutions, i.e. Brownian stopping times, some optimal embeddings can be constructed in "weak" sense, such as the solution provided by Hobson & Pedersen [62]. We also notice that it should be natural to consider the weak formulation to obtain the existence of optimizers in general, since the set of all "weak" solutions is compact under the weak convergence shown below.*

2.2 Duality results

We introduce two dual problems. Recall that $\mathbb{P}_0$ is the Wiener measure on Ω under which the canonical process B is a standard Brownian motion, $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t\geq 0}$ is the canonical filtration and $\mathbb{F}^a = (\mathcal{F}_t^a)_{t\geq 0}$ is the $\mathbb{P}_0-$augmented filtration. Denote by $\mathcal{T}^a$ the collection of all increasing families of $\mathbb{F}^a-$stopping times $\tau = (\tau_k)_{1\leq k\leq m}$ such that the process $B_{\tau_m \wedge \cdot}$ is uniformly integrable. Recall also that Λ is the space of continuous functions on $\mathbb{R}$ with linear growth and Λ^m is its $m-$product.

For $\boldsymbol{\lambda} = (\lambda_k)_{1\leq k\leq m} \in \Lambda^m$ and $(\omega, \theta) \in \overline{\Omega}$, we denote

$$\boldsymbol{\mu}(\boldsymbol{\lambda}) := \sum_{k=1}^m \mu_k(\lambda_k) \quad \text{and} \quad \lambda(\omega, \theta) := \sum_{k=1}^m \lambda_k(\omega_{\theta_k}).$$

Then the first dual problem for the optimal SEP (IV.2.3) is given by

$$\overline{D}_0(\boldsymbol{\mu}) := \inf_{\boldsymbol{\lambda}\in\Lambda} \Big\{ \boldsymbol{\mu}(\boldsymbol{\lambda}) + \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\big[\Phi(B,\tau) - \boldsymbol{\lambda}(B,\tau)\big]\Big\}. \tag{IV.2.4}$$

As for the second dual problem, we return to the enlarged space $\overline{\Omega}$. Given $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$, an $\overline{\mathbb{F}}-$optional process $S : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ is called a strong $\overline{\mathbb{P}}-$supermartingale if

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[S_{\tau_2}\middle|\overline{\mathcal{F}}_{\tau_1}\right] \;\; \leq \;\; S_{\tau_1},\ \overline{\mathbb{P}}\text{ - a.s.}$$

for all $\overline{\mathbb{F}}-$stopping times $\tau_1 \leq \tau_2$. Let $\mathbb{L}^2_{loc}$ be the space of all $\overline{\mathbb{F}}-$progressively measurable processes $\overline{H} : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ such that

$$\int_0^t \overline{H}_s^2 ds \;\; < \;\; +\infty \text{ for all } t \in \mathbb{R}_+,\ \overline{\mathbb{P}}\text{ - a.s. for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}.$$

For $\overline{H} \in \mathbb{L}^2_{loc}$, the stochastic integral $(\overline{H} \cdot B) := \int_0^\cdot \overline{H}_t dB_t$ is well defined $\overline{\mathbb{P}}$ - almost surely for all $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$. We introduce a subset of processes :

$$\overline{\mathcal{H}} \;\; := \;\; \left\{\overline{H} \in \mathbb{L}^2_{loc} :\ (\overline{H} \cdot B)_{T_m \wedge \cdot} \text{ is a } \mathbb{P} - \text{strong supermartingale for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}\right\}.$$

Denote further

$$\overline{\mathcal{D}} \;\; := \;\; \left\{(\boldsymbol{\lambda}, \overline{H}) \in \Lambda \times \overline{\mathcal{H}} :\ \boldsymbol{\lambda}(B, T) \; + \; (\overline{H} \cdot B)_{T_m} \;\; \geq \;\; \Phi(B, T),\ \overline{\mathbb{P}}\text{ - a.s. for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}\right\},$$

and the second dual problem is given by

$$\overline{D}(\boldsymbol{\mu}) \;\; := \;\; \inf_{(\boldsymbol{\lambda}, \overline{H}) \in \overline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda}). \tag{IV.2.5}$$

Remark IV.2.3. (i) *Notice that the stochastic integral depend a priori on the probability measure under consideration. However, it follows by Nutz [80] that it can be universally defined, i.e. there exist a process* $\int_0^\cdot \overline{H}_t dB_t$ *such that*

$$\int_0^t \overline{H}_s dB_s \;\; = \;\; \left(\int_0^t \overline{H}_s dB_s\right)^{\overline{\mathbb{P}}} \textit{ for all } t \in \mathbb{R}_+,\ \overline{\mathbb{P}} \textit{ - a.s. for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}.$$

(ii) *As in the martingale optimal transport problem, the two dual problems penalize respectively different constraints of the primal problem* (IV.2.3)*. By penalizing the marginal constraints, we obtain the first dual problem* (IV.2.4)*, where a multi-period optimal stopping problem appears for every fixed* $\boldsymbol{\lambda} \in \Lambda^m$*. Then the second dual problem* (IV.2.5) *follows by the resolution of the optimal stopping problem via the Snell envelope approach and the Doob-Meyer decomposition.*

Our main duality results require the following conditions.

Assumption IV.2.4. *The function* $\Phi : \overline{\Omega} \to \mathbb{R}$ *is Borel measurable, non-anticipative, bounded* *from above, and the map $\theta \mapsto \Phi(\omega, \theta)$ is upper semicontinuous for $\mathbb{P}_0$ - almost every $\omega \in \Omega$.*

Assumption IV.2.5. *One of the following conditions holds true :*

(i) $m = 1$ *;*

(ii) $m \geq 2$ *and* $\bar{\omega} \mapsto \Phi(\bar{\omega})$ *is upper semicontinuous ;*

(iii) $m \geq 2$ *and* Φ *admits the representation*

$$\Phi(\bar{\omega}) \quad = \quad \sum_{k=1}^{m} \Phi_k(\omega, \theta_1, \cdots, \theta_k),$$

where, for each $k = 1, \cdots, m$, $\Phi_k : \Omega \times \mathbb{R}_+^k \to \mathbb{R}$ *satisfies that*

$$\Phi_k(\omega, \theta_1, \cdots, \theta_k) \quad = \quad \Phi_k(\omega_{\theta_k \wedge \cdot}, \theta_1, \cdots, \theta_k),$$

and the map $(\theta_1, \cdots, \theta_{k-1}) \mapsto \Phi_k(\omega, \theta_1, \cdots, \theta_{k-1}, \theta_k)$ *is uniformly continuous for* $0 \leq \theta_1 \leq \cdots \leq \theta_{k-1} \leq \theta_k$, *uniformly in* θ_k.

Theorem IV.2.6. (i) *Under Assumption IV.2.4, there is some* $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ *such that*

$$\mathbb{E}^{\overline{\mathbb{P}}^*}\Big[\Phi(B, T)\Big] \quad = \quad \overline{P}(\boldsymbol{\mu}) \quad = \quad \overline{D}_0(\boldsymbol{\mu}).$$

(ii) *Suppose in addition that Assumption IV.2.5 holds true, then*

$$\overline{P}(\boldsymbol{\mu}) \quad = \quad \overline{D}_0(\boldsymbol{\mu}) \quad = \quad \overline{D}(\boldsymbol{\mu}).$$

A special case : separable reward function When Φ is of the form introduced in Assumption IV.2.5 (iii), we can consider a stronger dual formulation. Denote by $\mathcal{H}^0$ the collection of all $\mathbb{F}^0-$predictable processes $H : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ such that the stochastic integral $(H \cdot B)$ is a $\mathbb{P}_0-$martingale, and $(H \cdot B)_t \geq -C(1 + |B_t|)$ for all $t \in \mathbb{R}_+$ with some constant $C > 0$.

In the filtered space $(\Omega, \mathcal{F}^0, \mathbb{P}_0, \mathbb{F}^0)$, a process X is said to be of class (DL) if for all $t \in \mathbb{R}_+$, the family $\{X_\tau : \tau \leq t \text{ is a stopping time}\}$ is uniformly integrable, and an $\mathbb{F}^0-$optional process X of class (DL) is an $\mathbb{F}^0-$supermartingale if $X_\sigma \geq \mathbb{E}^{\mathbb{P}_0}[X_\tau | \mathcal{F}^0_\sigma]$ holds for all bounded stopping times $\sigma \leq \tau$. Denote further by $\mathcal{S}^0$ the set of all $\mathbb{F}^0-$supermartingales on $(\Omega, \mathcal{F}^0, \mathbb{P}_0)$ such that $|S_t| \leq C(1 + |B_t|)$ for all $t \in \mathbb{R}_+$ with some constant $C > 0$.

Define then

$$\overline{\mathcal{D}}' := \Big\{(\boldsymbol{\lambda}, H^1, \cdots, H^m) \in \Lambda^m \times (\mathcal{H}^0)^m : \ \boldsymbol{\lambda}(\omega,\theta) \ + \ \sum_{k=1}^m \int_{\theta_{k-1}}^{\theta_k} H_t^k dB_t \ \geq \ \Phi(\omega,\theta) \\ \text{for all } \theta \in \Theta \text{ and } \mathbb{P}_0 \text{ - a.e. } \omega \in \Omega\Big\},$$

and

$$\overline{\mathcal{D}}'' := \Big\{(\boldsymbol{\lambda}, S^1, \cdots, S^m) \in \Lambda^m \times (\mathcal{S}^0)^m : \ \boldsymbol{\lambda}(\omega,\theta) \ + \ \sum_{k=1}^m \big(S_{\theta_k}^k - S_{\theta_{k-1}}^k\big) \ \geq \ \Phi(\omega,\theta) \\ \text{for all } \theta \in \Theta \text{ and } \mathbb{P}_0 \text{ - a.e. } \omega \in \Omega\Big\}.$$

Proposition IV.2.7. *Let Assumptions IV.2.4 and IV.2.5* (iii) *hold true. Assume in addition that $\Phi_k(\omega, \theta_1, \cdots, \theta_k)$ depends only on (ω, θ_k) for all $k = 1, \cdots, m$. Then*

$$\overline{P}(\boldsymbol{\mu}) \ = \ \overline{D}'(\boldsymbol{\mu}) \ := \inf_{(\boldsymbol{\lambda},H^1,\cdots,H^m)\in\overline{\mathcal{D}}'} \boldsymbol{\mu}(\boldsymbol{\lambda}) \ = \ \overline{D}''(\boldsymbol{\mu}) \ := \inf_{(\boldsymbol{\lambda},S^1,\cdots,S^m)\in\overline{\mathcal{D}}''} \boldsymbol{\mu}(\boldsymbol{\lambda}). \quad \text{(IV.2.6)}$$

2.3 More discussions and examples

Remark IV.2.8. *The dual formulation* (IV.2.5) *has been initially provided in [5] for the one-marginal case, and been proved under the condition that $\bar\omega = (\omega, \theta) \mapsto \Phi(\bar\omega)$ is bounded from above and upper semicontinuous. Indeed, the upper semicontinuity of $\bar\omega \mapsto \Phi(\bar\omega)$ can be weakened : $\theta \mapsto \Phi(\omega, \theta)$ is upper semicontinuous for $\mathbb{P}_0$ - almost every $\omega \in \Omega$. This allows to include the case where Φ is a function of the local time of the Brownian motion at zero, since the local time is continuous in θ but has no regularity in ω. In particular, when Φ is a convex function of the local time, the optimal embedding is provided by Vallois's solution, see e.g. Cox, Hobson & Obłój [23] and Claisse, Guo & Henry-Labordère [20].*

For the multi-marginal case, when Φ has no regularity in ω, we need a uniform continuity condition in time variables $(\theta_1, \cdots, \theta_{m-1})$ but not in θ_m. The uniform continuity condition is a technical condition to aggregate a family of supermartingales appearing in the classical optimal stopping problem. We can next approximate a upper semicontinuous function by a sequence of Lipschitz functions. However, to keep the non-anticipative property of Φ, we need to assume that Φ is upper semicontinuous with respect to $\bar\omega$ in Assumption IV.2.5 (ii)*. This is also the main reason for the regularity conditions in ω in Källblad, Tan & Touzi [69], where the duality result is extended to the case of infinitely-many marginals.*

Remark IV.2.9. (i) *As stated at the beginning, a characterization of the optimizers, named* *monotonicity principle, has been provided in [5]. An alternative proof of this result based on the duality result is given in Chapter V.*

(ii) *For a general function Φ, the existence of optimizers for the dual problems is not ensured. Nevertheless, when the function Φ has a particular form, we do have the dual optimizer that can be explicitly constructed. For example, Hobson [57] provided the construction when Φ is an increasing function of the running maximum, Hobson & Klimmek [58] studied forward starting straddles, Cox, Hobson & Obłój [23] considered functions on local time, see also Brown, Hobson & Rogers [17, 18], Cox & Obłój [24], Davis, Obłój & Raval[30], etc. for more concrete cases.*

(iii) *Based on the dual problem $\overline{D}_0(\boldsymbol{\mu})$, a numerical algorithm has been obtained in Bonnans & Tan [14] for the optimal SEP.*

Indeed, several financial derivatives has a reward function Φ satisfying our assumptions.

Example IV.2.10. (i) *Let $\phi : (\mathbb{R}_+ \times \mathbb{R}^3)^m \to \mathbb{R}$ be continuous and bounded from above. Denote $\overline{\omega}_t := \sup_{0\le s\le t} \omega_s$ and $\underline{\omega}_t := \inf_{0\le s\le t} \omega_s$ for all $t \in \mathbb{R}_+$. Since $\omega \mapsto (\omega_t, \overline{\omega}_t, \underline{\omega}_t)$ is continuous, the function Φ defined by*

$$\Phi(\bar{\omega}) \;\; := \;\; \phi\Big(\theta_i,\ \omega_{\theta_i},\ \overline{\omega}_{\theta_i},\ \underline{\omega}_{\theta_i},\ i = 1, \cdots, m\Big)$$

satisfies clearly Assumptions IV.2.4 and IV.2.5 (ii)*.*

(ii) *Let $L^B = (L^B_t)_{t\ge0}$ be the local time of the Brownian motion at zero. We can choose L^B to be $\mathbb{F}^0-$predictable since any $\mathbb{F}^a-$predictable process is indistinguishable to an $\mathbb{F}^0-$predictable process. Then $t \mapsto L^B_t(\omega)$ is continuous and increasing for $\mathbb{P}_0$ - almost every $\omega \in \Omega$. Let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be continuous and bounded from above, then $\Phi(\bar{\omega}) := \phi(L^B_{\theta_m}(\omega))$ satisfies Assumptions IV.2.4 and IV.2.5* (iii)*.*

The following example shows that, in general, our weak formulation is not equivalent to the strong one.

Example IV.2.11. *Let $m = 1$, $\mu = \delta_{\{0\}}/3 + \delta_{\{1\}}/3 + \delta_{\{-1\}}/3$ and $\Phi(\bar{\omega}) = \mathbb{1}_{\{0\}}(\theta)$. Define*

$$\tau_0 \;\; := \;\; \inf\Big\{t \in \mathbb{R}_+ :\ |B_t| \ \ge\ 1\Big\} \;\; \text{and} \;\; \overline{\mathbb{P}}_0 \;\; := \;\; \frac{1}{3}\mathbb{P}_0 \circ (B, 0)^{-1} \; + \; \frac{2}{3}\mathbb{P}_0 \circ (B, \tau_0)^{-1},$$

then $\overline{\mathbb{P}}_0 \in \overline{\mathcal{P}}(\mu)$ and $\mathbb{E}^{\overline{\mathbb{P}}}[\Phi(B, T)] = 1/3$. Further, for any $\tau \in \mathcal{T}^a$ such that $\mathbb{P}_0 \circ B_\tau^{-1} = \mu$, one has $\mathbb{P}_0[\tau > 0] > 0$. Since the augmented Brownian filtration satisfies Blumenthal's zero-one law, *then* $\mathbb{P}_0[\tau > 0] = 1$. *It follows that*

$$\sup_{\tau\in\mathcal{T}^a:\ \mathbb{P}_0\circ B_\tau^{-1}=\mu} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi(B,\tau)\Big] \;=\; 0 \;<\; \frac{1}{3} \;\le\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big].$$

We finally provide an example where the duality fails when Φ has no regularity in θ.

Example IV.2.12. *Let* $m = 1$, $\mu = \delta_{\{1\}}/2 + \delta_{\{-1\}}/2$ *and* $\Phi(\bar{\omega}) = \mathbb{1}_{\mathbb{Q}}(\theta)$. *We first notice that* $\overline{\mathcal{P}}(\mu)$ *has only one element, which is the probability measure induced by* (B, τ_0), *where* $\tau_0 := \inf\{t \in \mathbb{R}_+ : |B_t| \ge 1\}$. *Indeed, for any* $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)$, *one has* $\mathbb{E}^{\overline{\mathbb{P}}}[T] = \mathbb{E}^{\overline{\mathbb{P}}}[B_T^2] = \mathbb{E}^{\overline{\mathbb{P}}}[\tau_0]$ *and* $T \ge \tau_0$, $\overline{\mathbb{P}}$ *- almost surely. Moreover, since the hitting time* τ_0 *has a continuous distribution on* $\mathbb{R}_+$, *then*

$$\overline{P}(\mu) \;=\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big] \;=\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\mathbb{1}_{\mathbb{Q}}(\tau_0)\Big] \;=\; \mathbb{E}^{\mathbb{P}_0}\Big[\mathbb{1}_{\mathbb{Q}}(\tau_0)\Big] \;=\; 0.$$

As for the dual problem, we notice that $\lambda \in \Lambda$ *is a continuous function, and any stopping time can be approximated by stopping times taking value in* $\mathbb{Q}$, *then*

$$\sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\mathbb{1}_{\mathbb{Q}}(\tau) - \lambda(B_\tau)\Big] \;=\; \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[1 - \lambda(B_\tau)\Big] \text{ for all } \lambda \in \Lambda.$$

Then by its definition in (IV.2.4), $\overline{D}_0(\mu) = \inf_{\lambda\in\Lambda}\{\mu(\lambda) + \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}[1-\lambda(B_\tau)]\} = 1$. *Similarly, we can easily deduce that* $\overline{D}(\mu) = 1$ *and, it follows that*

$$\overline{P}(\mu) \;=\; 0 \;<\; 1 \;=\; \overline{D}_0(\mu) \;=\; \overline{D}(\mu).$$

3 Application to martingale optimal transport problems

In this section, we use the previous duality results of the optimal SEP to study a continuous-time martingale transport problem under multiple marginal constraints.

3.1 Martingale optimal transport on the space of continuous functions

Recall that $\mathbb{C}$ is the space of continuous funcitons $\omega = (\omega_t)_{0\le t\le 1}$ such that $\omega_0 = 0$, $X = (X_t)_{0\le t\le 1}$ is the canonical process, *i.e* $X_t(\omega) := \omega_t$ and $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ is its natural filtration, *i.e.* $\mathcal{F}_t := \sigma(X_s, s \le t)$. Recall also that $\mathcal{M}$ is the collection of all martingale measures $\mathbb{P}$, *i.e.* probability measures on $(\mathbb{C}, \mathcal{F}_1)$ under which X is a martingale. For a set of maturities

$\mathbb{T} = \{t_1 < \cdots < t_m = 1\} \subseteq (0,1]$ and a centered peacock $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m}$, define the set of martingale transport plans

$$\mathcal{M}(\boldsymbol{\mu}) \;:=\; \Big\{\mathbb{P} \in \mathcal{M}:\; \mathbb{P}\circ X_{t_k}^{-1} \;=\; \mu_k \text{ for all } k = 1,\cdots,m\Big\}.$$

For an $\mathcal{F}_1-$measurable function $\xi : \mathbb{C} \to \mathbb{R}$, the martingale transport problem is defined by

$$\mathsf{P}(\boldsymbol{\mu}) \;:=\; \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[\xi(X)\Big]. \tag{IV.3.1}$$

By Karandikar [70], there is a non-decreasing $\mathbb{F}-$progressive process $\langle X\rangle$ taking value in $[0,+\infty]$, such that $\langle X\rangle$ coincides with the quadratic variation of X, $\mathbb{P}$ - almost surely, for each martingale measure $\mathbb{P}\in\mathcal{M}$. Denote by $\mathcal{H}$ the collection of all $\mathbb{F}-$progressive processes $H : [0,1]\times\mathbb{C} \to \mathbb{R}$ such that

$$\int_0^1 H_t^2 d\langle X\rangle_t < +\infty,\; \mathbb{P}\text{ - a.s. and } (H\cdot X) := \int_0^{\cdot} H_t dX_t \text{ is a } \mathbb{P}-\text{supermartingale for all } \mathbb{P}\in\mathcal{M}.$$

Then the two dual problems are given by

$$\mathsf{D}_0(\boldsymbol{\mu}) \;:=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m}\Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) + \sup_{\mathbb{P}\in\mathcal{M}}\mathbb{E}^{\mathbb{P}}\Big[\xi(X) - \boldsymbol{\lambda}(X)\Big]\Big\} \quad\text{and}\quad \mathsf{D}(\boldsymbol{\mu}) \;:=\; \inf_{(\boldsymbol{\lambda},H)\in\mathcal{D}} \boldsymbol{\mu}(\boldsymbol{\lambda}), \tag{IV.3.2}$$

where

$$\boldsymbol{\lambda}(X) \;:=\; \sum_{k=1}^m \lambda_k(X_{t_k}), \qquad \boldsymbol{\mu}(\boldsymbol{\lambda}) \;:=\; \sum_{k=1}^m \mu_i(\lambda_i)$$

and

$$\mathcal{D} \;:=\; \Big\{(\boldsymbol{\lambda},H)\in\Lambda^m\times\mathcal{H}:\; \boldsymbol{\lambda}(X) \;+\; (H\cdot X)_1 \;\ge\; \xi(X),\; \mathbb{P}\text{ - a.s. for all } \mathbb{P}\in\mathcal{M}\Big\}.$$

It is easy to check that the weak dualities hold :

$$\mathsf{P}(\boldsymbol{\mu}) \;\le\; \mathsf{D}_0(\boldsymbol{\mu}) \;\le\; \mathsf{D}(\boldsymbol{\mu}). \tag{IV.3.3}$$

3.2 Duality results

Using the duality results in Theorem IV.2.6, we can establish the dualities for the above martingale transport problem. Denote $\langle X\rangle_t^{-1} := \inf\{s\in\mathbb{R}_+ : \langle X\rangle_s > t\}\wedge 1$ and

$$W_t \;:=\; X_{\langle X\rangle_t^{-1}}\mathbb{1}_{\{t<\langle X\rangle_1\}} \;+\; \Big(X_1 \;+\; \widehat{W}_{t-\langle X\rangle_1}\Big)\,\mathbb{1}_{\{t\geq\langle X\rangle_1\}}, \tag{IV.3.4}$$

where $\widehat{W}$ is an $\mathbb{F}-$independent Brownian motion [1]. Then it follows from the Dambis-Dubins-Schwarz theorem, see *e.g.* Revuz & Yor [92, Theorem 1.7, Chapter V], that the process W is a Brownian motion. Put $W(X) := (X_{\langle X\rangle_t^{-1}})_{t\geq 0}$, then we have the following theorem.

Theorem IV.3.1. *Assume that the function ξ admits the representation*

$$\xi(X) \;=\; \Phi\Big(W(X), \langle X\rangle_{t_1}, \cdots, \langle X\rangle_{t_m}\Big)$$

for some $\Phi : \overline{\Omega}\to\mathbb{R}$ satisfying Assumptions IV.2.4 and IV.2.5. Then

$$\mathsf{P}(\boldsymbol{\mu}) \;=\; \mathsf{D}_0(\boldsymbol{\mu}) \;=\; \mathsf{D}(\boldsymbol{\mu}).$$

Example IV.3.2. *Let $\phi : (\mathbb{R}_+\times\mathbb{R}^3)^m \to \mathbb{R}$ be continuous and bounded from above. Define $\xi : \mathbf{C}\to\mathbb{R}$ by*

$$\xi(X) \;:=\; \phi\Big(\langle X\rangle_{t_i},\; X_{t_i},\; \overline{X}_{t_i},\; \underline{X}_{t_i},\; i=1,\cdots m\Big), \tag{IV.3.5}$$

where $\overline{X}_t := \sup_{0\leq s\leq t} X_s$ and $\underline{X}_t := \inf_{0\leq s\leq t} X_s$ for all $t\in[0,1]$. Then it follows by a time-change from ξ to Φ defined by

$$\Phi(\bar{\omega}) \;:=\; \phi\Big(\theta_i,\; \omega_{\theta_i},\; \overline{\omega}_{\theta_i},\; \underline{\omega}_{\theta_i},\; i=1,\cdots,m\Big),$$

and thus ξ satisfies the conditions in Theorem IV.3.1. The form (IV.3.5) contains a big class of derivatives such as lookback option, barrier options and variance options.

Proof of Theorem IV.3.1. Combining the dualities $\overline{P}(\boldsymbol{\mu}) = \overline{D}_0(\boldsymbol{\mu}) = \overline{D}(\boldsymbol{\mu})$ in Theorem IV.2.6 and the weak dualities $\mathsf{P}(\boldsymbol{\mu}) \leq \mathsf{D}_0(\boldsymbol{\mu}) \leq \mathsf{D}(\boldsymbol{\mu})$, it is enough to prove

$$\overline{P}(\boldsymbol{\mu}) \;\leq\; \mathsf{P}(\boldsymbol{\mu}) \quad\text{and}\quad \overline{D}(\boldsymbol{\mu}) \;\geq\; \mathsf{D}(\boldsymbol{\mu}),$$

1. In general case, one needs to enlarge the space to obtain an independent Brownian motion $\widehat{W}$. However, in the following we will always consider a non-anticipative functional $\Phi(W_{\langle X\rangle_{t_m}\wedge\cdot}, \langle X\rangle_{t_1}, \cdots, \langle X\rangle_{t_m})$ for $0 < t_1\leq\cdots\leq t_m = 1$, which does not really depend on $\widehat{W}$.

where $\overline{P}(\boldsymbol{\mu})$ and $\overline{D}(\boldsymbol{\mu})$ are defined respectively in (IV.2.3) and (IV.2.5) relative to Φ.

(i) Define the process $M := (M_t)_{0\le t\le 1}$ by

$$M_t \;:=\; B_{\left(T_k+\frac{t-t_k}{t_{k+1}-t}\right)\wedge T_{k+1}} \text{ for all } t\in[t_k,t_{k+1}) \text{ and } 0\le k\le m-1,$$

with $T_0 = t_0 = 0$ and $M_1 = B_{T_m}$. It is clear that M is a continuous martingale under every probability $\overline{\mathbb{P}}\in\overline{\mathcal{P}}$ and $M_{t_k} = B_{T_k}$ for all $k = 1,\cdots,m$, which implies in particular $\overline{\mathbb{P}}\circ M_{t_k}^{-1} = \mu_k$ for every $\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\boldsymbol{\mu})$. Let $\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\boldsymbol{\mu})$ be arbitrary, then $\mathbb{P} := \overline{\mathbb{P}}\circ M^{-1}\in\mathcal{M}(\boldsymbol{\mu})$. Moreover, one finds $\overline{\mathbb{P}}$ - almost surely, $\langle M\rangle_{t_k} = T_k$ for all $k = 1,\cdots,m$ and $B_t = M_{\langle M\rangle_t^{-1}}$, which yields

$$\xi(M) \;=\; \Phi\left(B,\langle M\rangle_{t_1},...,\langle M\rangle_{t_m}\right) \;=\; \Phi(B,T),\quad \overline{\mathbb{P}}\text{ - a.s.}$$

Thus

$$\mathsf{P}(\boldsymbol{\mu}) \;\ge\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\xi(M)\Big] \;=\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big]. \tag{IV.3.6}$$

It follows that

$$\overline{P}(\boldsymbol{\mu}) \;\le\; \mathsf{P}(\boldsymbol{\mu}).$$

(ii) Let us now prove $\mathsf{D}(\boldsymbol{\mu})\le\overline{D}(\boldsymbol{\mu})$. Let $(\boldsymbol{\lambda},\overline{H})\in\overline{\mathcal{D}}$, *i.e.* $(\boldsymbol{\lambda},\overline{H})\in\Lambda^m\times\overline{\mathcal{H}}$ such that

$$\boldsymbol{\lambda}(B,T) \;+\; (\overline{H}\cdot B)_{T_m} \;\ge\; \Phi(B,T),\quad \overline{\mathbb{P}}\text{ - a.s. for all } \overline{\mathbb{P}}\in\overline{\mathcal{P}}.$$

For every $\mathbb{P}\in\mathcal{M}$, it follows by Dambis-Dubins-Schwarz theorem that the time-changed process W defined in (IV.3.4) is a Brownian motion with respect to the time-changed filtration $(\mathcal{F}_{\langle X\rangle_t^{-1}})_{t\ge 0}$ under $\mathbb{P}$ and

$$X_t \;=\; W_{\langle X\rangle_t} \text{ for all } t\in[0,1],\ \mathbb{P}\text{ - a.s.}$$

Moreover, $\langle X\rangle_{\mathbb{T}} := (\langle X\rangle_{t_k})_{1\le k\le m}$ are stopping times with respect to $(\mathcal{F}_{\langle X\rangle_t^{-1}})_{t\ge 0}$. Let us define $\overline{\mathbb{P}} := \mathbb{P}\circ(W,\langle X\rangle_{t_1},...,\langle X\rangle_{t_m})^{-1}$, then $\overline{\mathbb{P}}\in\overline{\mathcal{P}}$ and thus we have $\mathbb{P}$ - almost surely,

$$\boldsymbol{\lambda}(W,\langle X\rangle_{\mathbb{T}}) \;+\; (\overline{H}\cdot W)_{\langle X\rangle_1} \;\ge\; \Phi(W,\langle X\rangle_{t_1},...,\langle X\rangle_{t_m}).$$

Define

$$H_t(X) \;:=\; \overline{H}_{\langle X\rangle_t}\Big(W, \langle X\rangle_{t_1}, ..., \langle X\rangle_{t_m}\Big) \text{ for all } t\in[0,1],$$

then it follows by Propositions V.1.4 and V.1.5 of Revuz & Yor [92] that H is $\mathbb{F}-$progressively measurable such that

$$\int_0^1 H_t^2 d\langle X\rangle_t \;=\; \int_0^{\langle X\rangle_1} \overline{H}_t^2 dt \;<\; +\infty,\; \mathbb{P}\text{ - a.s.},$$

and

$$\Big(\overline{H}\cdot W\Big)_{\langle X\rangle_t} \;=\; (H\cdot X)_t \text{ for all } t\in[0,1],\; \mathbb{P}\text{ - a.s.}$$

Hence

$$\boldsymbol{\lambda}(X) \;+\; (H\cdot X)_1 \;\geq\; \Phi\Big(W, \langle X\rangle_{t_1}, ..., \langle X\rangle_{t_m}\Big) \;=\; \xi(X),\; \mathbb{P}\text{ - a.s.} \tag{IV.3.7}$$

Notice that $(\overline{H}\cdot W)_{T_m\wedge\cdot}$ is a strong $\mathbb{P}-$supermartingale, which implies by the time-change argument that $(H\cdot X)$ is a $\mathbb{P}-$supermartingale. Hence $H\in\mathcal{H}$ and further $(\boldsymbol{\lambda}, H)\in\mathcal{D}$. It follows that $\mathsf{D}(\boldsymbol{\mu})\leq\overline{D}(\boldsymbol{\mu})$, which concludes the proof. □

4 Proof of Theorem IV.2.6

To prove our main results in Theorem IV.2.6, we start with some technical lemmas in Section 4.1. Then in Section 4.2, we show the existence of an optimizer $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ and the first duality $\overline{P}(\boldsymbol{\mu}) = \overline{D}_0(\boldsymbol{\mu})$ in Theorem IV.2.6, where the main argument is the compactness and the Fenchel-Moreau theorem.

Finally, in Section 4.3, we complete the proofs for the second duality $\overline{P}(\boldsymbol{\mu}) = \overline{D}(\boldsymbol{\mu})$ in Theorem IV.2.6 and $\overline{P}(\boldsymbol{\mu}) = \overline{D}'(\boldsymbol{\mu}) = \overline{D}''(\boldsymbol{\mu})$ in Proposition IV.2.7, using classical results of the optimal stopping theory. The superhedging strategy in the second dual formulation can be obtained directly from the Doob-Meyer decomposition of the Snell envelop of the stopping problem in $\overline{D}_0(\boldsymbol{\mu})$, and the martingale representation theorem. The reasoning is better illustrated in the proof of Proposition IV.2.7.

4.1 Technical lemmas

Recall that $\overline{\mathcal{P}}(\overline{\Omega})$ is the space of probability measures on the Polish space $\overline{\Omega}$. Denote by $C_b(\overline{\Omega})$ the collection of bounded continuous functions on $\overline{\Omega}$, and by $B_{mc}(\overline{\Omega})$ the collection of bounded measurable function ϕ, such that $\theta \mapsto \phi(\omega,\theta)$ is continuous for all $\omega \in \Omega$. Notice that the weak convergence topology on $\overline{\mathcal{P}}(\overline{\Omega})$ is defined as the coarsest topology under which $\overline{\mathbb{P}} \mapsto \mathbb{E}^{\overline{\mathbb{P}}}[\zeta]$ is continuous for all $\zeta \in C_b(\overline{\Omega})$. Following Jacod & Mémin [65], we introduce the stable convergence topology on $\overline{\mathcal{P}}(\overline{\Omega})$ as the coarsest topology under which $\overline{\mathbb{P}} \mapsto \mathbb{E}^{\overline{\mathbb{P}}}[\zeta]$ is continuous for all $\zeta \in B_{mc}(\overline{\Omega})$. Recall that every probability measure in $\overline{\mathcal{P}}$ has the same marginal law on Ω. Then as an immediate consequence of Proposition 2.4 of [65], we have the following result.

Lemma IV.4.1. *The weak convergence topology and the stable convergence topology coincide on the space $\overline{\mathcal{P}}$.*

Let $\mathbf{P}_0^{\preceq}$ be the collection of all centered peacocks $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m}$ on $\mathbb{R}$. We say, a sequence of centered peacocks $(\boldsymbol{\mu}^n)_{n\ge1} \subseteq \mathbf{P}_0^{\preceq}$ converges to $\boldsymbol{\mu}^0 \in \mathbf{P}_0^{\preceq}$ if μ_k^n converges to μ_k^0 under the Wasserstein metric $\mathcal{W}_1$ for all $k = 1,\cdots,m$. This convergence is denoted by $\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^m} \boldsymbol{\mu}^0$.

Lemma IV.4.2. *Let $(\boldsymbol{\mu}^n)_{n\ge1}$ be a sequence of centered peacocks such that $\boldsymbol{\mu}^n \xrightarrow{\mathcal{W}_1^m} \boldsymbol{\mu}^0$, and $(\overline{\mathbb{P}}_n)_{n\ge1}$ be a sequence of probability measures with $\overline{\mathbb{P}}_n \in \overline{\mathcal{P}}(\boldsymbol{\mu}^n)$ for all $n \ge 1$. Then $(\overline{\mathbb{P}}_n)_{n\ge1}$ is relatively compact under the weak convergence topology. Moreover, any limit point of $(\overline{\mathbb{P}}_n)_{n\ge1}$ belongs to $\overline{\mathcal{P}}(\boldsymbol{\mu}^0)$.*

Proof. (i) For any $\varepsilon > 0$, there exists a compact set $D \subseteq \Omega$ such that $\overline{\mathbb{P}}_n[D\times\Theta] = \mathbb{P}_0[D] \ge 1-\varepsilon$ for all $n \ge 1$. In addition, by Proposition 7 of Monroe [79], one has for any constant $C > 0$,

$$\overline{\mathbb{P}}_n\Big[T_m \ge C\Big] \;\le\; C^{-1/3}\Big(1 \;+\; \big(\mu_m^n(|x|)\big)^2\Big) \;\le\; C^{-1/3}\Big(1 \;+\; \big(\sup_{n\ge1}\; \mu_m^n(|x|)\big)^2\Big).$$

Choose the cube $[0,C]^m$ large enough such that $\overline{\mathbb{P}}_n\Big[T \in [0,C]^m\Big] \ge 1-\varepsilon$ for all $n \ge 1$. The tightness of $(\overline{\mathbb{P}}_n)_{n\ge1}$ follows by

$$\overline{\mathbb{P}}_n\Big[D\times[0,C]^m\Big] \;\ge\; \overline{\mathbb{P}}_n\Big[D\times\Theta\Big] \;+\; \overline{\mathbb{P}}_n\Big[\Omega\times[0,C]^m\Big] \;-\; 1 \;\ge\; 1 \;-\; 2\varepsilon \text{ for all } n\ge1.$$

It follows by Prokhorov's theorem that the tightness yields the relative compactness.

(ii) Let $\overline{\mathbb{P}}_0$ be any limit point. By possibly subtracting a subsequence, we assume without loss of generality that $\overline{\mathbb{P}}_n \to \overline{\mathbb{P}}_0$. Notice that B is an $\overline{\mathbb{F}}-$Brownian motion under each $\overline{\mathbb{P}}_n$ and thus the process $\varphi(B_t) - \int_0^t \frac{1}{2}\varphi''(B_s)ds$ is an $\overline{\mathbb{F}}-$martingale under $\overline{\mathbb{P}}_n$ whenever φ is bounded, smooth and of bounded derivatives. Notice that the map $(\omega, t) \mapsto \varphi(\omega_t) - \int_0^t \varphi''(\omega_s)ds$ is also bounded continuous, then

$$\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Big(\varphi(B_t) \;-\; \varphi(B_r) \;-\; \int_r^t \frac{1}{2}\varphi''(B_u)du\Big)\;\psi\Big] \;=\; 0,$$

for every $s < r < t$ and $\overline{\mathcal{F}}_r-$measurable random variable ψ that is bounded and continuous. Taking the limit $n \to \infty$, it follows that

$$\mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\Big(\varphi(B_t) \;-\; \varphi(B_r) \;-\; \int_r^t \frac{1}{2}\varphi''(B_u)du\Big)\;\psi\Big] \;=\; 0, \tag{IV.4.1}$$

for all $\overline{\mathcal{F}}_r-$measurable random variables ψ that are bounded and continuous. Since $\overline{\mathcal{F}}_s \subseteq \overline{\mathcal{F}}_{r-}$, where $\overline{\mathcal{F}}_{r-}$ is generated by the class of all $\overline{\mathcal{F}}_r-$measurable random variables that are bounded and continuous, see Lemma IV.5.1, it follows that (IV.4.1) is still true for every bounded and $\overline{\mathcal{F}}_s-$measurable ψ. Letting $r \to s$, by the dominated convergence theorem, it follows that (IV.4.1) holds for every $s < t$ and bounded $\overline{\mathcal{F}}_s-$measurable random variable ψ. This implies that B is an $\overline{\mathbb{F}}-$Brownian motion under $\overline{\mathbb{P}}_0$.

(iii) We next assume that $\overline{\mathbb{P}}_n \in \overline{\mathcal{P}}(\boldsymbol{\mu}^n)$ and prove

$$B_{T_m \wedge \cdot} \text{ is uniformly integrable under } \overline{\mathbb{P}}_0. \tag{IV.4.2}$$

The convergence of $(\boldsymbol{\mu}^n)_{n\geq 1}$ to $\boldsymbol{\mu}^0$ implies in particular

$$\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Big(\Big|B_{T_m}\Big| - R\Big)^+\Big] \;=\; \mu_m^n\Big((|x| - R)^+\Big) \;\longrightarrow\; \mu_m^0\Big((|x| - R)^+\Big) \;<\; +\infty.$$

Therefore, for every $\varepsilon > 0$, there exists $R_\varepsilon > 0$ large enough such that $\mu_m^n\Big((|x| - R_\varepsilon)^+\Big) < \varepsilon$ for all $n \geq 1$. It follows by Jensen's inequality and $|x|\mathbb{1}_{\{|x|>2R\}} \leq 2(|x| - R)^+$ that

$$\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Big|B_{T_m\wedge t}\Big|\mathbb{1}_{\{|B_{T_m\wedge t}|>2R_\varepsilon\}}\Big] \;\leq\; 2\mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Big(\Big|B_{T_m}\Big| - R_\varepsilon\Big)^+\Big] \;\leq\; 2\varepsilon \;\text{ for all } t \geq 0.$$

Notice also that the function $|x|\mathbb{1}_{\{|x|>2R_\varepsilon\}}$ is lower semicontinuous and we obtain by Fatou's lemma

$$\mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\Big|B_{T_m\wedge t}\Big|\mathbb{1}_{\{|B_{T_m\wedge t}|>2R_\varepsilon\}}\Big] \;\leq\; \liminf_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Big|B_{T_m\wedge t}\Big|\mathbb{1}_{\{|B_{T_m\wedge t}|>2R_\varepsilon\}}\Big] \;\leq\; 2\varepsilon,$$

which justifies the claim (IV.4.2). Moreover, since the map $(\omega, \theta) \mapsto \omega_{\theta_k}$ is continuous, it follows that $\overline{\mathbb{P}}_0 \circ B_{T_k}^{-1} = \mu_k^0$ for all $k = 1, \cdots, m$. Therefore, $\overline{\mathbb{P}}_0 \in \overline{\mathcal{P}}(\boldsymbol{\mu}^0)$, which concludes the proof. □

4.2 Proof of $\overline{P} = \overline{D}_0$

We now provide the proof of the first duality result in Theorem IV.2.6. We proceed as same as in Section 4 of Chapter III : show that $\boldsymbol{\mu} \mapsto \overline{P}(\boldsymbol{\mu})$ is concave and upper semicontinuous and then use the Fenchel-Moreau theorem.

Lemma IV.4.3. *Under Assumption IV.2.4, the map $\boldsymbol{\mu} \in \mathbf{P}_0^{\preceq} \mapsto \overline{P}(\boldsymbol{\mu}) \in \mathbb{R}$ is concave and upper semicontinuous under $\mathcal{W}_1^m$. Moreover, for every $\boldsymbol{\mu} \in \mathbf{P}_0^{\preceq}$, there is some $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ such that $\mathbb{E}^{\overline{\mathbb{P}}^*}[\Phi] = \overline{P}(\boldsymbol{\mu})$.*

Proof. (i) Let $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathbf{P}_0^{\preceq}$ and $\alpha \in (0,1)$. Take $\overline{\mathbb{P}}_1 \in \overline{\mathcal{P}}(\boldsymbol{\mu}^1)$ and $\overline{\mathbb{P}}_2 \in \overline{\mathcal{P}}(\boldsymbol{\mu}^2)$, then by their definition, one has $\alpha\overline{\mathbb{P}}_1 + (1-\alpha)\overline{\mathbb{P}}_2 \in \overline{\mathcal{P}}(\alpha\boldsymbol{\mu}^1 + (1-\alpha)\boldsymbol{\mu}^2)$. It follows immediately that the map $\boldsymbol{\mu} \mapsto \overline{P}(\boldsymbol{\mu})$ is concave.

(ii) Let $(\boldsymbol{\mu}^n)_{n\geq 1} \subseteq \mathbf{P}_0^{\preceq}$ and $\boldsymbol{\mu}^n \to \boldsymbol{\mu}^0 \in \mathbf{P}_0^{\preceq}$ under $\mathcal{W}_1^m$. After possibly passing to a subsequence, we can get a sequence $(\overline{\mathbb{P}}_n)_{n\geq 1}$ such that

$$\overline{\mathbb{P}}_n \;\in\; \overline{\mathcal{P}}(\boldsymbol{\mu}^n) \;\text{ and }\; \limsup_{n\to\infty} \overline{P}(\boldsymbol{\mu}^n) \;=\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Phi(B,T)\Big].$$

By Lemma IV.4.2, we may find a subsequence still denoted by $(\overline{\mathbb{P}}_n)_{n\geq 1}$, which converges weakly to some $\overline{\mathbb{P}}_0 \in \overline{\mathcal{P}}(\boldsymbol{\mu}^0)$. Moreover, it follows by Lemma IV.4.1 that the map $\overline{\mathbb{P}} \mapsto \mathbb{E}^{\overline{\mathbb{P}}}[\Phi(B,T)]$ is upper semicontinuous on $\overline{\mathcal{P}}$ with respect to the weak convergence topology under Assumption IV.2.4. We then obtain by Fatou's lemma that

$$\limsup_{n\to\infty} \overline{P}(\boldsymbol{\mu}^n) \;=\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Phi(B,T)\Big] \;\leq\; \mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\Phi(B,T)\Big] \;\leq\; \overline{P}(\boldsymbol{\mu}^0).$$

(iii) Let $\boldsymbol{\mu} \in \mathbf{P}_0^{\preceq}$, taking $\boldsymbol{\mu}^n = \boldsymbol{\mu}$ and repeating the same reasoning, it follows immediately that there is some $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ such that $\mathbb{E}^{\overline{\mathbb{P}}^*}[\Phi] = \overline{P}(\boldsymbol{\mu})$. □

The results in Lemma IV.4.3 together with the Fenchel-Moreau theorem implies the first duality in Theorem IV.2.6. Before providing the proof, we consider the optimal stopping problem arising in the dual formulation (IV.2.4). Denote for every $\boldsymbol{\lambda} \in \Lambda^m$,

$$\Phi^{\boldsymbol{\lambda}}(\omega,\theta) \;:=\; \Phi(\omega,\theta) \;-\; \boldsymbol{\lambda}(\omega,\theta) \text{ for all } (\omega,\theta) \in \overline{\Omega}. \tag{IV.4.3}$$

Recall that $\mathcal{T}^a$ denotes the collection of all increasing families of $\mathbb{F}^a$−stopping times $\tau = (\tau_k)_{1\leq k\leq m}$ such that $B_{\tau_m\wedge\cdot}$ is uniformly integrable. Given $N > 0$, denote by $\mathcal{T}^a_N \subseteq \mathcal{T}^a$ the subset of families $\tau = (\tau_k)_{1\leq k\leq m}$ such that $\tau_m \leq N$, $\mathbb{P}_0$ - almost surely. Denote further by $\overline{\mathcal{P}}_N \subseteq \overline{\mathcal{P}}$ the subset of $\overline{\mathbb{P}}$ such that $T_m \leq N$, $\overline{\mathbb{P}}$ - almost surely.

Lemma IV.4.4. *Let Φ be bounded, then for every $\boldsymbol{\lambda} \in \Lambda$,*

$$\begin{aligned}\sup_{\tau \in \mathfrak{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] &= \lim_{N\to\infty} \sup_{\tau\in\mathfrak{T}^a_N} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] \\ &= \lim_{N\to\infty} \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_N} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big] = \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big].\end{aligned} \tag{IV.4.4}$$

In particular, let $\lambda \in \Lambda$ and denote by λ^{conc} its concave envelope, one has

$$\sup_{\tau\in\mathfrak{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\lambda(B_{\tau_m})\Big] = \lambda^{conc}(0). \tag{IV.4.5}$$

Proof. (i) Given $\boldsymbol{\lambda} \in \Lambda$, there is some constant $C > 0$ such that

$$\Big|\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big| \le C\Big(1 + \sum_{k=1}^{m} \Big|B_{\tau_k}\Big|\Big). \tag{IV.4.6}$$

Let $\tau \in \mathfrak{T}^a$, define $\tau^N := (\tau^N_k)_{1\le k\le m}$ with $\tau^N_k := \tau_k \wedge N$, then it is clear that $\lim_{N\to\infty} \Phi^{\boldsymbol{\lambda}}(B,\tau^N) = \Phi^{\boldsymbol{\lambda}}(B,\tau)$, $\mathbb{P}_0$ - almost surely. By the domination in (IV.4.6) and the fact that $B_{\tau_m\wedge\cdot}$ is uniformly integrable, we have $\lim_{N\to\infty} \mathbb{E}^{\mathbb{P}_0}[\Phi^{\boldsymbol{\lambda}}(B,\tau^N)] = \mathbb{E}^{\mathbb{P}_0}[\Phi^{\boldsymbol{\lambda}}(B,\tau)]$. It follows by the arbitrariness of $\tau \in \mathfrak{T}^a$ and the fact $\mathfrak{T}^a_N \subseteq \mathfrak{T}^a$ that

$$\sup_{\tau\in\mathfrak{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] = \lim_{N\to\infty} \sup_{\tau\in\mathfrak{T}^a_N} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big].$$

By the same arguments, it is clear that we also have

$$\sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big] = \lim_{N\to\infty} \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_N} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big].$$

(ii) We now apply Lemma IV.5.7 to prove that for every fixed constant $N > 0$,

$$\sup_{\tau\in\mathfrak{T}^a_N} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] = \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_N} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big]. \tag{IV.4.7}$$

First, we assume $m = 1$. Let $\overline{\mathbb{P}} \in \overline{\mathcal{P}}_N$, denote $Y_t := \Phi^{\lambda}(B, t\wedge N)$, it is clear that

$$\mathbb{E}^{\overline{\mathbb{P}}}\Big[\sup_{t\ge 0} Y_t\Big] < +\infty.$$

Recall that $\overline{\mathbb{F}}^B$ is the filtration generated by B on $\overline{\Omega}$. Denote by $\overline{\mathbb{F}}^{\overline{\mathbb{P}}} = (\overline{\mathcal{F}}^{\overline{\mathbb{P}}}_t)_{t\ge 0}$ the augmented filtration of $\overline{\mathbb{F}}$ under $\overline{\mathbb{P}}$, and by $\overline{\mathbb{F}}^{B,\overline{\mathbb{P}}} = (\overline{\mathcal{F}}^{B,\overline{\mathbb{P}}}_t)_{t\ge 0}$ the augmented filtration of $\overline{\mathbb{F}}^B$ under $\overline{\mathbb{P}}$. It is clear that $\overline{\mathcal{F}}_t^{B,\overline{\mathbb{P}}} \subseteq \overline{\mathcal{F}}_t^{\overline{\mathbb{P}}}$ for all $t \in \mathbb{R}_+$. More importantly, by the fact that B is an $\overline{\mathbb{F}}^{\overline{\mathbb{P}}}-$Brownian motion under $\overline{\mathbb{P}}$, it is easy to check that the probability space $(\overline{\Omega}, \overline{\mathcal{F}}^{\overline{\mathbb{P}}}, \overline{\mathbb{P}})$ together with the filtration $\overline{\mathbb{F}}^{\overline{\mathbb{P}}}$ and $\overline{\mathbb{F}}^{B,\overline{\mathbb{P}}}$ satisfies Assumption IV.5.6, where $\overline{\mathcal{F}}^{\overline{\mathbb{P}}} := \bigvee_{t\geq 0} \overline{\mathcal{F}}_t^{\overline{\mathbb{P}}}$. Then by Lemma IV.5.7, $\mathbb{E}^{\overline{\mathbb{P}}}[Y_T] \leq \sup_{\tau\in\mathcal{T}_N^a} \mathbb{E}^{\mathbb{P}_0}[Y_\tau]$ and hence $\sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}_N} \mathbb{E}^{\overline{\mathbb{P}}}[\Phi^\lambda(B,T)] \leq \sup_{\tau\in\mathcal{T}_N^a} \mathbb{E}^{\mathbb{P}_0}[\Phi^\lambda(B,\tau)]$. We then have equality (IV.4.7) since the inverse inequality is clear.

Finally, when $m > 1$, it is enough to use the same arguments together with the induction to prove (IV.4.7).

(iii) To prove (IV.4.5) it suffices to set $\Phi \equiv 0$. Then by (IV.4.4), it follows that

$$\sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\lambda(B_{\tau_m})\Big] \;=\; \lim_{N\to\infty} \sup_{\tau\in\mathcal{T}_N^a} \mathbb{E}^{\mathbb{P}_0}\Big[\lambda(B_{\tau_m})\Big] \;\leq\; \phi^{conc}(0).$$

The inverse inequality is obvious by considering the exiting time of the Brownian motion from an open interval. We hence conclude the proof. □

We prove Theorem IV.2.6 by exactly the same reasoning of Theorem III.3.12. Recall that $\mathbf{M}$ is the space of finite signed measures μ on $\mathbb{R}$ such that

$$\int_{\mathbb{R}} \Big(1 + |x|\Big)|\mu|(dx) \;<\; +\infty.$$

Then it follows by the Fenchel-Moreau theorem that, the dual space of $\mathbf{M}^m$ is Λ^m.

Proof of Theorem IV.2.6. (i) The existence of optimizers is already proved in Lemma IV.4.3. For the first duality result, we shall use the Fenchel-Moreau theorem. Let us first extend the map $\boldsymbol{\mu} \mapsto \overline{P}(\boldsymbol{\mu})$ from $\mathbf{P}_0^{\preceq}$ to $\mathbf{M}^m$ by setting that $\overline{P}(\boldsymbol{\mu}) = -\infty$, for every $\boldsymbol{\mu} \in \mathbf{M}^m \setminus \mathbf{P}_0^{\preceq}$. It is easy to check, using Lemma IV.4.3, that the extended map $\boldsymbol{\mu} \mapsto \overline{P}(\boldsymbol{\mu})$ on $\mathbf{M}^m$ to $\mathbb{R}$ is still concave and upper semicontinuous. Then it follows from the Fenchel-Moreau theorem that

$$\begin{aligned}
\overline{P}(\boldsymbol{\mu}) \;=\; \overline{P}^{**}(\boldsymbol{\mu}) \;&=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\boldsymbol{\nu}\in\mathbf{P}^{\preceq}} \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\boldsymbol{\nu})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big]\Big\} \\
&=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi^{\boldsymbol{\lambda}}(B,T)\Big]\Big\} \\
&=\; \inf_{\boldsymbol{\lambda}\in\Lambda^m} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}\Big(B,\tau\Big)\Big]\Big\},
\end{aligned}$$

where the last equality follows by (IV.4.4). Hence we have $\overline{P}(\boldsymbol{\mu}) = \overline{D}_0(\boldsymbol{\mu})$. □

Remark IV.4.5. *Assume further that Φ is bounded, then we may show that*

$$\overline{D}_0(\boldsymbol{\mu}) \;=\; \inf_{\boldsymbol{\lambda}\in\Lambda^+} \Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big]\Big\}, \tag{IV.4.8}$$

where

$$\Lambda^+ \ := \ \Big\{\boldsymbol{\lambda} = (\lambda_k)_{1\le k\le m} \in \Lambda^m : \ \lambda_k \ge 0 \ \textit{for all } k = 1, \cdots, m\Big\}.$$

Indeed, using (IV.4.5)*, it is easy to see that in the definition of* $\overline{D}_0(\boldsymbol{\mu})$*, it is enough to take the infimum over the class of all functions* $\boldsymbol{\lambda} \in \Lambda^+$ *such that the convex envelope* $\lambda_k^{conv}(0) > -\infty$ *for all* $k = 1, \cdots, m$*, since by* (IV.4.5) *and the boundedness of* Φ*,* $\sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}[\Phi^{\lambda}] = +\infty$ *whenver* $(-\lambda_k)^{conc}(0) = +\infty$ *for some* k*. Hence the infimum is taken among all* $\boldsymbol{\lambda} \in \Lambda$ *such that* $\lambda_k^{conv}(0) > -\infty$ *for all* $k = 1, \cdots, m$*, and consequently* λ_k *is dominated from below by some affine function. Since* $\mathbb{E}^{\mathbb{P}_0}[B_{\tau_k}] = 0$ *for all* $\tau \in \mathcal{T}^a$*, then by possibly subtracting from* λ_k *some affine function, it is enough to take infimum over the class* Λ^+*.*

4.3 Proof of $\overline{D}_0 = \overline{D}$

We now prove the second duality $\overline{P}(\boldsymbol{\mu}) = \overline{D}(\boldsymbol{\mu})$ in Theorem IV.2.6 (ii), and $\overline{P}(\boldsymbol{\mu}) = \overline{D}'(\boldsymbol{\mu}) = \overline{D}''(\boldsymbol{\mu})$ in Proposition IV.2.7. The main technique is to use the Snell envelope characterization of the optimal stopping problem, together with the Doob-Meyer decomposition. We will provide the proof progressively. In Section 4.3.1, we prove a weak duality result in the context of Theorem IV.2.6 and Proposition IV.2.7. Then in Section 4.3.2, we show that it is enough to prove Theorem IV.2.6 (ii) and Proposition IV.2.7, assuming that Φ is bounded. Next, in Section 4.3.3, we provide the proof of Proposition IV.2.7, which implies immediately Theorem IV.2.6 (ii) under Assumptions IV.2.4 and IV.2.5 (i). Finally, we complete the proof of Theorem IV.2.6 (ii) under Assumptions IV.2.4 and IV.2.5 (ii) or (iii) in Sections 4.3.5 and 4.3.6.

Throughout this section, we say that a process X, defined on the filtered space $(\Omega, \mathcal{F}^0, \mathbb{P}_0, \mathbb{F}^a)$, is of class (DL) if for every $t \in \mathbb{R}_+$, the family $\{X_\tau : \tau \text{ is an } \mathbb{F}^a - \text{stopping time with } \tau \le t\}$ is uniformly integrable; an $\mathbb{F}^a-$optional process X of class (DL) is a supermartingale if $X_\sigma \ge \mathbb{E}^{\mathbb{P}_0}[X_\tau|\mathcal{F}^a_\sigma]$ holds for all bounded stopping times $\sigma \le \tau$.

4.3.1 Weak dualities

In the context of Theorem IV.2.6 and Proposition IV.2.7, we can easily deduce the weak dualities by definition.

Lemma IV.4.6. *Let* $\Phi : \overline{\Omega} \to \mathbb{R}$ *be Borel measurable and non-anticipative, then*

$$\overline{P}(\boldsymbol{\mu}) \ \le \ \overline{D}(\boldsymbol{\mu}), \quad \overline{P}(\boldsymbol{\mu}) \ \le \ \overline{D}'(\boldsymbol{\mu}) \ \ \textit{and} \ \ \overline{P}(\boldsymbol{\mu}) \ \le \ \overline{D}''(\boldsymbol{\mu}).$$

Proof. (i) For any $(\boldsymbol{\lambda}, \overline{H}) \in \overline{\mathcal{D}}$, one has by definition

$$\boldsymbol{\lambda}(B,T) \;+\; (\overline{H}\cdot B)_{T_m} \;\geq\; \Phi(B,T),\; \overline{\mathbb{P}}\text{ - a.s. for all } \overline{\mathbb{P}} \in \overline{\mathcal{P}}.$$

For any $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$, we obtain

$$\boldsymbol{\mu}(\boldsymbol{\lambda}) \;=\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\boldsymbol{\lambda}(B,T)\Big] \;\geq\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big],$$

since $(\overline{H}\cdot B)_{T_m\wedge\cdot}$ is a strong $\overline{\mathbb{P}}-$supermartingale. This yields $\overline{P}(\boldsymbol{\mu}) \leq \overline{D}(\boldsymbol{\mu})$ by their arbitrariness.
(ii) We next prove $\overline{P}(\boldsymbol{\mu}) \leq \overline{D}'(\boldsymbol{\mu})$. For any $(\boldsymbol{\lambda}, H^1, \cdots, H^m) \in \overline{\mathcal{D}}'$, one has

$$\boldsymbol{\lambda}(\omega,\theta) \;+\; \sum_{k=1}^{m}\int_{\theta_{k-1}}^{\theta_k} H^k_s dB_s \;\geq\; \Phi(\omega,\theta) \text{ for all } \theta\in\Theta \text{ and } \mathbb{P}_0 \text{ - a.e. } \omega\in\Omega.$$

Moreover, there exists some $C>0$ such that $(H^k\cdot B)_t \geq -C(1+|B_t|)$ for all $k=1,\cdots,m$. Take an arbitrary $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$, then it follows that $B_{T_m\wedge\cdot}$ is uniformly integrable. Hence we obtain by Fatou's lemma

$$\mathbb{E}^{\overline{\mathbb{P}}}\Big[\int_{T_k}^{T_{k+1}} H_s dB_s\Big] \;\leq\; 0 \text{ for all } k=0,\cdots,m-1,$$

which yields $\overline{P}(\boldsymbol{\mu}) \leq \overline{D}'(\boldsymbol{\mu})$.
(iii) Similarly, one can easily show $\overline{P}(\boldsymbol{\mu}) \leq \overline{D}''(\boldsymbol{\mu})$. □

4.3.2 Reduction of Φ to be bounded

Proposition IV.4.7. *To prove Theorem IV.2.6* (ii) *and Proposition IV.2.7, it is enough to prove the results under the additional condition that* Φ *is bounded.*

Proof. We will prove it only for Theorem IV.2.6 (ii), since the arguments in the context of Proposition IV.2.7 are the same.

Assume that the duality $\overline{P}(\boldsymbol{\mu}) = \overline{D}(\boldsymbol{\mu})$ holds true whenever Φ is bounded and satisfies Assumptions IV.2.4 and IV.2.5. We now turn to the case where Φ is bounded from above. Let $\Phi^n := \Phi\vee(-n)$ or $\Phi^n := \sum_{k=1}^m \Phi_k \vee (-n)$, according to Assumption IV.2.5 (ii) or (iii), then Φ^n is bounded and satisfies Assumptions IV.2.4 and IV.2.5. Denote by $\overline{P}^n(\boldsymbol{\mu})$ and $\overline{D}^n(\boldsymbol{\mu})$ the corresponding primal and dual values associated to Φ^n, then we have the duality

$$\overline{P}^n(\boldsymbol{\mu}) \;=\; \overline{D}^n(\boldsymbol{\mu}).$$

Further, since $\Phi^n \geq \Phi$, one has $\overline{P}^n(\boldsymbol{\mu}) = \overline{D}^n(\boldsymbol{\mu}) \geq \overline{D}(\boldsymbol{\mu}) \geq \overline{P}(\boldsymbol{\mu})$, where the last inequality follows by Lemma IV.4.6. Then it is enough to show that

$$\limsup_{n\to\infty} \overline{P}^n(\boldsymbol{\mu}) \;\; \leq \;\; \overline{P}(\boldsymbol{\mu}).$$

Let $\overline{\mathbb{P}}_n \in \overline{\mathcal{P}}(\boldsymbol{\mu})$ such that $\limsup_{n\to\infty} \overline{P}^n(\boldsymbol{\mu}) = \limsup_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}[\Phi^n]$. Then after possibly passing to a subsequence we may assume that $\limsup_{n\to\infty} P^n(\boldsymbol{\mu}) = \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}[\Phi^n]$. By Lemma IV.4.2, we know that $(\overline{\mathbb{P}}_n)_{n\geq 1}$ is relative compact and every limit point belongs to $\overline{\mathcal{P}}(\boldsymbol{\mu})$. Let $\overline{\mathbb{P}}_0$ be a limit point of $(\overline{\mathbb{P}}_n)_{n\geq 1}$, and label again the convergent subsequence by n. Then by the monotone convergence theorem

$$\begin{aligned} \overline{P}(\boldsymbol{\mu}) \;\; \geq \;\; \mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\Phi\Big] \;\; &= \;\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_0}\Big[\Phi^n\Big] \;\; = \;\; \lim_{n\to\infty} \Big(\lim_{k\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_k}\Big[\Phi^n\Big]\Big) \\ &\geq \;\; \lim_{n\to\infty} \Big(\lim_{k\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_k}\Big[\Phi^k\Big]\Big) \;\; = \;\; \limsup_{k\to\infty} \overline{P}^k(\boldsymbol{\mu}), \end{aligned}$$

which is the required result. $\square$

4.3.3 Proof of Proposition IV.2.7

By Proposition IV.4.7, we can assume in the following that Φ is bounded without loss of generality. Then given the first duality $\overline{P}(\boldsymbol{\mu}) = \overline{D}_0(\boldsymbol{\mu})$, it suffices to study the optimal stopping problem

$$\sup_{\tau\in\mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] \;\; = \;\; \lim_{N\to\infty} \sup_{\tau\in\mathcal{T}^a_N} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big], \tag{IV.4.9}$$

for a given $\boldsymbol{\lambda} \in \Lambda^+$, see Remark IV.4.5. Notice that in this case, there is some $C > 0$ such that

$$-C\Big(1 + \sum_{k=1}^m |\omega_{\theta_k}|\Big) \;\; \leq \;\; \Phi^{\boldsymbol{\lambda}}(\bar{\omega}) \;\; \leq \;\; C. \tag{IV.4.10}$$

Assume first $m = 1$, then $\mathcal{T}^a = \mathcal{T}^0$ and $\mathcal{T}^a_N = \mathcal{T}^0_N$, where $\mathcal{T}^0$ is the set of $\mathbb{F}^a-$stopping times and $\mathcal{T}^0_N \subseteq \mathcal{T}^0$ is the subset of $\mathbb{F}^a-$stopping times bounded by N. then by Lemma IV.5.3, there is an $\mathbb{F}^a-$optional làdlàg process $(Z^{1,N}_t)$, for every $N \in \mathbb{N}$, which is an $\mathbb{F}^a-$supermartingale and the Snell envelope of the optimal stopping problem $\sup_{\tau\in\mathcal{T}^0_N} \mathbb{E}[\Phi^{\lambda}(B,\tau)]$. Clearly, $Z^{1,N}$ increases in N. Moreover, since $Z^{1,N}$ dominates Φ^{λ}, then it follows by (IV.4.10) that $-C(1 + |B_t|) \leq Z^{1,N}_t \leq C$. Then by the dominated convergence theorem together with Lemma IV.4.4,

$Z^1 := \sup_{N\in\mathbb{N}} Z^{1,N}$ is still a làdlàg $\mathbb{F}^a-$supermartingale, of class (DL), such that

$$Z^1_0 \;=\; \sup_{\tau\in\mathcal{T}^0} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^\lambda(B,\tau)\Big] \quad\text{and}\quad Z^1_t \;\geq\; \Phi^\lambda(B,t) \text{ for all } t\in\mathbb{R}_+,\; \mathbb{P}_0 \text{ - a.s.}$$

Now, by the Doob-Meyer decomposition, see Lemma IV.5.4 below, for supermartingales of class (DL) without right-continuity, together with the martingale representation theorem, there is an $\mathbb{F}^a-$predictable process H such that

$$\lambda(B_t) \;+\; (H\cdot B)_t \;\geq\; \Phi(B,t) \text{ for all } t\in\mathbb{R}_+,\; \mathbb{P}_0 \text{ - a.s.}$$

Further, since any $\mathbb{F}^a-$predictable process, or equivalently $\mathbb{F}^a-$optional process, is indistinguishable to an $\mathbb{F}^0-$predictable process, see Theorem IV.78 and Remark IV.74 of Dellacherie & Meyer [34], then we can also choose H to be $\mathbb{F}^0-$predictable. This implies in particular that

$$\overline{D}'(\mu) \;\leq\; \overline{D}_0(\mu) \;=\; \overline{P}(\mu) \quad\text{and}\quad \overline{D}''(\mu) \;\leq\; \overline{D}_0(\mu) \;=\; \overline{P}(\mu).$$

Combining with the weak dualities $\overline{P}(\mu) \leq \overline{D}'(\mu)$ and $\overline{P}(\mu) \leq \overline{D}''(\mu)$ in Lemma IV.4.6, we obtain

$$\overline{P}(\mu) \;=\; \overline{D}_0(\mu) \;=\; \overline{D}'(\mu) \;=\; \overline{D}''(\mu).$$

Assume now $m=2$, we first consider the optimal stopping problem

$$\sup_{\tau_2\in\mathcal{T}^0_N} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi_2(B,\tau_2) \;-\; \lambda_2(B_{\tau_2})\Big].$$

Clearly, its Snell envelope is given by $Z^{2,N}$ by Lemma IV.5.3, where in particular $-C(1+|B_t|) \leq Z^{2,N}_t \leq C$ for some constant $C>0$ independent of N, and

$$Z^{2,N}_{\theta_2} \;\geq\; \Phi_2(B,\theta_2) \;-\; \lambda_2(B_{\theta_2}) \text{ for all } \theta_2\leq N,\; \mathbb{P}_0 \text{ - a.s.}$$

Then the multiple optimal stopping problem (IV.4.9) is reduced to the case $m=1$, *i.e.*

$$\sup_{\tau\in\mathcal{T}^a_N} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] \;=\; \sup_{\tau_1\in\mathcal{T}^0_N} \mathbb{E}^{\mathbb{P}_0}\Big[Z^{2,N}_{\tau_1} \;+\; \Phi_1(B,\tau_1) \;-\; \lambda_1(B_{\tau_1})\Big].$$

Using again the procedure for the case $m=1$, we obtain a new Snell envelop, denoted by $Z^{1,N}$, such that $Z^{1,N}_t \geq -C(1+|B_t|)$ for some $C>0$.

Thus, $Z^{1,N}$ and $Z^{2,N}$ are supermartingales of class (D), bounded from above by C, and dominated from below by $-C(1+|B_t|)$ for some constant $C>0$ independent of N. More importantly, we have $Z_0^{1,N} = \sup_{\tau\in\mathcal{T}_N^a}\mathbb{E}^{\mathbb{P}_0}[\Phi^{\boldsymbol{\lambda}}(B,\tau)]$, and

$$Z_{\theta_1}^{1,N} \;+\; \Big(Z_{\theta_2}^{2,N} - Z_{\theta_1}^{2,N}\Big) \;\geq\; \Phi^{\boldsymbol{\lambda}}(B,\theta_1,\theta_2) \text{ for all } \theta_1\leq\theta_2\leq N,\ \mathbb{P}_0 \text{ - a.s.}$$

Since $Z^{1,N}$ and $Z^{2,N}$ both increase in N, define $Z^1 := \sup_N Z^{1,N}$ and $Z^2 := \sup_N Z^{2,N}$, it follows by the dominated convergence theorem that Z^1 and Z^2 are both supermartingales of class (DL). Moreover, it follows from Lemma IV.4.4 that $Z_0^1 = \sup_{\tau\in\mathcal{T}^a}\mathbb{E}^{\mathbb{P}_0}[\Phi^{\boldsymbol{\lambda}}(B,\tau)]$ and

$$Z_{\theta_1}^{1} \;+\; \Big(Z_{\theta_2}^{2} - Z_{\theta_1}^{2}\Big) \;\geq\; \Phi^{\boldsymbol{\lambda}}(B,\theta_1,\theta_2) \text{ for all } \theta_1\leq\theta_2,\ \mathbb{P}_0 \text{ - a.s.}$$

Then $(S^1,S^2) := (Z^1,Z^2)$ are the required supermartingale in dual formulation $\overline{\mathfrak{D}}''$. Further, using the Doob-Meyer decomposition, together with the martingale representations of Z^1 and Z^2, we obtain the processes H^1 and H^2 needed in the dual formulation $\overline{\mathfrak{D}}'$.

Finally, the case $m>2$ can be handled by exactly the same recursive arguments as for the case $m=2$. □

4.3.4 Proof of Theorem IV.2.6 (ii) under Assumption IV.2.5 (i)

When $m=1$, Theorem IV.2.6 is an immediate consequence of Proposition IV.2.7.

4.3.5 Proof of Theorem IV.2.6 (ii) under Assumption IV.2.5 (iii)

Given $N>0$, we first study the multiple optimal stopping problem

$$\sup_{\tau\in\mathcal{T}_N^a}\mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big] \;=\; \sup_{\tau\in\mathcal{T}_N^a}\mathbb{E}^{\mathbb{P}_0}\Big[\sum_{k=1}^m\Big(\Phi_k(B,\tau_1,\cdots,\tau_k) - \lambda_k(B_{\tau_k})\Big)\Big], \quad \text{(IV.4.11)}$$

where $\boldsymbol{\lambda}\in\Lambda^+$. Hence

$$-C\Big(1+\sum_{k=1}^m|\omega_{\theta_k}|\Big) \;\leq\; \Phi^{\boldsymbol{\lambda}}(\bar\omega) \;\leq\; C, \quad \text{(IV.4.12)}$$

for some constant $C>0$. Denote $v_{m+1}^N(\omega,\theta_1,\cdots,\theta_m,\theta_m) := \Phi^{\lambda}(\omega,\theta_1,\cdots,\theta_m)$.

Lemma IV.4.8. *There are functionals $(v_k^N)_{1\leq k\leq m}$ with $v_k^N:\Omega\times\mathbb{R}_+^k\to\mathbb{R}$, such that*

$$v_1^N(\omega,0) \;=\; \sup_{\tau\in\mathcal{T}_N^a}\mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B,\tau)\Big],$$

and, for all $k = 1, \cdots, m$ and $\theta_1 \le \cdots \le \theta_{k-1}$, the process

$$\begin{aligned}\theta \ \mapsto \ v_k^N(B, \theta_1, \cdots, \theta_{k-1}, \theta) \text{ is a } \mathbb{P}_0 - \text{supermartingale satisfying}\\ v_k^N(B, \theta_1, \theta_{k-1}, \theta) \ \ge \ v_{k+1}^N(B, \theta_1, \cdots, \theta_{k-1}, \theta, \theta), \ \mathbb{P}_0 \text{ - a.s.}\end{aligned}$$

Moreover, v_k^N increases in N and satisfies $-C(1+\sum_{i=1}^k |\omega_{\theta_i}|) \le v_k^N(\omega, \theta_1, \cdots, \theta_k) \le C$ for some constant $C > 0$ independent of N.

Proof of Theorem IV.2.6 (ii). Recall that, by Remark IV.4.5 and Proposition IV.4.7, Φ_k is bounded for $k = 1, \cdots, m$ and $\boldsymbol{\lambda} \in \Lambda^+$ in the dual formulation $\overline{D}_0(\boldsymbol{\mu})$.

(i) Let v_k^N be given by Lemma IV.4.8, we define further

$$v_k(\cdot) \ := \ \sup_N \ v_k^N(\cdot).$$

Thus $v_1(\omega, 0) = \sup_{\tau \in \mathcal{T}^a} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B, \tau)\Big]$. It follows from the dominated convergence theorem that, for all $k = 1, \cdots, m$ and $0 = \theta_0 \le \theta_1 \le \cdots \le \theta_{k-1}$, the process $\Big(v_k(B, \theta_1, \cdots, \theta_{k-1}, t)\Big)_{t \ge \theta_{k-1}}$ is an $\mathbb{F}^a-$supermartingale and

$$v_k(B, \theta_1, \cdots, \theta_{k-1}, t) \ \ge \ v_{k+1}(B, \theta_1, \cdots, \theta_{k-1}, t, t) \text{ for all } t \ge \theta_{k-1}, \ \mathbb{P}_0 \text{ - a.s.}$$

(ii) By the Doob-Meyer decomposition, see Lemma IV.5.4 below, and the martingale representation theorem, it follows that for each $k = 1, \cdots, m$, there is some $\mathbb{F}^a-$predictable process $H^k(\omega) := H^k(\omega, \theta_1, ..., \theta_{k-1})$ such that

$$\begin{aligned} v_k(\omega, \theta_1, ..., \theta_{k-1}, \theta_{k-1}) \ + \int_{\theta_{k-1}}^{\theta_k} H_t^k dB_t \ &\ge \ v_k(\omega, \theta_1, ...\theta_{k-1}, \theta_k)\\ &\ge \ v_{k+1}(\omega, \theta_1, ...\theta_{k-1}, \theta_k, \theta_k), \ \mathbb{P}_0 \text{ - a.s.} \end{aligned} \tag{IV.4.13}$$

(iii) Next, following the pathwise construction in (IV.5.2) of the quadratic co-variation of a supermartingale and a continuous martingale, one has a Borel version of the quadratic co-variation $\langle v_k(B, \theta_1, \cdots, \theta_{k-1}, \cdot), B\rangle_t$. Then by Lemma IV.5.5, the process H^k defined below is $\mathbb{F}^0-$predictable,

$$\begin{aligned} & H_t^k(B, \theta_1, \cdots, \theta_{k-1})\\ := \ & \limsup_{\varepsilon \to 0} \frac{\langle v_k(B, \theta_1, \cdots, \theta_{k-1}, \cdot), B\rangle_t \ - \ \langle v_k(B, \theta_1, \cdots, \theta_{k-1}, \cdot), B\rangle_{t-\varepsilon}}{\varepsilon}. \end{aligned}$$

In particular, the map $(\omega, \theta_1, \cdots, \theta_k) \mapsto H^k_{\theta_k}(\omega, \theta_1, \cdots, \theta_{k-1})$ is Borel measurable, and

$$\int_{\theta_{k-1}}^{t} \Big(H^k_s(\cdot, \theta_1, \cdots, \theta_{k-1})\Big)^2 ds \;<\; +\infty \text{ for all } t \geq \theta_{k-1},\ \mathbb{P}_0 \text{ - a.s.} \tag{IV.4.14}$$

(iv) Next, we define a process $\overline{H} : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ by

$$\overline{H}_t(\bar{\omega}) \;:=\; \sum_{k=1}^{m} \mathbb{1}_{(\theta_{k-1},\theta_k]}(t) H^k_t(\omega, \theta_1, ..., \theta_{k-1}) \text{ for all } \bar{\omega} = (\omega, \theta) \in \overline{\Omega}.$$

Moreover, since

$$(\omega, \theta_1, \cdots, \theta_k) \;\mapsto\; H^k_{\theta_k}(\omega, \theta_1, \cdots, \theta_{k-1}) \text{ is Borel measurable,}$$

and one has clearly that $H^k_{\theta_k}(\omega, \theta_1, \cdots, \theta_{k-1}) = H^k_{\theta_k}(\omega_{\theta_k \wedge \cdot}, \theta_1, \cdots, \theta_{k-1})$, then the process $\overline{H}$ is $\overline{\mathbb{F}}-$optional by Lemma IV.5.2 in Appendix.

(v) Now, let us take an arbitrary $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$ and consider a family of r.c.p.d. (regular conditional probability distributions) $(\overline{\mathbb{P}}_{\bar{\omega}})_{\bar{\omega}\in\overline{\Omega}}$ of $\overline{\mathbb{P}}$ with respect to $\overline{\mathcal{F}}_{T_k}$ for all $k = 0, \cdots, m-1$, see Lemma IV.5.2 for the existence of r.c.p.d. Then for $\overline{\mathbb{P}}$ - almost every $\bar{\omega} \in \overline{\Omega}$, under the conditional probability $\overline{\mathbb{P}}_{\bar{\omega}}$, the process $t \mapsto B_t$ for $t \geq T_k$ is still a Brownian motion. Moreover, we have $\overline{\mathbb{P}}_{\bar{\omega}}[T_k = \theta_k \text{ and } B_{T_k \wedge \cdot} = \omega_{\theta_k \wedge \cdot}] = 1$. Then it follows by (IV.4.13) that

$$\begin{aligned} v_{k+1}(B, T_1, ..., T_k, T_k) \;&\leq\; v_k(B, T_1, ..., T_k) \\ &\leq\; v_k(B, T_1, ..., T_{k-1}, T_{k-1}) \;+\; \int_{T_{k-1}}^{T_k} H^k_s dB_s,\ \overline{\mathbb{P}}_{\bar{\omega}} \text{ - a.s.} \end{aligned}$$

This means that the set $A_k := \{v_{k+1} \leq v_k + \int_{T_{k-1}}^{T_k} H^k_s dB_s\}$ is of full measure under $\overline{\mathbb{P}}_{\bar{\omega}}$ for $\overline{\mathbb{P}}$ - almost every $\bar{\omega} \in \overline{\Omega}$, and hence by the tower property $\overline{\mathbb{P}}[A_k] = 1$ for all $k = 0, \cdots, m$. Here we refer to [27] for some more discussion on the measurability of A_k under $\overline{\mathbb{P}}_{\bar{\omega}}$. This yields that

$$\Phi^{\lambda}(B, T) \;=\; v_{m+1}(B, T_1, ..., T_m, T_m) \;\leq\; v_1(B, 0) \;+\; (\overline{H} \cdot B)_{T_m},\ \overline{\mathbb{P}} \text{ - a.s.} \tag{IV.4.15}$$

(vi) To conclude the proof, it suffices to check that $\overline{H} \in \overline{\mathcal{H}}$. First, for any probability measure $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$, by taking the r.c.p.d and using (IV.4.14), it is clear that

$$\int_0^t \overline{H}^2_s ds \;<\; +\infty \text{ for all } t \in \mathbb{R}_+,\ \overline{\mathbb{P}} \text{ - a.s.}$$

Notice also that (IV.4.15) holds true for all $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$, and by the tower property, it follows that for any $\overline{\mathbb{F}}-$stopping time τ, we have for all $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$,

$$(\overline{H} \cdot B)_{T_m \wedge \tau} \quad \geq \quad -C\Big(1 + \sup_{1 \leq k \leq m} |B_{T_k \wedge \tau}|\Big), \ \overline{\mathbb{P}} \text{ - a.s.},$$

where the r.h.s. is uniformly integrable under $\overline{\mathbb{P}}$. Using Fatou's Lemma, we conclude that $(\overline{H} \cdot B)_{T_m \wedge \cdot}$ is a strong $\overline{\mathbb{P}}-$supermartingale for all $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$. □

Proof of Lemma IV.4.8. We provide here a proof for the case $m = 2$ for ease of presentation. The general case can be treated by exactly the same backward iterative procedure. We will use the aggregation procedure in the optimal stopping theory, see *e.g.* El Karoui [38], Peskir & Shiryaev [87], Karatzas & Shreve [71] and Kobylanski, Quenez & Rouy-Mironescu [73].

(i) Recall that $\mathcal{T}^0_N$ is the subset of $\mathbb{F}^a-$stopping times bounded by N. For each $\tau_1 \in \mathcal{T}^0_N$, we first consider the optimal stopping problem

$$\sup_{\tau_2 \in \mathcal{T}^0_N:\ \tau_2 \geq \tau_1} \ \mathbb{E}^{\mathbb{P}_0}\Big[\Phi_2(B, \tau_1, \tau_2) \ - \ \lambda_2(B_{\tau_2})\Big],$$

whose Snell envelope is denoted by $(Z^{2,N}_{\tau_1,t})_{\tau_1 \leq t \leq N}$. We shall prove in (ii) below that the above process can be aggregated into a function $u^{2,N}(\omega, \theta_1, \theta_2)$ which increases in N and is Borel measurable as a map from $\Omega \times \mathbb{R}^2_+ \to \mathbb{R}$,

$$\text{uniformly continuous in } \theta_1 \text{ and } u^{2,N}(\cdot, \tau_1, \tau_2) = Z^{2,N}_{\tau_1,\tau_2}, \ \mathbb{P}_0 \text{ - a.s. for all } \tau_1 \leq \tau_2 \leq N. \quad \text{(IV.4.16)}$$

Set

$$v^{2,N}(\omega, \theta_1, \theta_2) \ := \ u^{2,N}(\omega, \theta_1, \theta_2) \ + \ \Phi_1(\omega, \theta_1) \ - \ \lambda_1(\omega_{\theta_1}),$$

and consider the optimal stopping problem

$$\sup_{\tau_1 \in \mathcal{T}^0_N} \ \mathbb{E}^{\mathbb{P}_0}\Big[v^{2,N}(\cdot, \tau_1, \tau_1)\Big] \ = \ \sup_{\tau \in \mathcal{T}^a_N} \ \mathbb{E}^{\mathbb{P}_0}\Big[\Phi^{\boldsymbol{\lambda}}(B, \tau)\Big]. \quad \text{(IV.4.17)}$$

Denote by $(Z^{1,N}_t)_{0 \leq t \leq N}$ the corresponding Snell envelop, which is $\mathbb{F}^a-$optional, or equivalently $\mathbb{F}^a-$predictable since $\mathbb{F}^a$ is the augmented Brownian filtration, then $Z^{1,N}_t$ can be chosen to be $\mathbb{F}^0-$predictable, see *e.g.* Dellacherie & Meyer [34, Theorem IV.78 and Remark IV.74]. Moreover, in view of (IV.4.12), by truncating it with $-C(1 + |\omega_t|)$ from below and with C from above, we can assume that $Z^{1,N}_t$ is bounded between $-C(1 + |\omega_t|)$ and C. Further, since $u^{2,N}$ increases in N, then for every $N_1 < N_2$, we know $Z^{1,N_2} \vee Z^{1,N_1}$ is still a Snell envelop of problem

(IV.4.17), then we can assume without loss of generality that $Z^{1,N}$ is increasing in N. Define $v^{1,N}(\omega,\theta_1) := Z^{1,N}(\omega,\theta_1)$, it follows that $v^{1,N}(\cdot)$ and $v^{2,N}(\cdot)$ are the required functionals.

(ii) We now construct the measurable map $u^{2,N}$ satisfying (IV.4.16). Let $\tau_1 \le \tau_2 \in \mathcal{T}_N^0$, define a random variable

$$\overline{Z}^{2,N}_{\tau_1,\tau_2} \;:=\; \operatorname*{ess\,sup}_{\tau_3\in\mathcal{T}^0_N:\ \tau_3\ge\tau_2} \mathbb{E}^{\mathbb{P}_0}\Big[\Phi_2\big(B,\tau_1,\tau_3\big) - \lambda_2(B_{\tau_3})\,\Big|\mathcal{F}^a_{\tau_2}\Big]. \tag{IV.4.18}$$

Then, for every fixed τ_1, $(\overline{Z}^{2,N}_{\tau_1,\tau_2})_{\tau_2\ge\tau_1}$ can be aggregated into a supermartingale, denoted by $Z^{2,N}_{\tau_1,t}$, see Lemma IV.5.3, such that $\overline{Z}^{2,N}_{\tau_1,\tau_2} = Z^{2,N}_{\tau_1,\tau_2}$, $\mathbb{P}_0$ - almost surely, for each $\tau_2 \ge \tau_1$. Notice that $Z^{2,N}_{\tau_1,t}$ is $\mathbb{F}^a-$optional and equivalently $\mathbb{F}^a-$predictable, we can choose $Z^{2,N}_{\tau_1,t}$ to be $\mathbb{F}^0-$predictable by [34, Theorem IV.78 and Remark IV.74]. Moreover, since $\overline{Z}^{2,N}_{\tau_1,\tau_2}$ is increasing in N, $\mathbb{P}_0$ - almost surely, then for any $N_1 \le N_2$, $Z^{2,N_1} \vee Z^{2,N_2}$ is also an aggregated supermartingale for $(\overline{Z}^{2,N_2}_{\tau_1,\tau_2})_{\tau_2\ge\tau_1}$, and hence we can assume in addition that $Z^{2,N}$ is increasing in N. Further, in view of (IV.4.12), by truncation, we can assume in addition that $-C(1+|\omega_{t_1}|+|\omega_{t_2}|) \le Z^{2,N}_{t_1,t_2} \le C$.

Notice also that for two stopping times τ_1^1 and τ_1^2 smaller than τ_2, we have

$$\overline{Z}^{2,N}_{\tau_1^1,\tau_2} \;=\; \overline{Z}^{2,N}_{\tau_1^2,\tau_2},\ \mathbb{P}_0 \text{ - a.s. on } A \;=\; \{\tau_1^1 = \tau_1^2\}. \tag{IV.4.19}$$

Further, since $\Phi_2(\omega,\theta_1,\theta_2)$ is uniformly continuous in θ_1, denote by ρ the continuity modulus. Then it follows by its definition in (IV.4.18) that the family of random variables $\overline{Z}^{2,N}_{\tau_1,\tau_2}$ is uniformly continuous with respect to τ_1, in sense that

$$\Big|\overline{Z}^{2,N}_{\tau_1^1,\tau_2} - \overline{Z}^{2,N}_{\tau_1^2,\tau_2}\Big| \;\le\; \rho(|\tau_1^1-\tau_1^2|),\ \mathbb{P}_0 \text{ - a.s. for stopping times } \tau_1^i \le \tau_2.$$

We now define $u^{2,N}$ by

$$u^{2,N}(\omega,\theta_1,\theta_2) \;:=\; Z^{2,N}_{\theta_1,\theta_2}(\omega) \text{ for all } \theta_1 \in \mathbb{Q},$$

and

$$u^{2,N}(\omega,\theta_1,\theta_2) \;:=\; \limsup_{\mathbb{Q}\ni\theta_1'\to\theta_1} u^{2,N}(\omega,\theta_1',\theta_2) \text{ for all } \theta_1 \notin \mathbb{Q}.$$

It is clear that $u^{2,N}$ is Borel measurable with respect to each variable since $Z^{2,N}_{\theta_1,\theta_2}(\omega)$ is chosen to be $\mathbb{F}^0-$predictable. Furthermore, by (IV.4.19), we have $u^{2,N}(\omega,\tau_1,\theta_2) = Z^{2,N}(\omega,\tau_1,\theta_2)$ for all $\theta \ge \tau_1$, $\mathbb{P}_0$ - almost surely, for every stopping times τ_1 taking values in $\mathbb{Q}$. Since we can approximate any stopping time by stopping times taking values in $\mathbb{Q}$, then by the uniform continuity of $\overline{Z}^{2,N}_{\tau_1,\tau_2}$ with respect to τ_1, we obtain that

$$\overline{Z}^{2,N}_{\tau_1,\tau_2} \;=\; Z^{2,N}_{\tau_1,\tau_2} \;=\; u^{2,N}(\cdot,\tau_1,\tau_2) \text{ for all } (\tau_1,\tau_2)\in\mathcal{T}^a_N,\ \mathbb{P}_0 \text{ - a.s.}$$

In particular, $u^{2,N}(\omega,\theta_1,\theta_2)$ is uniformly continuous in θ_1, $\mathbb{P}_0$ - almost surely, which is the required functional in claim (IV.4.16). □

Remark IV.4.9. *A general multiple optimal stopping problem has been studied in Kobylanski, Quenez & Rouy-Mironescu [73], where the stopping times are not assumed to be ordered. In particular, they proved the existence of optimal multiple stopping times by a constructive method. Here we are in a specific context of Brownian motion and we are interested in finding a process $\overline{H}$ whose stochastic integral dominates the value process.*

4.3.6 Proof of Theorem IV.2.6 (ii) under Assumption IV.2.5 (ii)

Let Φ satisfy Assumptions IV.2.4 and IV.2.5 (ii), *i.e.* $\bar\omega\mapsto\Phi(\bar\omega)$ is upper semicontinuous and bounded from above. Define a metric $\overline{d}$ on $\overline{\Omega}$ by

$$\overline{d}(\bar\omega,\bar\omega') \;:=\; \sum_{k=1}^m \Big(|\theta_k-\theta'_k| \;+\; \|\omega_{\theta_k\wedge\cdot}-\omega'_{\theta'_k\wedge\cdot}\|\Big),$$

and then $\Phi^n:\overline{\Omega}\to\mathbb{R}$ by

$$\Phi^n(\bar\omega) \;:=\; \sup_{\bar\omega'\in\overline{\Omega}} \Big\{\Phi(\bar\omega') \;-\; n\overline{d}(\bar\omega,\bar\omega')\Big\}. \tag{IV.4.20}$$

Then Φ^n is $\overline{d}-$Lipschitz, and satisfies in particular Assumptions IV.2.4 and IV.2.5 (iii). Moreover, $\Phi^n(\bar\omega)$ decreases to $\Phi(\bar\omega)$ as n goes to infinity for all $\bar\omega\in\overline{\Omega}$.

Denote by $\overline{P}^n(\boldsymbol{\mu})$ and $\overline{D}^n(\boldsymbol{\mu})$ the corresponding primal and dual values associated to the function Φ^n. Since Φ^n satisfies Assumptions IV.2.4 and IV.2.5 (iii), we have proved in Section 4.3.5 the duality

$$\overline{P}^n(\boldsymbol{\mu}) \;=\; \overline{D}^n(\boldsymbol{\mu}).$$

Then by following the same reasoning in Proposition IV.4.7, we deduce that $\overline{P}(\boldsymbol{\mu})=\overline{D}(\boldsymbol{\mu})$. □

5 Appendix

5.1 Canonical filtration on $\overline{\Omega}$

We finally provide some properties of the canonical filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$ of canonical space $\overline{\Omega}$, where $\overline{\mathcal{F}}_t$ is generated by the processes $B_{t\wedge\cdot}$ and $\{T_k \leq s\}$ for all $s \in [0,t]$ and $k = 1, \cdots, m$. Recall also $\overline{\mathcal{F}} = \bigvee_{t\geq 0} \overline{\mathcal{F}}_t$.

Lemma IV.5.1. *The σ-field $\overline{\mathcal{F}}$ is the Borel σ-filed of $\overline{\Omega}$. Moreover, the σ-field $\overline{\mathcal{F}}_{t-} := \bigvee_{s<t} \overline{\mathcal{F}}_s$ is generated by the class of all bounded, continuous and $\overline{\mathcal{F}}_t-$measurable functions on $\overline{\Omega}$.*

Proof. (i) Denote by $\mathcal{B}(\overline{\Omega})$ the Borel σ-filed of $\overline{\Omega}$. Since T_k and B are all $\mathcal{B}(\overline{\Omega})-$measurable, one has $\overline{\mathcal{F}} \subseteq \mathcal{B}(\overline{\Omega})$. On the other hand, the process $(B_t)_{t\geq 0}$ generates the Borel σ-field $\mathcal{B}(\Omega)$ of Ω and the collection of sets $\{T_k \leq s\}$ generates the Borel σ-filed $\mathcal{B}(\Theta)$ of Θ, then it follows that $\mathcal{B}(\overline{\Omega}) = \mathcal{B}(\Omega) \otimes \mathcal{B}(\Theta) \subseteq \overline{\mathcal{F}}$.
(ii) Let $t \geq 0$, recall that $\overline{\mathcal{F}}^B_t = \sigma(B_s, 0 \leq s \leq t)$ and denote $\overline{\mathcal{F}}^{T_k}_t := \sigma(\{T_k \leq s\}, s \in [0,t])$ and by $\mathcal{G}^{T_k}_t$ the σ-field generated by all bounded, continuous and $\overline{\mathcal{F}}^{T_k}_t-$measurable functions. First, for every $s < t$, it is clear that $\overline{\mathcal{F}}^{T_k}_s \subseteq \mathcal{G}^{T_k}_t$, thus $\overline{\mathcal{F}}^{T_k}_{t-} \subseteq \mathcal{G}^{T_k}_t$. Further, let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be a bounded continuous function such that $\phi(T_k)$ is $\overline{\mathcal{F}}^{T_k}_t-$measurable, then we have $\phi(t_1) = \phi(t_2)$ for every $t_1 \geq t_2 \geq t$. It follows that $\phi(T_k)$ is $\overline{\mathcal{F}}^{T_k}_{t-}-$measurable. Therefore, we have $\overline{\mathcal{F}}^{T_k}_{t-} = \mathcal{G}^{T_k}_t$. Besides, it is well known that $\overline{\mathcal{F}}^B_{t-} = \overline{\mathcal{F}}^B_t$ is the σ-field generated by all bounded, continuous and $\overline{\mathcal{F}}^B_t-$measurable functions. It follows that $\overline{\mathcal{F}}_{t-} = \cup_{k=1}^m \overline{\mathcal{F}}^{T_k}_{t-} \cup \overline{\mathcal{F}}^B_{t-}$ is in fact the σ-field generated by all bounded, continuous and $\overline{\mathcal{F}}_t-$measurable functions. □

We consider now the filtration $\overline{\mathbb{F}}$. Let $t \geq 0$ and $\bar{\omega} = (\omega, \theta_1, \cdots, \theta_m) \in \overline{\Omega}$, we introduce $[\bar{\omega}]_t = (\omega_{t\wedge\cdot}, [\theta_1]_t, \cdots, [\theta_m]_t)$ by $[\theta_k]_t := \theta_k \mathbb{1}_{\theta_k \leq t} + \infty \mathbb{1}_{\theta_k > t}$.

Lemma IV.5.2. (i) *$Y : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ is $\overline{\mathbb{F}}-$optional if and only if it is $\mathcal{B}(\mathbb{R}_+ \times \overline{\Omega})-$measurable and satisfies*

$$Y_t(\bar{\omega}) \quad = \quad Y_t([\bar{\omega}]_t) \text{ for all } t \in \mathbb{R}_+ \text{ and } \bar{\omega} \in \overline{\Omega}. \qquad \text{(IV.5.1)}$$

(ii) *Consequently, $\overline{\mathcal{F}}_{T_k}$ is countably generated and every probability measure $\overline{\mathbb{P}}$ on $(\overline{\Omega}, \overline{\mathcal{F}})$ admits a family $(\overline{\mathbb{P}}_{\bar{\omega}})_{\bar{\omega}\in\overline{\Omega}}$ with respect to $\overline{\mathcal{F}}_{T_k}$ such that*

a) $(\overline{\mathbb{P}}_{\bar{\omega}})_{\bar{\omega}\in\overline{\Omega}}$ is a family of conditional probabilities of $\overline{\mathbb{P}}$ with respect to $\overline{\mathcal{F}}_{T_k}$,

b) $\overline{\mathbb{P}}_{\bar{\omega}}[T_k = \theta_k$ and $B_{T_k\wedge\cdot} = \omega_{T_k\wedge\cdot}] = 1$ for all $\bar{\omega} = (\omega, \theta_1, \cdots, \theta_m) \in \overline{\Omega}$.

Proof. (i) If Y is $\overline{\mathbb{F}}-$optional, *i.e.* Y is measurable and $\overline{\mathbb{F}}-$adapted, then Y_t is $\overline{\mathcal{F}}_t-$measurable. Since $\overline{\mathcal{F}}_t$ is generated by $\bar{\omega} \mapsto (\omega_{t\wedge\cdot}, [\theta_1]_t, \cdots, [\theta_m]_t)$, it follows that (IV.5.1) holds true. On the other hand, the process $(t, \bar{\omega}) \mapsto (\omega_{t\wedge\cdot}, [\theta_1]_t, \cdots, [\theta_m]_t)$ is adapted and càdlàg, and hence $\overline{\mathbb{F}}-$optional. Therefore, every measurable process Y satisfying (IV.5.1) is $\overline{\mathbb{F}}-$optional.

(ii) Notice that $\mathcal{B}(\overline{\Omega})$ is countably generated. Thus by (IV.5.1), the $\overline{\mathbb{F}}-$optional σ-field is generated by the map $(t, \bar{\omega}) \in \mathbb{R}_+ \times \overline{\Omega} \mapsto [\bar{\omega}]_t \in \overline{\Omega}$, and hence is also countably generated. Moreover, in view of Dellacherie & Meyer [34, Theorem IV-64], we have

$$\overline{\mathcal{F}}_{T_k} \;=\; \sigma\Big(B_{T_k\wedge\cdot},\ T_1, \cdots, T_k\Big),$$

and hence $\overline{\mathcal{F}}_{T_k}$ is countably generated. Therefore, it follows by Stroock & Varadhan [98, Theorem 1.1.6] that, every probability measure $\overline{\mathbb{P}}$ on $(\overline{\Omega}, \overline{\mathcal{F}})$ admits a family of r.c.p.d. with respect to the σ-field $\overline{\mathcal{F}}_{T_k}$ satisfying the condition of (ii). $\square$

5.2 On the optimal stopping problem

We next recall some useful results from the classical optimal stopping theory, see *e.g.* El Karoui [38], Peskir & Shiryaev [87], Karatzas & Shreve [71]. Let $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ be an abstract complete probability space, which is equipped with a filtration $\mathbb{F}^* = (\mathcal{F}^*_t)_{t\ge 0}$ satisfy the usual conditions. Denote $\mathcal{F}^*_\infty := \bigvee_{t\ge 0} \mathcal{F}^*_t$ and by $\mathcal{T}^*$ the class of all $\mathbb{F}^*-$stopping times taking value in $[0, +\infty)$. Let Y be an $\mathbb{F}^*-$optional process defined on Ω^* of class (D).

For every $\tau \in \mathcal{T}^*$, we denote by $\mathcal{T}^*_\tau$ the collection of all stopping times σ in $\mathcal{T}^*$ such that $\sigma \ge \tau$. We then define a family of random variables

$$Z^0_\tau \;:=\; \operatorname*{ess\,sup}_{\sigma\in\mathcal{T}^*_\tau} \mathbb{E}\Big[Y_\sigma\Big|\mathcal{F}_\tau\Big] \text{ for all } \tau \in \mathcal{T}^*.$$

Then by the dynamic programming principle, the family $(Z^0_\tau)_{\tau\in\mathcal{T}^*}$ is a supermartingale system, *i.e.* $Z^0_\sigma \ge \mathbb{E}[Z^0_\tau|\mathcal{F}^*_\sigma]$ for all stopping times $\sigma, \tau \in \mathcal{T}^*$ such that in $\sigma \le \tau$. Using Dellacherie & Lenglart [32, Theorem 6 and Remark 7c)], it follows that one can find a làdlàg optional process $Z = (Z_t)_{t\ge 0}$ which aggregates the family $(Z^0_\tau)_{\tau\in\mathcal{T}^*}$, *i.e.*

$$Z_\tau \;=\; Z^0_\tau \text{ for all } \tau \in \mathcal{T}^*,\ \mathbb{P}^* \text{ - a.s.}$$

In particular, $Z = (Z_t)_{t\ge 0}$ is a strong supermartingale of class (D), and it is called the Snell envelope of process Y, or equivalently the minimum strong supermartingale dominating the optional process Y, *i.e.* $Z_0 = \operatorname{ess\,sup}_{\tau\in\mathcal{T}^*} \mathbb{E}[Y_\tau|\mathcal{F}_0]$ and $Z_\tau \ge Y_\tau$, $\mathbb{P}^*$ - almost surely for all

$\tau \in \mathcal{T}^*$. Using the optional cross-section theorem, see [34, Theorem IV.86], it follows that

$$Z_t \;\geq\; Y_t \text{ for all } t \in \mathbb{R}_+,\; \mathbb{P}^* \text{ - a.s.}$$

We summarize the above facts in the following lemma.

Lemma IV.5.3. *Let Y be an $\mathbb{F}^*-$optional process of class (D), then there is an $\mathbb{F}^*-$optional làdlàg process Z, which is the smallest strong supermartingale such that $Z_0 = \operatorname{ess\,sup}_{\tau\in\mathcal{T}^*} \mathbb{E}[Y_\tau|\mathcal{F}_0]$ and $Z_t \geq Y_t$ for all $t \geq 0$, $\mathbb{P}^*$ - almost surely. In particular, one has $\mathbb{E}[Z_0] = \sup_{\tau\in\mathcal{T}^*} \mathbb{E}[Y_\tau]$.*

We next recall the Doob-Meyer decomposition for supermartingales without right continuity, see *e.g.* [34, Theorem 20, Appendix I] or Mertens [77, Theorem T3].

Lemma IV.5.4. *Let $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ be a probability space equipped with a filtration $\mathbb{F}^* = (\mathcal{F}^*_t)_{t\geq 0}$ satisfying the usual conditions, $X = (X_t)_{t\geq 0}$ be an $\mathbb{F}^*-$optional process class (DL) and an $\mathbb{F}^*-$supermartingale*[2]*. Then X has a unique decomposition $X = X_0 + M - A$, where $M_0 = A_0 = 0$, M is a càdlàg $\mathbb{F}^*-$martingale, A is an $\mathbb{F}^*-$predictable increasing process.*

The above decomposition allows to define the quadratic co-variation of a làdlàg supermartingale with a continuous martingale in a pathwise way, as in Karandikar [70]. Let us stay in the context of Lemma IV.5.4, and assume that W is a continuous martingale defined on $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*, \mathbb{F}^*)$. We denote by $\mathbb{F}^{X,W} = (\mathcal{F}^{X,W}_t)_{t\geq 0}$ the filtration generated by (X, W), *i.e.* $\mathcal{F}^{X,W}_t := \sigma(X_s, W_s\ s \leq t)$. Next, define

$$X^+_t \;:=\; \lim_{\mathbb{Q}\ni s\searrow t} X_s \;=\; X_0 \;+\; M_t \;+\; A^+_t \text{ with } A^+_t \;:=\; \lim_{\mathbb{Q}\ni s\searrow t} A_s.$$

Then X^+ is clearly still an $\mathbb{F}^*-$supermartingale, and has càdlàg paths almost surely. Let $\tau^n_0 := 0$, $\tau^n_{i+1} := \inf\{t \geq \tau^n_i\ : |X^+_t - X^+_{\tau^n_i}| \geq 2^{-n} \text{ or } |W_{\tau^n_i} - W_t| \geq 2^{-n}\}$,

$$Q^n_t \;:=\; \sum_{i=0}^{\infty} \Big(X^+_{\tau^n_{i+1}\wedge t} - X^+_{\tau^n_i\wedge t}\Big)\Big(W_{\tau^n_{i+1}\wedge t} - W_{\tau^n_i\wedge t}\Big) \quad\text{and}\quad Q_t \;:=\; \limsup_{n\to\infty} Q^n_t. \qquad \text{(IV.5.2)}$$

Define finally $Q^-_0 := Q_0$ and $Q^-_t := \lim_{\mathbb{Q}\ni s\nearrow t} Q_s$ for all $t > 0$.

Lemma IV.5.5. *The process Q^- is indistinguishable from the quadratic co-variation $\langle X, W\rangle$ of X and W. Moreover, Q^- is $\mathbb{F}^{X,W}-$predictable.*

2. Here X may not be of class (D), and X is an $\mathbb{F}^*-$supermartingale if $\mathbb{E}[X_\tau|\mathcal{F}_\sigma] \leq X_\sigma$ for all bounded $\mathbb{F}^*-$stopping times $\sigma \leq \tau$.

Proof. Notice that X^+ is càdlàg, $\mathbb{P}^*$ - almost surely, then following Theorem 3 of Karandikar [70], the process $(Q_t)_{t\geq 0}$, taking values in $(-\infty, +\infty]$, is indistinguishable from the quadratic co-variation between X^+ and W. Since A^+ has finite variation and W is continuous, then Q is also the quadratic co-variation between X and W, and moreover Q is continuous, $\mathbb{P}^*$ - almost surely. Then Q^- and Q are indistinguishable.

Further, by its construction, it is clear that Q^- is $\mathbb{F}^{X,W+}-$adapted, where $\mathbb{F}^{X,W+} = (\mathcal{F}_t^{X,W+})_{t\geq 0}$ is the right-continuous filtration defined by $\mathcal{F}_t^{X,W+} := \lim_{s\searrow t} \mathcal{F}_s^{X,W}$. Since Q^- is left-continuous, it follows that Q^- is $\mathbb{F}^{X,W+}-$predictable, or equivalently, $\mathbb{F}^{X,W}-$predictable. □

We finally provide an equivalence result of the optimal stopping problems. Let $\mathbb{G}^* = (\mathcal{G}_t^*)_{t\geq 0}$ be another filtration on $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$ satisfying the usual conditions such that $\mathcal{F}_t^* \subseteq \mathcal{G}_t^*$ for all $t \in \mathbb{R}_+$. Denote also by $\mathcal{T}_{\mathbb{G}^*}$ the collection of all $\mathbb{G}^*-$stopping times. Let Y be an $\mathbb{F}^*-$optional process defined on Ω^* of class (D).

Assumption IV.5.6. *For every $t \in \mathbb{R}_+$ and every $\mathcal{G}_t-$measurable bounded random variable X, the following equality holds :*

$$\mathbb{E}\Big[X\Big|\mathcal{F}_t^*\Big] \;=\; \mathbb{E}\Big[X\Big|\mathcal{F}_\infty^*\Big],\ \mathbb{P}^* \text{ - } a.s.$$

Lemma IV.5.7. *Under Assumption IV.5.6 we have*

$$\sup_{\tau\in\mathcal{T}^*} \mathbb{E}\Big[Y_\tau\Big] \;=\; \sup_{\tau\in\mathcal{T}_{\mathbb{G}^*}} \mathbb{E}\Big[Y_\tau\Big].$$

Proof. The result follows by Theorem 5 of Szpirglas & Mazziotto [99]. Notice that in [99], Y is assumed to be làdlàg, it can be easily generalized for any optional process by considering the Snell envelop of Y with respect to the filtration $\mathbb{F}^*$, since its Snell envelop is làdlàg, $\mathbb{P}^*$ - almost surely. Further, Y is also assumed to be positive in [99], which induces immediately the same result when Y is of class (D) since and process of class (D) can be dominated from below by a uniformly integrable martingale. □

Chapitre V

Monotonicity principle and stability

1 Introduction

This chapter is a continuation of the previous one. The main contribution is twofold : We provide an alternative proof of the monotonicity principle introduced by Beiglböck, Cox & Huesmann [5] and show a stability result using Lévy-Prokhov and Wasserstein metrics.

It is well known that, numerous solutions of the SEP enjoy an optimality property with respect to some criterion. Beiglböck, Cox & Huesmann [5] introduced the monotonicity principle which characterizes the optimal embeddings by their geometric support, such as Azéma-Yor's solution and Root's solution. To the best of our knowledge, all well known solutions to the SEP with optimality properties can be interpreted through this unifying principle. Our argument follows the classical proof of the monotonicity principle for the classical optimal transport problem, see Villani [104, Chapter 5] and the corresponding adaptation by Zaev [105, Theorem 3.6] for the martingale version in Beiglböck & Juillet [9]. Based on the duality in the previous chapter, we prove the principle by a delicate application of the optional cross-section theorem, and a clever conditioning argument introduced in [5].

Next, we consider another optimization problem similar to the optimal SEP. Motivated by the model-independent pricing of derivatives consistent with the real market, we consider a new optimal embedding problem, where the embedded Brownian motion needs only to reproduce a finite number of prices of call options. A stability result is provided, *i.e.* when more and more call options are quoted, the optimization problem converges to an optimal Skorokhod embedding problem. In addition, by means of different metrics on the space of probability measures, a convergence rate analysis is provided under suitable conditions.

Chapter V is organized as follows. In Section 2, we formulate the monotonicity principle in terms of probability measures on $\overline{\Omega}$. Then we provide all the proofs in Section 3. Next, we turn to another optimization problem in Section 4, where we show its convergence to the optimal

SEP. Finally, a convergence rate analysis is given in Section 5.

2 Monotonicity principle of optimal Skorokhod embedding problem

2.1 Preliminaries

Recall that Ω is the space of continuous functions $\omega = (\omega_t)_{t\geq 0}$ on $\mathbb{R}_+$ such that $\omega_0 = 0$, $B = (B_t)_{t\geq 0}$ is the canonical process, and $\mathbb{F}^0 = (\mathcal{F}^0_t)_{t\geq 0}$ is its canonical filtration. Notice that Ω is a Polish space under the compact convergence topology, and its Borel σ-field is given by $\mathcal{F}^0 := \bigvee_{t\geq 0} \mathcal{F}^0_t$. Denote by $\mathcal{P}(\Omega)$ the space of probability measures on $(\Omega, \mathcal{F}^0)$ and by $\mathbb{P}_0 \in \mathcal{P}(\Omega)$ the Wiener measure.

Recall also that $\overline{\Omega} := \Omega \times \mathbb{R}_+$ ($\Theta = \mathbb{R}_+$ when $m = 1$) denotes the enlarged space, (B, T) denotes the canonical elements defined by $B_t(\bar{\omega}) := \omega_t$ and $T(\bar{\omega}) := \theta$ for all $\bar{\omega} = (\omega, \theta) \in \overline{\Omega}$ and $t \in \mathbb{R}_+$, and $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\geq 0}$ denotes the canonical filtration defined by $\overline{\mathcal{F}}_t := \sigma(B_s, s \leq t) \bigvee \sigma(\{T \leq s\}, s \leq t)$. In particular, the canonical variable T is an $\overline{\mathbb{F}}-$stopping time and $\overline{\mathcal{F}}_T$ is a σ-field on $\overline{\Omega}$. Recall that $\overline{\mathcal{F}}^B := \sigma(B_t, t \geq 0)$ is the σ-field on $\overline{\Omega}$ generated by B. Under the product topology, $\overline{\Omega}$ is still a Polish space, and its Borel σ-field is given by $\overline{\mathcal{F}} := \bigvee_{t\geq 0} \overline{\mathcal{F}}_t$.

Recall that $\overline{\mathcal{P}}(\overline{\Omega})$ is the set of probability measures on $(\overline{\Omega}, \overline{\mathcal{F}})$ and $\overline{\mathcal{P}} \subseteq \overline{\mathcal{P}}(\overline{\Omega})$ is the subset defined by

$$\overline{\mathcal{P}} \;\; := \;\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) : \; B \text{ is an } \overline{\mathbb{F}} - \text{Brownian motion and } B_{T\wedge\cdot} \text{ is UI under } \overline{\mathbb{P}}\Big\}. \quad \text{(V.2.1)}$$

For a centered distribution μ on $\mathbb{R}$, the set of embeddings is given by

$$\overline{\mathcal{P}}(\mu) \;\; := \;\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}} : \; \overline{\mathbb{P}} \circ B_T^{-1} \; = \; \mu\Big\}.$$

Let $\Phi : \overline{\Omega} \to \mathbb{R}$ be Borel measurable and non-anticipative, then the optimal SEP is defined by

$$\overline{P}(\mu) \;\; := \;\; \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big]. \quad \text{(V.2.2)}$$

To introduce the monotonicity principle, we need the following notations. For all $\bar{\omega} = (\omega, \theta), \bar{\omega}' = (\omega', \theta') \in \overline{\Omega}$, we define the concatenation $\bar{\omega} \otimes \bar{\omega}' \in \overline{\Omega}$ by

$$\bar{\omega} \otimes \bar{\omega}' \;\; := \;\; \Big(\omega \otimes_\theta \omega', \; \theta + \theta'\Big),$$

where

$$(\omega \otimes_\theta \omega')_t \;\; := \;\; \omega_t \mathbb{1}_{[0,\theta)}(t) \; + \; \Big(\omega_\theta \; + \; \omega'_{t-\theta}\Big) \, \mathbb{1}_{[\theta,+\infty)}(t), \;\text{ for all } t \in \mathbb{R}_+.$$

Recall that Λ is the space of all continuous functions on $\mathbb{R}$ with linear growth and $\mathcal{S}^0$ is the space of $\mathbb{F}^0-$supermartingales $S = (S_t)_{t \geq 0}$ (which are automatically làdlàg $\mathbb{P}_0$ - almost surely) defined on $(\Omega, \mathcal{F}^0, \mathbb{P}_0)$ such that $S_0 = 0$ and for some $C > 0$,

$$|S_t(\omega)| \;\; \leq \;\; C(1 + |\omega_t|) \text{ for all } (\omega, t) \in \overline{\Omega}. \tag{V.2.3}$$

Recall also

$$\overline{\mathcal{D}}'' \;\; := \;\; \Big\{(\lambda, S) \in \Lambda \times \mathcal{S}^0 : \; \lambda(\omega_t) \; + \; S_t(\omega) \; \geq \; \Phi(\omega, t) \text{ for all } t \in \mathbb{R}_+, \; \mathbb{P}_0 \text{ - a.s.}\Big\},$$

and the dual problem is given by

$$\overline{D}''(\mu) \;\; := \;\; \inf_{(\lambda,S) \in \overline{\mathcal{D}}''} \; \mu(\lambda). \tag{V.2.4}$$

2.2 Monotonicity principle

We now introduce the monotonicity principle formulated in Beiglböck, Cox & Huesmann[5], which provides a geometric characterization of the optimal embedding of (V.2.2) in terms of its support.

Given a set $\overline{\Gamma} \subseteq \overline{\Omega}$, we define $\overline{\Gamma}^<$ by

$$\overline{\Gamma}^< \;\; := \;\; \Big\{\bar{\omega} = (\omega, \theta) \in \overline{\Omega} : \; \omega_{\theta \wedge \cdot} \; = \; \omega'_{\theta \wedge \cdot} \text{ for some } \bar{\omega}' = (\omega', \theta') \in \overline{\Gamma} \text{ with } \theta' > \theta\Big\}.$$

Definition V.2.1. *A pair $(\bar{\omega}, \bar{\omega}') \in \overline{\Omega} \times \overline{\Omega}$ is said to be a stop-go pair if $\omega_\theta = \omega'_{\theta'}$ and*

$$\Phi(\bar{\omega}) \; + \; \Phi(\bar{\omega}' \otimes \bar{\omega}'') \;\; > \;\; \Phi(\bar{\omega} \otimes \bar{\omega}'') \; + \; \Phi(\bar{\omega}') \text{ for all } \bar{\omega}'' \in \overline{\Omega}^+,$$

where $\overline{\Omega}^+ := \{\bar{\omega} = (\omega, \theta) \in \overline{\Omega} : \theta > 0\}$. Denote by SG *the set of all stop-go pairs.*

The following monotonicity principle is proved in [5].

Theorem V.2.2. *Let $\Phi : \overline{\Omega} \to \mathbb{R}$ be Borel measurable and non-anticipative. Assume that the optimal Skorokhod embedding problem* (V.2.2) *admits an optimizer $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\mu)$, and the duality*

$\overline{P}(\mu) = \overline{D}''(\mu)$ *holds. Then there exists a Borel subset* $\overline{\Gamma}^* \subseteq \overline{\Omega}$ *such that*

$$\overline{\mathbb{P}}^*\Big[\overline{\Gamma}^*\Big] \;=\; 1 \quad \textit{and} \quad \mathrm{SG} \,\cap\, \Big(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\Big) \;=\; \emptyset.$$

Remark V.2.3. *The above monotonicity principle has been proved in [5], without using the duality. However, their conditions to ensure the existence of optimizers yield automatically the above duality.*

3 Proof of the monotonicity principle

Throughout this section, let $\overline{\mathbb{P}}^*$ be an optimizer of problem (V.2.2) in the context of Theorem V.2.2.

3.1 A heuristic proof

We start with a purely heuristic argument to illustrate the essential idea in our proof. Suppose that there exists a dual optimizer (λ^*, S^*) of (V.2.4), *i.e.*

$$\lambda^*(\omega_t) \;+\; S^*_t(\omega) \;\geq\; \Phi(\omega, t) \text{ for all } t \in \mathbb{R}_+,\ \mathbb{P}_0 \text{ - a.s. and } \mu(\lambda^*) \;=\; \mathbb{E}^{\overline{\mathbb{P}}^*}\Big[\Phi(B, T)\Big], \quad \text{(V.3.1)}$$

which implies that $\overline{\Gamma} := \{(\omega, \theta) : \lambda^*(\omega_\theta) + S^*_\theta(\omega) = \Phi(\omega, \theta)\}$ has full measure under $\overline{\mathbb{P}}^*$. Assume also for simplicity that S^* is a $\mathbb{P}_0-$martingale, then we claim that $(\overline{\Gamma}^< \times \overline{\Gamma}) \cap \mathrm{SG} = \emptyset$. Otherwise, any pair $(\bar\omega, \bar\omega') \in (\overline{\Gamma}^< \times \overline{\Gamma}) \cap \mathrm{SG}$ satisfies the condition

$$\Phi(\bar\omega) \;+\; \Phi(\bar\omega' \otimes \bar\omega'') \;\;>\;\; \Phi(\bar\omega \otimes \bar\omega'') \;+\; \Phi(\bar\omega') \text{ for all } \bar\omega'' \in \overline{\Omega}^+.$$

Let $\overline{\mathbb{Q}}^*_{\bar\omega}$ be the conditional probability of $\overline{\mathbb{P}}^*$ given $\{B_{\theta\wedge\cdot} = \omega_{\theta\wedge\cdot} \text{ and } T > \theta\}$. Then it follows that

$$\xi(\bar\omega) \;+\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega' \otimes \cdot)\Big] \;\;>\;\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega \otimes \cdot)\Big] \;+\; \xi(\bar\omega').$$

On the other hand, notice that the marginal distribution of $\overline{\mathbb{Q}}^*_{\bar\omega}$ on Ω is still a Wiener measure. Then denoting $(S^* + \lambda^*)(\omega, \theta) := S^*_\theta(\omega) + \lambda^*(\omega_\theta)$, one has from (V.3.1) that

$$\Phi(\bar\omega) \;+\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega' \otimes \cdot)\Big] \;\;\leq\;\; (S^* + \lambda^*)(\bar\omega) \;+\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[(S^* + \lambda^*)(\bar\omega' \otimes \cdot)\Big].$$

Since S^* is a martingale and $\omega_\theta = \omega'_{\theta'}$ by the definition of SG, it follows that

$$(S^*+\lambda^*)(\bar\omega) \;+\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[(S^*+\lambda^*)(\bar\omega'\otimes\cdot)\Big] \;=\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[(S^*+\lambda^*)(\bar\omega\otimes\cdot)\Big] \;+\; (S^*+\lambda^*)(\bar\omega').$$

Finally, using again the definitions of SG and $\overline{\mathbb{Q}}^*_{\bar\omega}$, one knows that $\overline{\mathbb{Q}}^*_{\bar\omega}[\{\bar\omega'' : \bar\omega\otimes\bar\omega''\in\overline{\Gamma}\}] = 1$ and $\bar\omega'\in\overline{\Gamma}$, then

$$\begin{aligned}\Phi(\bar\omega) \;+\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega'\otimes\cdot)\Big] \;&\le\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[(S^*+\lambda^*)(\bar\omega\otimes\cdot)\Big] \;+\; (S^*+\lambda^*)(\bar\omega')\\ &=\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega\otimes\cdot)\Big] \;+\; \Phi(\bar\omega'),\end{aligned}$$

which is a contradiction. □

Remark V.3.1. *The main technical problem in the above proof arises from the conditional probability $\overline{\mathbb{Q}}^*_{\bar\omega}$ of $\overline{\mathbb{P}}^*$ given $\{B_{\theta\wedge\cdot} = \omega_{\theta\wedge\cdot}$ and $T>\theta\}$, which should be defined with respect to a σ-field in an almost sure way, creating too many $\overline{\mathbb{P}}^*-$null sets to control.*

3.2 An enlarged stop-go set

Notice that by Definition V.2.1, the set SG is a universally measurable set, or co-analytic set, but not a Borel set *a priori*. To overcome some measurability difficulty, we will consider as in [5] another set $\mathrm{SG}^*\subseteq\overline{\Omega}\times\overline{\Omega}$ that is Borel.

For the optimizer $\overline{\mathbb{P}}^*$ of the problem (V.2.2), one has a family of regular conditional probability distributions (r.c.p.d.), see *e.g.* Stroock & Varadhan [98], $(\overline{\mathbb{P}}^*_{\bar\omega})_{\bar\omega\in\overline{\Omega}}$ with respect to $\overline{\mathcal{F}}^B = \sigma(B_t, t\ge 0)$ on $\overline{\Omega}$. Notice that for all $\bar\omega = (\omega,\theta)$, the measure $\overline{\mathbb{P}}^*_{\bar\omega}$ is independent of θ, and thus we may denote this family by $(\overline{\mathbb{P}}^*_{\omega})_{\omega\in\Omega}$. In particular, one has $\overline{\mathbb{P}}^*_{\omega}[B=\omega]=1$ for all $\omega\in\Omega$. Next, for every $\bar\omega\in\overline{\Omega}$, define a probability $\overline{\mathbb{Q}}^1_{\bar\omega}$ on $(\overline{\Omega},\overline{\mathcal{F}})$ by

$$\overline{\mathbb{Q}}^1_{\bar\omega}[\overline{A}] \;:=\; \int_\Omega \overline{\mathbb{P}}^*_{\omega\otimes_\theta\omega'}[\overline{A}]\mathbb{P}_0(d\omega') \text{ for all } \overline{A}\in\overline{\mathcal{F}}. \tag{V.3.2}$$

Intuitively, $\overline{\mathbb{Q}}^1_{\bar\omega}$ is the conditional probability with respect to the event $\{B_{\theta\wedge\cdot}=\omega_{\theta\wedge\cdot}\}$. We next define, for every $\bar\omega\in\overline{\Omega}$, a probability $\overline{\mathbb{Q}}^2_{\bar\omega}$ by

$$\overline{\mathbb{Q}}^2_{\bar\omega}[\overline{A}] \;:=\; \overline{\mathbb{Q}}^1_{\bar\omega}\Big[\overline{A}\Big|T>\theta\Big]\mathbb{1}_{\{\overline{\mathbb{Q}}^1_{\bar\omega}[T>\theta]>0\}} \;+\; \mathbb{P}^{\theta,\omega}_0\otimes\delta_{\{\theta\}}[\overline{A}]\mathbb{1}_{\{\overline{\mathbb{Q}}^1_{\bar\omega}[T>\theta]=0\}} \tag{V.3.3}$$

for all $\overline{A}\in\overline{\mathcal{F}}$, where $\mathbb{P}^{\theta,\omega}_0$ is the shifted Wiener measure on $(\Omega,\mathcal{F}^0)$ defined by

$$\mathbb{P}^{\theta,\omega}_0[A] \;:=\; \mathbb{P}_0\Big[\omega\otimes_\theta B\in A\Big] \text{ for all } A\in\mathcal{F}^0.$$

We finally introduce a shifted probability $\overline{\mathbb{Q}}^*_{\bar\omega}$ by

$$\overline{\mathbb{Q}}^*_{\bar\omega}[\overline{A}] \;:=\; \overline{\mathbb{Q}}^2_{\bar\omega}\Big[\bar\omega \otimes (B,T) \in \overline{A}\Big] \text{ for all } \overline{A} \in \overline{\mathcal{F}},$$

and then define a new set SG^* by

$$\mathrm{SG}^* \;:=\; \Big\{(\bar\omega,\bar\omega') : \; \omega_\theta = \omega'_{\theta'} \text{ and } \Phi(\bar\omega) + \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega' \otimes \cdot)\Big] \;>\; \mathbb{E}^{\overline{\mathbb{Q}}^*_{\bar\omega}}\Big[\Phi(\bar\omega \otimes \cdot)\Big] + \Phi(\bar\omega')\Big\}. \tag{V.3.4}$$

Lemma V.3.2. (i) *The set* $\mathrm{SG}^* \subseteq \overline{\Omega} \times \overline{\Omega}$ *defined by* (V.3.4) *is* $\overline{\mathcal{F}}_T \otimes \overline{\mathcal{F}}_T-$*measurable.*

(ii) *Let* $\tau \leq T$ *be an* $\overline{\mathbb{F}}-$*stopping time, then the family* $(\hat{\mathbb{P}}_{\bar\omega})_{\bar\omega \in \overline{\Omega}}$ *defined by*

$$\hat{\mathbb{P}}_{\bar\omega} \;:=\; \mathbb{1}_{\{\tau(\bar\omega)<\theta\}} \overline{\mathbb{Q}}^2_{(\omega,\tau(\bar\omega))} \;+\; \mathbb{1}_{\{\tau(\bar\omega)=\theta\}} \mathbb{P}_0^{\tau(\bar\omega),\omega} \otimes \delta_{\{\theta\}}$$

is a family of r.c.p.d. of $\overline{\mathbb{P}}^*$ *with respect to* $\overline{\mathcal{F}}_\tau$*, i.e.* $\bar\omega \mapsto \hat{\mathbb{P}}_{\bar\omega}$ *is* $\overline{\mathcal{F}}_\tau-$*measurable, and for all bounded* $\overline{\mathcal{F}}-$*measurable random variable* ζ*, one has* $\mathbb{E}^{\overline{\mathbb{P}}^*}[\zeta|\overline{\mathcal{F}}_\tau](\bar\omega) = \hat{\mathbb{P}}_{\bar\omega}[\zeta]$ *for* $\overline{\mathbb{P}}^*$ *- almost every* $\bar\omega \in \overline{\Omega}$*.*

Proof. (i) Let us denote $[\omega]_t := \omega_{t\wedge\cdot}$, $[\theta]_t := \theta\mathbb{1}_{\{\theta\leq t\}} + \infty\mathbb{1}_{\{\theta>t\}}$ and $[\bar\omega]_t := ([\omega]_t, [\theta]_t)$. Then by Lemma IV.5.2, a process $Y : \mathbb{R}_+ \times \overline{\Omega} \to \mathbb{R}$ is $\overline{\mathbb{F}}-$optional, if and only if it is $\mathcal{B}(\mathbb{R}_+) \otimes \overline{\mathcal{F}}-$measurable and $Y_t(\bar\omega) = Y_t([\bar\omega]_t)$ for all $\bar\omega \in \overline{\Omega}$. Further, using Theorem IV-64 of Dellacherie & Meyer [34], a random variable X is $\overline{\mathcal{F}}_T-$measurable if and only if it is $\overline{\mathcal{F}}-$measurable and $X(\bar\omega) = X([\omega]_\theta, \theta)$ for all $\bar\omega \in \overline{\Omega}$.

Next, by the definitions of $\overline{\mathbb{Q}}^1_{\bar\omega}$, $\overline{\mathbb{Q}}^2_{\bar\omega}$ and $\overline{\mathbb{Q}}^*_{\bar\omega}$, it is easy to see that the maps $\bar\omega \mapsto (\overline{\mathbb{Q}}^1_{\bar\omega}, \overline{\mathbb{Q}}^2_{\bar\omega}, \overline{\mathbb{Q}}^*_{\bar\omega})$ are $\overline{\mathcal{F}}-$measurable and $\overline{\mathbb{Q}}^*_{\bar\omega} = \overline{\mathbb{Q}}^*_{([\omega]_{\theta},\theta)}$ for all $\bar\omega \in \overline{\Omega}$. Then $\bar\omega \mapsto \overline{\mathbb{Q}}^*_{\bar\omega}$ is $\overline{\mathcal{F}}_T-$measurable, and hence SG^* is $\overline{\mathcal{F}}_T \otimes \overline{\mathcal{F}}_T-$measurable by (V.3.4).

(ii) Let $\tau \leq T$ be an $\overline{\mathbb{F}}-$stopping time, we claim that

$$\text{there is some } \mathbb{F}^0 - \text{stopping time } \tau_0 \text{ on } (\Omega, \mathcal{F}^0) \text{ such that } \tau(\bar\omega) \;=\; \tau_0(\omega) \wedge \theta. \tag{V.3.5}$$

Moreover, again by Theorem IV-64 of [34], we have

$$\overline{\mathcal{F}}_\tau \;=\; \sigma\Big(B_{\tau\wedge t}, \; t \geq 0\Big) \bigvee \sigma\Big(T\mathbb{1}_{\{\tau=T\}}, \; \{\tau < T\}\Big). \tag{V.3.6}$$

Let $(\hat{\mathbb{P}}^0_{\bar\omega})_{\bar\omega\in\overline{\Omega}}$ be a family of r.c.p.d. of $\overline{\mathbb{P}}^*$ with respect to $\overline{\mathcal{F}}_\tau$, which implies that

$$\hat{\mathbb{P}}^0_{\bar\omega}\Big[B_{\tau\wedge\cdot} = \omega_{\tau(\bar\omega)\wedge\cdot}\Big] = 1 \text{ for all } \bar\omega \in \overline{\Omega} \quad \text{and} \quad \hat{\mathbb{P}}^0_{\bar\omega}[T = \theta] = 1 \text{ for all } \bar\omega \in \{\tau = T\}.$$

It follows that for $\overline{\mathbb{P}}^*$ - almost every $\bar{\omega} \in \{\tau = T\}$, one has $\hat{\mathbb{P}}^0_{\bar{\omega}} = \hat{\mathbb{P}}_{\bar{\omega}} := \mathbb{P}_0^{\tau(\bar{\omega}),\omega} \otimes \delta_{\{\theta\}}$.

We next focus on the set $\{\tau < T\}$. Recall that $\overline{\mathbb{P}}^*_\omega$ is a family of r.c.p.d. of $\overline{\mathbb{P}}^*$ with respect to $\overline{\mathcal{F}}^B$ and $\overline{\mathbb{Q}}^1_{\bar{\omega}}$ is defined by (V.3.2). Then $(\overline{\mathbb{Q}}^1_{(\omega,\tau_0(\omega))})_{\bar{\omega}\in\overline{\Omega}}$ is a family of conditional probability measures of $\overline{\mathbb{P}}^*$ with respect to $\sigma(B_{\tau_0\wedge t}, t \geq 0)$. Further, by the representation of $\overline{\mathcal{F}}_\tau$ in (V.3.6), it follows that for $\overline{\mathbb{P}}^*$ - almost every $\bar{\omega} \in \{\tau < T\}$, one has $\hat{\mathbb{P}}^0_{\bar{\omega}} = \overline{\mathbb{Q}}^2_{\bar{\omega}}$.

We now prove claim (V.3.5). For all $\omega \in \Omega$ and $t \in \mathbb{R}_+$, we denote $\overline{A}_{\omega,t} := \{\bar{\omega}' \in \overline{\Omega} : \omega'_{t\wedge\cdot} = \omega_{t\wedge\cdot} \text{ and } \theta' > t\}$. Then it is clear that $\overline{A}_{\omega,t}$ is an atom in $\overline{\mathcal{F}}_t$, *i.e.* for any set $C \in \overline{\mathcal{F}}_t$, one has either $\overline{A}_{\omega,t} \subseteq C$ or $\overline{A}_{\omega,t} \cap C = \emptyset$. Let $\bar{\omega} \in \overline{\Omega}$ and $\theta' \in \mathbb{R}_+$ such that $\tau(\bar{\omega}) < \theta < \theta'$, then $\bar{\omega} \in \overline{A}_{\omega,t}$ and $(\omega, \theta') \in \overline{A}_{\omega,t}$ for all $t < \theta$. Let $t_0 := \tau(\bar{\omega})$, then $\bar{\omega} \in \overline{A}_{\omega,t_0}$ and $\bar{\omega} \in \{\tau = t_0\} \in \overline{\mathcal{F}}_{t_0}$, which implies that $(\omega,\theta') \in \overline{A}_{\omega,t_0} \subseteq \{\tau = t_0\}$ since $\overline{A}_{\omega,t_0}$ is an atom in $\overline{\mathcal{F}}_{t_0}$. It follows that $\tau(\omega,\theta') = \tau(\bar{\omega})$ for all $\theta' > \theta$ and $\bar{\omega} \in \overline{\Omega}$ such that $\tau(\bar{\omega}) < \theta$. Notice that for all $t \in \mathbb{R}_+$, $\{\bar{\omega} \in \overline{\Omega} : \tau(\bar{\omega}) \leq t\}$ is $\overline{\mathcal{F}}_t-$measurable, then by Doob's functional representation theorem, there exists some Borel measurable function $f : \Omega \times (\mathbb{R}_+ \cup \{+\infty\}) \to \mathbb{R}$ such that $\mathbb{1}_{\{\tau(\bar{\omega}) \leq t\}} = f([\omega]_t, [\theta]_t)$. It follows that for all $\theta_0 \in \mathbb{R}_+$, $\{\omega \in \overline{\Omega} : \tau(\omega,\theta_0) \leq t\}$ is $\mathcal{F}_t-$measurable, and hence $\tau(\omega,\theta_0)$ is an $\mathbb{F}^0-$stopping time on $(\Omega, \mathcal{F}^0)$. Then the random variable $\tau_0 : \Omega \to \mathbb{R}_+$ defined by $\tau_0(\omega) := \sup_{n\in\mathbb{N}} \tau(\omega, n)$ is the required $\mathbb{F}^0-$stopping time of claim (V.3.5).

Finally, we notice that by the definition, $\bar{\omega} \mapsto \hat{\mathbb{P}}_{\bar{\omega}}$ is $\overline{\mathcal{F}}-$measurable and $\hat{\mathbb{P}}_{\bar{\omega}} = \hat{\mathbb{P}}_{[\bar{\omega}]_\theta}$ for all $\bar{\omega} \in \overline{\Omega}$. Moreover, we have proved that $\hat{\mathbb{P}}^0_{\bar{\omega}} = \hat{\mathbb{P}}_{\bar{\omega}}$ for $\overline{\mathbb{P}}^*$ - almost every $\bar{\omega} \in \overline{\Omega}$, where $(\hat{\mathbb{P}}^0_{\bar{\omega}})_{\bar{\omega}\in\overline{\Omega}}$ is a family of r.c.p.d. of $\overline{\mathbb{P}}^*$ with respect to $\overline{\mathcal{F}}_\tau$. Therefore, $(\hat{\mathbb{P}}_{\bar{\omega}})_{\bar{\omega}\in\overline{\Omega}}$ is a family of conditional probability measures of $\overline{\mathbb{P}}^*$ with respect to $\overline{\mathcal{F}}_\tau$. $\square$

To prove Theorem V.2.2, we will first prove a closely related result as in [5].

Theorem V.3.3. *Let $\Phi : \overline{\Omega} \to \mathbb{R}$ be Borel measurable and non-anticipative. Assume that the optimal SEP (V.2.2) admits an optimizer $\overline{\mathbb{P}}^* \in \overline{\mathcal{P}}(\mu)$ and the duality $\overline{P}(\mu) = \overline{D}''(\mu)$ holds. Then there exists a Borel subset $\overline{\Gamma}^* \subseteq \overline{\Omega}$ such that*

$$\overline{\mathbb{P}}^*\left[\overline{\Gamma}^*\right] \;=\; 1 \quad \textit{and} \quad \mathrm{SG}^* \;\cap\; \left(\overline{\Gamma}^{*<} \times \overline{\Gamma}^*\right) \;=\; \emptyset.$$

3.3 Technical results

We define a projection operator $\Pi_S : \overline{\Omega} \times \overline{\Omega} \to \overline{\Omega}$ by

$$\Pi_S\big(\mathcal{K}\big) \;:=\; \Big\{\bar{\omega} : \text{ there exists some } \bar{\omega}' \in \overline{\Omega} \text{ such that } (\bar{\omega}, \bar{\omega}') \in \mathcal{K}\Big\}.$$

Proposition V.3.4. *Let the conditions in Theorem V.2.2 hold true and $\overline{\mathbb{P}}^*$ be the fixed optimizer. Then there is some Borel set $\overline{\Gamma}^*_0 \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}^*_0] = 1$ and for all $\overline{\mathbb{F}}-$stopping time*

$\tau \leq T$, *one has*

$$\overline{\mathbb{P}}^*\Big[\tau < T \quad and \quad (B_{\tau\wedge\cdot}, T) \in \Pi_S\Big(\mathrm{SG}^* \cap (\overline{\Omega} \times \overline{\Gamma}_0^*)\Big)\Big] \;=\; 0. \tag{V.3.7}$$

Proof. (i) Let us start with the duality result $\overline{P}(\mu) = \overline{D}''(\mu)$. By definition, we may find a minimizing sequence $\{(\lambda^n, S^n)\}_{n\geq 1} \subseteq \overline{\mathcal{D}}''$, such that $\mu(\lambda_n) \to \overline{D}''(\mu) = \overline{P}(\mu)$ as $n \to \infty$. Then, there is some $\Gamma_0 \subseteq \Omega$ such that $\mathbb{P}_0[\Gamma_0] = 1$ and

$$\eta^n(\bar{\omega}) \;:=\; \lambda^n(\omega_t) \;+\; S^n_t(\omega) \;-\; \Phi(\bar{\omega}) \;\geq\; 0 \text{ for all } \bar{\omega} \in \Gamma_0 \times \mathbb{R}_+. \tag{V.3.8}$$

Notice that $(S^n_t)_{t\geq 0}$ are all strong supermartingales on $(\Omega, \mathcal{F}^0, \mathbb{P}_0)$ satisfying (V.2.3), and thus strong supermartingales on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}^*)$ with respect to $\overline{\mathbb{F}}$. It follows that

$$0 \;\leq\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}^*}\Big[\eta^n\Big] \;=\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}^*}\Big[\lambda^n(B_T) \;+\; S^n_T \;-\; \Phi\Big] \;\leq\; \lim_{n\to\infty} \mu(\lambda^n) \;-\; \overline{P}(\mu) \;=\; 0. \tag{V.3.9}$$

Therefore, we can find some $\overline{\Gamma}_0 \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}_0] = 1$, and after possibly passing to a subsequence,

$$\lim_{n\to\infty} \eta^n(\bar{\omega}) \;=\; 0 \text{ for all } \bar{\omega} \in \overline{\Gamma}_0.$$

Moreover, since S^n are $\overline{\mathbb{F}}-$strong supermartingales on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}^*)$, then there is some Borel set $\overline{\Gamma}_1 \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}_1] = 1$, and for all $\bar{\omega} = (\omega, \theta) \in \overline{\Gamma}_1$, $\mathbb{P}_0[\omega \otimes_\theta B \in \Gamma_0] = 1$ and $(S^n_{\theta+t}(\omega \otimes_\theta \cdot))_{t\geq 0}$ are $\mathbb{P}_0-$strong supermartingales. Set $\overline{\Gamma}_0^* := \overline{\Gamma}_0 \cap \overline{\Gamma}_1$, and we show next that $\overline{\Gamma}_0^*$ is the required Borel set.

(ii) For a fixed pair

$$(\bar{\omega}, \bar{\omega}') \;\in\; \mathrm{SG}^* \;\cap\; \Big(\overline{\Omega} \times \overline{\Gamma}_0^*\Big),$$

let us define

$$\delta(\bar{\omega}'') \;:=\; \Big(\Phi(\bar{\omega}) \;+\; \Phi(\bar{\omega}' \otimes \bar{\omega}'')\Big) \;-\; \Big(\Phi(\bar{\omega} \otimes \bar{\omega}'') \;+\; \Phi(\bar{\omega}')\Big) \text{ for all } \bar{\omega}'' \in \overline{\Omega}.$$

By the definition of SG*, one has $\omega_\theta = \omega'_{\theta'}$. Then it follows that for all $\bar\omega'' \in \overline{\Omega}$,

$$\begin{aligned}
\delta(\bar\omega'') &= \lambda^n(\omega_\theta) + S^n_\theta(\omega) - \eta^n(\bar\omega) - \Big(\lambda^n(\omega'_{\theta'}) + S^n_{\theta'}(\omega') - \eta^n(\bar\omega')\Big) \\
&\quad + \lambda^n\Big(\omega'_{\theta'} + \omega''_{\theta''}\Big) + S^n_{\theta'+\theta''}(\omega' \otimes_{\theta'} \omega'') - \eta^n(\bar\omega' \otimes \bar\omega'') \\
&\quad - \Big(\lambda^n\Big(\omega_\theta + \omega''_{\theta''}\Big) + S^n_{\theta+\theta''}(\omega \otimes_\theta \omega'') - \eta^n(\bar\omega \otimes \bar\omega'')\Big) \\
&= S^n_\theta(\omega) - \eta^n(\bar\omega) + S^n_{\theta'+\theta''}(\omega' \otimes_{\theta'} \omega'') - \eta^n(\bar\omega' \otimes \bar\omega'') \\
&\quad - \Big(S^n_{\theta'}(\omega') - \eta^n(\bar\omega') + S^n_{\theta+\theta''}(\omega \otimes_\theta \omega'') - \eta^n(\bar\omega \otimes \bar\omega'')\Big) \\
&\le \Big(\eta^n(\bar\omega \otimes \bar\omega'') + \eta^n(\bar\omega')\Big) - \eta^n(\bar\omega' \otimes \bar\omega'') \\
&\quad + \Big(S^n_{\theta'+\theta''}(\omega' \otimes_{\theta'} \omega'') - S^n_{\theta'}(\omega')\Big) - \Big(S^n_{\theta+\theta''}(\omega \otimes_\theta \omega'') - S^n_\theta(\omega)\Big).
\end{aligned}$$

(iii) Let $\tau \le T$ be an $\overline{\mathbb{F}}-$stopping time and $(\hat{\mathbb{P}}_{\bar\omega})_{\bar\omega\in\overline{\Omega}}$ be the family of r.c.p.d. of $\overline{\mathbb{P}}^*$ with respect to $\overline{\mathcal{F}}_\tau$ introduced in Lemma V.3.2. Recall that $\hat{\mathbb{P}}^*_{\bar\omega} := \overline{\mathbb{Q}}^*_{(\omega,\tau(\bar\omega))}$ is the shifted probability measure for all $\bar\omega \in \{\tau < T\}$. By (V.3.9), and there is some set $\overline{\Gamma}^1_\tau$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}^1_\tau] = 1$ and

$$\lim_{n\to\infty} \mathbb{E}^{\hat{\mathbb{P}}_{\bar\omega}}\Big[\eta^n\Big] \;=\; 0 \text{ for all } \bar\omega \in \overline{\Gamma}^1_\tau. \tag{V.3.10}$$

Further, (V.3.9) implies that $0 \ge \mathbb{E}^{\overline{\mathbb{P}}^*}[S^n_T] \to 0$ as $n \to \infty$. Then it follows from the strong supermartingale property of S^n that

$$S^n_\tau - \mathbb{E}^{\overline{\mathbb{P}}^*}\Big[S^n_T\Big|\overline{\mathcal{F}}_\tau\Big] \ge 0,\ \overline{\mathbb{P}}^* \text{ - a.s. and } \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}^*}\Big[S^n_\tau - \mathbb{E}^{\overline{\mathbb{P}}^*}\Big[S^n_T\Big|\overline{\mathcal{F}}_\tau\Big]\Big] \le -\lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}^*}[S^n_T] = 0.$$

Hence there is some set $\overline{\Gamma}^2_\tau \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}^2_\tau] = 1$ and for all $\bar\omega \in \overline{\Gamma}^2_\tau$,

$$0 \le \lim_{n\to\infty} \Big(S^n_\tau(\bar\omega) - \mathbb{E}^{\hat{\mathbb{P}}_{\bar\omega}}[S^n_T]\Big) = 0, \tag{V.3.11}$$

after possibly taking some subsequence. Moreover, by the definition of $\overline{\mathcal{P}}$, B is an $\overline{\mathbb{F}}-$Brownian motion and $B_{T\wedge\cdot}$ is uniformy integrable under $\overline{\mathbb{P}}^*$, and the property still holds under the conditional probability measures. Then there is some measurable set $\overline{\Gamma}^3_\tau \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}^3_\tau] = 1$ and for all $\bar\omega \in \overline{\Gamma}^3_\tau \cap \{\tau < T\}$, one has

$$\hat{\mathbb{P}}^*_{\bar\omega}\Big[T > 0\Big] > 0,\ \hat{\mathbb{P}}^*_{\bar\omega}\Big[(\omega_{\tau(\bar\omega)\wedge\cdot}, \theta\wedge\tau(\bar\omega)) \otimes (B,T) \in \overline{\Gamma}^*_0\Big] = 1 \text{ and } \hat{\mathbb{P}}^*_{\bar\omega} \in \overline{\mathcal{P}}. \tag{V.3.12}$$

Set $\overline{\Gamma}^0_\tau := \overline{\Gamma}^1_\tau \cap \overline{\Gamma}^2_\tau \cap \overline{\Gamma}^3_\tau$, and in the rest of this proof, we show that

$$\Big((\overline{\Gamma}^0_\tau \cap \{\tau < T\}) \times \overline{\Omega}\Big) \cap \mathrm{SG}^* \cap \Big(\overline{\Omega} \times \overline{\Gamma}^*_0\Big) = \emptyset, \tag{V.3.13}$$

which justifies (V.3.7).
(iv) We finally prove (V.3.13) by contradiction. Let $(\bar{\omega}, \bar{\omega}') \in (\overline{\Gamma}^0_\tau \times \overline{\Omega}) \cap \mathrm{SG}^* \cap (\overline{\Omega} \times \overline{\Gamma}^*_0)$. Notice that $\bar{\omega}' = (\omega', \theta') \in \overline{\Gamma}^*_0 \subseteq \overline{\Gamma}_1$ and there is some constant C_n such that

$$\Big|S^n_{\theta'+T}(\omega' \otimes_\theta B)\Big| \;\leq\; C_n\Big|1 \;+\; \omega'_{\theta'} \;+\; B_T\Big|,$$

it follows by the supermartingale property, together with Fatou's lemma, that

$$\mathbb{E}^{\hat{\mathbb{P}}^*_{\bar{\omega}}}\Big[S^n_{\theta'+T}(\omega' \otimes_{\theta'} B)\Big] \;\leq\; S^n_{\theta'}(\omega').$$

Moreover, one has $\mathbb{E}^{\hat{\mathbb{P}}^*_{\bar{\omega}}}[\eta^n(\bar{\omega}' \otimes B)] \geq 0$. Further, using (V.3.12) then (V.3.9), (V.3.10) and (V.3.11), we obtain that

$$0 \;<\; \mathbb{E}^{\hat{\mathbb{P}}^*_{\bar{\omega}}}[\delta] \;\leq\; \mathbb{E}^{\hat{\mathbb{P}}^*_{\bar{\omega}}}\Big[\eta^n(\bar{\omega} \otimes (B,T))\Big] + \eta^n(\bar{\omega}') - \mathbb{E}^{\hat{\mathbb{P}}^*_{\bar{\omega}}}\Big[S^n_{\tau(\bar{\omega})+T}(\omega \otimes_{\tau(\bar{\omega})} B)\Big] - S^n_{\tau(\bar{\omega})}(\omega) \;\longrightarrow\; 0,$$

as $n \to \infty$, which is a contradiction. We hence conclude the proof. □

Assume that $\Pi_S(\mathrm{SG}^* \cap (\overline{\Omega} \times \overline{\Gamma}^*_0))$ is Borel measurable on $\overline{\Omega}$, then by Lemma IV.5.2 the set

$$\Big\{(t, \bar{\omega}, \bar{\omega}') \in \mathbb{R}_+ \times \overline{\Gamma}^*_0 \times \overline{\Gamma}^*_0 : \; t < \theta \text{ and } (\omega, t, \bar{\omega}') \in \mathrm{SG}^*\Big\}.$$

is $\overline{\mathbb{F}}-$optional. Using Proposition V.3.4 together with the classical optional cross-section theorem, see *e.g.* Theorem IV.86 of Dellacherie & Meyer [34], it follows immediately that there is some measurable set $\overline{\Gamma}^*_1 \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}^*_1] = 1$ and $\Pi_S(\mathrm{SG}^* \cap (\overline{\Omega} \times \overline{\Gamma}^*_0)) \cap \overline{\Gamma}^{*<}_1 = \emptyset$. However, when the set $\mathrm{SG}^* \cap (\overline{\Omega} \times \overline{\Gamma}^*_0)$ is a Borel set in $\overline{\Omega} \times \overline{\Omega}$, the projection set $\Pi_S(\mathrm{SG}^* \cap (\overline{\Omega} \times \overline{\Gamma}^*_0))$ is *a priori* $\mathcal{B}(\overline{\Omega})-$analytic in $\overline{\Omega}$, see Definition III.7 of [34]. Therefore, we need to adapt the arguments of the optional cross-section theorem to our context.

Denote by $\overline{\mathcal{O}}$ the optional σ-field with respect to the filtration $\overline{\mathbb{F}}$ on $\mathbb{R}_+ \times \overline{\Omega}$. Let E be some auxiliary space. For all $A \subseteq \mathbb{R}_+ \times \overline{\Omega} \times E$, we denote

$$\Pi_2(A) \;:=\; \Big\{\bar{\omega} : \; \text{there is some } (t,e) \in \mathbb{R}_+ \times E \text{ such that } (t, \bar{\omega}, e) \in A\Big\}$$

and

$$\Pi_{12}(A) \;:=\; \{(t, \bar{\omega}) : \; \text{there is some } e \in E \text{ such that } (t, \bar{\omega}, e) \in A\}.$$

Proposition V.3.5. *Let $\overline{\mathbb{P}}$ be a probability measure on $(\overline{\Omega}, \overline{\mathcal{F}})$ and $(E, \mathcal{E})$ be a Lusin measurable* *space*[1]*. Assume that $A \subseteq \mathbb{R}_+ \times \overline{\Omega} \times E$ is an $\overline{\mathcal{O}} \times \mathcal{E}-$measurable set. Then for every $\varepsilon > 0$, there is some $\overline{\mathbb{F}}-$stopping time τ such that $\overline{\mathbb{P}}[\tau < +\infty] \geq \overline{\mathbb{P}}[\Pi_2(A)] - \varepsilon$ and $(\tau(\bar{\omega}), \bar{\omega}) \in \Pi_{12}(A)$ whenever $\bar{\omega} \in \overline{\Omega}$ satisfies $\tau(\bar{\omega}) < +\infty$.*

Proof. We follow the argument of Theorem IV.84 in [34].

(i) Notice that every Lusin space is isomorphic to a Borel subset of $[0,1]$, see *e.g.* Theorem III.20 of [34], we can then suppose without loss of generality that $(E, \mathcal{E}) = ([0,1], \mathcal{B}([0,1]))$. Then the projection set $\Pi_{12}(A)$ is clearly $\overline{\mathcal{O}}-$analytic by Definition III.7 of [34].

(ii) Using the measurable section theorem, see Theorem III.44 of [34], there is an $\overline{\mathcal{F}}-$measurable random variable $R : \overline{\Omega} \to \mathbb{R} \cup \{+\infty\}$ such that $\overline{\mathbb{P}}[R < +\infty] = \overline{\mathbb{P}}[\Pi_2(A)]$ and $R(\bar{\omega}) < +\infty \Rightarrow (R(\bar{\omega}), \bar{\omega}) \in \Pi_{12}(A)$. The variable R is indeed a stopping time with respect to the completed filtration $\overline{\mathbb{F}}^{\overline{\mathbb{P}}}$, see *e.g.* Proposition 2.13 of [40], but not an $\overline{\mathbb{F}}-$stopping time *a priori*. We then need to modify R following the measure ν defined on $\mathcal{B}(\mathbb{R}_+) \otimes \overline{\mathcal{F}}$ by

$$\nu(G) \;:=\; \int_{\overline{\Omega}} \mathbb{1}_G(R(\bar{\omega}), \bar{\omega}) \mathbb{1}_{\{R<+\infty\}}(\bar{\omega}) \overline{\mathbb{P}}(d\bar{\omega}) \text{ for all } G \in \mathcal{B}(\mathbb{R}_+) \otimes \overline{\mathcal{F}}.$$

(iii) We continue by following the lines of the proof of Theorem IV.84 in [34]. Denote by ζ_0 the set of all stochastic intervals $[[\sigma, \tau[[$, where σ, τ are both $\overline{\mathbb{F}}-$stoping times such that $\sigma \leq \tau$. Denote further by ζ the closure of ζ_0 under finite union operation, then ζ is a Boolean algebra which generates the optional σ-filed $\overline{\mathcal{O}}$. Moreover, the debut of a set $C \in \zeta_\delta$ is almost surely equal to an $\overline{\mathbb{F}}-$stopping time, where ζ_δ denotes the smallest collection containing ζ and stable under countable intersection. Further, the projection set $\Pi_{12}(A)$ is $\overline{\mathcal{O}}-$analytic and hence $\overline{\mathcal{O}}-$universally measurable. Therefore, there exists a set $C \in \zeta_\delta$ contained in $\Pi_{12}(A)$ such that $\nu(C) \geq \nu(\Pi_{12}(A)) - \varepsilon$. Let τ_0 be the $\overline{\mathbb{F}}-$stopping time, which equals to the debut of C, $\overline{\mathbb{P}}$ - almost surely. Then define $\tau := \tau_0 \mathbb{1}_{\{(\tau_0(\bar{\omega}), \bar{\omega}) \in C\}}$, which is a new $\overline{\mathbb{F}}-$stopping time since $\{\bar{\omega} : (\tau_0(\bar{\omega}), \bar{\omega}) \in C\} \in \overline{\mathcal{F}}_{\tau_0}$ by Theorem IV.64 of [34]. Then τ is the required stopping time and we may conclude the proof. □

3.4 Proofs of Theorems V.3.3 and V.2.2

Proof of Theorem V.3.3. Let us define

$$A \;:=\; \Big\{(t, \bar{\omega}, \bar{\omega}') \in \mathbb{R}_+ \times \overline{\Gamma}_0^* \times \overline{\Gamma}_0^* : \; t < \theta \text{ and } ((\omega, t), \bar{\omega}')) \in \mathrm{SG}^*\Big\}. \tag{V.3.14}$$

1. A measurable space $(E, \mathcal{E})$ is said to be Lusin if it is isomorphic to a Borel subset of a compact metrizable space, see Definition III.16 of [34].

Since SG^* is $\overline{\mathcal{F}}_T \otimes \overline{\mathcal{F}}_T-$measurable, then it follows by Lemma IV.5.2 that, the set A satisfies the conditions in Proposition V.3.5 with $E = \overline{\Omega}$.

We next prove that $\Pi_2(A)$ has no mass under $\overline{\mathbb{P}}^*$. Indeed, if $\overline{\mathbb{P}}^*[\Pi_2(A)] > 0$, by Proposition V.3.5, there is some $\overline{\mathbb{F}}-$stopping time τ such that $(\tau(\bar{\omega}), \bar{\omega}) \in \Pi_{12}(A)$ for all $\bar{\omega} \in \{\tau < +\infty\}$. Notice that $\bar{\omega} \in \{\tau < +\infty\}$ implies that $(\tau(\bar{\omega}), \bar{\omega}) \in \Pi_{12}(A)$ and hence $\tau(\bar{\omega}) < T$ by the definition of A. Therefore, one has $\{\tau < +\infty\} = \{\tau < T\}$, and hence $\overline{\mathbb{P}}^*[\tau < +\infty] = \overline{\mathbb{P}}^*[\tau < T] > 0$. Notice further that $(\tau(\bar{\omega}), \bar{\omega}) \in \Pi_{12}(A)$ implies that $(\omega, \tau(\bar{\omega})) \in \Pi_S(\mathrm{SG}^*)$. We then have

$$0 \quad < \quad \overline{\mathbb{P}}^*[\tau < T] \quad = \quad \overline{\mathbb{P}}^*\Big[\tau < T \quad \text{and} \quad (B_{\tau\wedge\cdot}, T) \in \Pi_S\Big(\mathrm{SG}^* \cap \Big(\overline{\Omega} \times \overline{\Gamma}_0^*\Big)\Big)\Big].$$

This is a contradiction by Proposition V.3.4.

Since $\Pi_2(A)$ is a $\overline{\mathbb{P}}^*-$null set, then we may find a Borel set $\overline{\Gamma}_1^* \subseteq (\overline{\Omega} \setminus \Pi_2(A))$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}_1^*] = 1$ and $\Pi_S(\mathrm{SG}^*) \cap \overline{\Gamma}_1^{*<} = \emptyset$. Therefore, $\overline{\Gamma}^* := \overline{\Gamma}_0^* \cap \overline{\Gamma}_1^*$ is the required Borel subset of $\overline{\Omega}$ in Theorem V.3.3. $\square$

Proof of Theorem V.2.2. Let us define an $\overline{\mathbb{F}}-$optional process $Z : \mathbb{R}_+ \times \overline{\Omega}$ by

$$Z_t(\bar{\omega}) \quad := \quad \mathbb{1}_{\{t<\theta \text{ and } \overline{\mathbb{Q}}^*_{(\omega,t)}[T>0]=0\}}.$$

Let τ be an arbitrary $\overline{\mathbb{F}}-$stopping time, then $\overline{\mathbb{P}}^*[T > \tau] = 1$ implies that $\overline{\mathbb{Q}}^*_{(\omega,\tau(\bar{\omega}))}[T > 0] = 1$ for $\overline{\mathbb{P}}^*$ - almost every $\bar{\omega} \in \overline{\Omega}$. It follows that

$$Z_\tau \quad = \quad 0 \text{ for all } \overline{\mathbb{F}} - \text{stopping time } \tau, \ \overline{\mathbb{P}}^* \text{ - a.s.}$$

Using the optional cross-section theorem, one has a Borel set $\overline{\Gamma}_2^* \subseteq \overline{\Omega}$ such that $\overline{\mathbb{P}}^*[\overline{\Gamma}_2^*] = 1$ and

$$Z_t \quad = \quad 0 \text{ for all } t \in \mathbb{R}_+ \text{ and } \bar{\omega} \in \overline{\Gamma}_2^*.$$

It is clear that $\overline{\mathbb{Q}}^*_{\bar{\omega}}[\overline{\Omega}^+] > 0$ for all $\bar{\omega} \in \overline{\Omega}$, then it follows by definition that,

$$\mathrm{SG} \ \cap \ \Big(\overline{\Gamma}_2^{*<} \times \overline{\Gamma}_2^*\Big) \ \subseteq \ \mathrm{SG}^* \ \cap \ \Big(\overline{\Gamma}_2^{*<} \times \overline{\Gamma}_2^*\Big).$$

Hence, one can conclude the proof by setting $\overline{\Gamma}^* := \overline{\Gamma}_0^* \cap \overline{\Gamma}_1^* \cap \overline{\Gamma}_2^*$, where $\overline{\Gamma}_0^*$ and $\overline{\Gamma}_1^*$ are defined as same as in the proof of Theorem V.3.3. $\square$

Remark V.3.6. (i) *Proposition V.3.4 can be compared to Proposition 6.6 of [5], while the proofs are different. The proof of Proposition V.3.4 is in the same spirit of that for the monotonicity principle in classical optimal transport, see e.g. Chapter 5 of Villani [104], which is based on* *the duality.*

(ii) *Proposition V.3.5 should be compared to the filtered Kellerer lemma, see Proposition 6.7 of [5], where a key argument in their proof is Choquet's capacity theory. The proof of Proposition V.3.5 uses the optional cross-section theorem, which is based on a measurable section theorem, and the latter is also proved in [34] using Choquet's capacity theory, see also the review in [40].*

4 A new optimization problem

From Section 4 to the end of this chapter, we focus on another optimization problem. Indeed, from the financial viewpoint, call options are assumed to be ideally liquid in the above literature, *i.e.* the prices of call options are known for all strikes at some maturity, or equivalently, the underlying has the unique distribution of some maturity determined by the market. However, only a finite number of call options are quoted in practice, which leads to another optimization problem.

4.1 Another optimal embedding problem

Recall that $\overline{\mathcal{P}}(\overline{\Omega})$ is the space of probability measures on $(\overline{\Omega}, \overline{\mathcal{F}})$ and

$$\overline{\mathcal{P}} \;:=\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\overline{\Omega}) : B \text{ is an } \overline{\mathbb{F}}-\text{Brownian motion and } B_{T_m \wedge \cdot} \text{ is UI under } \overline{\mathbb{P}}\Big\}. \tag{V.4.1}$$

Then we define a new optimization problem, where the embedded Brownian motion is only required to be consistent with a finite number of market prices of call options. For the sake of clarity, we assume that the call options of different maturities have the same set of strikes. Let $\boldsymbol{K} := (K_i)_{1\le i\le n}$ be a vector of strikes with $K_1 < \cdots < K_n$, and $\boldsymbol{C} := (C_{i,j})_{1\le i\le m, 1\le j\le n}$ be the corresponding price matrix of call options, indexed by maturity and strike. Then a probability $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$ is called a $(\boldsymbol{K}, \boldsymbol{C})-$embedding if

$$\mathbb{E}^{\overline{\mathbb{P}}}\Big[(B_{T_i} - K_j)^+\Big] \;=\; C_{i,j}, \text{ for all } i = 1, \cdots, m \text{ and } j = 1, \cdots, n. \tag{V.4.2}$$

Denote by $\overline{\mathcal{P}}(\boldsymbol{K}, \boldsymbol{C})$ the collection of all $(\boldsymbol{K}, \boldsymbol{C})-$embeddings and by $\mathcal{A}(\boldsymbol{K}) \subset \mathbb{R}_+^{mn}$ the set of all matrices $\boldsymbol{C}$ such that $\overline{\mathcal{P}}(\boldsymbol{K}, \boldsymbol{C})$ is nonempty. Similarly, for each matrix $\boldsymbol{C} \in \mathcal{A}(\boldsymbol{K})$, we may define the optimal embedding problem by

$$\overline{P}(\boldsymbol{K}, \boldsymbol{C}) \;:=\; \sup_{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{K}, \boldsymbol{C})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B, T)\Big]. \tag{V.4.3}$$

Remark V.4.1. (i) *Davis & Hobson provided a characterization of $\mathcal{A}(\boldsymbol{K})$ in [29]. It follows by definition that $\mathcal{A}(\boldsymbol{K})$ is convex and $\boldsymbol{C} \mapsto \overline{P}(\boldsymbol{K}, \boldsymbol{C})$ is concave on $\mathcal{A}(\boldsymbol{K})$, then a classical result in convex analysis can be applied to obtain the corresponding dualities.*
(ii) *However, the set $\overline{\mathcal{P}}(\boldsymbol{K}, \boldsymbol{C})$ is generally not compact, see Example V.4.2 below.*

Example V.4.2. *Take $m = n = 1$, $K = 0$ and $C = 2$. Let $(\mu^k)_{k\geq 2}$ be a sequence of probability distributions defined by*

$$\mu^k \;:=\; \frac{1}{k}\delta_{\{-k\}} \;+\; \left(1 - \frac{2}{k}\right)\delta_{\{-\frac{k}{k-2}\}} \;+\; \frac{1}{k}\delta_{\{2k\}}.$$

It follows by a straightforward computation that

$$\mu^k(x) \;=\; 0 \quad and \quad \mu^k(x^+) \;=\; 2.$$

Moreover, μ^k converges weakly to the measure μ which puts the unit mass on -1. Take an arbitrary sequence of measures $(\overline{\mathbb{P}}_k)_{k\geq 2}$ with $\overline{\mathbb{P}}_k \in \overline{\mathcal{P}}(\mu^k) \subseteq \overline{\mathcal{P}}(0, 2)$, then it admits a convergent subsequence denoted again by $(\overline{\mathbb{P}}_k)_{k\geq 2}$, see e.g. Theorem V.4.4. Moreover, any limit point $\overline{\mathbb{P}}$ satisfies $\mathbb{E}^{\overline{\mathbb{P}}}[B_T] = -1$ and $\mathbb{E}^{\overline{\mathbb{P}}}[B_T^+] = 0$, which implies further $\overline{\mathbb{P}} \notin \overline{\mathcal{P}}(0, 2)$.

4.2 Convergence of optimal embedding problems

Next, let us study here the asymptotic behavior of the upper bound $\overline{P}(\boldsymbol{K}, \boldsymbol{C})$ with respect to the market information $(\boldsymbol{K}, \boldsymbol{C})$. Assume that the market is consistent with a centered peacock $\boldsymbol{\mu} = (\mu_k)_{1\leq k\leq m}$, then we ask : when more and more call options are traded, does the upper bound converge to $\overline{P}(\boldsymbol{\mu})$? Namely, given the vector of strikes $\boldsymbol{K}^n = (K_i^n)_{1\leq i\leq n}$, the prices of call options $\boldsymbol{C}^n = (C_{i,j}^n)_{1\leq i\leq m, 1\leq j\leq n}$ are determined by $C_{i,j}^n = \mu_i((x - K_j^n)^+)$ for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$. Define the bound and the mesh of $\boldsymbol{K}$ by

$$|\boldsymbol{K}^n| \;:=\; (K_1^n)^- \wedge (K_n^n)^+ \quad \text{and} \quad \Delta\boldsymbol{K}^n \;:=\; \max_{1<i\leq n}\left(K_i^n - K_{i-1}^n\right).$$

Loosely speaking, to capture asymptotically $\boldsymbol{\mu}$ by $(\boldsymbol{K}^n, \boldsymbol{C}^n)$, a necessary condition for the sequence $(\boldsymbol{K}^n)_{n\geq 1}$ is the following :

Assumption V.4.3. *The sequence $(\boldsymbol{K}^n)_{n\geq 1}$ satisfies*

$$\lim_{n\to\infty} |\boldsymbol{K}^n| \;=\; +\infty \quad and \quad \lim_{n\to\infty} \Delta\boldsymbol{K}^n \;=\; 0.$$

Motivated by financial applications, the sequence $(\boldsymbol{K}^n)_{n\geq 1}$ is assumed to be increasing, *i.e.* $\boldsymbol{K}^n \subseteq \boldsymbol{K}^{n+1}$ for all $n \geq 1$ if $\boldsymbol{K}^n$ and $\boldsymbol{K}^{n+1}$ are viewed as sets. It follows by definition that the map $n \mapsto \overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n)$ is non-increasing and $\overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n) \geq \overline{P}(\boldsymbol{\mu})$ for all $n \geq 1$.

Theorem V.4.4. *Let Assumption IV.2.4 hold. Then for any increasing sequence $(\boldsymbol{K}^n)_{n\geq 1}$ satisfying Assumption V.4.3, one has*

$$\lim_{n\to\infty} \overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n) \quad = \quad \overline{P}(\boldsymbol{\mu}),$$

where $\boldsymbol{C}^n$ is defined by $\boldsymbol{\mu}$ as above.

Proof. Notice that the sequence $\overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n)$ is non-increasing and thus the limit exists. Let $(\overline{\mathbb{P}}_n)_{n\geq 1}$ be a sequence such that $\overline{\mathbb{P}}_n \in \overline{\mathcal{P}}(\boldsymbol{K}^n, \boldsymbol{C}^n)$ and

$$\lim_{n\to\infty} \overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n) \quad = \quad \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}\Big[\Phi(B, T)\Big].$$

Then by the reasoning of Lemma IV.4.2, we deduce that the sequence $(\overline{\mathbb{P}}_n)_{n\geq 1}$ is tight and any limit point $\overline{\mathbb{P}}$ of $(\overline{\mathbb{P}}_n)_{n\geq 1}$ belongs to $\overline{\mathcal{P}}$. Without loss of generality, denote again by $(\overline{\mathbb{P}}_n)_{n\geq 1}$ the convergent subsequence with limit $\overline{\mathbb{P}}$, then it follows by Lemma V.4.5 that

$$\overline{\mathbb{P}} \circ B_{T_i}^{-1} \quad = \quad \mu_i \text{ for all } i = 1, \cdots, m,$$

which implies that $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$. The proof is fulfilled by Fatou's lemma :

$$\overline{P}(\boldsymbol{\mu}) \;\leq\; \lim_{n\to\infty} \overline{P}(\boldsymbol{K}^n, \boldsymbol{C}^n) \;=\; \lim_{n\to\infty} \mathbb{E}^{\overline{\mathbb{P}}_n}\left[\Phi(B, T)\right] \;\leq\; \mathbb{E}^{\overline{\mathbb{P}}}\left[\Phi(B, T)\right] \;\leq\; \overline{P}(\boldsymbol{\mu}).$$

□

Lemma V.4.5. *Let μ be a probability measure on $\mathbb{R}$ and set $c(K) := \mu\left((x-K)^+\right)$ for all $K \in \mathbb{R}$. Let $(\mu^n)_{n\geq 1}$ be a weakly convergent sequence of probability measures such that*

$$\mu^n\Big((x - K_i^n)^+\Big) \quad = \quad \mu\Big((x - K_i^n)^+\Big) \text{ for all } i = 1, \cdots, n,$$

where $(\boldsymbol{K}^n = (K_i^n)_{1\leq i\leq n})_{n\geq 1}$ is an increasing sequence satisfying Assumption V.4.3. Then

$$\lim_{n\to\infty} \mu^n \quad = \quad \mu.$$

Proof. Set $c^n(K) := \mu^n\left((x-K)^+\right)$ (resp. $c(K) := \mu\left((x-K)^+\right)$) for all $K \in \mathbb{R}$. Since $|(x-a)^+ - (x-b)^+| \leq |a-b|$, the function $K \mapsto c^n(K)$ (resp. $K \mapsto c(K)$) is Lipschitz.

Moreover, for every $K \in \cup_{n\geq 1}\boldsymbol{K}^n$, one has $c^n(K) = c(K)$ for n large enough, which implies that c^n converges uniformly to c as $\cup_{n\geq 1}\boldsymbol{K}^n$ is dense on $\mathbb{R}$.

Let $f : \mathbb{R} \to \mathbb{R}$ be twice differentiable with compact support, then it follows by Carr-Madan's formula that

$$f(x) \;=\; \int_{\mathbb{R}} f''(K)(x-K)^+ dK.$$

We obtain in view of Fubini's Theorem

$$\mu^n(f) \;=\; \int_{\mathbb{R}} f''(K)c^n(K)dK,$$

which implies that $\mu^n(f) \to \mu(f)$ as $n \to \infty$. As any continuous function with compact support can be uniformly approximated by a twice differentiable function with compact support, one has $\lim_{n\to\infty} \mu^n = \mu$. □

5 Convergence rate analysis

This section aims to estimate the convergence rate of the above upper bounds. For technical reasons, we consider in the following a subset $\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C})$:

$$\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C}) \;:=\; \Big\{\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{K},\boldsymbol{C}) : \; \mathbb{E}^{\overline{\mathbb{P}}}\Big[|B_{T_m}|^p\Big] \;=\; V\Big\},$$

where $p > 1$ and $V > 0$ are fixed throughout. The restriction of the embeddings comes from a new information : the Power option of the last maturity is observed in the market. Consequently, this implies the unknown peacock $\boldsymbol{\mu}$ must satisfy $\mu_m(|x|^p) = V$. Put similarly

$$\overline{P}^V(\boldsymbol{K},\boldsymbol{C}) \;:=\; \sup_{\overline{\mathbb{P}}\in\overline{\mathcal{P}}^V(\boldsymbol{K},\boldsymbol{C})} \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big].$$

First, notice by definition that

$$\overline{P}(\boldsymbol{\mu}) \;\leq\; \overline{P}^V(\boldsymbol{K}^n,\boldsymbol{C}^n) \;\leq\; \overline{P}(\boldsymbol{K}^n,\boldsymbol{C}^n) \text{ for all } n \geq 1,$$

which implies by Theorem V.4.4 that

$$\lim_{n\to\infty} \overline{P}^V(\boldsymbol{K}^n,\boldsymbol{C}^n) \;=\; \overline{P}(\boldsymbol{\mu}).$$

Throughout this section we focus on the asymptotic behavior $\overline{P}^V(\boldsymbol{K}^n, \boldsymbol{C}^n)$. Let $q > 1$ be the conjugate number of p, *i.e.* $1/p + 1/q = 1$.

5.1 Some metrics on P

In preparation, let us introduce two metrics that are used in the following. Denote respectively by $\rho(\cdot,\cdot)$ the Lévy-Prokhorov metric and by $\mathcal{W}_1(\cdot,\cdot)$ the Wasserstein metric on $\mathbf{P}$, *i.e.* for any two probability measures μ and $\nu \in \mathbf{P}$, one has

$$\begin{aligned} \rho(\mu,\nu) &:= \inf \Big\{\varepsilon > 0 : \ F_\mu(x-\varepsilon) - \varepsilon \le F_\nu(x) \le F_\mu(x+\varepsilon) + \varepsilon \text{ for all } x \in \mathbb{R}\Big\}, \\ \mathcal{W}_1(\mu,\nu) &:= \inf_{\mathbb{P}\in\mathcal{P}(\mu,\nu)} \mathbb{E}^{\mathbb{P}}\Big[|X_1 - X_0|\Big], \end{aligned}$$

where F_μ (resp. F_ν) denotes the cumulative distribution function of μ (resp. ν).

Remark V.5.1. *It is well known that, see e.g. Chapter 1.2 of Rachev & Rüschendorf [90]*

$$\begin{aligned} \mathcal{W}_1(\mu,\nu) &= \int_{\mathbb{R}} \Big|F_\mu(x) - F_\nu(x)\Big| dx \\ &= \inf \Big\{\mu(f) - \nu(f) : \ f : \mathbb{R} \to \mathbb{R} \text{ is } 1-\text{Lipschitz}\Big\}. \end{aligned}$$

Lemma V.5.2. *Let μ and ν be two probability measures supported on $[-R,R]$ for some fixed $R > 0$. Assume that there exists some $\varepsilon > 0$ such that*

$$\Big|\mu\Big((x-K)^+\Big) \ - \ \nu\Big((x-K)^+\Big)\Big| \ \le \ \varepsilon \text{ for all } K \in [-R,R].$$

Then

$$\rho(\mu,\nu) \ \le \ \sqrt{2\varepsilon} \quad \text{and} \quad \mathcal{W}_1(\mu,\nu) \ \le \ 4R\sqrt{\varepsilon}.$$

Proof. (i) Take an arbitrary $0 < \delta < \rho(\mu,\nu)$, then one has some $K \in [-R,R]$ such that $F_\mu(K-\delta) - F_\nu(K) > \delta$ or $F_\nu(K) - F_\mu(K+\delta) > \delta$. Take the first case without loss of generality, which yields

$$\int_{K-\delta}^{K} (F_\mu(x) - F_\nu(x))\, dx \ \ge \ \int_{K-\delta}^{K} (F_\mu(K-\delta) - F_\mu(K))\, dx \ > \ \delta^2.$$

In addition,

$$
\begin{aligned}
&\int_{K-\delta}^{K} \Big(F_\mu(x) - F_\nu(x)\Big)dx \\
=\ & \Big|\int_{K-\delta}^{+\infty} \Big(F_\mu(x) - F_\nu(x)\Big)dx \ - \ \int_{K}^{+\infty} \Big(F_\mu(x) - F_\nu(x)\Big)dx\Big| \\
=\ & \Big|\mu\Big((x-K+\delta)^+\Big) \ - \ \nu\Big((x-K+\delta)^+\Big) \ - \ \mu\Big((x-K)^+\Big) \ + \ \nu\Big((x-K+\delta)^+\Big)\Big|.
\end{aligned}
$$

It follows by assumption that $\delta^2 < 2\varepsilon$. That is,

$$
\rho(\mu,\nu) \ \leq \ \sqrt{2\varepsilon}.
$$

(ii) It follows by Theorem 1.1.8 in Rachev & Rüschendorf [90], see also Remark V.5.1, that

$$
\begin{aligned}
\mathcal{W}_1(\mu,\nu) \ &= \ \int_{\mathbb{R}} \Big|F_\mu(x) - F_\nu(x)\Big|dx \\
&= \ \int_{-R}^{R} \Big|F_\mu(x) - F_\nu(x)\Big|dx.
\end{aligned}
$$

In addition, it follows by definition that $|F_\mu(x) - F_\nu(x)| \leq \sqrt{2}\rho(\mu,\nu) \leq 2\sqrt{\varepsilon}$ for all $x \in [-R,R]$. Thus

$$
\mathcal{W}_1(\mu,\nu) \ \leq \ 4R\sqrt{\varepsilon}.
$$

□

Proposition V.5.3. *For each $n \geq 1$ and any $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\boldsymbol{K}^n, \boldsymbol{C}^n)$, set $\nu_i = \overline{\mathbb{P}} \circ (B_{T_i})^{-1}$ for all $i = 1,\cdots,m$. Then there exists a constant $C > 0$ depending only on V such that, for all $i = 1,\cdots,m$ one has*

$$
\rho(\mu_i,\nu_i) \ \leq \ C\left(\sqrt{\Delta \boldsymbol{K}^n} + |\boldsymbol{K}^n|^{-p/2q}\right) \quad \textit{and} \quad \mathcal{W}_1(\mu_i,\nu_i) \ \leq \ C|\boldsymbol{K}^n|\left(\sqrt{\Delta \boldsymbol{K}^n} + |\boldsymbol{K}^n|^{-p/2q}\right).
$$

Proof. Without loss of generality, it suffices to show the inequality for $i = m$. For the sake of simplicity, we write $\mu \equiv \mu_m$ and $\nu \equiv \nu_m$. The main idea is to approximate μ and ν by their truncated versions.

Set $R = |\boldsymbol{K}^n|$ and let μ' (resp. ν') be the truncated distribution of μ (resp. ν). Indeed, let X (resp. Y) denotes some random variable of law μ (resp. ν), then μ' (resp. ν') be the law of

$X' := (-R) \vee (R \wedge X)$ (resp. $Y' := (-R) \vee (R \wedge Y)$). Then one has for all $K \in [-R, R]$

$$\begin{aligned}
&\left|\mu'\left((x-K)^+\right) \;-\; \nu'\left((x-K)^+\right)\right| \;=\; \left|\mathbb{E}\left[(X'-K)^+\right] \;-\; \mathbb{E}\left[(Y'-K)^+\right]\right| \\
\le\;& \left|\mathbb{E}\left[(X-K)^+\right] \;-\; \mathbb{E}\left[(Y-K)^+\right]\right| \;+\; \mathbb{E}\Big[|X-X'|\Big] \;+\; \mathbb{E}\Big[|Y-Y'|\Big] \\
\le\;& \left|\mu\left((x-K)^+\right) \;-\; \nu\left((x-K)^+\right)\right| \;+\; 2\mathbb{E}\left[|X|\mathbb{1}_{\{|X|>R\}}\right] \;+\; 2\mathbb{E}\left[|Y|\mathbb{1}_{\{|Y|>R\}}\right]. \quad (*)
\end{aligned}$$

Notice that for each $K \in [-R, R]$ there exists $1 \le i < n$ such that $K \in [K_i^n, K_{i+1}^n]$, thus

$$\begin{aligned}
\mu\left((x-K)^+\right) \;-\; \nu\left((x-K)^+\right) \;&\le\; \mu\left((x-K_i^n)^+\right) \;-\; \nu\left((x-K_{i+1}^n)^+\right) \\
&=\; \mu\left((x-K_i^n)^+\right) \;-\; \mu\left((x-K_{i+1}^n)^+\right) \;\le\; \Delta\boldsymbol{K}^n.
\end{aligned}$$

Hence

$$\left|\mu\left((x-K)^+\right) \;-\; \nu\left((x-K)^+\right)\right| \;\le\; \Delta\boldsymbol{K}^n.$$

In addition,

$$\mathbb{E}\left[|X|\mathbb{1}_{\{|X|>R\}} \;+\; |Y|\mathbb{1}_{\{|Y|>R\}}\right] \;\le\; \mathbb{E}\left[\frac{|X|^p}{R^{p-1}}\mathbb{1}_{\{|X|>R\}} \;+\; \frac{|Y|^p}{R^{p-1}}\mathbb{1}_{\{|Y|>R\}}\right] \;\le\; \frac{2V}{R^{p-1}},$$

which yields by $(*)$ that

$$\left|\mu'\left((x-K)^+\right) \;-\; \nu'\left((x-K)^+\right)\right| \;\le\; \Delta\boldsymbol{K}^n \;+\; \frac{4V}{|\boldsymbol{K}^n|^{p/q}}.$$

It follows by Lemma V.5.2 that there exists some $C > 0$ such that

$$\rho(\mu', \nu') \;\le\; C\left(\sqrt{\Delta\boldsymbol{K}^n} + |\boldsymbol{K}^n|^{-p/2q}\right) \quad \text{and} \quad \mathcal{W}_1(\mu', \nu') \;\le\; C|\boldsymbol{K}^n|\left(\sqrt{\Delta\boldsymbol{K}^n} + |\boldsymbol{K}^n|^{-p/2q}\right).$$

It remains to estimate $\rho(\mu, \mu')$ (resp. $\rho(\nu, \nu')$) and $\mathcal{W}_1(\mu, \mu')$ (resp. $\mathcal{W}_1(\nu, \nu')$). It follows by definition

$$\rho(\mu, \mu') \text{ (resp. } \rho(\nu, \nu')) \;\le\; \frac{2V}{|\boldsymbol{K}^n|^p} \quad \text{and} \quad \mathcal{W}_1(\mu, \mu') \text{ (resp. } \mathcal{W}_1(\nu, \nu')) \;\le\; \frac{2V}{|\boldsymbol{K}^n|^{p/q}},$$

which yield the required inequalities by the triangle inequality. $\square$

Remark V.5.4. *Notice that, in order to ensure that $\mathcal{W}_1(\mu,\nu)$ converges to zero, we need $p>3$ and $|\boldsymbol{K}^n|\sqrt{\Delta \boldsymbol{K}^n}\to 0$ as $n\to\infty$.*

To estimate the convergence, we need more regularity on Φ. Let us formulate the assumption on Φ. Recall that $\overline{d}$ is the metric on $\overline{\Omega}$ defined by, for all $\bar{\omega}=(\omega,\theta_1,\cdots,\theta_m)$, $\bar{\omega}'=(\omega',\theta_1',\cdots,\theta_m')\in\overline{\Omega}$,

$$\overline{d}(\bar{\omega},\bar{\omega}') \;:=\; \sum_{i=1}^m\Big(|\theta_i-\theta_i'| \;+\; \|\omega_{\theta_i\wedge\cdot}-\omega'_{\theta_i'\wedge\cdot}\|\Big).$$

We end this section by the following assumption on Φ.

Assumption V.5.5. *$\Phi:\overline{\Omega}\to\mathbb{R}$ is bounded and $\overline{d}-$Lipschitz, where the uniform norm and Lipschitz constant are respectively denoted by $\|\Phi\|$ and L.*

5.2 One-marginal case

We start by the one-marginal case, where we may construct explicitly an approximation of a given martingale. Through this, we obtain the difference of the upper bounds corresponding to different target distributions under the following assumption.

Assumption V.5.6. *Φ is time-invariant, i.e. $\Phi(\omega,\theta)=\Phi\left(\omega_\varphi,\varphi(\theta)\right)$ holds for all $(\omega,\theta)\in\overline{\Omega}$ and increasing functions $\varphi:\mathbb{R}_+\to\mathbb{R}_+$.*

Proposition V.5.7. *Let Assumptions V.5.5 and V.5.6 hold. Then there exists a constant $C>0$ such that, for any centered probability distribution ν on $\mathbb{R}$ satisfying $\nu(|x|^p)\le V$, one has*

$$\left|\overline{P}(\mu)\;-\;\overline{P}(\nu)\right| \;\le\; C\rho(\mu,\nu)^{1/2q}.$$

Proof. Write $\rho:=\rho(\mu,\nu)$ for the sake of simplicity. Take a $\overline{\mathbb{P}}\in\overline{\mathcal{P}}(\mu)$, then one has by definition $\overline{\mathbb{P}}\circ B_T^{-1}=\mu$. It follows by Theorem 4 on page 358 in Shiryaev [95] and Theorem 1 in Skorokhod [96] that, there exist a measurable function $f:\mathbb{R}^2\to\mathbb{R}$ and a Gaussian random variable G that is independent of $\overline{\mathbb{F}}$ such that

$$M \;:=\; f(G,B_T)\;\overset{\overline{\mathbb{P}}}{\sim}\;\nu \quad\text{and}\quad \overline{\mathbb{P}}\left[|B_T-M|>\rho\right]\;\le\;\rho.$$

Recall that $\overline{\mathbb{F}}^B=(\overline{\mathcal{F}}_t^B)_{t\ge0}$ is the canonical filtration generated by B. Set $\mathbb{G}:=(\mathcal{G}_t)_{t\ge0}$ with $\mathcal{G}_t:=\sigma(\overline{\mathcal{F}}_t^B,G)$, then $\mathbb{G}$ is again a Brownian filtration and more importantly, B is a $\mathbb{G}-$Brownian motion. Take the continuous martingale $M = (M_t)_{t\geq 0}$ given by

$$M_t \;:=\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[M\Big|\mathcal{G}_t\Big] \text{ for all } t \in \mathbb{R}_+,$$

then it follows by Doob's martingale inequality that, for all $r > 0$,

$$\begin{aligned}
\overline{\mathbb{P}}\left[\sup_{t\leq T}|B_t - M_t| > \rho^r\right] \;&\leq\; \frac{1}{\rho^r}\mathbb{E}^{\overline{\mathbb{P}}}\left[|B_T - M|\right] \\
\leq\; \frac{1}{\rho^r}&\left\{\mathbb{E}^{\overline{\mathbb{P}}}\left[|B_T - M|\mathbb{1}_{|B_T-M|\leq\rho\}}\right] \;+\; \mathbb{E}^{\overline{\mathbb{P}}}\left[|B_T - M|\mathbb{1}_{|B_T-M|>\rho\}}\right]\right\} \\
\leq\; \frac{1}{\rho^r}&\left\{\rho \;+\; \left(\mathbb{E}^{\overline{\mathbb{P}}}\left[|B_T - M|^p\right]\right)^{1/p}\overline{\mathbb{P}}\left[|B_T - M| > \rho\right]^{1/q}\right\} \\
\leq\; \rho^{1-r}& \;+\; 2V^{1/p}\rho^{1/q-r}.
\end{aligned}$$

Notice that, in view of Lévy's theorem, there exists a Brownian motion $W = (W_t)_{t\geq 0}$ such that $M = \left(W_{\langle M\rangle_t}\right)_{t\geq 0}$ $\overline{\mathbb{P}}$ - almost surely. Then it follows from the time-invariance of Φ that $\Phi(M,T) = \Phi\left(W, \langle M\rangle_T\right)$, which yields that

$$\begin{aligned}
\mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big] \;-\; \overline{P}(\nu) \;\;&\leq\;\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big] \;-\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(M,T)\Big] \\
\leq\; \mathbb{E}^{\overline{\mathbb{P}}}&\left[|\Phi(B,T) - \Phi(M,T)|\,\mathbb{1}_{\{\sup_{t\leq T}|B_t-M_t|\leq\rho^r\}}\right] \;+\; 2\|\Phi\|\overline{\mathbb{P}}\left[\sup_{t\leq T}|B_t - M_t| > \rho^r\right] \\
\leq\; L\rho^r& \;+\; 2\|\Phi\|\rho^{1-r} \;+\; 4\|\Phi\|V^{1/p}\rho^{1/q-r}.
\end{aligned}$$

Optimizing with respect to $r > 0$ and $\overline{\mathbb{P}} \in \overline{\mathcal{P}}(\mu)$, one has a constant C depending on L, $\|\Phi\|$ and V such that

$$\overline{P}(\mu) \;-\; \overline{P}(\nu) \;\leq\; C\rho^{1/2q}.$$

Repeating the argument above by interchanging μ and ν, we deduce the required result. $\square$

Combing Propositions V.5.3 and V.5.7, we obtain immediately the estimation.

Theorem V.5.8. *Let Assumptions V.5.5 and V.5.6 hold. Then there exists a constant $C > 0$ depending on L, $\|\Phi\|$ and V such that*

$$0 \;\leq\; \overline{P}^V(\boldsymbol{K}^n, \boldsymbol{C}^n) \;-\; \overline{P}(\mu) \;\leq\; C\left((\Delta\boldsymbol{K}^n)^{1/4q} \;+\; |\boldsymbol{K}^n|^{-p/4q^2}\right).$$

5.3 Multi-marginal case

For the multi-marginal case, we will make use of a duality result to be specified later. The main idea here is to restrict the vector $\boldsymbol{\lambda} \in \Lambda^m$ in the subset $\Lambda_0^m \subseteq \Lambda^m$ defined by

$$\Lambda_0^m \quad := \quad \Big\{ \boldsymbol{\lambda} = (\lambda_k)_{1 \le k \le m} : \ \lambda_k \text{ is } (1+L)(1+\|\Phi\|) - \text{Lipschitz for all } k = 1, \cdots, m \Big\}.$$

Then we have another estimation of the difference between the upper bounds of different target distributions.

Proposition V.5.9. *Let Assumption V.5.5 hold and $p > 2$. Then there exists a constant $C > 0$ such that, for any centered peacock $\boldsymbol{\nu} = (\nu_k)_{1 \le k \le m}$ satisfying $\nu_m(|x|^p) \le V$, one has*

$$\Big| \overline{P}(\boldsymbol{\mu}) \ - \ \overline{P}(\boldsymbol{\nu}) \Big| \ \le \ C \sum_{k=1}^m \mathcal{W}_1(\mu_k, \nu_k)^{\frac{p-2}{p-1}}.$$

The strategy for the proof is to translate the embedding problem into a constrained transport problem between the Wiener measure $\mathbb{P}_0$ and the target $\boldsymbol{\mu}$. To this end we equip the enlarged canonical space $\Xi := \overline{\Omega} \times \mathbb{R}^m$ with the function $\xi : \Xi \to \mathbb{R}$ by

$$\xi\left(\omega, \theta = (\theta_i)_{1 \le i \le m}, x = (x_i)_{1 \le i \le m}\right) \quad := \quad \begin{cases} \Phi(\omega, \theta), & \text{if } \omega_{\theta_i} = x_i \text{ for all } i = 1, \cdots, m, \\ -\infty, & \text{otherwise.} \end{cases}$$

We introduce next the canonical elements (B, T, X) defined by $B(\omega, \theta, x) = \omega$, $T(\omega, \theta, x) = \theta$ and $X(\omega, \theta, x) = x$. It is clear that Ξ is a Polish space and denote by $\mathcal{Q}(\Xi)$ the set of all probability measures on Ξ. Define further by $\mathcal{Q}(\boldsymbol{\mu}) \subseteq \mathcal{Q}(\Xi)$ the subset consisting of measures $\mathbb{Q}$ such that

$$\mathbb{Q} \circ (B, T)^{-1} \in \mathcal{P}, \ \mathbb{Q} \circ X_i^{-1} \ = \ \mu_i \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}[T_i] \ \le \ \mu_i(|x|^2) \text{ for all } i = 1, \cdots, m.$$

Then the optimal transport problem is defined by

$$\mathsf{P}_e(\boldsymbol{\mu}) \quad := \quad \sup_{\mathbb{Q} \in \mathcal{Q}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{Q}} \Big[\xi(B, T, X) \Big]. \tag{V.5.1}$$

The problem (V.5.1) is an optimal transport problem with additional moment constraints, see *e.g.* Chapter 4.6.3 of Rachev & Rüschendorf [90] for more details. And the corresponding dual problem is defined as follows :

$$\mathsf{D}_e(\boldsymbol{\mu}) \;:=\; \inf_{(S,\boldsymbol{\lambda},\boldsymbol{k})\in\mathcal{D}_e} \boldsymbol{\mu}(\boldsymbol{\lambda}) \;+\; \boldsymbol{k}\cdot\Big(\mu_1(|x|^2),\cdots,\mu_m(|x|^2)\Big), \tag{V.5.2}$$

where $\mathcal{D}_e$ denotes the set of elements $(S,\boldsymbol{\lambda},\boldsymbol{k} := (k_1,\cdots,k_m)) \in \mathcal{S}^0 \times \Lambda_0^m \times \mathbb{R}^m$ such that

$$0 \;\le\; k_i \;\le\; L, \text{ for all } i=1,\cdots,m$$

and

$$S_{\theta_m} \;+\; \sum_{i=1}^m \Big(\lambda_i(x_i) - k_i(\theta_i - A_i)\Big) \;\ge\; \xi(\omega,\theta,x), \text{ for all } (\omega,\theta,x)\in\Xi.$$

Then Theorem 4.6.12 of [90] yields immediately the following duality.

Lemma V.5.10. *Let Assumption V.5.5 hold and $p\ge 2$. Then the following duality holds :*

$$\mathsf{P}_e(\boldsymbol{\mu}) \;=\; \mathsf{D}_e(\boldsymbol{\mu}).$$

In view of the above duality, it remains to estimate $\mu_i(|x|^2)-\nu_i(|x|^2)$ for all $i=1,\cdots,m$, which is achieved by the next lemma.

Lemma V.5.11. *Assume $p>2$, then there exists a constant $C>0$ such that, for any centered peacock $\boldsymbol{\nu}=(\nu_k)_{1\le k\le m}$ satisfying $\nu_m(|x|^p)\le V$, one has*

$$\Big|\mu_i(|x|^2) \;-\; \nu_i(|x|^2)\Big| \;\le\; C\mathcal{W}_1(\mu,\nu)^{\frac{p-2}{p-1}} \text{ for all } i=1,\cdots,m.$$

Proof. For the sake of simplicity we write $\mu\equiv\mu_i$ (resp. $\nu\equiv\nu_i$) for short. For all $R>0$ one has

$$\begin{aligned}
&\Big|\mu(|x|^2) \;-\; \nu(|x|^2)\Big| \\
\le\;& \Big|\mu(|x|^2\mathbb{1}_{\{|x|<R\}}) \;-\; \nu(|x|^2\mathbb{1}_{\{|x|<R\}})\Big| \;+\; \Big|\mu(|x|^2\mathbb{1}_{\{|x|\ge R\}}) \;-\; \nu(|x|^2\mathbb{1}_{\{|x|\ge R\}})\Big| \\
\le\;& 2R\mathcal{W}_1(\mu,\nu) \;+\; \frac{2V}{R^{p-2}}.
\end{aligned}$$

Taking in particular $R=\mathcal{W}_1(\mu,\nu)^{-1/(p-1)}$, one obtains some suitable constant C depending only on V such that

$$\Big|\mu(|x|^2) \;-\; \nu(|x|^2)\Big| \;\le\; C\mathcal{W}_1(\mu,\nu)^{\frac{p-2}{p-1}}.$$

□

Now let us turn to the proof of Proposition V.5.9 by a use of Lemmas V.5.10 and V.5.11.

Proof of Proposition V.5.9. First let us show $\overline{P}(\boldsymbol{\mu}) = \mathsf{P}_e(\boldsymbol{\mu})$. On the one hand, one has by definition $\overline{P}(\boldsymbol{\mu}) \le \mathsf{P}_e(\boldsymbol{\mu})$. On the other hand, for each $\mathbb{Q} \in \mathcal{Q}(\boldsymbol{\mu})$, it is easy to see $\mathbb{E}^{\mathbb{Q}}[\xi(B,T,X)] \neq -\infty$ if and only if $B_T \overset{\mathbb{Q}}{\sim} \boldsymbol{\mu}$, or equivalently, $\overline{\mathbb{P}} := \mathbb{Q} \circ (B,T)^{-1} \in \overline{\mathcal{P}}(\boldsymbol{\mu})$. Thus

$$\mathbb{E}^{\mathbb{Q}}\Big[\xi(B,T,X)\Big] \;=\; \mathbb{E}^{\overline{\mathbb{P}}}\Big[\Phi(B,T)\Big] \;\le\; \overline{P}(\boldsymbol{\mu}),$$

which yields the required identity. Next, let us turn to estimate $|\overline{P}(\boldsymbol{\mu}) - \overline{P}(\boldsymbol{\nu})|$. It follows by Lemma V.5.10 that there exists a sequence $(S^n, \boldsymbol{\lambda}^n, \boldsymbol{k}^n)$ such that

$$\lim_{n\to\infty} \boldsymbol{\nu}(\boldsymbol{\lambda}^n) \;+\; \boldsymbol{k}^n \cdot \boldsymbol{\nu}(|x|^2) \;=\; \overline{P}(\boldsymbol{\nu}),$$

which implies that

$$\begin{aligned}
\overline{P}(\boldsymbol{\mu}) - \overline{P}(\boldsymbol{\nu}) \;&=\; \overline{P}(\boldsymbol{\mu}) \;-\; \lim_{n\to\infty}\Big(\boldsymbol{\nu}(\boldsymbol{\lambda}^n) \;+\; \boldsymbol{k}^n \cdot \boldsymbol{\nu}(|x|^2)\Big) \\
&\le\; \sup_{n\ge1}\Big\{\boldsymbol{\mu}(\boldsymbol{\lambda}^n) \;+\; \boldsymbol{k}^n \cdot \boldsymbol{\mu}(|x|^2) \;-\; \boldsymbol{\nu}(\boldsymbol{\lambda}^n) \;+\; \boldsymbol{k}^n \cdot \boldsymbol{\nu}(|x|^2)\Big\} \\
&\le\; \sup_{n\ge1}\left(\boldsymbol{\mu}(\boldsymbol{\lambda}^n) \;-\; \boldsymbol{\nu}(\boldsymbol{\lambda}^n)\right) \;+\; \sup_{n\ge1}\Big(\boldsymbol{k}^n \cdot \boldsymbol{\mu}(|x|^2) \;-\; \boldsymbol{k}^n \cdot \boldsymbol{\nu}(|x|^2)\Big) \\
&\le\; (1+L)(1+\|\Phi\|)\sum_{i=1}^m \mathcal{W}_1(\mu_i,\nu_i) \;+\; L\sum_{i=1}^m \Big|\mu_i(|x|^2) - \nu_i(|x|^2)\Big| \\
&\le\; C\sum_{i=1}^m \mathcal{W}_1(\mu_i,\nu_i)^{\frac{p-2}{p-1}},
\end{aligned}$$

where C depends only on L, $\|\Phi\|$ and V. Repeating the reasoning above by interchanging $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ we get the required result. □

Combing Propositions V.5.3 and V.5.9, we obtain immediately the main result under the following assumption.

Assumption V.5.12. *The sequence $(\boldsymbol{K}^n)_{n\ge1}$ satisfies*

$$\lim_{n\to\infty} |\boldsymbol{K}^n| \;=\; +\infty \quad \textit{and} \quad \lim_{n\to\infty} |\boldsymbol{K}^n|\sqrt{\Delta \boldsymbol{K}^n} \;=\; 0.$$

Theorem V.5.13. *Let Assumption V.5.5 hold and $p > 3$. Then there exists a constant $C > 0$*

depending on L, $\|\Phi\|$ and V such that

$$0 \quad \leq \quad \overline{P}^V(\boldsymbol{K}^n, \boldsymbol{C}^n) \ - \ \overline{P}(\boldsymbol{\mu}) \quad \leq \quad C|\boldsymbol{K}^n|^{\frac{p-2}{p-1}} \left(\sqrt{\Delta \boldsymbol{K}^n} \ + \ |\boldsymbol{K}^n|^{-p/2q}\right)^{\frac{p-2}{p-1}}.$$

Chapitre VI

Robust hedging of options on local time

1 Introduction

As discussed in Chapters III, IV and V, the robust hedging problem is classically approached by means of the martingale optimal transport and optimal Skorokhod embedding. Recently, a new approach to study the robust hedging problem was developed by Galichon, Henry-Labordère & Touzi [43]. It is based on a dual representation of the robust hedging problem, which can be addressed by means of the stochastic control theory. It appears that the stochastic control approach is remarkably devised to provide candidates for the optimal hedging strategies. This is illustrated by the study of Henry-Labordère, Obłój, Spoida & Touzi [50], where they solve the robust hedging problem of lookback options given a finite number of marginals. To the best of our knowledge, at the exception of Brown, Hobson & Rogers [17] and Hobson & Neuberger [61] who considered the two-marginal case, it is the first paper to address the multi-marginal case. More importantly, it leads to the first solution of the multi-marginal SEP, which can be seen as a generalization of Azéma-Yor's solution, see Obłój & Spoida [84].

In this chapter, we consider the application of stochastic control to the robust hedging problem for options written on the local time of the underlying process. Financial derivatives on local time appear naturally when considering payoffs depending on the portfolio value of an at-the-money call option delta hedged with its intrinsic value as highlighted by Itô-Tanaka's formula. Chapter VI is organized as follows. We recall briefly in Section 2 the framework of robust hedging on the space of continuous functions and its relation with the martingale optimal transport. In Section 3, using the stochastic control approach, we provide explicit formulas for the bounds of no-arbitrage prices and the optimal hedging strategies when one marginal is known. In Section 4, we proceed by iteration to provide candidates for the optimal hedging strategies in the two-marginal case. Although we are not able to verify the hedging inequality in general, we provide examples where we can complete the proof. More importantly, we construct a new solution to the two-marginal SEP. Finally, we consider in Section 5 a special multi-marginal setting, where Vallois' solutions are well ordered. When finitely-many marginals are known, we solve the robust hedging problem in this case as it essentially reduces to the one-marginal case. In the full marginal setting, we construct a remarkable Markov martingale and compute explicitly its generator. As an application, we provide a new example of fake Brownian motion.

2 Formulation

We consider a financial market consisting of one risky asset, which may be traded at any time $t \in [0,1]$. Recall that $\mathbf{C}$ is the space of continuous functions $\omega = (\omega_t)_{0\le t\le 1}$ with $\omega_0 = 0$, $X = (X_t)_{0\le t\le 1}$ is the canonical process, *i.e.* $X_t(\omega) := \omega_t$, $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ is its natural filtration, *i.e.* $\mathcal{F}_t := \sigma(X_s, s \le t)$, and $\mathcal{M}$ is the set of martingale measures $\mathbb{P}$ on $(\mathbf{C}, \mathcal{F}_1)$, *i.e.* the process X is a $\mathbb{P}-$martingale. Given a set of maturities $\mathbb{T} = \{t_1 < \cdots < t_m = 1\} \subseteq (0,1]$, we assume that the calibrated market yields a centered peacock $\boldsymbol{\mu} = (\mu_k)_{1\le k\le m} \in \mathbf{P}^m$, *i.e.*

$$\mu_k(x) \;=\; 0 \text{ for all } k = 1, \cdots, m \quad \text{and} \quad \mu_k \;\preceq\; \mu_{k+1} \text{ for all } k = 1, \cdots, m-1.$$

Then, the corresponding collection of martingale transport plans is given by

$$\mathcal{M}(\boldsymbol{\mu}) \;:=\; \Big\{\mathbb{P} \in \mathcal{M} : \; \mathbb{P} \circ X_{t_k}^{-1} \;=\; \mu_k \text{ for all } k = 1, \ldots, m\Big\}.$$

Let $L = (L_t)_{0\le t\le 1}$ be the local time of X at zero. For a measurable function $F : \mathbb{R}_+ \to \mathbb{R}$, define the martingale optimal transport problems

$$\overline{\mathsf{P}}(\boldsymbol{\mu}) \;:=\; \sup_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[F(L_1)\Big], \tag{VI.2.1}$$

$$\underline{\mathsf{P}}(\boldsymbol{\mu}) \;:=\; \inf_{\mathbb{P}\in\mathcal{M}(\boldsymbol{\mu})} \mathbb{E}^{\mathbb{P}}\Big[F(L_1)\Big]. \tag{VI.2.2}$$

Then let us turn to the corresponding dual problems. Recall that $\overline{\mathcal{H}}$ (resp. $\underline{\mathcal{H}}$) is the collection of all $\mathbb{F}-$progressively measurable processes $H : [0,1] \times \mathbf{C} \to \mathbb{R}$ such that,

$$\int_0^1 H_t^2 d\langle X\rangle_t \;<\; +\infty, \; \mathbb{P}\text{ - a.s. and the stochastic integral } (H\cdot X) \;:=\; \int_0^{\cdot} H_t dX_t$$
$$\text{is a } \mathbb{P} - \text{supermartingale (resp. submartingale) for all } \mathbb{P} \in \mathcal{M}. \tag{VI.2.3}$$

Define further

$$\overline{\mathcal{D}} \text{ (resp. } \underline{\mathcal{D}}) := \Big\{(\boldsymbol{\lambda}, H) \in \Lambda^m \times \overline{\mathcal{H}} \text{ (resp. } \underline{\mathcal{H}}) : \ \boldsymbol{\lambda}(X) + (H \cdot X)_1 \\ \geq \text{ (resp. } \leq) \ F(L_1), \ \mathbb{P}\text{ - a.s. for all } \mathbb{P} \in \mathcal{M}\Big\}, \tag{VI.2.4}$$

where $\boldsymbol{\lambda}(X) := \sum_{k=1}^{m} \lambda_i(X_{t_i})$. Here, the process H stands for a dynamic trading strategy, where H_t corresponds to the number of shares of the underlying asset held by the investor at time t, and the vector $\boldsymbol{\lambda}$ defines static positions in Vanilla options with maturities $t_1, \cdots, t_m$. Then the dual problems are defined by

$$\overline{\mathsf{D}}(\boldsymbol{\mu}) \text{ (resp. } \underline{\mathsf{D}}(\boldsymbol{\mu})) := \inf_{(\boldsymbol{\lambda},H)\in\overline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda}) \text{ (resp. } \sup_{(\boldsymbol{\lambda},H)\in\underline{\mathcal{D}}} \boldsymbol{\mu}(\boldsymbol{\lambda})). \tag{VI.2.5}$$

Remark VI.2.1. *Recall that both the quadratic variation process and the stochastic integral depend a priori on the probability measure under consideration. However, it follows by Nutz [80] that they can be universally defined, i.e. there exist two processes $\langle X\rangle$ and $\int_0^\cdot H_t dX_t$ such that*

$$\langle X\rangle_t = \langle X\rangle_t^{\mathbb{P}} \quad \textit{and} \quad \int_0^t H_s dX_s = \Big(\int_0^t H_s dX_s\Big)^{\mathbb{P}} \textit{ for all } t \in [0,1], \ \mathbb{P}\textit{ - a.s. for all } \mathbb{P} \in \mathcal{M}.$$

Moreover, in view of Itô-Tanaka's formula

$$\frac{1}{2}L_t = X_t^+ - \int_0^t \mathbb{1}_{\{X_s > X_0\}} dX_s \textit{ for all } t \in [0,1],$$

the local time L can also be universally defined.

Assumption VI.2.2. *$F : \mathbb{R}_+ \to \mathbb{R}$ is Lipschitz and convex.*

3 One-marginal case

We start by considering the one-marginal case. Using the stochastic control approach, we reproduce the results for the robust superhedging problem — optimizers and duality — obtained in Cox, Hobson & Obłój [23] and Vallois [102]. In addition, we provide the corresponding results for the robust subhedging problem. Throughout this section, we work under the following assumption on the marginal μ.

Assumption VI.3.1. *μ is a centered probability distribution without mass at zero.*

3.1 Robust superhedging problem

In this section, under Assumptions VI.2.2 and VI.3.1, we provide the optimal measure for $\overline{\mathsf{P}}(\mu)$ as well as the optimal strategy for $\overline{\mathsf{D}}(\mu)$ and we show that there is no duality gap, *i.e.* $\overline{\mathsf{P}}(\mu) = \overline{\mathsf{D}}(\mu)$. The key idea is that for any suitable pair of monotone functions (ϕ_+, ϕ_-), we may construct a superhedging strategy (H, λ). The optimality then follows by taking the pair of functions given by Vallois that embeds the probability distribution μ.

Assumption VI.3.2. *$\phi_+ : \mathbb{R}_+ \to \mathbb{R}_+$ (resp. $\phi_- : \mathbb{R}_+ \to \mathbb{R}_-$) is right-continuous and nondecreasing (resp. nonincreasing) such that $\gamma(0+) = 0$ and $\gamma(+\infty) = +\infty$, where*

$$\gamma(l) \;:=\; \frac{1}{2}\int_0^l \left(\frac{1}{\phi_+(z)} \;-\; \frac{1}{\phi_-(z)}\right) dz \text{ for all } l > 0.$$

Under Assumption VI.3.2, we define the static strategy $\lambda : \mathbb{R} \to \mathbb{R}$ via (ϕ_+, ϕ_-) by

$$\lambda'(\pm x) \;:=\; \lambda'(0\pm) \;+\; \int_0^{\psi_\pm(\pm x)} \frac{dy}{\phi_\pm(y)} e^{\gamma(y)} \int_y^{+\infty} e^{-\gamma(z)} F''(dz) \text{ for all } x > 0, \tag{VI.3.1}$$

$$\lambda'(0\pm) \;:=\; \pm F'(0) \;\pm\; \int_0^{+\infty} e^{-\gamma(z)} F''(dz), \tag{VI.3.2}$$

$$\lambda(0) \;:=\; F(0). \tag{VI.3.3}$$

where $\psi_\pm$ denote the right-continuous inverses of $\phi_\pm$, F' is the right derivative of F and $F''(dz)$ is the second derivative of F in the sense of distributions. Notice that F' is uniformly bounded and the restriction of λ on $\mathbb{R}_+$ and $\mathbb{R}_-$ is convex.

3.1.1 Pathwise inequality

We start by showing a remarkable pathwise inequality, which is a key step in our analysis. Indeed, together with Vallois' solution, it leads to the solution of the robust superhedging problem.

Proposition VI.3.3. *Under Assumptions VI.2.2 and VI.3.2, the following inequality holds*

$$\int_0^1 H_t dX_t \;+\; \lambda(X_1) \;\geq\; F(L_1), \; \mathbb{P}\text{ - a.s. for all } \mathbb{P} \in \mathcal{M}, \tag{VI.3.4}$$

where the function λ is defined by (VI.3.1), (VI.3.2), (VI.3.3) and the process H is given by

$$H_t \;:=\; -A_+(L_t)\mathbb{1}_{\{X_t>0\}} \;-\; A_-(L_t)\mathbb{1}_{\{X_t\leq 0\}}, \tag{VI.3.5}$$

with $A_\pm : \mathbb{R}_+ \to \mathbb{R}$ *given by*

$$A_\pm(l) \quad := \quad \lambda'(0\pm) \;+\; \int_0^l \frac{dy}{\phi_\pm(y)} e^{\gamma(y)} \int_y^{+\infty} e^{-\gamma(z)} F''(dz). \tag{VI.3.6}$$

Proposition VI.3.3 implies that, in order to construct a super-replication, it suffices to consider a pair (ϕ_+, ϕ_-) satisfying Assumption VI.3.2. The duality and the optimality will follow once we find an optimal pair as it will be shown in the next section.

Remark VI.3.4. *A slightly different pathwise inequality is derived in Appendix A. Notice also that* λ' *coincides with* $A_+ \circ \psi_+$ *(resp.* $A_- \circ \psi_-$*) on* $\mathbb{R}_+$ *(resp.* $\mathbb{R}-$*) by definition and thus* $A_\pm = \lambda' \circ \phi_\pm$ *if* $\phi_\pm$ *are strictly monotone or, equivalently,* $\psi_\pm$ *are continuous.*

Before giving the proof of Proposition VI.3.3, we show that the pathwise inequality yields an upper bound for $\overline{\mathsf{D}}(\mu)$.

Corollary VI.3.5. *Under Assumptions VI.2.2 and VI.3.2, one has for any centered probability measure* μ*,*

$$\overline{\mathsf{P}}(\mu) \quad \le \quad \overline{\mathsf{D}}(\mu) \quad \le \quad \mu(\lambda).$$

Proof. It suffices to prove that $\overline{\mathsf{D}}(\mu) \le \mu(\lambda)$. This result follows immediately from Proposition VI.3.3 providing that $\lambda \in \Lambda$ and $H \in \overline{\mathcal{H}}$. The former is a straightforward consequence of the boundedness of λ' which is proved in Lemma VI.3.6 below. Let us show now that $H \in \overline{\mathcal{H}}$. Since H is bounded, then $\int_0^\cdot H_t dX_t$ is a local martingale under all $\mathbb{P} \in \mathcal{M}$. Further, the pathwise inequality (VI.3.4) implies that for all $\mathbb{P} \in \mathcal{M}$,

$$\int_0^t H_s dX_s \quad \ge \quad F(L_t) \;-\; \lambda(X_t) \quad \ge \quad -C(L_t \;+\; |X_t|) \text{ for all } t \in [0,1], \;\; \mathbb{P}\text{ - a.s.}$$

where $C := \|F'\| \vee \|H'\|$. The conclusion results from the fact that a local martingale bounded from below by a non-positive supermartingale is a supermartingale by Fatou's Lemma. $\square$

The rest of the section is devoted to the proof of Proposition VI.3.3. We start by establishing in the following lemma some technical results with regard to λ and $A_\pm$. $\square$

Lemma VI.3.6. *With the notations of Proposition VI.3.3, it holds for all* $l > 0$*,*

$$\frac{1}{2}\left(A_+(l) \;-\; A_-(l)\right) \quad = \quad F'(l) \;+\; e^{\gamma(l)} \int_l^{+\infty} e^{-\gamma(z)} F''(dz), \tag{VI.3.7}$$

$$\lambda(\phi_+(l)) \;-\; A_+(l)\phi_+(l) \quad = \quad \lambda(\phi_-(l)) \;-\; A_-(l)\phi_-(l). \tag{VI.3.8}$$

In addition, the functions $A_\pm$ are uniformly bounded on $\mathbb{R}_+$.

Proof. The first identity results easily from applying Fubini-Tonelli's theorem in the following relation :

$$\frac{1}{2}\left(A_+(l) - A_-(l)\right) \;=\; \frac{1}{2}\left(\lambda'(0+) - \lambda'(0-)\right) \;+\; \int_0^l \gamma'(y)e^{\gamma(y)}\int_y^{+\infty} e^{-\gamma(z)}F''(dz)dy.$$

Further, we have

$$\lambda'(0+) \;\leq\; A_+(l) \;\leq\; A_+(l) \;+\; \lambda'(0-) \;-\; A_-(l).$$

Using the relation (VI.3.7), we deduce that $|A_+(l)| \leq |\lambda'(0+)| + 2\|F'\|$. Repeating the same arguments, we can also show that $|A_-(l)| \leq |\lambda'(0-)| + 2\|F'\|$. Let us turn now to the identity (VI.3.8). By change of variable, we get

$$\lambda(\phi_+(l)) \;-\; \lambda(0) \;=\; \int_{[0,l]} A_+(z)\phi'_+(dz),$$

where ϕ'_+ denotes the first derivative of ϕ_+ in the sense of distributions. By integration by parts, we obtain

$$\lambda(\phi_+(l)) \;-\; A_+(l)\phi_+(l) \;=\; \lambda(0) \;-\; \int_0^l e^{\gamma(y)}\int_y^{+\infty} e^{-\gamma(z)}F''(dz)dy. \qquad \text{(VI.3.9)}$$

Repeating the same arguments, we can show that $\lambda(\phi_-(l)) - A_-(l)\phi_-(l)$ coincides with the r.h.s. above, which ends the proof. $\square$

Proof of Proposition VI.3.3. Let $u : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ be given by

$$u(x,l) \;:=\; -A_+(l)x^+ \;+\; A_-(l)x^- \;+\; A_+(l)\phi_+(l) \;-\; \lambda(\phi_+(l)) \;+\; F(l).$$

We notice first that

$$\lim_{x\nearrow\phi_\pm(l)} \lambda'(x) \;\leq\; A_\pm(l) \;\leq\; \lim_{x\searrow\phi_\pm(l)} \lambda'(x).$$

Using the convexity of λ and the relation (VI.3.8), we deduce that $u(x,l) \geq F(l) - \lambda(x)$. Let us prove now that

$$u\left(X_1, L_1\right) \;=\; \int_0^1 H_t dX_t,$$

which clearly yields the required inequality. Using successively Itô-Tanaka's formula and the relation (VI.3.7), we derive

$$
\begin{aligned}
&- A_+(L_1)X_1^+ \;+\; A_-(L_1)X_1^- \\
&\qquad\qquad = \int_0^1 H_t dX_t \;-\; \frac{1}{2}\int_0^{L_1} \left(A_+(l) - A_-(l)\right) dl \\
&\qquad\qquad = \int_0^1 H_t dX_t \;-\; F(L_1) \;+\; F(0) \;-\; \int_0^{L_1} e^{\gamma(l)} \int_l^{+\infty} e^{-\gamma(z)} F''(dz)dl.
\end{aligned}
$$

We deduce that

$$
\begin{aligned}
u(X_1, L_1) \;=\; & \int_0^1 H_t dX_t \;+\; A_+(L_1)\phi_+(L_1) \;-\; \lambda(\phi_+(L_1)) \\
& \;+\; F(0) \;-\; \int_0^{L_1} e^{\gamma(l)} \int_l^{+\infty} e^{-\gamma(z)} F''(dz)dl.
\end{aligned}
$$

The conclusion follows immediately by using the identity (VI.3.9). □

3.1.2 Optimality and duality

In view of Corollary VI.3.5, the duality is achieved once we find a suitable pair $(\overline{\phi}_+, \overline{\phi}_-)$ such that the corresponding static strategy $\overline{\lambda}$ satisfies $\mu(\overline{\lambda}) = \overline{\mathsf{P}}(\mu)$. In this section, we use Vallois' solution to the SEP to construct such a pair and to provide the optimizers for both $\overline{\mathsf{P}}(\mu)$ and $\overline{\mathsf{D}}(\mu)$.

Recall that $(B_t)_{t\geq 0}$ and $(L_t^B)_{t\geq 0}$ denote a Brownian motion and its local time at zero. Under Assumption VI.3.1, there exists a pair $(\overline{\phi}_+, \overline{\phi}_-)$ satisfying Assumption VI.3.2 such that $B_{\overline{\tau}\wedge\cdot}$ is uniformly integrable and $B_{\overline{\tau}} \sim \mu$, where

$$
\overline{\tau} \;:=\; \inf\left\{t > 0 : \quad B_t \notin (\overline{\phi}_-(L_t^B), \overline{\phi}_+(L_t^B))\right\}.
$$

We refer to Vallois [102] and Cox, Hobson & Obłój [23] for the proof.

Remark VI.3.7. *If μ admits a positive density $\mu(x)dx$ with respect to the Lebesgue measure, it holds*

$$
\overline{\phi}'_\pm(l) \;=\; \frac{1 - \mu\left([\overline{\phi}_-(l), \overline{\phi}_+(l)]\right)}{2\overline{\phi}_\pm(l)\mu(\overline{\phi}_\pm(l))} \;\; \textit{for all } l > 0.
$$

If we assume further that μ is symmetric, then $\overline{\phi}_{\pm} = \pm\overline{\phi}$ and

$$\overline{\psi}(x) \quad = \quad \int_0^x \frac{y\mu(y)}{\mu\left([y,+\infty)\right)}dy \text{ for all } x \in \mathbb{R}_+,$$

where $\overline{\psi}$ denotes the inverse of $\overline{\phi}$.

The following theorem, which is the main result of this section, shows that the pair given by Vallois yields the duality and the optimizers for both $\overline{\mathsf{P}}(\mu)$ and $\overline{\mathsf{D}}(\mu)$.

Theorem VI.3.8. *Under Assumptions VI.2.2 and VI.3.1, there is no duality gap, i.e.,*

$$\overline{\mathsf{P}}(\mu) \quad = \quad \overline{\mathsf{D}}(\mu) \quad = \quad \mu(\overline{\lambda}),$$

where $\overline{\lambda}$ is constructed by (VI.3.1), (VI.3.2) and (VI.3.3) from $(\overline{\phi}_+, \overline{\phi}_-)$. In addition, there exists an optimizer $\overline{\mathbb{P}}$ for $\overline{\mathsf{P}}(\mu)$ such that

$$\overline{\lambda}(X_1) \;+\; \int_0^1 \overline{H}_t dX_t \quad = \quad F(L_1),\ \overline{\mathbb{P}}\text{ - a.s.,} \tag{VI.3.10}$$

where the process $\overline{H}$ is given by (VI.3.5) with $(\overline{\phi}_+, \overline{\phi}_-)$.

Proof. Denote by $\overline{\mathbb{P}}$ the law of the process $Z = (Z_t)_{0\le t\le 1}$ given by

$$Z_t \quad := \quad B_{\overline{\tau}\wedge\frac{t}{T-t}} \text{ for all } t \in [0,1].$$

Since Z is a continuous martingale with respect to its natural filtration such that $Z_1 \sim \mu$, then $\overline{\mathbb{P}}$ clearly belongs to $\mathcal{M}(\mu)$. Let us turn first to the proof of (VI.3.10). Define $\overline{u} : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ by

$$\overline{u}(x,l) \quad := \quad -\overline{A}_+(l)x^+ \;+\; \overline{A}_-(l)x^- \;+\; \overline{A}_+(l)\overline{\phi}_+(l) \;-\; \overline{\lambda}(\overline{\phi}_+(l)) \;+\; F(l).$$

where $\overline{A}_\pm$ is given by (VI.3.6) with $(\overline{\phi}_+, \overline{\phi}_-)$. Since the local time is invariant under time change, we have $X_1 = \overline{\phi}_\pm(L_1)$ almost surely under $\overline{\mathbb{P}}$. Thus, it clearly holds

$$\overline{u}(X_1, L_1) \quad = \quad F(L_1) \;-\; \overline{\lambda}(X_1),\ \overline{\mathbb{P}}\text{ - a.s.}$$

By repeating the arguments of Proposition VI.3.3, we obtain

$$\overline{\lambda}(X_1) \;+\; \int_0^1 \overline{H}_t dX_t \quad = \quad F(L_1),\ \overline{\mathbb{P}}\text{ - a.s.}$$

To conclude, it remains to show that $\mathbb{E}^{\overline{\mathbb{P}}}[F(L_T)] = \mu(\overline{\lambda})$. This will be achieved in Lemma VI.3.9 below by proving that

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[\int_0^1 \overline{H}_t dX_t\right] \;=\; 0.$$

□

Lemma VI.3.9. *With the notations of Theorem VI.3.8, one has*

$$\mathbb{E}^{\overline{\mathbb{P}}}\left[\int_0^1 \overline{H}_t dX_t\right] \;=\; 0.$$

Proof. First, we notice that the desired result is equivalent to

$$\mathbb{E}\left[\int_0^{\overline{\tau}} \left(\overline{A}_+(L_s^B)\mathbb{1}_{\{B_s>0\}} \;+\; \overline{A}_-(L_s^B)\mathbb{1}_{\{B_s\leq 0\}}\right) dB_s\right] \;=\; 0,$$

since the local time is invariant under time change. For the sake of clarity, we omit the overline of the notations in the rest of the proof. Let $\sigma_n := \inf\{t\in\mathbb{R}_+ : |B_t|\geq n\}$, $\rho_m := \inf\{t\in\mathbb{R}_+ : L_t^B \geq m\}$ and $\tau_{n,m} := \tau\wedge\sigma_n\wedge\rho_m$. Denote $M = (M_t)_{t\geq 0}$ by

$$M_t \;:=\; \int_0^t \left(A_+(L_s^B)\mathbb{1}_{\{B_s>0\}} \;+\; A_-(L_s^B)\mathbb{1}_{\{B_s\leq 0\}}\right) dB_s.$$

From Itô-Tanaka's formula, it follows that

$$M_t \;=\; A_+(L_t^B)B_t^+ \;-\; A_-(L_t^B)B_t^- \;-\; \frac{1}{2}\int_0^t \left(A_+(L_s^B) - A_-(L_s^B)\right) dL_s^B.$$

We deduce that the stopped local martingale $M_{\tau_{m,n}\wedge\cdot}$ is bounded. Hence, it is a uniformly integrable martingale and we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{2}\int_0^{\tau_{n,m}} \left(A_+(L_s^B) - A_-(L_s^B)\right) dL_s^B\right] \\
&\qquad = \;\mathbb{E}\left[A_+(L_{\tau_{n,m}}^B)B_{\tau_{n,m}}^+ \;-\; A_-(L_{\tau_{n,m}}^B)B_{\tau_{n,m}}^-\right] \\
&\qquad = \;\mathbb{E}\left[\left(A_+(L_{\tau_{n,\infty}}^B)B_{\tau_{n,\infty}}^+ \;-\; A_-(L_{\tau_{n,\infty}}^B)B_{\tau_{n,\infty}}^-\right)\mathbb{1}_{\{\tau\wedge\sigma_n\leq\rho_m\}}\right].
\end{aligned}$$

It follows that

$$\mathbb{E}\left[\frac{1}{2}\int_0^{\tau_{n,m}}\left(\left(A_+(L_s^B)-H'(0+)\right) \;-\; \left(A_-(L_s^B)-H'(0-)\right)\right)dL_s^B\right]$$
$$= \;\; \mathbb{E}\left[\left(\left(A_+(L^B_{\tau_{n,\infty}})-H'(0+)\right)B^+_{\tau_{n,\infty}} \;-\; \left(A_-(L^B_{\tau_{n,\infty}})-H'(0-)\right)B^-_{\tau_{n,\infty}}\right)\mathbb{1}_{\{\tau\wedge\sigma_n\le\rho_m\}}\right].$$

By the monotone convergence theorem, as m tends to infinity, we obtain

$$\mathbb{E}\left[\frac{1}{2}\int_0^{\tau_{n,\infty}}\left(\left(A_+(L_s^B)-H'(0+)\right) \;-\; \left(A_-(L_s^B)-H'(0-)\right)\right)dL_s^B\right]$$
$$= \;\; \mathbb{E}\left[\left(A_+(L^B_{\tau_{n,\infty}})-H'(0+)\right)B^+_{\tau_{n,\infty}} \;-\; \left(A_-(L^B_{\tau_{n,\infty}})-H'(0-)\right)B^-_{\tau_{n,\infty}}\right].$$

Then, as n tends to infinity, the l.h.s. converges, again by the monotone convergence theorem, to

$$\mathbb{E}\left[\frac{1}{2}\int_0^{\tau}\left(\left(A_+(L_s^B)-H'(0+)\right) \;-\; \left(A_-(L_s^B)-H'(0-)\right)\right)dL_s^B\right].$$

As for the r.h.s., it converges to

$$\mathbb{E}\left[\left(A_+(L^B_{\tau})-H'(0+)\right)B^+_{\tau} \;-\; \left(A_-(L^B_{\tau})-H'(0-)\right)B^-_{\tau}\right] \;\;<\;\; +\infty,$$

since H' is bounded and $(B^+_{t\wedge\tau})_{t\ge0}$ and $(B^-_{t\wedge\tau})_{t\ge0}$ are uniformly integrable. Hence, we obtain

$$\mathbb{E}\left[\frac{1}{2}\int_0^{\tau}\left(A_+(L_s^B)-A_-(L_s^B)\right)dL_s^B\right] \;\;=\;\; \mathbb{E}\left[A_+(L^B_\tau)B^+_\tau - A_-(L^B_\tau)B^-_\tau\right],$$

where both sides are finite. This ends the proof. □

3.2 Robust subhedging problem

In this section, we address the robust subhedging problem. In contrast with the results on the superhedging problem which are contained in Cox, Hobson & Obłój [23]., the results of this section are new to the literature. We proceed along the lines of Section 3.1, but we reverse the monotonicity assumption on the functions ϕ_+ and ϕ_-.

Assumption VI.3.10. *$\phi_+ : \mathbb{R}_+^* \to \mathbb{R}_+^*$ (resp. $\phi_- : \mathbb{R}_+^* \to \mathbb{R}_-^*$) is right-continuous and nonincreasing (resp. nondecreasing).*

Under Assumption VI.3.2, we define the static strategy $\lambda : \mathbb{R} \to \mathbb{R}$ via (ϕ_+, ϕ_-) by

$$\lambda'(\pm x) \;:=\; \lambda'(0\pm) \;-\; \int_{\psi_\pm(\pm x)}^{+\infty} \frac{dy}{\phi_\pm(y)} e^{\gamma(y)} \int_y^{+\infty} e^{-\gamma(z)} F''(dz) \text{ for all } x > 0, \tag{VI.3.11}$$

$$\lambda'(0\pm) \;:=\; \pm F'(+\infty), \tag{VI.3.12}$$

$$\lambda(0) \;:=\; F(0) \;-\; \int_0^{+\infty} e^{\gamma(y)} \int_y^{+\infty} e^{-\gamma(z)} F''(dz) dy, \tag{VI.3.13}$$

where $\psi_\pm$ denote the right-continuous inverses of $\phi_\pm$, F' is the right derivative of F, $F''(dz)$ is the second derivative of F in the sense of distributions and

$$\gamma(l) \;:=\; \frac{1}{2} \int_0^l \left(\frac{1}{\phi_+(z)} \;-\; \frac{1}{\phi_-(z)} \right) dz \text{ for all } l > 0.$$

Notice that the restriction of λ on $\mathbb{R}_+$ and $\mathbb{R}_-$ is now concave.

3.2.1 Pathwise inequality

We start by showing the pathwise inequality corresponding to the subhedging problem. Together with the second solution provided by Vallois to the SEP, it leads to the relative results for the robust subhedging problem.

Proposition VI.3.11. *Under Assumptions VI.2.2 and VI.3.10, the following inequality holds*

$$\lambda(X_1) \;+\; \int_0^1 H_t dX_t \;\;\le\;\; F(L_1), \;\; \mathbb{P}\text{ - a.s. for all } \mathbb{P} \in \mathcal{M}, \tag{VI.3.14}$$

where the function λ is defined by (VI.3.11), (VI.3.12), (VI.3.13) and the process $H = (H_t)_{0\le t\le 1}$ is given by

$$H_t \;:=\; -A_+(L_t)\mathbb{1}_{\{X_t>0\}} \;-\; A_-(L_t)\mathbb{1}_{\{X_t\le 0\}}, \tag{VI.3.15}$$

with $A_\pm : \mathbb{R}_+ \to \mathbb{R}$ given by

$$A_\pm(l) \;:=\; \lambda'(0\pm) \;-\; \int_l^{+\infty} \frac{dy}{\phi_\pm(y)} e^{\gamma(y)} \int_y^{+\infty} e^{-\gamma(z)} F''(dz). \tag{VI.3.16}$$

Corollary VI.3.12. *Under Assumptions VI.2.2 and VI.3.10, one has for any centered probability measure μ,*

$$\underline{\mathsf{P}}(\mu) \;\;\ge\;\; \underline{\mathsf{D}}(\mu) \;\;\ge\;\; \mu(\lambda).$$

The proof of Corollary VI.3.12 is identical to the proof of Corollary VI.3.5. However, the proof of the pathwise inequality (VI.3.14) is not completely straightforward and thus we provide some details below.

Proof of Proposition VI.3.11. Let us show first

$$\begin{aligned} \frac{1}{2}\left(A_+(l) - A_-(l)\right) &= F'(l) + e^{\gamma(l)}\int_l^{+\infty} e^{-\gamma(z)}F''(dz), && \text{(VI.3.17)}\\ \lambda(\phi_+(l)) - A_+(l)\phi_+(l) &= \lambda(\phi_-(l)) - A_-(l)\phi_-(l). && \text{(VI.3.18)}\end{aligned}$$

The first identity follows from the same arguments as in the proof of Lemma VI.3.6. As for the second one, by change of variable and integration by parts, we have

$$\begin{aligned} \lambda(\phi_+(+\infty)) - \lambda(\phi_+(l)) &= \int_{(l,+\infty)} A_+(z)\phi_+'(dz)\\ &= A_+(+\infty)\phi_+(+\infty) - A_+(l)\phi_+(l)\\ &\qquad - \int_l^{+\infty} e^{\gamma(y)}\int_y^{+\infty} e^{-\gamma(z)}F''(dz)dy.\end{aligned}$$

Further, by the definition of λ, we have $\lambda(\phi_+(+\infty)) - \lambda(0) = \lambda'(0+)\phi_+(+\infty) = A_+(+\infty)\phi_+(+\infty)$. Hence, we obtain

$$\lambda(\phi_+(l)) - A_+(l)\phi_+(l) = F(0) - \int_0^l e^{\gamma(y)}\int_y^{+\infty} e^{-\gamma(z)}F''(dz)dy.$$

Similarly, we can show that $\lambda(\phi_-(l)) - A_-(l)\phi_-(l)$ coincides with the r.h.s. above. The rest of the proof follows by repeating the arguments of Proposition VI.3.3. □

3.2.2 Optimality and duality

Using another solution to the SEP provided by Vallois, we derive the duality and provide the optimizers for $\underline{\mathrm{P}}(\mu)$ and $\underline{\mathrm{D}}(\mu)$.

Under Assumption VI.3.1, there exists a pair $(\underline{\phi}_+, \underline{\phi}_-)$ satisfying Assumption VI.3.10 such that $B_{\underline{\tau}\wedge\cdot}$ is uniformly integrable and $B_{\underline{\tau}} \sim \mu$, where

$$\underline{\tau} := \inf\left\{t > 0 : B_t \notin (\underline{\phi}_-(L_t^B), \underline{\phi}_+(L_t^B))\right\}.$$

We refer to Vallois [102] for the proof.

Theorem VI.3.13. *Under Assumptions VI.2.2 and VI.3.1, there is no duality gap, i.e.,*

$$\underline{\mathsf{P}}(\mu) \quad = \quad \underline{\mathsf{D}}(\mu) \quad = \quad \mu(\underline{\lambda}),$$

where $\underline{\lambda}$ is constructed by (VI.3.11), (VI.3.12) and (VI.3.13) from $(\underline{\phi}_+, \underline{\phi}_-)$. In addition, there exists an optimizer $\underline{\mathbb{P}}$ for $\underline{\mathsf{D}}(\mu)$ such that

$$\underline{\lambda}(X_1) \ + \ \int_0^1 \underline{H}_t dX_t \quad = \quad F(L_1), \ \underline{\mathbb{P}} \text{ - a.s.},$$

where the process $\underline{H}$ is given by (VI.3.15) with $(\underline{\phi}_+, \underline{\phi}_-)$.

The proof of this result is identical to the proof of Theorem VI.3.8.

Remark VI.3.14. *The assumption that μ has no mass at zero can be dropped in Theorem VI.3.13. In this setting, the embedding functions given by Vallois can reach zero and we need to modify slightly the definitions of $A_\pm$ and H by replacing the upper bound $+\infty$ in the integral by $\psi_+(0)$.*

3.3 A special case : Ocone martingales

In practice, the practitioners have a belief on the dynamic of Vanillas. For instance, they may expect that the price X_1 is symmetric around the spot X_t for any future date $t \in (0,1)$. In particular, this is true if the process X is a continuous Ocone martingale, *i.e.*, $X_t = B_{\langle X \rangle_t}$ where B is a Brownian motion independent of $\langle X \rangle$.
For any symmetric distribution μ satisfying Assumption VI.3.1, we introduce the set $\mathcal{O}(\mu) \subseteq \mathcal{M}(\mu)$ of all probability measures $\mathbb{P}$ under which X is an Ocone martingale. Then it follows by Skorokhod's lemma that

$$L_1 \quad = \quad \sup_{0 \le t \le 1} \ -\int_0^t \operatorname{sgn}(X_s) dX_s.$$

Further, the processes $-\int_0^\cdot \operatorname{sgn}(X_s) dX_s$ and X are identical in law by the definition of Ocone martingales, see Ocone [85]. It follows by the reflection principle that

$$L_1 \ \overset{\mathbb{P}}{\sim} \ \sup_{0 \le t \le 1} X_t \ \overset{\mathbb{P}}{\sim} \ |X_1| \text{ for all } \mathbb{P} \in \mathcal{O}(\mu).$$

Hence, one has for every $\mathbb{P} \in \mathcal{O}(\mu)$,

$$\mathbb{E}^{\mathbb{P}}\Big[F(L_T)\Big] \quad = \quad 2\int_{\mathbb{R}_+} F(x)\mu(dx).$$

This shows that, by adding a belief on the dynamic of Vanillas, the interval of no-arbitrage prices $[\underline{\mathsf{P}}(\mu), \overline{\mathsf{P}}(\mu)]$ can be tightened. In the present case, it is reduced to a singleton. The duality and the determination of optimal robust hedging strategies are left for further research.

4 Two-marginal case

Using the stochastic control approach, we aim at extending the analysis of Section 3 to the two-marginal case. Although we are not able to solve completely the robust hedging problems in this setting, we give candidates for the optimal strategies and we provide a new solution to the two-marginal SEP as an extension of Vallois' solution. In the sequel, we consider only the robust superhedging problem for the sake of simplicity but the arguments can be easily adapted to the robust subhedging problem.

4.1 Two-marginal Skorokhod embedding

In this section, we provide a new solution to the two-marginal SEP as an extension of Vallois' solution. Namely, given a centered peacock $\boldsymbol{\mu} = (\mu_1, \mu_2)$, we want to construct a pair of stopping rules based on the local time such that $\tau_1 \leq \tau_2$, $B_{\tau_2 \wedge \cdot}$ is uniformly integrable and $B_{\tau_i} \sim \mu_i$ for $i = 1, 2$. A natural idea is to take Vallois' embeddings corresponding to μ_1 and μ_2. However these stopping times are not ordered in general and thus we need to be more careful. For technical reasons, we make the following assumption on the marginals.

Assumption VI.4.1. $\boldsymbol{\mu} = (\mu_1, \mu_2)$ *is a centered peacock such that* μ_1 *and* μ_2 *are symmetric and equivalent to the Lebesgue measure.*

4.1.1 Generalization of Vallois' solution

For the first stopping time, we take the solution given by Vallois [102] that embeds μ_1, *i.e.*,

$$\tau_1 \;\; := \;\; \inf \Big\{ t > 0 : \; |B_t| \; \geq \; \phi_1(L_t^B) \Big\},$$

where $\phi_1 : \mathbb{R}_+ \to \mathbb{R}_+$ is the inverse of ψ_1 given by

$$\psi_1(x) \;\; := \;\; \int_0^x \frac{y\mu_1(y)}{\mu_1([y, +\infty))} dy \text{ for all } x \in \mathbb{R}_+. \qquad \text{(VI.4.1)}$$

For the second stopping time, we look for an increasing function $\phi_2 : \mathbb{R}_+ \to \mathbb{R}_+$ such that $B_{\tau_2 \wedge \cdot}$ is uniformly integrable and $B_{\tau_2} \sim \mu_2$, where

$$\tau_2 \;\; := \;\; \inf \Big\{ t \geq \tau_1 : \; |B_t| \;\geq\; \phi_2(L_t^B) \Big\}.$$

As before, ϕ_2 is defined through its inverse ψ_2. First, we set

$$\psi_2(x) \;\; := \;\; \int_0^x \frac{y\mu_2(y)}{\mu_2([y, +\infty))} dy \text{ for all } x \in [0, x_1], \tag{VI.4.2}$$

where x_1 is the smallest element such that the r.h.s. above exceeds ψ_1. To ensure that $x_1 > 0$, we need to assume that $\delta\mu := \mu_2 - \mu_1 \leq 0$ on a neighborhood of zero. If $x_1 = +\infty$, the construction is over. This corresponds to the case when Vallois's solutions are well ordered. Otherwise, we proceed by induction as follows :
(i) if $x_{2i-1} < +\infty$, we set for all $x \in (x_{2i-1}, x_{2i}]$,

$$\psi_2(x) \;\; := \;\; \psi_2(x_{2i-1}) \; + \; \int_{x_{2i-1}}^x \frac{y\delta\mu(y)}{\delta\mu([y, +\infty))} dy, \tag{VI.4.3}$$

where x_{2i} is the smallest element such that ψ_1 exceeds the r.h.s. above ;
(ii) if $x_{2i} < +\infty$, we set for all $x \in (x_{2i}, x_{2i+1}]$,

$$\psi_2(x) \;\; := \;\; \psi_2(x_{2i}) \; + \; \int_{x_{2i}}^x \frac{y\mu_2(y)}{\mu_2([y, +\infty))} dy, \tag{VI.4.4}$$

where x_{2i+1} is the smallest element such that the r.h.s. above exceeds ψ_1.
To ensure that $\psi_2^{\boldsymbol{\mu}}$ is well-defined and increasing, we need to make the following assumption.

Assumption VI.4.2. (i) *$\delta\mu \leq 0$ and $\delta\mu \not\equiv 0$ on a neighborhood of zero ;*
(ii) *$\delta\mu > 0$ whenever $\psi_1 < \psi_2$;*
(iii) *$x_i = +\infty$ for some $i \geq 1$.*

Theorem VI.4.3. *Under Assumptions VI.4.1 and VI.4.2, if ψ_2 is given by* (VI.4.3) *and* (VI.4.4)*, then $B_{\tau_2 \wedge \cdot}$ is uniformly integrable and $B_{\tau_2} \sim \mu_2$.*

The proof of this result will be performed in the next section. Notice that both (i) and (ii) of Assumption VI.4.2 are to a certain extent necessary to ensure the existence of an increasing function ψ_2 that solves the two-marginal SEP as above. See Remark VI.4.4 below for more details. This suggests that one needs to relax the monotonicity assumption on ψ_2 in order to iterate Vallois' solution. However, our approach does not allow to compute the function ψ_2

in this general setting. Notice also that (iii) ensures that ψ_2 is well defined and satisfies an integrable condition. In view of the proof below, it is clear that our construction works in a more general setting.

4.1.2 Proof of Theorem VI.4.3

(i) We start by showing that the distribution of B_{τ_2} admits a density with respect to the Lebesgue measure. We claim that ψ_2 is increasing and $\gamma_2(+\infty) = +\infty$, which is proved in (iii). Notice first that the distribution of B_{τ_2} is symmetric by construction. By the strong Markov property and Lemma VI.6.3, if we take the function $\lambda(y) = \mathbb{1}_{\{y \le x\}}$ for some $x \in \mathbb{R}_+$, it holds

$$\mathbb{E}\Big[\lambda\left(|B_{\tau_2}|\right)\Big] \;=\; \mathbb{E}\Big[f\left(B_{\tau_1}, L^B_{\tau_1}\right)\Big] \;=\; \int_0^{+\infty} \frac{f\left(\phi_1(l), l\right)}{\phi_1(l)} e^{-\gamma_1(l)} dl,$$

where

$$f(x,l) \;=\; \begin{cases} -c'(l)|x| \;+\; c(l), & \text{if } |x| < \phi_2(l), \\ \lambda(|x|), & \text{otherwise,} \end{cases} \quad \text{with } c(l) \;:=\; e^{\gamma_2(l)} \int_l^{+\infty} \frac{\lambda(\phi_2(z))}{\phi_2(z)} e^{-\gamma_2(z)} dz.$$

By a straightforward calculation, we get

$$c(l) \;=\; \left(1 - e^{\gamma_2(l) \;-\; \gamma_2(\psi_2(x))}\right) \mathbb{1}_{\{l \le \psi_2(x)\}},$$

and

$$\frac{f\left(\phi_1(l), l\right)}{\phi_1(l)} \;=\; \begin{cases} \left(\frac{1}{\phi_1(l)} \;+\; e^{\gamma_2(l) - \gamma_2(\psi_2(x))} \left(\frac{1}{\phi_2(l)} - \frac{1}{\phi_1(l)}\right)\right) \mathbb{1}_{\{l \le \psi_2(x)\}}, & \text{if } \phi_1(l) < \phi_2(l), \\ \frac{1}{\phi_1(l)} \mathbb{1}_{\{l \le \psi_1(x)\}}, & \text{otherwise.} \end{cases}$$

Hence, we obtain

$$\begin{aligned} \mathbb{P}\Big[\,|B_{\tau_2}| \;\le\; x\Big] \;=\; \sum_{i \ge 0} \Bigg(& e^{-\gamma_2(\psi_2(x))} \int_{l_{2i}}^{l_{2i+1}} \mathbb{1}_{\{l \le \psi_2(x)\}} de^{\gamma_2(l) - \gamma_1(l)} \\ & - \int_{l_{2i}}^{l_{2i+1}} \mathbb{1}_{\{l \le \psi_2(x)\}} de^{-\gamma_1(l)} \;-\; \int_{l_{2i+1}}^{l_{2i+2}} \mathbb{1}_{\{l \le \psi_1(x)\}} de^{-\gamma_1(l)} \Bigg). \end{aligned}$$

where $l_i := \psi_1(x_i) = \psi_2(x_i)$. We deduce by the direct differentiation of the identity above that the distribution of B_{τ_2} admits a density ν with respect to the Lebesgue measure given by

$$\nu(x) \quad = \quad \begin{cases} \dfrac{\psi_2'(x)}{x} S_{2i}(x), & \text{if } x \in (x_{2i}, x_{2i+1}], \\ \dfrac{\psi_1'(x)}{2x} e^{-\gamma_1(\psi_1(x))} \; + \; \dfrac{\psi_2'(x)}{x} S_{2i+1}(x), & \text{if } x \in (x_{2i+1}, x_{2i+2}], \end{cases}$$

where

$$S_i(x) \quad := \quad \frac{e^{-\gamma_2(\psi_2(x))}}{2} \sum_{j=0}^{i} (-1)^j e^{\gamma_2(l_j)-\gamma_1(l_j)}.$$

(ii) Let us show now that ν coincides with μ_2. Notice first that the relation (VI.4.1) yields

$$e^{-\gamma_1(\psi_1(x))} \quad = \quad 2\mu_1([x, +\infty)) \text{ for all } x \in \mathbb{R}_+.$$

Using further the identities (VI.4.3) and (VI.4.4), it follows that

$$\nu(x) \quad = \quad \begin{cases} \dfrac{\mu_2(x)}{\mu_2([x, +\infty))} S_{2i}(x), & \text{if } x \in (x_{2i}, x_{2i+1}], \\ \mu_1(x) \; + \; \dfrac{\delta\mu(x)}{\delta\mu([x, +\infty))} S_{2i+1}(x), & \text{if } x \in (x_{2i+1}, x_{2i+2}]. \end{cases}$$

To conclude, it remains to show that $S_{2i}(x) = \mu_2([x, +\infty))$ for all $x \in (x_{2i}, x_{2i+1}]$ and $S_{2i+1}(x) = \delta\mu([x, +\infty))$ for all $x \in (x_{2i+1}, x_{2i+2}]$. For $i = 0$, it follows from the relation (VI.4.2) that

$$S_0(x) \quad = \quad \frac{e^{-\gamma_2(\psi_2(x))}}{2} \quad = \quad \mu_2([x, +\infty)) \text{ for all } x \in [0, x_1].$$

In addition, Assumption VI.4.1 (i) ensures that

$$\delta\mu([x_1, +\infty)) \quad = \quad \frac{e^{-\gamma_2(l_1)} - e^{-\gamma_1(l_1)}}{2} \quad > \quad 0.$$

Assume that $S_{2i}(x) = \mu_2([x, +\infty))$ for all $x \in (x_{2i}, x_{2i+1}]$. It results from the relation (VI.4.3) that

$$e^{-\gamma_2(\psi_2(x))+\gamma_2(l_{2i+1})} \quad = \quad \frac{\delta\mu([x, +\infty))}{\delta\mu([x_{2i+1}, +\infty))} \text{ for all } x \in (x_{2i+1}, x_{2i+2}].$$

Hence, we deduce that for all $x \in (x_{2i+1}, x_{2i+2}]$,

$$S_{2i+1}(x) \quad = \quad e^{-\gamma_2(\psi_2(x))+\gamma_2(l_{2i+1})}\left(S_{2i}\left(x_{2i+1}\right) - e^{-\gamma_1(l_{2i+1})}\right) \quad = \quad \delta\mu([x,+\infty)).$$

In particular, it follows that $\delta\mu([x,+\infty)) > 0$ for all $x \in (x_{2i+1}, x_{2i+2}]$. Further, it results from the relation (VI.4.4) that

$$e^{-\gamma_2\circ\psi_2(x)+\gamma_2(l_{2i+2})} \quad = \quad \frac{\mu_2([x,+\infty))}{\mu_2([x_{2i+2},+\infty))} \text{ for all } x \in (x_{2i+2}, x_{2i+3}].$$

Hence, we deduce that for all $x \in (x_{2i+2}, x_{2i+3}]$,

$$S_{2i+2}(x) \quad = \quad e^{-\gamma_2(\psi_2(x))+\gamma_2(l_{2i+2})}\left(S_{2i+1}\left(x_{2i+2}\right) + e^{-\gamma_1(l_{2i+2})}\right) \quad = \quad \mu_2([x,+\infty)).$$

In addition, it holds

$$\delta\mu([x_{2i+3},+\infty)) \quad = \quad S_{2i+2}(x_{2i+3}) \; - \; \frac{e^{-\gamma_1(l_{2i+3})}}{2} \quad > \quad 0.$$

(iii) We are now in a position to complete the proof. The relation (VI.4.4) clearly imposes that ψ_2 is increasing on every interval such that $\psi_1 > \psi_2$. Under Assumption VI.4.2, the relation (VI.4.3) ensures that ψ_2 is increasing on every interval such that $\psi_1 < \psi_2$ since

$$\delta\mu\left([x,+\infty)\right) \; > \; 0 \text{ for all } x \in [x_{2i+1}, x_{2i+2}].$$

In addition, one may easily check that $\gamma_2(+\infty) = +\infty$ in view of Assumption VI.4.1 (iii). It remains to prove that the stopped process $B_{\tau_2\wedge\cdot}$ is uniformly integrable. Since $|B_{\tau_2\wedge t}| \leq |B_{\tau_2}|$ for all $t \in \mathbb{R}_+$ by construction, the uniform integrability follows immediately from the assumption $\mu_2(|x|) < +\infty$. □

Remark VI.4.4. *The first step of the proof does not rely on the specific form of the function ψ_2. Namely, for any absolutely continuous and increasing function ψ, it ensures that the distribution of B_τ — where τ is defined as τ_2 with ψ instead of ψ_2 and ϕ instead of ψ — has a density ν with respect to the Lebebesgue measure given by*

$$\nu(x) \quad = \quad \begin{cases} \dfrac{\psi'(x)}{x}\left(S(x) \; + \; e^{-\gamma_1(\psi(x))}\right), & \textit{if } \psi_1 > \psi, \\ \mu_1(x) \; + \; \dfrac{\psi'(x)}{x}S(x), & \textit{otherwise,} \end{cases}$$

where

$$S(x) \quad := \quad -e^{\gamma(\psi(x))}\int_{\{\phi_1<\phi\}} \mathbb{1}_{\{\phi(l)<x\}} de^{\gamma(l)-\gamma_1(l)}.$$

The relation above sheds new light on Assumption VI.4.2. For instance, if ψ_1 is below ψ near zero, then we see that ν has to coincide with μ_1 near zero. In particular, one can not expect to solve the two-marginal SEP as above, if $\delta\mu \geq 0$ and $\delta\mu \not\equiv 0$ on a neighborhood of zero. Moreover, since $S \geq 0$ by construction, we deduce that ψ is increasing if and only if $\nu - \mu_1 > 0$ whenever $\psi_1 < \psi$.

4.1.3 A numerical example

To illustrate our construction by an example, we consider the symmetric densities μ_1 and μ_2 given by

$$\mu_1(x) \;=\; e^{-2x} \quad \text{and} \quad \mu_2(x) \;=\; \begin{cases} \frac{5}{2}x^3 e^{-\frac{5x^4}{4}}, & \text{if } x \in [0,1], \\ e^{-2x} \;+\; \delta\mu(1)x^{\gamma-2}e^{\frac{-\gamma(x^{\gamma-1}-1)}{\gamma-1}}, & \text{if } x \in (1,+\infty), \end{cases}$$

where γ is a parameter satisfying $\delta\mu(1) = \gamma\delta\mu([1,+\infty))$. One can easily check by direct calculation that the marginals μ_1 and μ_2 are increasing in convex order. The corresponding embedding maps ψ_1 and ψ_2 are given by

$$\psi_1(x) \;=\; x^2 \quad \text{and} \quad \psi_2(x) \;=\; \begin{cases} x^5, & \text{if } x \in [0,1] \\ x^\gamma, & \text{if } x \in (1,+\infty). \end{cases}$$

All the assumptions in Theorem VI.4.3 are satisfied as can be seen in Figure VI.1.

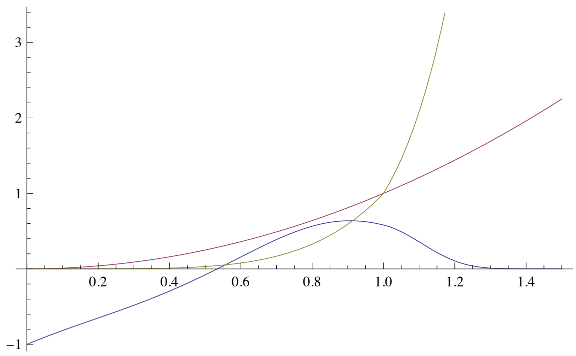

Figure VI.1 – Red : ψ_1 ; Yellow : ψ_2 ; Blue : $\delta\mu$

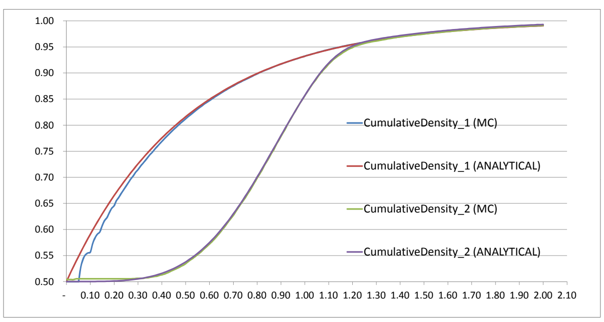

Figure VI.2 – Cumulative distributions : Monte-Carlo / Analytical

In Figure VI.2, we provide a comparison of the analytical cumulative distributions of B_{τ_1} and B_{τ_2} with their Monte-Carlo estimations using 2^{17} realizations. We find a very good match except on a neighborhood of zero. Notice that the simulation of τ_1 and τ_2 are quite difficult as we need to simulate the Brownian local time, which is a highly irregular object. We have chosen to simulate the local time $L^B_{k\Delta t}$ at $k\Delta t$ using

$$L^B_{k\Delta t} \;-\; L^B_{(k-1)\Delta t} \;=\; \frac{\Delta t}{2\varepsilon}\mathbb{1}_{\{B_{(k-1)\Delta t}\in[-\varepsilon,\varepsilon]\}}$$

with $\varepsilon = 0.04$ and $\Delta t = 1/4000$. Since the derivatives of ϕ_1 and ϕ_2 are infinite at zero, the accuracy of our Monte-Carlo estimations near zero depend strongly on the discretization of the local time, which explains the small mismatch in Figure VI.2.

4.2 Robust superhedging problem

In this section, we extend the stochastic control approach to provide candidates for the optimal superhedging strategy in the two-marginal case. Although we are not able to verify the superhedging inequality in general, we provide a couple of examples where we can complete the proof.

4.2.1 Stochastic control approach

Following the inspiration of Henry-Labordère, & Obłój, Spoida & Touzi [50], we extend the stochastic control approach to the two-marginal setting. We consider the following intermediate problem

$$\overline{\mathsf{D}}_0(\boldsymbol{\mu}) \;=\; \inf_{(\lambda_1,\lambda_2)\in\Lambda^2} \Big\{u(0,0) \;+\; \mu_1(\lambda_1) \;+\; \mu_2(\lambda_2)\Big\},$$

where $u : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ is given by

$$u(x,l) \;:=\; \sup_{(\tau_1,\tau_2)\in\mathcal{T}_2} \mathbb{E}_{x,l}\Big[F(L^B_{\tau_2}) \;-\; \lambda_2(B_{\tau_2}) \;-\; \lambda_1(B_{\tau_1})\Big].$$

Here $\mathcal{T}_2$ is the collection of ordered stopping times $\tau_1 \le \tau_2$ such that $B_{\tau_2\wedge\cdot}$ is uniformly integrable and $\mathbb{E}_{x,l}$ denotes the conditional expectation knowing that $B_0 = x$ and $L^B_0 = l$. Then, we introduce the functions u_0, u_1 and u_2 as follows : $u_2(x,l) := F(l)$ and for $i = 0,1$,

$$u_i(x,l) \;:=\; \sup_{\tau\in\mathcal{T}} \mathbb{E}_{x,l}\Big[u_{i+1}(B_\tau, L^B_\tau) \;-\; \lambda_{i+1}(B_\tau)\Big],$$

where $\mathfrak{T}$ is the set of stopping times τ such that $B_{\tau\wedge\cdot}$ is uniformly integrable. The dynamic programming principle induces formally $u_0 = u$. In addition, for $i = 0, 1$, each function u_i should satisfy the corresponding HJB equation, that reads formally

$$\max\Big(u_{i+1} \; - \; \lambda_{i+1} \; - \; u_i, \; \frac{1}{2}\partial_{xx}u_i \; + \; \delta(x)\partial_l u_i\Big) \;\; = \;\; 0. \tag{VI.4.5}$$

For $i = 0, 1$, we look for a solution u_i of the form

$$u_i(x,l) \;\; := \;\; \begin{cases} a_i(l)x^+ \; + \; b_i(l)x \; + \; c_i(l), & \text{if } |x| < \phi_{i+1}(l)), \\ u_{i+1}(x,l) \; - \; \lambda_{i+1}(x), & \text{otherwise,} \end{cases}$$

where $\phi_{i+1} : \mathbb{R}_+ \to \mathbb{R}_+$ is a right-continuous and increasing function such that $\gamma_{i+1}(0+) = 0$ and $\gamma_{i+1}(+\infty) = +\infty$ with

$$\gamma_{i+1}(l) \;\; := \;\; \int_0^l \frac{1}{\phi_{i+1}(z)} dz \text{ for all } l > 0.$$

For the sake of simplicity, we assume that all the functions are sufficiently smooth to allow the calculation sketched below. Using the continuity and smooth fit of u_i at the boundary, we deduce that

$$u_1(x,l) \;\; = \;\; \lambda_2'(\phi_2(l))\,(\phi_2(l) - |x|)^+ \; + \; F(l) \; - \; \lambda_2(|x| \vee \phi_2(l))$$

and

$$\begin{aligned} u_0(x,l) \;\; = \;\; & \lambda_1'(\phi_1(l))\,(\phi_1(l) - |x|)^+ \; + \; \lambda_2'(\bar\phi(l))\,\big(\bar\phi(l) - |x|\big)^+ \\ & + \; F(l) \; - \; \lambda_1(|x| \vee \phi_1(l)) \; - \; \lambda_2(|x| \vee \bar\phi(l)), \end{aligned}$$

where $\bar\phi := \phi_1 \vee \phi_2$. In addition, the functions λ_1 and λ_2 are even and satisfy

$$\lambda_2' \circ \phi_2(l) \;\; = \;\; F'(l) \; + \; \phi_2(l)(\lambda_2' \circ \phi_2)'(l), \tag{VI.4.6}$$

$$\lambda_1' \circ \phi_1(l) \;\; = \;\; \phi_1(l)(\lambda_1' \circ \phi_1)'(l), \qquad \text{if } \phi_1(l) \le \phi_2(l), \tag{VI.4.7}$$

$$(\lambda_1 + \lambda_2)' \circ \phi_1(l) \;\; = \;\; F'(l) \; + \; \phi_1(l)((\lambda_1 + \lambda_2)' \circ \phi_1)'(l), \qquad \text{if } \phi_1(l) > \phi_2(l). \tag{VI.4.8}$$

It follows that

$$\lambda_2'(\phi_2(l)) \;=\; e^{\gamma_2(l)} \int_l^{+\infty} \frac{F'(z)}{\phi_2(z)} e^{-\gamma_2(z)} dz. \tag{VI.4.9}$$

Similarly, on each interval such that $\phi_1(l) < \phi_2(l)$, there exists a constant $C \in \mathbb{R}$ such that

$$\lambda_1'(\phi_1(l)) \;=\; Ce^{\gamma_1(l)} \tag{VI.4.10}$$

and, on each interval such that $\phi_1(l) > \phi_2(l)$, there exists a constant $C \in \mathbb{R}$ such that

$$(\lambda_1 \;+\; \lambda_2)'(\phi_1(l)) \;=\; e^{\gamma_1(l)} \left(\int_l^{+\infty} \frac{F'(z)}{\phi_1(z)} e^{-\gamma_1(z)} dz \;+\; C \right). \tag{VI.4.11}$$

If we assume that $\phi_2 - \phi_1$ changes sign finitely many times, we can determine all the constants involved in the expression of λ_1' above by using the continuity of λ_1' at the intersection points of ϕ_1 and ϕ_2 and by imposing that λ_1' is bounded.

Provided that F is convex, λ_2 is also convex on $\mathbb{R}_+$ and $\mathbb{R}_-$ and one easily checks that u_1 solves the equation (VI.4.5) for $i = 1$. However, it is still an open problem to show that u_0 is actually a solution to the equation (VI.4.5) for $i = 0$. The issue is to prove that $u_0 \geq u_1 - \lambda_1$.

4.2.2 Robust superhedging strategy

Under the assumption that $u_0 \geq u_1 - \lambda_1$, we explain how to derive the pathwise inequality. Given the maturities $0 < t_1 < t_2 = 1$, we can show by the arguments of Proposition VI.6.1 that for all $\mathbb{P} \in \mathcal{M}$

$$\begin{aligned} F(L_{t_2}) \;-\; \lambda_2(X_{t_2}) \;&\leq\; u_1(X_{t_2}, L_{t_2}) \;\leq\; u_1(X_{t_1}, L_{t_1}) \;+\; \int_{t_1}^{t_2} \partial_x u_1 \left(X_t, L_t\right) dX_t,\; \mathbb{P}\text{ - a.s.} \\ u_1(X_{t_1}, L_{t_1}) \;-\; \lambda_1(X_{t_1}) \;&\leq\; u_0(X_{t_1}, L_{t_1}) \;\leq\; \int_0^{t_1} \partial_x u_0 \left(X_t, L_t\right) dX_t,\; \mathbb{P}\text{ - a.s.,} \end{aligned}$$

where λ_1 and λ_2 are given by (VI.4.9), (VI.4.10), (VI.4.11) and $\lambda_1(0) + \lambda_2(0) \;=\; F(0)$. This yields the following pathwise inequality

$$F(L_{t_2}) \;\leq\; \int_0^{t_2} H_t dX_t \;+\; \lambda_1(X_{t_1}) \;+\; \lambda_2(X_{t_2}),\; \mathbb{P}\text{ - a.s. for all } \mathbb{P} \in \mathcal{M}, \tag{VI.4.12}$$

where

$$H_t \;:=\; \begin{cases} -\Big(\lambda_1'\big(\phi_1(L_t)\vee|X_t|\big) \;+\; \lambda_2'\big(\bar{\phi}(L_t)\vee|X_t|\big)\Big)\,\mathrm{sgn}(X_t), & \text{if } 0\le t\le t_1,\\ -\lambda_2'\big(\phi_2(L_t)\vee|X_t|\big)\mathrm{sgn}(X_t), & \text{if } t_1<t\le t_2. \end{cases}$$

Equivalently, $(\lambda_1, \lambda_2, H)$ is a robust superhedging strategy.

Assume further that $\boldsymbol{\mu}$ satisfies Assumptions VI.4.1 and VI.4.2, then we can take the functions (ϕ_1, ϕ_2) given in Theorem VI.4.3 and the corresponding stopping times (τ_1, τ_2). Denote by $\mathbb{P}^* \in \mathcal{M}(\boldsymbol{\mu})$ the law of the process $(Z_t)_{0\le t\le 1}$ given by

$$Z_t \;:=\; \begin{cases} B_{\tau_1\wedge\frac{t}{t_1-t}}, & \text{if } 0\le t\le t_1,\\ B_{\tau_2\wedge(\tau_1\vee\frac{t-t_1}{t_2-t})}, & \text{if } t_1<t\le t_2. \end{cases}$$

Following the arguments of Theorem VI.3.8, we deduce that there is no duality gap, *i.e.*,

$$\overline{\mathsf{P}}(\boldsymbol{\mu}) \;=\; \overline{\mathsf{D}}_0(\boldsymbol{\mu}) \;=\; \overline{\mathsf{D}}(\boldsymbol{\mu}) \;=\; \mu_1(\lambda_1) \;+\; \mu_2(\lambda_2).$$

In addition, $\mathbb{P}^*$ is a maximizer for $\overline{\mathsf{P}}(\boldsymbol{\mu})$ such that

$$F(L_{t_2}) \;=\; \int_0^{t_2} H_t dX_t \;+\; \lambda_1(X_{t_1}) \;+\; \lambda_2(X_{t_2}),\ \mathbb{P}^* \text{ - a.s.},$$

where

$$H_t \;:=\; \begin{cases} -\Big(\lambda_1'\left(\phi_1(L_t)\right) \;+\; \lambda_2'\left(\bar{\phi}(L_t)\right)\Big)\,\mathrm{sgn}(X_t), & \text{if } 0\le t\le t_1,\\ -\lambda_2'\left(\phi_2(L_t)\right)\mathrm{sgn}(X_t), & \text{if } t_1<t\le t_2. \end{cases}$$

As a by-product, we see that our extension of Vallois' solution maximizes the expectation of a convex function of the local time among all solutions to the two-marginal SEP.

Example VI.4.5. *Assume that there exists $l^* > 0$ such that*

$$\phi_1 \;\le\; \phi_2 \text{ on } [0, l^*] \quad \text{and} \quad \phi_1 \;\ge\; \phi_2 \text{ on } (l^*, +\infty). \tag{VI.4.13}$$

If $F''(l) = 0$ for all $l \geq l^*$, the pathwise inequality (VI.4.12) reduces to

$$\lambda_2(X_{t_2}) \;-\; \int_0^{t_1} \lambda_2'\left(\bar{\phi}(L_t) \vee |X_t|\right) sgn(X_t) dX_t \\ -\; \int_{t_1}^{t_2} \lambda_2'\left(\phi_2(L_t) \vee |X_t|\right) sgn(X_t) dX_t \;\;\geq\;\; F(L_{t_2}),\; \mathbb{P}\text{ - } a.s. \text{ for all } \mathbb{P} \in \mathcal{M}.$$

where λ_2 is given by (VI.4.9) with $\lambda_2(0) = F(0)$. This relation is now an easy consequence of Proposition VI.6.1. Thus, if the maps ϕ_1 and ϕ_2 satisfy (VI.4.13), the duality holds and the corresponding optimizers can be derived as above. We conclude that the information about the t_1-marginal does not tighten the bound in this setting.

4.2.3 More applications

We follow the main idea described above in order to derive a remarkable estimation of the moments of the local time. To this end, we take

$$F(l) \;=\; l^p, \;\; \phi_1(l) \;=\; a_1 l \;\; \text{and} \;\; \phi_2(l) \;=\; a_2 l,$$

with $p > 1$, $a_1 > 0$ and $a_2 > 0$. Notice that we are no longer in the setting of the previous sections because $1/\phi_1$ and $1/\phi_2$ are not locally integrable at zero. However, if we assume that $1-(p-1)a_1 > 0$ and $1-(p-1)a_2 > 0$, we can still define λ_1 and λ_2 through the relations (VI.4.9), (VI.4.10) and (VI.4.11). We find

$$\lambda_1(x) \;=\; c_1|x|^p \quad \text{and} \quad \lambda_2(x) \;=\; c_2|x|^p,$$

where

$$c_1 \;=\; \mathbb{1}_{\{a_1 \geq a_2\}} \left(\frac{a_1^{1-p}}{1-(p-1)a_1} \;-\; \frac{a_2^{1-p}}{1-(p-1)a_2} \right) \quad \text{and} \quad c_2 \;=\; \frac{a_2^{1-p}}{1-(p-1)a_2}.$$

Proposition VI.4.6. *With the notations above, if we assume further that either $c_1 \geq 0$, then the following pathwise inequality holds : for all $\mathbb{P} \in \mathcal{M}$,*

$$L_{t_2}^p \;\leq\; c_1|X_{t_1}|^p \;+\; c_2|X_{t_2}|^p \;-\; \int_0^{t_2} p\Big(\mathbb{1}_{\{t \leq t_1\}} c_1 a_1^{p-1} + c_2 a_2^{p-1}\Big) L_t^{p-1} \mathrm{sgn}(X_t) dX_t,\; \mathbb{P}\text{ - } a.s.$$

Further, one has

$$\mathbb{E}^{\mathbb{P}}\Big[L_{t_2}^p\Big] \;\leq\; c_1 \mathbb{E}^{\mathbb{P}}\Big[|X_{t_1}|^p\Big] \;+\; c_2 \mathbb{E}^{\mathbb{P}}\Big[|X_{t_2}|^p\Big]. \tag{VI.4.14}$$

Proof. Let $v_0 : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ and $v_1 : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ be given by

$$\begin{aligned} v_1(x,l) &:= \lambda_2'\Big(\phi_2(l)\Big)\Big(\phi_2(l)-|x|\Big) \;+\; F(l) \;-\; \lambda_2\Big(\phi_2(l)\Big),\\ v_0(x,l) &:= \Big(\lambda_1'\Big(\phi_1(l)\Big)-\partial_x v_1\Big(\phi_1(l),l\Big)\Big)\Big(\phi_1(l)-|x|\Big) \;+\; v_1\Big(\phi_1(l),l\Big) \;-\; \lambda_1\Big(\phi_1(l)\Big). \end{aligned}$$

By the convexity of λ_1 and λ_2, it holds

$$v_0(x,l) \;\ge\; v_1(x,l) \;-\; \lambda_1(x) \quad\text{and}\quad v_1(x,l) \;\ge\; F(l) \;-\; \lambda_2(x).$$

Then the pathwise inequality follows by applying Itô-Tanaka's formula to $v_0(X_t,L_t)$ and $v_1(X_t,L_t)-v_1(X_{t_1},L_{t_1})$ as in Proposition VI.3.3. Further, if we denote by $(\sigma_n)_{n\ge1}$ a sequence of stopping times that reduces the stochastic integral $\int_0^\cdot L_t^{p-1}\mathrm{sgn}(X_t)dX_t$, then one has by Jensen's inequality

$$\mathbb{E}^{\mathbb{P}}\Big[L^p_{t_2\wedge\sigma_n}\Big] \;\le\; c_1\mathbb{E}^{\mathbb{P}}\Big[|X_{t_1\wedge\sigma_n}|^p\Big] \;+\; c_2\mathbb{E}^{\mathbb{P}}\Big[|X_{t_2\wedge\sigma_n}|^p\Big] \;\le\; c_1\mathbb{E}^{\mathbb{P}}\Big[|X_{t_1}|^p\Big] \;+\; c_2\mathbb{E}^{\mathbb{P}}\Big[|X_{t_2}|^p\Big],$$

which yields the inequality (VI.4.14) by the monotone convergence theorem. □

The infimum in the r.h.s. of (VI.4.14) over all pairs (a_1,a_2) is attained for $a_1=a_2=1/p$, which yields the well-known inequality

$$\mathbb{E}^{\mathbb{P}}\Big[L^p_{t_2}\Big] \;\le\; p^p\mathbb{E}^{\mathbb{P}}\Big[|X_{t_2}|^p\Big],$$

see, *e.g.* Vallois [103]. The proposition above has a clear financial interpretation in terms of robust superhedging of power options on the local time. It provides a robust super-replication using Vanilla options with payoffs $\lambda(x)=p^px^p$. Since the minimal bound is achieved by holding only $p^p|X_{t_2}|^p$, it implies that the information about the intermediate marginal does not tighten the bound. However, if we impose further some liquidity constraints on the Vanilla options, the optimal bound is given from Proposition VI.4.6 by

$$\mathbb{E}^{\mathbb{P}}\Big[L^p_{t_2}\Big] \;\le\; \inf_{(a_1,a_2):\;|c_1|\le C_1,|c_2|\le C_2} \Big\{c_1\mathbb{E}^{\mathbb{P}}\Big[|X_{t_1}|^p\Big] \;+\; c_2\mathbb{E}^{\mathbb{P}}\Big[|X_{t_2}|^p\Big]\Big\},$$

where C_1 and C_2 are the constants describing the liquidity constraints. This will lead to a superhedging strategy depending on t_1-Vanillas and t_2-Vanillas.

5 A special multi-marginal case

In this section, we consider a special multi-marginal case when the stopping times given by Vallois are well ordered. In the case of finitely-many marginals, we show that the robust superhedging problem essentially reduces to the one-marginal case. Then we exhibit a remarkable Markov martingale under the assumption that all the marginals of the process are known.

5.1 Robust superhedging under finitely-many marginal constraints

Given $\boldsymbol{\mu} = (\mu_i)_{1\le i\le m}$ a vector of centered probability distributions without mass at zero, we denote by τ_i the stopping time and by (ϕ^i_+, ϕ^i_-) the pair of functions given by Vallois that embeds the probability distribution μ_i.

Assumption VI.5.1. *The map $i \mapsto \tau_i$ is increasing or, equivalently,*

$$\phi^j_- \;\le\; \phi^i_- \;\le\; \phi^i_+ \;\le\; \phi^j_+ \text{ for all } 1 \le i \le j \le m.$$

The next theorem shows that the robust superhedging problem essentially reduces to the one-marginal case in this setting. The corresponding result for the robust subhedging problem is obvious.

Theorem VI.5.2. *Under Assumptions VI.2.2 and VI.5.1, there is no duality gap, i.e.,*

$$\overline{\mathsf{P}}(\boldsymbol{\mu}) \;=\; \overline{\mathsf{D}}(\boldsymbol{\mu}) \;=\; \mu_m(\lambda_m),$$

where λ_m is constructed from (VI.3.1), (VI.3.2) and (VI.3.3) with (ϕ^m_+, ϕ^m_-). In addition, there exists an optimizer $\overline{\mathbb{P}}$ for $\overline{\mathsf{P}}(\boldsymbol{\mu})$ such that

$$\overline{\lambda}_m(X_1) \;+\; \int_0^1 \overline{H}_t dX_t \;=\; F(L_1),\; \overline{\mathbb{P}}\text{ - a.s.},$$

where the process $\overline{H}$ is given by (VI.3.5) with (ϕ^m_+, ϕ^m_-).

Proof. Denote by $\overline{\mathbb{P}}$ the law of the process Z given by

$$Z_t \;:=\; B_{\tau_i \wedge (\tau_{i-1} \vee \frac{t-t_{i-1}}{t_i - t})} \text{ for all } t \in (t_{i-1}, t_i].$$

where $\tau_0 = t_0 = 0$. Under Assumption VI.5.1, it is clear that $\overline{\mathbb{P}}$ belongs to $\mathcal{M}(\boldsymbol{\mu})$. The rest of the proof is identical to the proof of Theorem VI.3.8. Indeed, since $X_1 \in \{\phi^m_-(L_1), \phi^m_+(L_1)\}$, $\overline{\mathbb{P}}$

- almost surely, then it holds

$$\overline{\lambda}_m(X_1) \;+\; \int_0^1 \overline{H}_t dX_t \;\;=\;\; F(L_1),\; \overline{\mathbb{P}}\text{ - a.s.}$$

Then we can complete the proof by Lemma VI.3.9. $\square$

5.2 A Markov martingale given full marginals

In this section, we consider that all the marginals $(\mu_t)_{0\le t\le 1}$ of the process are known. For the sake of simplicity, we assume that μ_t is symmetric and equivalent to the Lebesgue measure for all $t \in [0,1]$. Denote by τ_t the stopping time and by ϕ_t the function given by Vallois relative to μ_t.

Assumption VI.5.3. (i) *For all $0 \le s \le t \le 1$, $\tau_s \le \tau_t$, or equivalently, $\phi_s \le \phi_t$;*
(ii) *The map $(t,x) \mapsto \psi_t(x)$ has the first-order partial derivatives, where ψ_t denotes the right-continuous inverse of ϕ_t.*

The next result gives the generator of the Markov process $(B_{\tau_t})_{0\le t\le 1}$. It is analogous to the study of Madan & Yor [76] with the Azéma-Yor's solution to the SEP. In particular, the process $(B_{\tau_t})_{0\le t\le 1}$ is a pure jump process, which corresponds to an example of local Lévy's model introduced in Carr, Geman, Madan & Yor [19].

Theorem VI.5.4. *Under Assumption VI.5.3, $(B_{\tau_t})_{0\le t\le 1}$ is an inhomogeneous Markov martingale whose generator is given by*

$$\mathcal{L}_t f(x) \;\;=\;\; \frac{\partial_t \psi_t(|x|)}{\partial_x \psi_t(|x|)} \left(\frac{e^{\gamma_t\circ\psi_t(|x|)}}{2|x|} \int_{|x|}^{+\infty} \Big(f(y)+f(-y)-2f(x)\Big) de^{-\gamma_t\circ\psi_t(y)} \;-\; f'(x)\right),$$

for all smooth bounded functions $f:\mathbb{R}\to\mathbb{R}$.

Proof. The process $(B_{\tau_t})_{0\le t\le 1}$ is clearly an inhomogeneous Markov martingale, see [76] for more details. It remains to compute the generator. For each $0 \le t < u \le 1$, we denote

$$\mathbb{E}\Big[f(B_{\tau_u})|B_{\tau_t}=x\Big] \;\;=\;\; \mathbb{E}[f(B_{\tau_u})|B_{\tau_t}=x, L^B_{\tau_t}=\psi_t(|x|)] \;\;=:\;\; v\Big(x,\psi_t(|x|)\Big).$$

Then it follows from Lemma VI.6.3 that

$$v\Big(x,\psi_t(|x|)\Big) \;\;=\;\; a_u\circ\psi_t(|x|)x^+ \;+\; b_u\circ\psi_t(|x|)x \;+\; c_u\circ\psi_t(|x|)$$

with

$$\begin{aligned} a_u(l) &:= \frac{f\Big(\phi_u(l)\Big) + f\Big(-\phi_u(l)\Big) - 2c_u(l)}{\phi_u(l)}, \\ b_u(l) &:= \frac{c_u(l) - f\Big(-\phi_u(l)\Big)}{\phi_u(l)}, \\ c_u(l) &:= e^{\gamma_u(l)} \int_l^{+\infty} e^{-\gamma_u(z)} \frac{f\Big(\phi_u(z)\Big) + f\Big(-\phi_u(z)\Big)}{2\phi_u(z)} dz. \end{aligned}$$

By definition, the generator is given by

$$\mathcal{L}_t f(x) = \partial_t a_t \circ \psi_t(|x|)x^+ + \partial_t b_t \circ \psi_t(|x|)x + \partial_t c_t \circ \psi_t(|x|) \text{ for all } x \in \mathbb{R}.$$

A straightforward calculation yields that for all $x \geq 0$,

$$\begin{aligned} \partial_t a_t \circ \psi_t(x)x &= -2\partial_t c_t \circ \psi_t(x) + \left\{\frac{f(x) + f(-x) - 2c_t \circ \psi_t(x)}{x} - f'(x) + f'(-x)\right\}\frac{\partial_t \psi_t(x)}{\partial_x \psi_t(x)}, \\ \partial_t b_t \circ \psi_t(x)x &= \partial_t c_t \circ \psi_t(x) + \left\{\frac{c_t \circ \psi_t(x) - f(-x)}{x} - f'(-x)\right\}\frac{\partial_t \psi_t(x)}{\partial_x \psi_t(x)}. \end{aligned}$$

and for all $x < 0$,

$$\partial_t b_t \circ \psi_t(-x)x = -\partial_t c_t \circ \psi_t(-x) + \left\{\frac{c_t \circ \psi_t(-x) - f(x)}{x} - f'(x)\right\}\frac{\partial_t \psi_t(-x)}{\partial_x \psi_t(-x)}.$$

Hence, we obtain

$$\mathcal{L}_t f(x) = \frac{\partial_t \psi_t(|x|)}{\partial_x \psi_t(|x|)}\left(\frac{f(x) - c_t \circ \psi_t(|x|)}{|x|} - f'(x)\right).$$

The desired result follows by the fact

$$\begin{aligned} c_t \circ \psi_t(|x|) &= e^{\gamma_t \circ \psi_t(|x|)} \int_{|x|}^{+\infty} e^{-\gamma_t \circ \psi_t(y)} \frac{f(y) + f(-y)}{2y} \partial_x \psi_t(y) dy, \\ &= -\frac{e^{\gamma_t \circ \psi_t(|x|)}}{2} \int_{|x|}^{+\infty} \Big(f(y) + f(-y)\Big) de^{-\gamma_t \circ \psi_t(y)}. \end{aligned}$$

□

As an application, we provide a new example of fake Brownian motion. If $(\mu_t)_{0 \leq t \leq 1}$ is a continuous Gaussian peacock, *i.e.*

$$\mu_t(x) \quad = \quad \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \text{ for all } t \in [0,1] \text{ and } x \in \mathbb{R},$$

then it satisfies Assumption VI.5.3 in view of Lemma VI.5.5 below. Then the process $(B_{\tau_t})_{0\leq t\leq 1}$ is a fake Brownian motion, *i.e.* a Markov martingale with the same marginal distributions as a Brownian motion that is not a Brownian motion.

Lemma VI.5.5. *If $(\mu_t)_{0\leq t\leq 1}$ is a continuous Gaussian peacock, then the map $t \mapsto \psi_t(x)$ is decreasing for all $x > 0$.*

Proof. For the sake of clarity, we denote $R_t(x) := \mu_t([x, +\infty))$ in this proof. By integration by parts, it holds for all $x \in \mathbb{R}_+$,

$$\psi_t(x) \quad = \quad \int_0^x \frac{y\mu_t(y)}{R_t(y)} dy \quad = \quad \int_0^x \log\left(\frac{R_t\,(y)}{R_t\,(x)}\right) dy.$$

Further, by change of variable, we have $R_t(x) = R_1(\frac{x}{\sqrt{t}})$. To conclude, it is clearly enough to prove that the map $t \mapsto \frac{R_1(ty)}{R_1(tx)}$ is increasing for all $x > 0$ and $0 < y < x$. By direct differentiation, we see that its derivative has the same sign as

$$xe^{-\frac{t^2x^2}{2}} \int_{ty}^{+\infty} e^{-\frac{z^2}{2}} dz \;-\; ye^{-\frac{t^2y^2}{2}} \int_{tx}^{+\infty} e^{-\frac{z^2}{2}} dz \;=\; \int_{txy}^{+\infty} \left(e^{-\frac{1}{2}\left(\frac{z^2}{x^2} + t^2x^2\right)} \;-\; e^{-\frac{1}{2}\left(\frac{z^2}{y^2} + t^2y^2\right)} \right) dz.$$

It remains to observe that the quantity above is positive since $\frac{z^2}{x^2} + t^2x^2 < \frac{z^2}{y^2} + t^2y^2$ for all $z > txy$, $0 < y < x$ and $x > 0$. □

6 Appendix

In the one-marginal case, the stochastic control approach induces naturally a pathwise inequality slighty different from the pathwise inequality of Proposition VI.3.3. The next proposition ensures that it is also valid.

Proposition VI.6.1. *Under Assumptions VI.2.2 and VI.3.2, assume further that $\phi_\pm$ are strictly monotone, then the following inequality holds*

$$\lambda(X_1) \;+\; \int_0^1 H_t dX_t \;\;\geq\;\; F(L_1),\;\; \mathbb{P}\text{ - } a.s. \text{ for all } \mathbb{P} \in \mathcal{M},$$

where the function λ is defined by (VI.3.1), (VI.3.2), (VI.3.3) and the process H is given by

$$H_t \;:=\; -\lambda'(X_t \vee \phi_+(L_t))\mathbb{1}_{\{X_t>0\}} - \lambda'(X_t \wedge \phi_-(L_t))\mathbb{1}_{\{X_t\le 0\}}. \tag{VI.6.1}$$

Proof. Consider the function $v : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ given by

$$v(x,l) \;:=\; \begin{cases} u(x,l), & \text{if } x \in (\phi_-(l), \phi_+(l)), \\ F(l) \;-\; \lambda(x), & \text{otherwise.} \end{cases}$$

where $u(x,l)$ was defined in the proof of Proposition VI.3.3. Since we have $v(x,l) \ge F(l) - \lambda(x)$, it suffices to show that

$$v\left(X_1, L_1\right) \;\le\; \int_0^1 H_t dX_t.$$

First, we notice that $v(x,\cdot)$ is absolutely continuous and its (right-continuous) derivative is given by

$$\partial_l v(x,l) \;=\; \begin{cases} e^{\gamma(l)} \int_l^{+\infty} e^{-\gamma(z)} F''(dz) \left(-\frac{x^+}{\phi_+(l)} + \frac{x^-}{\phi_-(l)} + 1\right) \;+\; F'(l), & \text{if } x \in (\phi_-(l), \phi_+(l)), \\ F'(l), & \text{otherwise,} \end{cases}$$

and $v(\cdot,l)$ has a left derivative given by

$$\partial_x v(x,l) \;=\; \begin{cases} -\lambda'(\phi_+(l))\mathbb{1}_{\{x>0\}} \;+\; \lambda'(\phi_-(l))\mathbb{1}_{\{x\le 0\}}, & \text{if } x \in (\phi_-(l), \phi_+(l)), \\ \lambda'(x), & \text{otherwise.} \end{cases}$$

We are going to apply Itô's formula to a smooth approximation of v in order to recover the desired result at the limit. This idea is very similar to the proof of Itô-Tanaka's formula in Karatzas & Shreve [71]. Let us define the smooth function ρ as follows :

$$\rho(x) \;:=\; C \exp\left(\frac{1}{(x-1)^2 - 1}\right)\mathbb{1}_{\{0<x<2\}},$$

where C is the appropriate constant such that $\int_{\mathbb{R}} \rho(x)dx = 1$. Using the mollifiers $(\rho_n)_{n\ge 1}$ given by $\rho_n(x) := n\rho(nx)$, we construct the smooth approximations of v by convolution :

$$v_{n,k}(x,l) \;:=\; \int_{\mathbb{R}\times\mathbb{R}_+} \rho_n(x-y)\rho_k(z-l)v(y,z)dydz \;=\; \int_{[0,2]^2} \rho(y)\rho(z)v(x-\frac{y}{n}, l+\frac{z}{k})dydz.$$

Notice that $v_{n,k}$, $\partial_x v_{n,k}$ and $\partial_l v_{n,k}$ converge pointwise to v, $\partial_x v$ and $\partial_l v$ respectively. Now Itô's formula yields

$$\begin{aligned} v_{n,k}(X_1, L_1) \; - \; v_{n,k}(0,0) \;\; &= \;\; \int_0^1 \partial_x v_{n,k}(X_t, L_t)\, dX_t \\ &\quad + \int_0^1 \partial_l v_{n,k}(0, L_t)\, dL_t \; + \; \frac{1}{2}\int_0^1 \partial_{xx} v_{n,k}(X_t, L_t)\, d\langle X\rangle_t. \end{aligned}$$

Sending n and k to infinity, we deduce that

$$\lim_{n,k\to\infty} \frac{1}{2}\int_0^1 \partial_{xx} v_{n,k}(X_t, L_t)\, d\langle X\rangle_t \;\; = \;\; v(X_1, L_1) \; - \; \int_0^1 H_t dX_t \; - \; \int_0^1 \partial_l v(0, L_t)\, dL_t,$$

in the sense of convergence in probability. In addition, we have by the convexity of λ

$$\int_{\mathbb{R}} \partial_{xx}\rho_n(x-y)v(y,l)dy \;\; = \;\; \int_{\mathbb{R}} \rho_n(x-y)\partial_{xx}v(dy,l) \;\; \leq \;\; -\Big(\lambda'(\phi_+(l)) - \lambda'(\phi_-(l))\Big)\rho_n(x).$$

where $\partial_{xx}v(\cdot,l)$ denotes the second derivative of $v(\cdot,l)$ in the sense of distributions. Using Lemma VI.6.2 below, we derive

$$\lim_{n,k\to\infty} \int_0^1 \partial_{xx} v_{n,k}(X_t, L_t)\, d\langle X\rangle_t \;\; \leq \;\; \int_0^1 \Big(\lambda'(\phi_+(L_t)) - \lambda'(\phi_-(L_t))\Big)dL_t.$$

Hence, we obtain

$$v(X_1, L_1) \;\; \leq \;\; \int_0^1 H_t dX_t \; + \; \int_0^1 \Big(\partial_l v(0, L_t) \; - \; \frac{1}{2}\Big(\lambda'(\phi_+(L_t)) - \lambda'(\phi_-(L_t))\Big)\Big)\, dL_t.$$

The conclusion follows immediately by using the identity (VI.3.7). □

Lemma VI.6.2. *With the notations of the proof of Proposition VI.6.1, if $f : \mathbb{R}_+ \to \mathbb{R}$ is absolutely continuous, then it holds for any $\mathbb{P} \in \mathcal{M}$,*

$$\lim_{n\to+\infty} \int_0^1 f(L_t)\, \rho_n(X_t) d\langle X\rangle_t \;\; = \;\; \int_0^1 f(L_t)\, dL_t, \;\; \mathbb{P}\text{ - } a.s.$$

Proof. By integration by parts, we have

$$\int_0^1 f(L_t)\, \rho_n(X_t) d\langle X\rangle_t \;\; = \;\; f(L_1)\int_0^1 \rho_n(X_t) d\langle X\rangle_t \; - \; \int_0^1 f'(L_t)\int_0^t \rho_n(X_s) d\langle X\rangle_s dL_t.$$

Further, the occupation time formula yields

$$\int_0^t \rho_n(X_s)d\langle X\rangle_s \quad = \quad \int_{\mathbb{R}} \rho_n(x)L_t(x)dx \text{ for all } t\in\mathbb{R}_+,$$

where $(L_t(x))_{t\geq 0}$ denotes the local time of X at $x\in\mathbb{R}$. Since X is a martingale, there exists a version of the local time making the map $(t,x)\mapsto L_t(x)$ continuous, see *e.g.* Theorem VI.1.7 in Revuz & Yor [92]. Hence, we obtain

$$\lim_{n\to+\infty}\int_0^t \rho_n(X_s)d\langle X\rangle_s \quad = \quad L_t \text{ for all } t\in\mathbb{R}_+.$$

By the dominated convergence theorem, we further deduce that

$$\lim_{n\to\infty}\int_0^1 f\left(L_t\right)\rho_n(X_t)d\langle X\rangle_t \quad = \quad f(L_1)L_1 \;-\; \int_0^1 f'\left(L_t\right)L_t dL_t.$$

The conclusion follows by integration by parts. □

The lemma below gives a characterization of the distribution of (B_τ, L^B_τ) where

$$\tau \quad := \quad \inf\left\{t>0:\; B_t\notin\left(\phi_-(L^B_t),\phi_+(L^B_t)\right)\right\}.$$

It is a key step in the proofs of Theorems VI.4.3 and VI.5.4.

Lemma VI.6.3. *Under Assumption VI.3.2, for all $f:\mathbb{R}\times\mathbb{R}_+\to\mathbb{R}$ such that $l\mapsto f(\phi_\pm(l),l)$ is a function of bounded variation, it holds*

$$\mathbb{E}_{x,l}\left[f(B_\tau,L^B_\tau)\right] \quad = \quad \begin{cases} \frac{f(\phi_+(l),l)-c(l)}{\phi_+(l)}x^+ \;-\; \frac{f(\phi_-(l),l)-c(l)}{\phi_-(l)}x^- \;+\; c(l), & \text{if } x\in(\phi_-(l),\phi_+(l)) \\ f(x,l), & \text{otherwise,} \end{cases}$$

where $\mathbb{E}_{x,l}$ denotes the conditional expectation operator $\mathbb{E}\left[\cdot|B_0=x,L^B_0=l\right]$ and

$$c(l) \quad = \quad \frac{1}{2}\int_l^{+\infty}\left(\frac{f(\phi_+(z),z)}{\phi_+(z)} \;-\; \frac{f(\phi_-(z),z)}{\phi_-(z)}\right)e^{\gamma(l)-\gamma(z)}dz \text{ for all } l\geq 0.$$

Proof. Let $(M_t)_{t\geq 0}$ be the process given by

$$M_t \quad = \quad \frac{f\left(\phi_+(L^B_t),L^B_t\right)-c(L^B_t)}{\phi_+(L^B_t)}B^+_t \;-\; \frac{f\left(\phi_-(L^B_t),L^B_t\right)-c(L^B_t)}{\phi_-(L^B_t)}B^-_t \;+\; c(L^B_t).$$

By applying Itô-Tanaka's formula, we deduce that

$$M_t \;=\; M_0 + \int_0^t \left(\frac{f\left(\phi_+(L_s^B), L_s^B\right) - c(L_s^B)}{\phi_+(L_s^B)} \mathbb{1}_{\{B_s>0\}} \;-\; \frac{f\left(\phi_-(L_s^B), L_s^B\right) - c(L_s^B)}{\phi_-(L_s^B)} \mathbb{1}_{\{B_s\leq 0\}} \right) dB_s.$$

Hence, the process M is a local martingale. Further, the stopped process $M_{\tau\wedge\cdot}$ is bounded by $5\|f\|$ since $\|c\| \leq \|f\|$ and $|B^{\pm}_{\tau\wedge t}| \leq \phi_{\pm}(L^B_{\tau\wedge t})$. It follows that

$$\mathbb{E}_{x,l}\left[M_\tau\right] \;=\; M_0 \;=\; \frac{f\left(\phi_+(l), l\right) - c(l)}{\phi_+(l)} x^+ \;-\; \frac{f\left(\phi_-(l), l\right) - c(l)}{\phi_-(l)} x^- \;+\; c(l).$$

To conclude, it remains to see that $M_\tau = f(B_\tau, L_\tau^B)$ by definition. $\square$

Acknowledgement

First of all, I would express my sincere gratitude to Professor Nizar Touzi for his patient guide during the past three years. I thank him for investing a great amount of time in our discussions, and sharing a great deal of precious experience and ideas. Last but not least, I always appreciate his example as diligence and modesty.

Also, I would thank Professor Xiaolu Tan and Pierre Henry-Labordère, for their inspiring help in my research and their great effort on our collaboration.

The past three years were great, thanks to all the colleagues in CMAP, especially all members of the financial mathematics team. Thanks to Stefano De Marco, Plamen Turkedjievm and Stefano Pagliarani, interesting group discussions are organized every week. I would thank Zhenjie Ren, Julien Claisse, Sigrid Källblad, Hadrien De Marc, Gang Liu, Ankush Agarwal, Guillaume Royer, Christian Litterer, Yiqing Lin, Nicolas El Karoui, Mathieu Rosenbaum, Alexander Cox, Marcel Nutz, Jiatu Cai, Yi Lu, Yiyi Zou, Dylan Possamaï, and all others with whom I had curious discussions.

I also thank the secretaries and the IT support of CMAP, in particular Násséra Naar and Sylvain Ferrand. Without their kind help, I could not easily handle all the administrative issues.

Finally, I would like to thank personally my parents, my brother as well as my girlfriend for their permanent support and their presence during my thesis.

Bibliography

[1] B. Acciaio, M. Beiglböck, F. Penkner, W. Schachermayer and J. Temme : *A Trajectorial Interpretation of Doob's Martingale Inequalities*, Ann. Appl. Probab. 23 (2013), no. 4, 1494-1505.

[2] M. Avellaneda, A. Levy and A. Parás : *Pricing and hedging derivative securities in markets with uncertain volatilities*, Applied Mathematical Finance, 2(2), 73-88, 1995.

[3] J. Azéma and M. Yor : *Une solution simple au problème de Skorokhod*, Séminaire de Probabilités, XIII, volume 721 of Lecture Notes in Math., 90-115. Springer, Berlin, 1979.

[4] J. Azéma and M. Yor : *Le problème de Skorokhod : compléments à "Une solution simple au problème de Skorokhod"*, Séminaire de Probabilités, XIII, volume 721 of Lecture Notes in Math., p. 625-633.

[5] M. Beiglböck, A. Cox and M. Huesmann : *Optimal Transport and Skorokhod Embedding*, arXiv preprint arXiv :1307.3656, 2013.

[6] M. Beiglböck, A. Cox, M. Huesmann, N. Perkowski and J. Prömel : *Robust super-replication for time-invariant derivatives*, preprint, 2015.

[7] M. Beiglböck, P. Henry-Labordère and F. Penkner : *Model-independent bounds for option prices : a mass transport approach*, Finance and Stochastics, 17(3), 477-501, 2013.

[8] M. Beiglböck, P. Henry-Labordère and N. Touzi : *Monotone Martingale Transport Plans and Skorokhod Embedding*, preprint, 2015.

[9] M. Beiglböck and N. Juillet : *On a problem of optimal transport under marginal martingale constraints*, arXiv preprint arXiv :1208.1509, 2012.

[10] M. Beiglböck, M. Nutz and N. Touzi : *Robust super-replication for time-invariant derivatives*, preprint, 2015.

[11] M. Beiglböck, M. Nutz and N. Touzi : *Complete duality for martingale optimal transport on the line*, arXiv preprint arXiv :1507.00671, 2015.

[12] S. Biagini, B. Bouchard, C. Kardaras and M. Nutz : *Robust fundamental theorem for continuous processes*, Mathematical Finance, 2015.

[13] P. Billingsley : *Convergence of Probability Measures*, John Wiley & Sons, 2013.

[14] J. Bonnans and X. Tan : *A model-free no-arbitrage price bound for variance options*, Applied Mathematics & Optimization, Vol. 68, Issue 1, 43-73, 2013.

[15] B. Bouchard and M. Nutz : *Arbitrage and duality in nondominated discrete-time models*, Annals of Applied Probability, vol. 25, no 2, p. 823-859, 2015.

[16] D. Breeden and R. Litzenberger : *Prices of state-contingent claims implicit in options prices*, Journal of business, p. 621-651, 1978.

[17] H. Brown, D. Hobson and L. Roggers : *Robust hedging of barrier options*, Mathematical Finance, 11(3), 285-314, 2001.

[18] H. Brown, D. Hobson and L. Roggers : *The maximum maximum of a martingale constrained by an intermediate law*, Probability theory and related fields, 119 :558-578, 2001.

[19] P. Carr, H. Geman, D. Madan, M. Yor : *From local volatility to local Lévy models*, Quantitative Finance, Volume 4, Issue 5 (2004).

[20] J. Claisse, G. Guo and P. Henry-Labordère : *Robust hedging of options on local time*, preprint, 2015.

[21] A. Cox and D. Hobson : *Local martingales, bubbles and option prices*, *Finance & Stochastics*, 9 : 477-492, 2005.

[22] A. Cox and D. Hobson : *Skorokhod embeddings, minimality and non-centred target distributions*, Probability Theory and Related Fields, 135(3) :395-414, 2006.

[23] A. Cox, D. Hobson and J. Obłój : *Pathwise inequalities for local time : applications to Skorokhod embeddings and optimal stopping*, Ann. Appl. Probab., 18(5) : 1870–1896, 2008.

[24] A. Cox and J. Obłój : *Robust pricing and hedging of double no-touch options*, Finance and Stochastics, 15(3) : 573–605, 2011.

[25] A. Cox and J. Obłój : *Robust hedging of double touch barrier options*, SIAM Journal on Financial Mathematics, 2 : 141–182, 2011.

[26] A. Cox, J. Obłój and N. Touzi : *The Root solution to the multi-marginal embedding problem : an optimal stopping and time-reversal approach*, arXiv preprint arXiv :1505.03169, 2015.

[27] J. Claisse, X. Tan and D. Talay : *A pseudo-Markov property for controlled diffusion processes*, SIAM Journal on Control and Optimization, to appear.

[28] A. Cox and J. Wang : *Root's barrier : Construction, optimality and applications to variance options*, Annals of Applied Probability (2013) 23 (3), 859-894.

[29] M. Davis and D. Hobson : *The Range of Traded Option Prices*, Mathematical Finance 17(1) 1-14, 2007.

[30] M. Davis, J. Obłój and V. Raval : *Arbitrage Bounds for Weighted Variance Swap Prices*, Mathematical Finance 24(4) : 821-854.

[31] F. Delbaen and W. Schachermayer : *A general version of the fundamental theorem of asset pricing*, Mathematische Annalen, Vol. 300 (1994), pp. 463-520.

[32] C. Dellacherie and E. Lenglart : *Sur des problèmes de régularisation, de recollement et d'interpolation en théorie des processus*, Séminaire de Probabilités de Strasbourg XVI, pages 298-313. Springer, 1982.

[33] C. Dellacherie and P. Meyer : *Probabilities and potential A*, Vol. 29 of *North-Holland Mathematics Studies*, North-Holland Publishing Co., Amsterdam, 1978.

[34] L. Denis and C. Martini : *A theoretical framework for the pricing of contingent claims in the presence of model uncertainty*, Ann. Appl. Probab. Volume 16, Number 2 (2006), 827-852.

[35] J. Deuschel and D. Stroock : *Large Deviations*, American Mathematical Society, 978-0821827574, 2001.

[36] Y. Dolinsky and H. Soner : *Robust hedging and martingale optimal transport in continuous time*, Probability Theory and Related Fields, 2014, vol. 160, no 1-2, p. 391-427.

[37] Y. Dolinsky and H. Soner : *Martingale optimal transport in the Skorokhod space*, Stochastic Processes and their Applications, 2015, vol. 125, no 10, p. 3893-3931.

[38] N. El Karoui : *Les aspects probabilistes du contrôle stochastique*, Ecole d'Eté de Probabilités de Saint Flour IX, Lecture Notes in Mathematics 876, Springer Verlag, 1979.

[39] N. El Karoui, J. Lepeltier and A. Millet : *A probabilistic approach to the reduite in optimal stopping*, Probab. Math. Statist. 13(1) :97-121, 1992.

[40] N. El Karoui and X. Tan : *Capacities, measurable selection and dynamic programming Part I : abstract framework*, arXiv preprint arXiv :1310.3363, 2013.

[41] N. El Karoui and X. Tan : *Capacities, measurable selection and dynamic programming Part II : application in stochastic control problems*, arXiv preprint arXiv :1310.3364, 2013.

[42] H. Föllmer and D. Kramkov : *Optional decompositions under constraints*, Prob. Theory Relat. Fields, vol. 109, 1-25 (1997).

[43] A. Galichon, P. Henry-Labordère and N. Touzi : *A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options*, Annals of Applied Probability, Volume 24, Number 1 (2014), 312-336.

[44] P. Gassiat, H. Oberhauser and G. dos Reis : *Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs*, preprint, 2013.

[45] G. Guo : *Stability of optimal Skorokhod embedding*, preprint, 2016.

[46] G. Guo, A. Jacquier, C. Martini and L. Neufcourt : *Generalised arbitrage-free SVI volatility surfaces*, SIAM Journal on Financial Mathematics, 2016.

[47] G. Guo, X. Tan and N. Touzi : *Tightness and duality of martingale transport on the Skorokhod space*, Stochastic Processes and their Applications, 2016.

[48] G. Guo, X. Tan and N. Touzi : *Optimal Skorokhod embedding under finitely-many marginal constraints*, SIAM Journal on Control and Optimization, 2016.

[49] G. Guo, X. Tan and N. Touzi : *On the monotonicity principle of optimal Skorokhod embedding problem*, SIAM Journal on Control and Optimization, 2016.

[50] P. Henry-Labordère, J. Obłój, P. Spoida and N. Touzi : *The maximum maximum of a martingale with given n marginals*, Annals of Applied Probability, 2016, vol. 26, no 1, p. 1-44.

[51] P. Henry-Labordère, X. Tan and N. Touzi : *An Explicit Martingale Version of the One-dimensional Brenier's Theorem with Full Marginals Constraint*, SSRN 2335969, 2014.

[52] P. Henry-Labordère and N. Touzi : *An Explicit Martingale Version of the one-dimensional Brenier Theorem*, SSRN 2218488, 2013.

[53] F. Hirsch, C. Profeta, B. Roynette and M. Yor : *Peacocks and Associated Martingales, with Explicit Constructions*, Springer Science & Business Media, 2011.

[54] F. Hirsch and B. Roynette : *A new proof of Kellerer's theorem*, ESAIM : Probability and Statistics, Volume 16, January 2012, pp 48-60.

[55] F. Hirsch and B. Roynette : *On $\mathbb{R}^d$-valued peacocks*, ESAIM : Probability and Statistics, Volume 17, January 2013, pp 444-454.

[56] D. Hobson : *Robust hedging of the lookback option*, Finance and Stochastics, 2, pp329-347, 1998.

[57] D. Hobson : *The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices*, Paris-Princeton Lectures on Mathematical Finance, 2010. Springer, LNM 2003.

[58] D. Hobson and M. Klimmek : *Model independent hedging strategies for variance swaps*, Finance and Stochastics 16, 611-649, 2012 (DOI) 10.1007/s00780-012-0190-3.

[59] D. Hobson and M. Klimmek : *Maximising functionals of the joint law of the maximum and terminal value in the Skorokhod embedding problem*, Annals of Applied Probability Vol. 23, No. 5, p2020-2052, 2013.

[60] D. Hobson and M. Klimmek : *Robust price bounds for the forward starting straddle*, Finance and Stochastics Vol 19, Issue 1, 189-214, 2015.

[61] D. Hobson and A. Neuberger : *Robust bounds for forward start options*, Mathematical Finance 22, 31-56, 2012.

[62] D. Hobson and J. Pedersen : *The minimum maximum of a continuous martingale with given initial and terminal laws*, Annals of probability, 978-999, 2002.

[63] Z. Hou and J. Obłój : *On robust pricing-hedging duality in continuous time*, arXiv preprint arXiv :1503.02822, 2015.

[64] P. Jäckel and C. Kahl : *Hyp hyp hooray*, *Wilmott Magazine* : 70-81, 2008.

[65] J. Jacod and J. Mémin : *Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité*, Séminaire de Probabilité XV 1979/80, Lecture Notes in Mathematics, Vol. 850, 529-546, 1981.

[66] J. Jacod and A. Shiryaev : *Limit theorems for stochastic processes*, volume 228 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer-Verlag, Berlin, 1987.

[67] A. Jakubowski : *A non-Skorohod topology on the Skorohod space*, Electronic journal of probability, 2(4) :1-21, 1997.

[68] A. Jakubowski : *The a.s. Skorohod representation for subsequences in non metric spaces*, Theory Probab. Appl. 42 :2141-2153, 1997.

[69] S. Källblad, X. Tan and N. Touzi : *Optimal Skorokhod embedding given full marginals and Azéma-Yor peacocks*, Annals of Applied Probability, to appear, 2016.

[70] R. Karandikar : *On pathwise stochastic integration*, Stoch. Proc. and their App. 57(1) :11-18, 1995.

[71] I. Karatzas and S. Shreve : *Methods of Mathematical Finance*, Springer Verlag, 1998.

[72] H. Kellerer : *Markov-Komposition und eine Anwendung auf Martingale*, Mathematische Annalen (1972).

[73] M. Kobylanski, M. Quenez and E. Rouy-Mironescu : *Optimal multiple stopping time problem*, Ann. Appl. Prob. 21(4) :1365-1399, 2011.

[74] D. Kramkov : *Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets*, Probability Theory and Related Fields, Nov., 1995.

[75] T. Lyons : *Uncertain volatility and the risk-free synthesis of derivatives*, Applied Mathematical Finance Volume 2, Issue 2, 1995.

[76] D. Madan and M. Yor : *Making Markov martingales meet marginals : with explicit constructions* Bernoulli 8(4) :509-536, 2002.

[77] J. Mertens : *Théorie des processus stochastiques généraux applications aux surmartingales* Probability Theory and Related Fields, 22(1) :45-68, 1972.

[78] P. Meyer and W. Zheng : *Tightness criteria for laws of semimartingales*, Annales de l'IHP, section B, 20(4) :353-372, 1984.

[79] I. Monroe : *On embedding right continuous martingales in Brownian motion*, The Annals of Mathematical Statistics, 43(4) :1293-1311, 1972.

[80] M. Nutz : *Robust Superhedging with Jumps and Diffusion*, Stochastic Processes and their Applications, 2015, vol. 125, no 12, p. 4543-4555.

[81] A. Neufeld and M. Nutz : *Measurability of Semimartingale Characteristics with Respect to the Probability Law*, Stochastic Processes and their Applications, Vol. 124, No. 11, pp. 3819-3845, 2014.

[82] A. Neufeld and M. Nutz : *Nonlinear Lévy Processes and their Characteristics*, arXiv preprint arXiv :1401.7253, 2014.

[83] J. Obłój : *The Skorokhod embedding problem and its offspring*, Probability Surveys, 1 : 321–392, 2004.

[84] J. Obłój and P. Spoida : *An Iterated Azéma-Yor Type Embedding for Finitely Many Marginals*, arXiv preprint arXiv :1304.0368, 2013.

[85] D. Ocone : *A symmetry characterization of conditionally independent increment martingales*, Progr. Probab., Vol. 32, 147–167, 1993.

[86] B. Pass : *Multi-marginal optimal transport and multi-agent matching problems : uniqueness and structure of solutions*, Discrete Contin. Dyn. Syst. 34 :1623-1639, 2014.

[87] G. Peskir and A. Shiryaev : *Optimal Stopping and Free-Boundary Problems*, Birkhauser, Basel, 2006.

[88] D. Possamaï, G. Royer and N. Touzi : *On the robust superhedging of measurable claims*, Electronic Communications in Probability, Vol. 18, 95 (2013).

[89] P. Protter : *Stochastic Integration and Differential Equations*, Springer-Verlag, Berlin, 1990.

[90] S. Rachev and L. Rüschendorf : *Mass Transportation Problems Volume 1 : Theory*, Springer Science & Business Media, 1998.

[91] E. Renault and N.Touzi : *Option hedging and implied volatilities in a stochastic volatility model*, *Mathematical Finance*, 6(3) : 279-302, 1996.

[92] D. Revue and M. Yor : *Continuous martingales and Brownian motion*, Grundlehren der Mathematischen Wissenschaften, Vol. 293, Springer, 1999.

[93] D. Root : *The existence of certain stopping times on Brownian motion*, Ann. Math. Statist., 40 :715-718, 1969.

[94] H. Rost : *Skorokhod stopping times of minimal variance*, Séminaire de Probabilités, X, 194-208. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

[95] A. Shiryaev : *Probability*, Springer, New York (1984).

[96] A. Skorokhod : *On a representation of random variables*, Theory Probab. Appl, 1976.

[97] J. Seebach and L. Steen : *Counterexamples in Topology*, New York : Springer, 1978.

[98] D. Stroock and S. Varadhan : *Multidimensional Diffusion Processes*, Grundlehren der Mathematischen Wissenschaften, Vol. 233, Springer, 1997.

[99] J. Szpirglas and G. Mazziotto : *Théorème de séparation dans le problème d'arrêt optimal*, Séminaires de Probabilités, XIII : 32, 378-384, LNM 721, 1979.

[100] X. Tan and N. Touzi : *Optimal Transportation under Controlled Stochastic Dynamics*, Annals of Probability, Vol. 41, No. 5, 3201-3240, 2013.

[101] M. Tehranchi : *Symmetric martingales and symmetric smiles*, Stochastic Processes and Their Applications 119(10) : 3785–3797. (2009)

[102] P. Vallois : *Le problème de Skorokhod sur $\mathbb{R}$: une approche avec le temps local*, Séminaire de probabilités de Strasbourg, 17 (1983), p. 227-239.

[103] P. Vallois : *Quelques inégalité avec le temps local en zéro du mouvement Brownien*, Stochastic Processes and their Applications 41 (1992) 117-155.

[104] C. Villani : *Optimal Transport : Old and New*, Grundlehren der mathematischen Wissenschaften, 978-3540710493, 2008.

[105] D. Zaev : *On the Monge-Kantorovich problem with additional linear constraints*, Mathematical Notes, 2015, vol. 98, no 5-6, p. 725-741.

universite PARIS-SACLAY
École doctorale de mathématiques Hadamard (EDMH)

Titre : Transport Optimal Martingale en Temps Continu et Plongement de Skorokhod Optimal

Mots clés : transport optimal martingale, couverture robuste, dualité, plongement de Skorokhod optimal, principe de monotonie, contrôle stochastique, solution de Vallois

Résumé: Cette thèse s'articule en trois principaux sujets de recherche. La première partie considère le problème de transport optimal martingale dans l'espace de Skorokhod, et nous établissons une dualité de type Kantorovich en utilisant la $S-$topologie. La deuxième partie de cette thèse est consacrée à trois contributions au plongement de Skorokhod optimal. Nous prouvons les dualités, retrouvons le principe de monotonie par une preuve alternative et simplifiée, et fournissons un résultat de stabilité. Dans la dernière partie, on considère l'application de contrôle stochastique au transport optimal martingale dans l'espace des fonctions continues avec la fonction de paiement dépendant uniquement du temps local. On produit les optimiseurs explicites pour les problèmes primitifs et duaux à l'aide des solutions de Vallois dans le cas d'une marginale. Nous étendons ensuite l'analyse au cas de deux marginales et construisons une généralisation de la solution de Vallois.

Title : Continuous-time Martingale Optimal Transport and Optimal Skorokhod embedding

Keywords : martingale optimal transport, robust hedging, duality, optimal Skorokhod embedding, monotonicity principle, stochastic control, Vallois' solution

Abstract : The Phd dissertation presents three independent research topics. The first part considers the continuous-time martingale optimal transport problem on the Skorokhod space, and we establish a Kantorovich duality via the $S-$topology. The second part of the thesis is devoted to three contributions for the optimal Skorokhod embedding problem. We prove the dualities, recover the monotonicity principle by an alternative and simplified proof, and provide a stability result. In the last part, we consider the application of stochastic control to the martingale optimal transport on the space of continuous functions with a payoff depending on the local time. For the one-marginal case, we recover the explicit optimizers for both primal and dual problems using Vallois' solutions. We then extend the analysis to the two-marginal case and construct a generalization of Vallois' solution.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France