A Robust Approach to Pricing–Hedging Duality and Related Problems in Mathematical Finance

Zhaoxu Hou

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2016

Acknowledgements

I would like to sincerely thank my supervisor, Jan Obłój, for introducing me to this interesting topic, for his patient guidance and continuous support of my DPhil study, and for his inspiration and enthusiasm in our research. This thesis also owes much to his thoughtful and helpful comments. I am also thankful to the Oxford-Man Institute of Quantitative Finance, Balliol College in Oxford and ERC Starting Grant RobustFinMath 335421 for their financial support.

I further express my gratitude to Alexander Cox, Marco Maggis, Matteo Burzoni, Anna Aksamit who work closely with me on different projects, for their encouragement and collaboration. Thanks to Pietro Siorpaes, Peter Spoida and all the other participants of Obłój's weekly group meetings for their insightful advices and tremendous help in various aspects throughout my whole DPhil study. The group meetings provide me with a deep understanding of many closely related subjects, as well as the internal seminars of the Mathematical Finance group in Oxford and the Oxford-Man Institute of Quantitative Finance, and Nomura seminars introduce to me a wider range of different topics in the frontier of probability, stochastic calculus and financial mathematics. I would also like to thank Matthias Qian, Xuejie Ding, Arthur Wu, Evan Cai, Siran Li, Walter Liu, Guy Flint and all my other friends in Oxford for making the place full of fun and excitement. Thank you to Jiabao Xu, for all her love and support.

My deepest appreciation goes to my parents, for everything they have done for me, for their faith in me and the protection they have given. Without them, this thesis would not be possible.

Contents

Chapter 1

Introduction

The era of modern quantitative finance began in 1900, when Louise Bachelier developed a theory of option pricing in his Doctoral thesis, based on the assumption that stock prices follow a driftless Brownian motion. Although the importance of this work did not receive immediate recognition outside the mathematical community, and continued to be largely ignored by economists for almost half a century, the early 1950's saw a surging interest in Louise Bachelier's thesis, eventually leading to the seminal works of Samuelson [94] and Black and Scholes [10], which established an explicit formula for the price of call option based on geometric Brownian motion. Since then, this approach to option pricing and hedging through finding the expected value of payoff under a risk neutral measure, called the *model-specific approach*, has revolutionised the financial industry and became the dominant paradigm for both practitioners and researchers in quantitative finance. Accordingly, we refer to it as the *classical approach*.

The ability to obtain unique prices and hedging strategies, which is the strength of the model-specific approach, relies on its primary weakness – the necessity to postulate a fixed probability measure $\mathbb{P}$ giving a full probabilistic description of future market dynamics. In particular, model mis-specification, a type of risk that arises from this fundamental weakness, can wreak havoc on a trading business due to the model's failure to capture some known risk, or complete ignorance of some unknown risk – which is almost inevitable and even more dangerous. Realising the shortcomings of this approach, significant efforts have been made to either generalise this pricing–hedging framework or to develop new ways of thinking. Such activities reached the peak in the wake of financial crisis. Among them, there is a growing research interest in the *model-independent approach*, a pricing–hedging paradigm which is in stark contrast to the model specific approach, as it is only "based on assumptions sufficiently weak to gain universal support[1]". This idea can be traced back to Merton [77] but was not considered practically relevant at the time, as it produced outputs which were too imprecise.

In this thesis, we are concerned with a robust approach to pricing and hedging, which is along the lines of the model-independent approach. In particular, we are interested in establishing general robust pricing–hedging duality.

1.1 Model-independent approach

The model-independent approach originated from Merton's seminal contribution [77], in which a number of results concerning necessary price restrictions on option pricing formulas were established based on the weak economic assumption that investors prefer more to less. As one might expect, there is usually no uniquely agreed option pricing rule among agents and in fact the resulting price bounds of options tend to be wide. Hence, this ideologically appealing framework always gives way to model-specific approach.

1973 was a milestone year in both the financial industry and the academic community. The Black Scholes model became the Nobel Prize-winning solution of option pricing problem, and the first listed options exchange opened its doors in Chicago. At that time, option pricing studies, focusing on call and put options (which were once considered exotic), were still at a preliminary stage, and the amount of time and space devoted to the development of a pricing theory was constantly questioned "because options are specialized and relatively unimportant financial securities[1]". Since then, the financial landscape has transformed dramatically. Within a decade, option trading had turned itself into a multi-billion industry, with expansion seen from both the volume and the range of contracts traded. These numbers continue to grow and according to the estimates of The Wall Street Journal in 2007, global market capitalisation of the derivatives markets exceeded 450 trillion dollars. Moreover, in contrast to the time when Merton [77] was examining the model-independent approach, at present call and put options are very liquidly traded in the market, and the bid-ask spreads are almost negligible due to high competition among market makers.

[1]Merton [77]

Accordingly, it is becoming increasingly more demanding for a bank to calibrate its models to European prices, where calibration here refers to the procedure of finding the optimal parameters of a parametric family of measures $\{P_\theta\}_{\theta\in\Theta}$ to fit the observed prices of target instruments. This has very significant implications for the theory of option pricing and hedging, and the re-examination of the roles European options play in the pricing and hedging framework has been considered for some time. As argued by many academic researchers (including Dupire [42, 43] and Hobson [60]), liquidly traded European options should be treated as *inputs* and hedging instruments, rather than *outputs*, the receiving end of a pricing–hedging framework. It has the advantage of reducing the possible universe of no-arbitrage scenarios and pushes model-independent approaches to produce results which are closer to being practically relevant. A series of endeavours to putting this idea on a more rigorous and well-defined footing began when Hobson [60] in his pioneering work considered an idealised scenario where all European options for a given maturity are available, and showed in this case how to use the model-independent approach to compute pricing bounds and hedge strategies for lookback options. Hobson [60] used the observation in Breeden and Litzenberger [13] that if many (all) European options for a given maturity trade then this is equivalent to fixing the marginal distribution of the stock under any equivalent martingale measure (EMM) in a classical setting. As a result, the option pricing problem could be approached using the Skorokhod embedding techniques, which had been well studied. In particular, if there is a single stock and we are ready to assume that its associated price process is continuous, then the pricing problem of any time invariant option is in fact equivalent to a Skorokhod embedding problem as any continuous martingale is a time change of a Brownian motion by the Dubins–Schwartz Theorem, see Section 1 of Hobson [59]. Following this idea, other exotic options were analysed in subsequent works, see Brown et al. [14], Cox and Wang [27], Cox and Obłój [25]. In these papers, optimal solutions to the pricing and hedging problems were established and consequently *pricing–hedging duality* were shown: the infimum of prices of super-hedging strategies is equal to the supremum of expectations of the payoff under calibrated martingale measures. Note that the superhedging property here is understood in a pathwise sense and typically the optimal strategies involve only buy-and-hold positions in options and simple dynamic trading in the underlying. The associated hedging strategies were shown to perform remarkably well when compared to traditional delta-vega hedging, see Obłój and Ulmer [84], despite the fact that the resulting no-arbitrage price bounds could still be too wide even for market making.

These papers above could be seen as studies of particular examples by constructing pathwise inequalities and solutions to Skorokhod embedding problems that exhibit some desirable optimality properties. While this approach has proven successful, there were no systematic ways of studying general pricing–hedging duality and robust version of the fundamental theorem of asset pricing (FTAP), until the ground-breaking work of Beiglböck et al. [6], which re-interpreted the study of upper price bounds on option as a primal problem related to the *martingale optimal transport problem*:

$$P_{\vec{\mu}}(G) := \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})],$$

where $\mathbb{S}$ denotes the canonical process on $\mathbb{R}_+^{n+1}$, G is an upper-semi continuous payoff function and $\mathcal{M}_{\vec{\mu}}$ is the set of martingale measures $\mathbb{P}$ starting in a given point s_0 and having fixed marginal distribution(s) as $\vec{\mu}$, which are derived from the option prices via the Breeden and Litzenberger [13] formula. The dual elements here are super-hedging strategies of the option payoff, i.e., a static position (a portfolio of call options) X plus a dynamic hedging strategy $(H_i)_{i\leq n}$ such that

$$\Psi_{X,H}(S) = X(S) + \sum_{j=0}^{n-1} H_j(S)(S_{j+1} - S_j) \geq G(S), \quad \forall S \in \Omega := \{\omega \in \mathbb{R}_+^{n+1} : \omega_0 = s_0\},$$

with the initial cost of financing the operation equal to the cost of establishing the static position, which we here denote by $\mathcal{P}(X)$. The main result of Beiglböck, Henry-Labordère and Penkner [6] is a general discrete-time pricing–hedging duality as an analogue of the Kantorovich duality in the optimal transport: for each upper-semi continuous payoff function G

$$P_{\vec{\mu}}(G) = V_{\vec{\mu}}(G)$$

where $V_{\vec{\mu}}(G)$ is the minimal super-hedging cost (market price) of G, i.e.,

$$V_{\vec{\mu}}(G) = \inf\{\mathcal{P}(X) \,:\, \Psi_{X,H} \geq G \text{ on } \Omega \text{ for some admissible } (X, H)\}.$$

Also in discrete time, Acciaio et al. [1] considered a frictionless market where there is an arbitrary set of market traded options with prices known at time zero, and obtained pricing–hedging duality and robust FTAP under significant technical assumptions. In continuous time, a duality result that is analogous to the one in Beiglböck et al. [6], was first obtained by Dolinsky and Soner [39], in which they considered a single-maturity case and assumed the underlying has continuous price paths. Here the primal elements are martingale measures $\mathbb{P}$ on the space of continuous functions and the dual elements again consist of a static hedging part X and dynamic hedging strategy γ which has the super-replicating property:

$$X(\mathbb{S}) + (\gamma \circ \mathbb{S})_T \geq G(\mathbb{S})$$

where $\gamma \circ \mathbb{S}$ is a well-defined pathwise integral. More recently, Dolinsky and Soner [41] and Guo et al. [51] enlarged the path space to the Skorokhod space and established a pricing–hedging duality to include multiple maturities and price processes in higher dimensions.

1.2 Developments in the model-specific approach

Since the original model of Black and Scholes was invented in 1973, the model-specific approach to option pricing and hedging has undergone rapid development, mainly in two directions. On one hand, as explained above, the complexity of financial markets has been increasing. To be able to calibrate the model to many European options, the model itself needs to be flexible enough, especially if these European options have different maturities. This gave rise to the ground-breaking work of Dupire [43]. By assuming that the underlying price process is a state-dependent time-inhomogeneous diffusion process and that call options for all possible maturities and strikes trade, he showed that, under regularity conditions, such a diffusion process is uniquely determined and hence unique pricing and hedging of options is preserved. This unique process is often called the local volatility model and has the advantage of matching the entire implied volatility surface. While appealing at first, it has been criticised heavily for producing the wrong dynamics for the price process and hence is not suitable for pricing and hedging exotic options e.g. cliquet options. As a solution to this, adding stochastic volatility is often considered. Such a generalisation usually leads to market incompleteness and lack of unique rational warrant prices. Nevertheless, no-arbitrage pricing and hedging was fully characterised in the body of works on the Fundamental Theorem of Asset Pricing (FTAP), culminating in the work of Delbaen and Schachermayer [34]. The feasible prices for a contingent claim correspond to expectations of the (discounted) payoff under equivalent martingale measures (EMM) and form an interval. The bounds of the interval are also given by the super- and sub- hedging prices. Put differently, the supremum of expectations of the payoff under EMMs is equal to the infimum of prices of super-hedging strategies. We refer to this fundamental result as the *pricing–hedging duality*.

The other direction of extension is mainly concerned with model uncertainty, also called *Knightian uncertainty*, see Knight [71]. Accordingly, researchers extended the classical setup to one where many measures $\{\mathbb{P}_\alpha : \alpha \in \Lambda\}$ are simultaneously deemed feasible. This can be seen as weakening assumptions and going back from the model-specific towards model-independent. The pioneering works considered *uncertain volatility*, see Lyons [75] and Avellaneda et al. [3]. More recently, a systematic approach based on quasi-sure analysis was developed with stochastic integration based on the capacity theory in Denis and Martini [37] and on the aggregation method in Soner et al. [99].In addition, Galichon et al. [49] applied the methods of stochastic control to deduce the model-independent prices and hedges, see also Henry-Labordère et al. [56]. In discrete time a corresponding generalisation of the FTAP and the pricing–hedging duality was obtained by Bouchard and Nutz [12] and in continuous time by Biagini et al. [8], see also references therein. We also mention that setups with frictions, e.g. trading constraints, were considered, see Bayraktar and Zhou [5].

1.3 Robust approach: model-independent approach with beliefs

In this thesis, we propose a *robust approach* which interpolates between the model-specific setting and the model-independent setting, the two ends of the spectrum considered by Merton [77]. In contrast to the quasi-sure approach, which can be seen as moving from model-specific towards model-independent by adding more and more *possible* scenarios $\{\mathbb{P}_\alpha : \alpha \in \Lambda\}$, the robust approach subsumes the model-independent setting. However, instead of starting from a "universally accepted" setting as proposed by Merton [77], inspired by Mykland [78]'s idea of incorporating a prediction set of paths into pricing and hedging problems, we here design a pricing and hedging framework which allows the modeller to include modelling beliefs by specifying a set of paths $\mathfrak{P}$ to be considered. More precisely, the modeller can exclude certain paths which are deemed *impossible* from the analysis: the superhedging property is only required to hold on the remaining set of paths $\mathfrak{P}$.

On a philosophical level, we can start with a "universally acceptable" setting, e.g. weak economic assumptions and path space. We can then consider an intermediate step: gathering all public information or easily accessible information to form information space, e.g. when the price processes contain both underlying assets and derivatives, any possible price path should respect future payoff constraints. The last step is to include our beliefs by *ruling out* more and more scenarios as *impossible*, see also Cassese [22]. We may proceed in this way until we end up with paths supporting a unique martingale measure, e.g. a geometric Brownian motion, giving us essentially a model-specific setting. This can be seen as an extreme example of how the robust approach can connect with the model-specific approach. In this case, they have the same pricing problem. In contrast, the hedging part, which we emphasise here, is the key difference between these two approaches. For robust hedging, the sub- or super-hedging arguments are always required for all the paths which remain under consideration, and a (strong) arbitrage would be a strategy which makes positive profit for all remaining paths, see also the recent work of Burzoni et al. [20]. In contrast, the quasi-sure framework is purely probabilistic, leading to probabilistic (quasi-sure) hedging and different notions of no arbitrage, see Bouchard and Nutz [12].

1.4 Trading restrictions

In an ideal market, trading can take place at any time, at any volume and at a unique price without additional costs. This is usually the starting point when considering pricing and hedging problems. In reality, a market is much more complicated. Not only trading frequently at high volume results in significant expenses due to transaction costs and bid-ask spreads, it is usually not possible to trade freely according to one's will as there are all types of trading restrictions imposed on financial assets. These, among many others, pose extra challenges for practitioners and academic researchers, and to produce results which are closer to being useful, these frictions in the market ought to be taken into account. In this thesis, with the main focus remaining on the case of markets where there are no trading restrictions, we will make an endeavour to extend the frictionless robust pricing–hedging framework to include trading restrictions.

Among all types of trading constraints, a ban on short selling has always been the centre of study due to its popularity among regulators. This is reflected in the fact that fewer than half of the more than 150 financial exchanges worldwide allow short selling. In addition, during times of financial crisis, the short-selling ban has been commonly used by developed economies as a regulatory tool to stabilise markets and curb speculation. On the other hand, it is often criticised by academic researchers and some central bankers because of the role it has played in magnifying the decline of asset prices and creating financial bubbles. The controversy around this ban makes it particularly interesting to study. In Chapter 3, we will study robust pricing and hedging under a ban on short selling and provide some insights into the impact of short selling bans on stocks and derivatives from a robust pricing–hedging perspective.

In a classical pricing–hedging framework, the problem of finding superhedging prices under various constraints on the set of admissible portfolios has been extensively studied. Questions of this type arise in Cvitanić and Karatzas [28], where convex constraints in the hedging problem lead to a dual problem where the auxiliary markets are a modification of the original markets reflecting the trading constraints. In the special case of markets where participants may not short sell assets, the class of auxiliary markets corresponds to the class of supermartingale measures. Further results in this direction were established in Jouini and Kallal [68]; Cvitanić et al. [29]; Pham and Touzi [86]; Pulido [89], among many others. In Cvitanić and Karatzas [28], only deterministic convex constraints were considered. It was later extended by Föllmer and Kramkov [48] to include general constraints for semi-martingales.

In contrast to the model-specific world, the model-independent approach to pricing and hedging with frictions is much less studied, even if it has undergone rapid development in recent years. Several authors have considered trading constraints in discrete time, see Fahim and Huang [46]. We also note that a version of pricing–hedging duality theorem in a discrete-time setting with transactions costs was established in Dolinsky and Soner [38]. More recently Dolinsky and Soner [40] have extended their original results to continuous time via an asymptotic approximation.

1.5 An overview of the thesis

The subsequent content in this thesis is concerned with robust fundamental theorem of asset pricing, pricing–hedging duality and their applications. We will assume that Ω is a Polish space with d_m being one of the possible metrics that can give rise to the topology, and $\mathcal{F} = \mathcal{B}_\Omega$ is the Borel σ-algebra induced by the metric d_m. The market has d risky asset $S := (S^{(1)}, \dots, S^{(d)})$ and a numeraire $S^{(0)}$ (e.g. the money market account) with the initial price S_0 known. All prices are denominated in the units of the numeraire, so that $S^{(0)} \equiv 1$. We fix a finite time horizon $[0, T]$ and a set I to index the time when trading stocks can take place. Examples of I include $I = \{0, 1, \cdots, n\}$ and $I = [0, T]$. In the former, S is traded in discrete time, say at $0 = T_0 < T_1 < \cdots < T_n$, while in the later S can be traded continuously in time and we let $T_t = t$ for every $t \in [0,T]$. The price process $\mathbb{S} := (\mathbb{S}_t)_{t\in I}$ is a $\mathbb{R}^d$-valued (or $\mathbb{R}^d_+$-valued) process on $(\Omega, \mathcal{B}_\Omega)$. To model the information stream, we can enrich the financial market by the specification of a σ-algebra $\mathcal{F}$ and a filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t\in I} \subseteq \mathcal{F}$ with the minimum requirement that the process $\mathbb{S}$ is $\mathbb{F}$-adapted. Therefore, a simple model of a financial market is given by the quadruple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{S})$.

In a robust setting, a typical choice of the quadruple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{S})$ is to have Ω be the canonical space with a fixed starting point, $\mathbb{S} = (\mathbb{S}_t)_{t\in I}$ be the canonical process and $\mathbb{F} = (\mathcal{F}^{\mathbb{S}}_t)_{t\in I}$ be its natural filtration. In discrete time, $\Omega = \{(\omega_0, \dots, \omega_n) \in (\mathbb{R}^d_+)^{n+1} : \omega_0 = s_0\}$ and in continuous time $\Omega = \{\omega \in C([0,T], \mathbb{R}^d_+) : \omega_0 = s_0\}$, and $\mathbb{S}$ is a canonical process, i.e.:

$$\mathbb{S}_t : \Omega \to \mathbb{R}^d_+,\ \mathbb{S}_t(\omega) = \omega_t,\ t \in I.$$

The specification of $\mathbb{F}$ is related to the set of admissible dynamic hedging strategies. With $H^{(i)}_t$ denoting the number of $S^{(i)}$ that we hold at time T_t, $H := (H_t)_{t\in I}$ with $H_t = (H^{(1)}_t, \dots, H^{(d)}_t)$ is a trading strategy, and the wealth process from trading S according to H is denoted by $(H \circ \mathbb{S})$. In discrete time, there are no technical difficulties in defining $(H \circ \mathbb{S})$ pathwise. Indeed, we simply put

$$((H \circ \mathbb{S})_{T_m}) = (H \circ \mathbb{S})_m := \sum_{t=0}^{m-1} \sum_{j=1}^{d} H^{(j)}_t (\mathbb{S}^{(j)}_{t+1} - \mathbb{S}^{(j)}_t) = \sum_{t=0}^{m-1} H_t \cdot \Delta \mathbb{S}_{t+1} \quad \forall m \in I.$$

However, in continuous time, to define this integral properly, we need to impose some regularity condition on the dynamic trading strategy. For example, we might follow Dolinsky and Soner [39] and consider H that is of finite variation for which, using integration by parts, for any continuous path S we set

$$(H \circ S)_t := \int_0^t H_u(S) \mathrm{d}S_u = H_t \cdot S_t - H_0 \cdot S_0 - \int_0^t S_u \mathrm{d}H_u,$$

where the last term on the right hand side is a Stieltjes integral. In either case, we require H to be measurable and adapted to $\mathbb{F}$. That means trading in S at time t can only be based on information available at time t. There are also admissibility conditions on H, where the exact meaning of admissibility usually depends on the context and whether there are trading restrictions. For example, in a discrete-time setting where there is a short selling ban on $\mathbb{S}$, apart from being adapted, $H = (H_t)_{t\in I}$ also needs to be a nonnegative process, while in continuous time, it is not uncommon that $(H \circ \mathbb{S})$ is required to be bounded below by another process or simply uniformly. Finally, we use $\mathcal{A}(\mathbb{F})$ to denote the set of admissible dynamic trading strategies.

In addition to S which may include some derivative products that are very liquidly traded, we assume that there are some other options which may be less liquidly traded with prices $\mathcal{P}(X)$, $X \in \mathcal{X}$, known at time zero, are only available for static hedging (trading at the initial time). These are European derivatives which can be seen as $\mathcal{F}_T$-measurable $\mathbb{R}$-valued random variables on Ω. The set of market options available for static trading is denoted by $\mathcal{X}$ and we assume they are linearly independent and can be traded frictionlessly so that in particular $\mathcal{P}$ is a linear operator on $\mathcal{X}$.

An admissible (semi-static) trading strategy is a pair (X, H) where

$$X = a_0 + \sum_{i=1}^{m} a_i X_i \quad \text{and} \quad H \in \mathcal{A}(\mathbb{F})$$

for some $m \in \mathbb{N}$, $X_i \in \mathcal{X}$ and $\mathcal{F}_0$-measurable random variables $a_i : \Omega \to \mathbb{R}$, $i = 0. \ldots, m$. The total payoff associated to (X, H) is given by

$$\Psi_{X,H}(\omega) := X + (H \circ \mathbb{S})_{T_n}(\omega),$$

and the cost of following such a trading strategy is equal to the cost of setting up its static part, i.e., of buying the options at time zero, and is equal to $\mathcal{P}(X) = a_0 + \sum_{i=1}^{m} a_i \mathcal{P}(X_i)$. We denote the class of admissible (semi-static) trading strategies by $\mathcal{A}_{\mathcal{X}}(\mathbb{F})$.

We also include in our setup modelling beliefs by allowing the specification of a set of paths $\mathfrak{P}$ to be considered, i.e., paths we deem feasible and for which the hedging strategies are required to work.

We call the quadruple $(\mathcal{X}, \mathcal{P}, \mathfrak{P}, \mathbb{F})$ of market traded options $\mathcal{X}$, their prices, prediction set and filtration the *robust modelling inputs*. The fundamental financial notions defined in this thesis, e.g. the arbitrage or the super-replication price, are implicitly relative to these inputs.

One of our aims here is to have a better understanding of arbitrage. There are various notions of arbitrage but in principle arbitrage opportunities are situations where it is possible to construct a portfolio $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ with zero initial costs, i.e., $\mathcal{P}(X) = 0$, such that the final wealth is always nonnegative, i.e., $\Psi_{X,H}(\omega) \geq 0$ for all $\omega \in \Omega$ (or $\omega \in \mathfrak{P}$), and shows positive gains in some scenarios, i.e., $\Psi_{X,H}(\omega) > 0$ for some $\omega \in \Omega$. The disagreement among notions of arbitrage usually lies in how "positive gains in some scenarios" is defined. It is clear that the more ω are required for $\Psi_{X,H}$ to be positive, the stronger is the notion of arbitrage. It is then interesting to understand what no-arbitrage condition is equivalent to the existence of a reasonable pricing system via a calibrated risk-neutral probability measure.

Our prime interest of this thesis is to study the robust pricing–hedging duality problem in both discrete and continuous time. We always focus on the case that $\mathcal{F}_0$ is trivial, i.e., $\mathcal{F}_0 = \{\varnothing, \Omega\}$. Note that in this case, $\mathcal{P}(X)$ is a constant and we can define the (minimal) super-replicating cost of a payoff function G on $\mathfrak{P}$ by

$$V^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \inf\{\mathcal{P}(X) \,:\, \exists (X,H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F}) \text{ such that } \Psi_{X,H} \geq G \text{ on } \mathfrak{P}\}.$$

On the other hand, in a market where trading can take place subject to no restriction, the pricing problem is to find the supremum of the expected value of the derivative's payoff over all measures $\mathbb{P}$ on $\mathfrak{P}$ calibrated to $(\mathcal{X},\mathcal{P})$ such that the canonical process S is an $\mathbb{F}$-martingale. Denoting $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})$ by the class of such measures, the primal value or the (maximal) modelling price is defined as

$$P^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})} \mathbb{E}_{\mathbb{P}}[G].$$

In the presence of portfolio constraints, the classification of the appropriate space of dual problem is more involved. Analogous to the classical pricing problem under portfolio constraint, it amounts to finding the supremum of the expected value of the derivative's payoff in a class of auxiliary markets, where the auxiliary markets are a modification of the original markets reflecting the trading constraints. The special type of constraint that we focus on in this thesis is a short selling ban on the underlying asset. In this case, the class of auxiliary markets correspond to the class of supermartingale measures, and hence

$$P^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \sup_{\mathbb{P}\in\mathcal{M}^{-}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})} \mathbb{E}_{\mathbb{P}}[G],$$

where $\mathcal{M}^{-}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})$ is the class of supermartingale measures $\mathbb{P}$ on $\mathfrak{P}$ calibrated to $(\mathcal{X},\mathcal{P})$.

Throughout the work, when the filtration $\mathbb{F}$ is chosen to be $\mathbb{F}^{\mathbb{S}}$, which is the filtration generated by $\mathbb{S}$, we drop $\mathbb{F}$ from any notation which is dependent on filtration. For example, in this case, we write $\mathcal{A}_{\mathcal{X}}$ and $P_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ instead of $\mathcal{A}_{\mathcal{X}}(\mathbb{F})$ and $P^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ respectively.

1.5.1 Chapter 2: Robust pricing and hedging in discrete time

In this chapter, we consider an abstract discrete-time setting, with the price process of the risky assets defined on a Polish space. As mentioned above, we assume that the risky assets, which can include stocks and options, are available for dynamic trading, and in addition there are some options which can only be traded statically. All the trading is frictionless. We first review several notions of arbitrage in the model-independent finance literature. In particular, our focus is on the so called Arbitrage de la classe $\mathcal{S}$, which is a flexible notion of arbitrage, generated by a class $\mathcal{S}$ of significant sets. It was first introduced by Burzoni et al. [20] in a setting where only dynamic hedging is allowed, and shown therein that with the flexibility of specifying $\mathcal{S}$, it can include many notions of arbitrage that already appeared in the literature. We then discuss another important contribution therein, which is a characterisation of the existence of a martingale measure through a decomposition of Ω into Ω^* and $(\Omega^*)^{\complement}$ with $(\Omega^*)^{\complement}$ identified as the largest polar set for martingale measures. It is maximal in the sense that every $\omega \in (\Omega^*)^{\complement}$ can not be charged by a finitely supported martingale measure, i.e.

$$(\Omega^*)^{\complement} := \{\omega^* \in \Omega : \mathbb{P}(\{\omega^*\}) = 0 \quad \forall \mathbb{P} \in \mathcal{M}^f\},$$

where $\mathcal{M}^f := \{\mathbb{P} \in \mathcal{M} : \text{supp}(\mathbb{P}) \text{ is finite}\}$. Following this, we look at a robust pricing–hedging duality result in Burzoni et al. [18], which states that for any measurable contingent claim G,

$$P_\Omega(G) = V_{\Omega^*}(G).$$

Namely, the minimal capital required for a strategy to superreplicate the contingent claim on the compliment of the largest polar set (which is hence the maximal support of martingale measures) is equal to the supremum of the expected value of the contingent claim over all the martingale measures.

The rest of Chapter 2 is concerned with several important generalisations of the results from Burzoni et al. [18, 20]. Firstly, a prediction set $\mathfrak{P}$ is introduced into the framework of Burzoni et al. [20], and the FTAP and pricing–hedging duality result from Burzoni et al. [17] are generalised accordingly. Secondly, static hedging in options is included in the pricing and hedging framework, which leads to the consideration of pricing and hedging problem with semi-static strategies. Although in the absence of static hedging, due to Burzoni et al. [20, 18] the topic is now well understood, in the presence of static hedging, a generalised FTAP and a pricing–hedging duality theorem are still missing, and our main contribution here is to fill the gap in the literature. Our results are presented in Sections 2.4, 2.5 and 2.6[2].

1.5.2 Chapter 3: Robust pricing and hedging under trading restrictions and the emergence of local martingale models

In Chapter 2, we considered the robust approach to pricing and hedging in the case that only finitely many options are available. In this chapter[3], we pursue the same approach but focus on the case that call or put options with different maturities and all strikes can be traded initially at their market prices. In addition, the underlying asset S, whose price process is modelled as the canonical process on the canonical space with a fixed starting point $\Omega = \{(\omega_0, \ldots, \omega_n) \in \mathbb{R}_+^{n+1} : \omega_0 = s_0\}$, may be traded at any time, however there is a *no short-selling ban* on it. We also allow for inclusion of robust modelling assumptions by specifying the set of feasible paths.

In Chapter 2, a version of discrete-time robust fundamental theorem of asset pricing is established when finitely many options are traded statically. However, with infinitely many call options or put options traded, a different notion of arbitrage is needed. We borrow the notion of *weak free lunch with vanishing risk* (WFLVR) from Cox and Obłój [25] and show that in the absence of WFLVR, if only call options trade, the robust pricing–hedging duality is preserved:

$$P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \quad \text{for upper-semi continuous } G.$$

Here primal elements are calibrated super-martingale models. By denoting by $\mathcal{M}^-_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ the set of such models,

$$P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \sup_{\mathbb{P}\in\mathcal{M}^-_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G]$$

is the modelling/fundamental price of G. For the super-hedging problem,

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \inf\{\mathcal{P}(X) \,:\, \Psi_{X,H} \geq G \text{ on } \mathfrak{P} \text{ for some admissible } (X,H)\}$$

[2] The original research presented in Sections 2.4, 2.5 and 2.6 will be published as part of a joint work in Burzoni et al. [17].

[3] The original research presented in this chapter will be published as part of a joint work in Cox et al. [24].

is the minimal super-hedging cost (market price) of G, where X is a semi-static position (a portfolio of options), H is a nonnegative dynamic hedging strategy and

$$\Psi_{X,H} = X + \sum_{j=0}^{n-1} H_j(S_{j+1} - S_j).$$

In contrast, if only put options are traded a duality gap may appear, i.e., $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) > P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ for some G. This misalignment of market price and modelling price has rather interesting financial interpretations. This is related to the notion of *rational bubbles* in asset pricing and economic literature and suggests that a rational bubble could arise when market prices of options are misaligned while an arbitrage does not arise because of trading restriction. Subsequently, embedding the results into a continuous-time framework, we show that the duality gap can be again interpreted as a financial bubble and link it to strict local martingales. This provides an intrinsic justification of strict local martingales as models for financial bubbles arising from a combination of trading restrictions and current market prices.

1.5.3 Chapter 4: Robust pricing–hedging duality without options and dynamic programming principle

We shift our focus from discrete time to continuous time. In this chapter, we consider a very simple market where only stocks are available for trading, and we also assume that those stocks have continuous price paths and can be traded without frictions. We work on the canonical space, which is the space of nonnegative continuous functions, and the stock price process is modelled as the canonical process. In addition,we only consider the natural filtration, which is the one generated by the price process.

Our first aim is to show an “unconstrained” pricing–hedging duality. In the absence of options and beliefs, the infimum of the capital needed to superreplicate an option with payoff G is

$$\inf \Big\{ x \,:\, \exists \gamma \in \mathcal{A} \text{ s.t. } \gamma \text{ superreplicates } G - x \text{ on } \Omega \Big\},$$

and we denote it by $\mathbf{V}(G)$. On the other hand, the robust price for G is given as

$$\mathbf{P}(G) := \sup_{\mathbb{P} \in \mathcal{M}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})],$$

where the supremum is taken over all martingale measures for $\mathbb{S}$. It can be shown that for any bounded and uniformly continuous G

$$\mathbf{P}(G) = \mathbf{V}(G).$$

This "unconstrained" duality is proved by a discretisation method which is inspired by Dolinsky and Soner [39, 38]. The duality result can serve as a crucial building block for the general pricing–hedging duality that we will establish in Chapter 5 via a variational approach as in Galichon et al. [49].

Our second contribution in this chapter is to establish a dynamic programming principle (DPP), which relates the time-0 superhedging cost (resp. model price) to any later time-t supherhedging cost (resp. model price). We prove that looking at the pricing and the hedging problems on $[0, T]$ is the same as first looking at the pricing and the hedging problems on $[T_1, T]$ and then on $[0, T_1]$.

1.5.4 Chapter 5: Robust pricing–hedging duality with options in continuous time

Following Chapter 4, we consider in this chapter a continuous-time setting where there are some underlying assets and options available for dynamic trading without trading constraints, and a further set of European options $\mathcal{X}$, possibly with varying maturities available for static trading. As before, we assume assets have continuous price paths and can be traded frictionlessly. Following the philosophy of robust approach, described in Section 1.3, we first form an information space $\mathcal{I}$ by excluding those price paths which do not respect future payoff constraints, and then include beliefs – a subset of paths $\mathfrak{P} \subseteq \mathcal{I}$ to be considered, by further excluding paths which are deemed impossible. By assuming a suitable structure for $\mathcal{X}$ and existence of market calibrated models, we obtain an approximated pricing–hedging duality result:

$$\widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \quad \text{for uniformly continuous and bounded } G\ .$$

Here

$$\widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \lim_{\eta \searrow 0} \sup_{\mathbb{P} \in \mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G]$$

is the approximated modelling pricing, where $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ is the set of martingale measure calibrated to $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$ subject to a calibration error η, and $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ is the limit of minimal super-hedging cost of G on an η neighbourhood of $\mathfrak{P}$ as η tends to zero. In particular, when $\mathfrak{P} = \mathcal{I}$ and $\mathcal{X}$ is finite, we can show that

$$P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \widetilde{V}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G).$$

Moreover, when $\mathcal{X}$ is the set of all put options, the pricing problem is connected to the martingale optimal transport problem and our duality results in this paper include the martingale optimal transport duality of Dolinsky and Soner [39] and extend it to multiple maturities and multiple assets. Furthermore, in this case, under a stronger assumption on payoff functions and $\mathfrak{P}$, we show that

$$P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G).$$

Chapter 2

Robust pricing and hedging in discrete time

In this chapter, we consider robust pricing and hedging in discrete time, where we take the idealised assumption on the market that trading financial assets is always frictionless. First, we discuss the simplest case, that is every financial asset can be traded discretely at any time without any constraint. Later on, we extend our analysis to include a finite number of financial derivatives which can only be traded at time zero. In both cases, the main focuses of our discussion are on a generalised version of the robust fundamental theorem of asset pricing and the pricing–hedging duality.

Mathematically, an idealised financial market can be simply described by $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{S})$, where Ω is the event space, $\mathbb{S} = (\mathbb{S}_t)_{t=0}^n$ is used to model to price process of financial assets and $\mathbb{F}$ is the filtration, which is a mathematical way of modelling the information stream on which any investment decision is based. Additionally, the information stream is always required to contain the price information given by $\mathbb{S}$, or equivalently $\mathbb{S}$ is adapted to $\mathbb{F} = (\mathcal{F}_t)_{t=0}^n$. The way of describing a market through $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{S})$ has always been the basis of any asset pricing theory.

Among all the studies of asset pricing theory in the risk-neutral framework, the study of no arbitrage lies in the centre. Although there are many variations in the notions of arbitrage, conceptually an arbitrage opportunity is a strategy which requires no initial capital or new funds in succeeding period and can yield, through a combination of buying and selling financial securities, a positive gain in some circumstances without a countervailing treat of loss in other circumstances. Such a strategy, if one exists, represent a plan of generating profit without any risk, and would be exploited by arbitrageurs who are actively seeking such opportunities and hence can only live for a short period. Therefore, a market containing arbitrage opportunities would not be sustainable and hence cannot be the one in which an economic equilibrium exists. Such argument can be formalised to form a version of the Fundamental Theorem of Asset Pricing, which is usually the foundation of any pricing theory. Roughly speaking, the Fundamental Theorem of Asset Pricing asserts that the absence of a (suitably defined) arbitrage opportunity is equivalent to the existence of a reasonable price system, which can be established under some (risk-neutral) probability measure.

The theory was first developed for the case of no static hedging in options. A first intuition for this equivalence can be traced back to De Finetti for his work on *coherence* and *previsions* (see De Finetti [33]), while the first systematic approach for understanding the deep relation among no-arbitrage pricing and risk-neutral pricing can be found in the work of Ross on Arbitrage Pricing Theory (see e.g. Ross [92, 93]) and further developments in Huberman [62]. Later on in the case of Ω being a finite set of events a version of the FTAP was proved in Harrison and Pliska [54] (see also Harrison and Kreps [53]; Kreps [72]) using geometric arguments and separation in finite dimensional spaces.

Theorem 2.0.1 (Theorem 2.7 in Harrison and Pliska [54])**.** Assume Ω is finite. Let $\mathbb{F}$ be a filtration such that $\mathcal{F}_0 = \{\varnothing, \Omega\}$ and $\mathcal{F}_n$ is the set of all subsets, and $\mathbb{S} = (\mathbb{S}^{(1)}, \cdots, \mathbb{S}^{(d)})$ be the price process such that $\mathbb{S}$ is adapted to $\mathbb{F}$.

$$\begin{array}{c} \nexists\ \mathbb{F}\text{-adapted } H \text{ such that} \\ (H \circ \mathbb{S})_n(\omega) \geq 0\ \forall \omega \in \Omega \text{ and} \\ (H \circ \mathbb{S})_n(\omega) > 0 \text{ for some } \omega \in \Omega \end{array} \iff \begin{array}{c} \exists\ \mathbb{Q} \in \Pi \text{ such that} \\ \mathbb{S} \text{ is an } \mathbb{F}\text{-martingale under } \mathbb{Q} \\ \text{and } \mathbb{Q}(\omega) > 0\ \forall \omega \in \Omega \end{array} \quad (2.0.1)$$

where $(H \circ \mathbb{S})_n := \sum_{t=0}^{n-1} \sum_{j=1}^{d} H_t^{(j)} (\mathbb{S}_{t+1}^{(j)} - \mathbb{S}_t^{(j)}) = \sum_{t=0}^{n-1} H_t \cdot \Delta \mathbb{S}_{t+1}$ and Π is the set of probability measures on Ω.

The result above requires no specification of a *reference probability measure* and hence fits into the robust framework. However, in the classical setting, when Ω is uncountable, then a *reference probability measure* $\mathbb{P}$ is introduced, and is used to define an arbitrage as an $\mathbb{F}$-adapted trading strategy H with $\mathbb{P}((H \circ \mathbb{S})_n(\omega) \geq 0) = 1$ and $\mathbb{P}((H \circ \mathbb{S})_n(\omega) > 0) > 0$. This notion of arbitrage is well known and here we refer to it as $\mathbb{P}$-*Classical Arbitrage*. Notice that in the case that Ω is finite, if $\mathbb{P}$ is chosen such that $\mathbb{P}(\omega) > 0$ for any $\omega \in \Omega$, then the left hand side of equivalence above is in fact absence of $\mathbb{P}$-Classical Arbitrage, and Theorem 2.0.1 can be restated as

$$\text{There is No } \mathbb{P}\text{-Classical Arbitrage} \iff \exists \mathbb{Q} \sim \mathbb{P} \text{ such that } \mathbb{S} \text{ is a martingale under } \mathbb{Q} \quad (2.0.2)$$

Later, Dalang et al. [30] extended the FTAP to the case of a general infinite dimensional Ω, using measurable selection arguments. However, for a infinite dimensional Ω, the equivalence between arbitrage in the classical setting and robust setting is lost since it is well known that on an uncountable dimensional Ω it is impossible to find a single measure $\mathbb{Q}$ such that $\mathbb{Q}(\{\omega\}) > 0$ for any $\omega \in \Omega$.

The classical approach to option pricing and hedging problems gained popularity after the seminal works of Samuelson [94] and Black and Scholes [10]. Since then, the use of a reference measure has became a dominant approach across the whole field of Mathematical Finance, which leads to substantial developments in the no-arbitrage theory along this direction. We already mentioned above the discrete-time extension of FTAP in Dalang et al. [30]. In continuous time, no-arbitrage pricing and hedging was fully characterised in the body of works on the Fundamental Theorem of Asset Pricing (FTAP), culminating in the work of Delbaen and Schachermayer [34] where the famous *no free lunch with vanishing risk* was introduced. In recent years, we also saw generalisation of the FTAP in both discrete-time and continuous-time to settings where more than one reference probability measures are considered simultaneously. See Bouchard and Nutz [12] and Biagini et al. [8], and also references therein.

In contrast, there had been relatively fewer studies of the generalization of (2.0.1) without a reference probability measure, until the recent years which have witnessed an increasing interest in the robust pricing and hedging. An interesting contribution to the generalisation was given by Riedel [90], where the existence of a 1p-arbitrage was studied in a one-period setting. We say a strategy trading H is a 1p-arbitrage if $(H \circ \mathbb{S})_1(\omega) \geq 0$ for all $\omega \in \Omega$ and $(H \circ \mathbb{S})_1(\omega) > 0$ for some $\omega \in \Omega$. The idea there is to replace the use of a reference probability by the use of topology of Ω. Under the assumption that Ω is a Polish space and $\mathbb{S}$ is a continuous function of ω, it was shown in Riedel [90] that the absence of 1p-arbitrage is equivalent to the existence of a martingale measure with full support. Clearly "no 1p-arbitrage" condition is the strongest that one can assume in the robust framework. Indeed, it was shown in proposition 17 of Burzoni et al. [20] that the absence of 1p-arbitrage automatically implies not only the existence of a martingale measure but also the existence of a martingale measure which has a positive probability on every open set. However, the inverse implication is not true in general. Even if continuity of $\mathbb{S}$ is assumed, in a multi-period setting, the existence of a martingale measure with full support can only guarantee the absence of the so called open arbitrage but not 1p-arbitrage, see section 4.1 and section 6 in Burzoni et al. [20] for counterexamples and explanation of open arbitrage.

In fact, before Riedel [90], some stronger notions of arbitrage were already suggested and considered by other authors. Among them, perhaps the most popular one is the model-independent arbitrage, which is also called *robust strong arbitrage* here. It is defined as a portfolio H such that $(H \circ \mathbb{S})_n(\omega) > 0$ for all $\omega \in \Omega$. In particular, in contrast to 1p-arbitrage, model-independent arbitrage has been mainly studied in the settings where there are some vanilla options available for static hedging, and the aims of the research are usually to derive no-arbitrage conditions on the prices of these vanilla options and the implications of these additional options on robust pricing–hedging bounds, see Davis and Hobson [31], Cox and Obłój [25], Acciaio et al. [1] and the references therein. In spite of the differences of settings and research focuses, many of the implications of the results from those papers shall be universal, considering that in a one-period setup, there is no difference between dynamic and static hedging in options. In those papers, the results usually state that under some conditions concerning either the initial prices of the options or the existence of some "special " options, absence of model-independent arbitrage is equivalent to the existence of a market calibrated martingale measure. However, as shown in Davis and Hobson [31], for the case of a market with a finite collection of call options, if the call options fail to satisfy some conditions on the initial prices, then there might be situations where there is no market calibrated martingale measure, but at the same time model-independent arbitrage is also impossible [1]. In fact, it was pointed out by the authors that in this case for every reference probability $\mathbb{P}$, $\mathbb{P}$-Classical Arbitrage exists, but the trading strategy which can realize it depends on $\mathbb{P}$ and it is not possible to aggregate them when considering canonical filtration. In more financial terms, it is the exact situation where every agent agrees that certain inefficiencies exist in the market but they cannot agree on the trading strategies to realize the arbitrage opportunities. On the other hand, in a canonical setup, it was shown in Acciaio et al. [1] that when there is an option with convex and superlinear payoff in the market, the existence of a market calibrated model is equivalent to the absence of model-independent arbitrage.

We just reviewed two extreme notions of arbitrage in the robust framework: 1p-arbitrage and model-independent arbitrage. On one hand, 1p-arbitrage is too weak

[1] The model-independent arbitrage defined in Davis and Hobson [31] and Cox and Obłój [25] is required to be uniform in outcomes and hence in fact slightly stronger than the model-independent arbitrage defined in Acciaio et al. [1] and here. To distinguish it from model-independent arbitrage defined here, we call it *robust uniformly strong arbitrage*.

in the sense that the absence of 1p-arbitrage is too strong to assume for the existence of a martingale measure, even for the existence of a martingale measure with full support. On the other hand, model-independent arbitrage is too strong, which can be reflected on the fact that the absence of model-independent arbitrage cannot guarantee the existence of a martingale measure in general. However, for different notions of arbitrage, including the two aforementioned, specific analysis can be applied to show that the no arbitrage condition is equivalent to the existence of a market calibrated model under a suitably chosen setting, in spite of the fact that the equivalence might fail in a different setting. This peculiar phenomenon seems to suggest that in general there is still a lack of fundamental understanding of arbitrage in our ways of thinking.

In the absence of static hedging, this phenomenon was noted and rigorously examined in Burzoni et al. [20] with many examples and new insights into the topic. In addition, the authors provided a new way of understanding arbitrage. The novelty there is to require a specification of subsets of the event space on which a positive gain everywhere is deemed significant. Then with the specification of significant sets, arbitrage is defined as a strategy H such that $(H \circ \mathbb{S})_n \geq 0$ and there exists a significant set such that $(H \circ \mathbb{S})_n > 0$ on it. The definition of arbitrage there is relative to the significant sets and has great flexibility. As shown in Burzoni et al. [20], many existing notions of arbitrage are nothing but special cases corresponding to different choices of significant sets, and the main result in the paper states the equivalence of absence of arbitrage with respected to a properly enlarged filtration is equivalent to the existence of martingale measures such that none of the significant sets are polar sets for the martingale measures.

The enlargement of filtration there is both technical and important. It is somewhat technical as the filtration is only "slightly" enlarged so that the set of pricing models – martingale measures with respect to the canonical or the enlarged filtration, stays unchanged. At the same time, it plays a very crucial role in allowing the arbitrage strategies which only work on parts of Ω to be aggregated into a single strategy, while as explained before, it is not always possible to do that when the filtration is canonical. To define an enlargement of filtration, it could be done by first defining an multifunction H denominated *Universal Arbitrage Aggregator*, whose task as suggested by its name is to capture all the inefficiencies of the market. Then the filtration could be enlarged progressively with respect to H and hence H is adapted to the enlarged one.

One of the important implications of the FTAP in Burzoni et al. [20] is that the possible significant sets which can guarantee the absence of arbitrage relative to them is determined by the class of polar sets for the martingale measures. In Burzoni et al. [20], the maximal $\mathcal{M}$-polar set was also studied. The complement of it is the maximal support of $\mathcal{M}$ and is denoted by Ω^*. As shown in the subsequent paper Burzoni et al. [18], this set is very important for studying the pricing–hedging duality problem — a problem that is often closely related to the FTAP. For an introduction to the pricing–hedging duality problem, we refer the reader to Chapter 1. In Burzoni et al. [18], instead of considering superhedging on the whole space Ω, superhedging is only required on Ω^*. In this case, the main result states that in the absence of static hedging there is a pricing–hedging duality for any general measurable payoff function G. In addition, the minimal superhedging cost could be attained by some optimal strategy. Moreover, a dynamic programming principle for the pricing and hedging problems was implicitly shown, despite not being explicitly stated.

It is not unusual to consider superhedging not on the whole space Ω but on a subset of it, as can be found in Mykland [78], Beiglböck et al. [7], Nadtochiy and Obłój [79], Hou and Obłój [61] and Spoida [100]. However the motivations there are different. In those papers, the consideration of a subset $\mathfrak{P}$ is to incorporate agent's beliefs into the robust framework. Through specifying the set $\mathfrak{P} \subset \Omega$ of "possible paths", i.e., paths the agent deems feasible and for which the hedging strategies are required to work, the maximal support of the plausible models is specified. The same consideration could also be taken in the framework of Burzoni et al. [20] and Burzoni et al. [18]. In fact, when the prediction set $\mathfrak{P} \subset \Omega$ is analytic, like in the case of no beliefs, $\mathfrak{P}$ can be decomposed into $\mathfrak{P}^*$ and $(\mathfrak{P}^*)^\complement$, with $\mathfrak{P}^*$ being the maximal support of martingale measures which are supported on $\mathfrak{P}$. $\mathfrak{P}^*$ has the natural financial interpretation as the efficient beliefs, since, using their previous results, it could be shown that among all the subset of $\mathfrak{P}$ it is the largest one on which there is no 1p-arbitrage opportunity[2].

In this Chapter, we have two objectives. First, we give a detailed review of the existing notions of arbitrage and the results in Burzoni et al. [20]. We shall also see how a prediction set $\mathfrak{P}$ could be incorporated into the framework of Burzoni et al. [20]. In particular, a generalised version of FTAP and a pricing–hedging duality result will be discussed in details. Our second objective is to fill a gap in the current literature. Although in the absence of static hedging, due to the contributions of Burzoni et al.

[2]The generalisation of the FTAP and pricing–hedging duality result from Burzoni et al. [17] Burzoni et al. [20] to include a prediction set is recently discussed in Burzoni et al. [19]. The results will be published as part of a joint work in Burzoni et al. [17].

[20, 18] the topic is well understood, in the presence of static hedging, a generalised FTAP and a pricing–hedging duality theorem are still missing. These are developed in Sections 2.4, 2.5 and 2.6[3].

This chapter is organised as follows. Section 2.1 discusses the general robust framework in discrete time. Sections 2.2 reviews the existing notions of arbitrage in the literature. In particular, it includes a discussion of Arbitrage *de la classe* $\mathcal{S}$ which was introduced in Burzoni et al. [20]. Section 2.3 specialises to robust FTAP and pricing–hedging duality without static hedging in the presence of prediction set, which is an extension of the results in Burzoni et al. [18]. Subsequently Sections 2.4 and 2.5 focus on the contribution that we made — extend the FTAP and pricing–hedging duality theorem to the case that a finite number of options are allowed for static hedging. Section 2.6 is devoted to FTAP and pricing–hedging duality theorem in the case that an arbitrary number of options are allowed for static hedging. Several proofs are relegated to Section 2.7.

2.1 Market setup

Our setting is largely similar to Burzoni et al. [20] but we adopt different notation for consistency with other chapters. We consider a financial market with d risky asset $S := (S^{(1)}, \ldots, S^{(d)})$ and a numeraire $S^{(0)}$ (e.g. the money market account). All prices are denominated in the units of the numeraire. In particular, $S^{(0)} \equiv 1$. We assume initially that S is traded discretely in time at $0 = T_0 < T_1 < T_2 < \ldots < T_n = T$ with the initial price S_0 known. Here, instead of having Ω as the canonical space and the prices of risky assets as the canonical process, we consider a more general and abstract setting, where the space Ω is Polish with d_m being one of the metrics that can give rise to the topology, $\mathcal{B}_\Omega$ be the Borel σ-algebra induced by d_m and $\mathbb{S} = (\mathbb{S}_i)_{i=0}^n$ is a $\mathbb{R}^d$-valued process on $(\Omega, \mathcal{B}_\Omega)$. In order to model the information stream, we shall then enrich the financial market by specifying a σ-algebra $\mathcal{F}$ and a filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t=0}^n \subseteq \mathcal{F}$ with the requirement that the process $\mathbb{S} := (\mathbb{S}_t)_{t=0}^n$ is $\mathbb{F}$-adapted. In particular, the requirement is the same as that $\mathbb{F}$ is larger than $\mathbb{F}^S := \{\mathcal{F}_t^{\mathbb{S}}\}_{t=0}^n$ – the natural filtration generated by $\mathbb{S}$. Later, we shall see that in some situations it is crucial to consider a filtration $\mathbb{F}$ that is strictly larger than $\mathbb{F}^{\mathbb{S}}$. Therefore, a simple financial market is modelled by the quadruple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{S})$.

[3]The original research presented in Sections 2.4, 2.5 and 2.6 will be published as part of a joint work in Burzoni et al. [17].

In the model independence literature, a popular choice of the quadruple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{S})$ to study the pricing and hedging problem, which can be found in [1; 46; 24] among many others, is to have Ω as the canonical space with a fixed starting point $\Omega = \{(\omega_0, \ldots, \omega_n) \in (\mathbb{R}_+^d)^{n+1} : \omega_0 = s_0\}$ and $\mathbb{S} = (\mathbb{S}_t)_{t=0}^n$ being the canonical process, i.e.,

$$\mathbb{S}_t : \Omega \to \mathbb{R}_+^d,\ \mathbb{S}_t(\omega_0, \omega_1, \ldots, \omega_n) = \omega_t,\ t = 0, \ldots, n,$$

with $\mathbb{F} = (\mathcal{F}_t^{\mathbb{S}})_{t=1}^n$ being its natural filtration.

The specification of $\mathbb{F}$ is often related to the set of admissible dynamic hedging. With $H_t^{(i)}$ denoting the number of $\mathbb{S}^{(i)}$ that we hold at time T_t, $H := (H_t)_{t=0}^{n-1}$ with $H_t = (H_t^{(1)}, \ldots, H_t^{(d)})$ yields a capital gain from trading $\mathbb{S}$ by

$$(H \circ \mathbb{S})_n := \sum_{t=0}^{n-1} \sum_{j=1}^{d} H_t^{(j)} (\mathbb{S}_{t+1}^{(j)} - \mathbb{S}_t^{(j)}) = \sum_{t=0}^{n-1} H_t \cdot \Delta \mathbb{S}_{t+1}.$$

We require H to be adapted to $\mathbb{F}$. That means trading in S can only take into account information available to us up-to-date. Finally, we use $\mathcal{A}(\mathbb{F})$ to denote the set of admissible dynamic trading strategies.

In addition to $\mathbb{S}$, which may include some derivative products that are liquidly traded, we assume that there are some other options which may be less liquidly traded with prices $\mathcal{P}(X)$, $X \in \mathcal{X}$, known at time zero, which are only available for static hedging (trading at the initial time). These are European derivatives which can be seen as $\mathcal{F}_T$-measurable $\mathbb{R}$-valued random variables on Ω. The set of market options available for static trading is denoted by $\mathcal{X}$ and we assume they are linearly independent and can be traded frictionlessly so that as discussed before $\mathcal{P}$ is a linear operator on $\mathcal{X}$.

An admissible (semi-static) trading strategy is a pair (X, H) where

$$X = a_0 + \sum_{i=1}^{m} a_i X_i \quad \text{and} \quad H \in \mathcal{A}(\mathbb{F})$$

for some $m \in \mathbb{N}$, $X_i \in \mathcal{X}$ and $\mathcal{F}_0$-measurable random variables $a_i : \Omega \to \mathbb{R}$, $i = 0. \ldots, m$. The total payoff associated to (X, H) is given by

$$\Psi_{X,H}(\omega) := X(\omega) + (H \circ \mathbb{S})_n(\omega),$$

and the cost of following such a trading strategy is equal to the cost of setting up its static part, i.e., of buying the options at time zero, and is equal to $\mathcal{P}(X) = a_0 + \sum_{i=1}^m a_i \mathcal{P}(X_i)$. We denote the class of admissible (semi-static) trading strategies by $\mathcal{A}_{\mathcal{X}}(\mathbb{F})$.

We also incorporate beliefs into the robust framework through specifying the prediction set $\mathfrak{P} \subset \Omega$, i.e., paths we deem feasible and for which the hedging strategies are required to work.

We call the quadruple $(\mathcal{X}, \mathcal{P}, \mathfrak{P}, \mathbb{F})$ of market traded options $\mathcal{X}$, their prices, prediction set and filtration, the *robust modelling inputs*. To present the existing notions of arbitrage in literature and results in Burzoni et al. [18, 20], we shall rewrite them in our notations defined below, which are implicitly relative to these inputs.

Definition 2.1.1. In the case that $\mathcal{F}_0$ is trivial, the *superreplication cost* of a derivative given by a payoff function $G : \Omega \to \mathbb{R}$, denoted by $V^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, is the smallest initial capital required to finance an admissible semi-static trading strategy which superreplicates G for every path in $\mathfrak{P}$, i.e.,

$$V^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \inf \Big\{ \mathcal{P}(X) \,:\, \exists (X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F}) \text{ s.t } \Psi_{X,H}(\omega) \geq G(\omega) \, \forall \omega \in \mathfrak{P} \Big\}.$$

In the case that $\mathcal{F}_0$ is trivial, $\mathcal{P}(X)$ is a constant and hence $V^{\mathbb{F}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ is well defined. However, it is not straightforward to extend the definition to non-trivial $\mathcal{F}_0$. Also, there had also been very little discussion of the superreplication cost in the robust framework, until the recent paper Aksamit et al. [2], where a general framework to express the superhedging and market model prices for an initially enlarged filtration was established. The motivation there is to study the informed agent problem and the framework allows quantification of the value of new information. In this chapter, for the superhedging problem, we shall focus on the case that $\mathcal{F}_0$ is trivial and hence we refrain from introducing the general framework to define the superreplication cost of a derivative when $\mathcal{F}_0$ is non-trivial. Our aim is to understand when a pricing–hedging duality holds, i.e., when the superreplication price can be computed through the supremum of expectations of the payoff over a suitable class of probabilistic models.

Definition 2.1.2. We let Π be the set of all probability measures on $(\Omega, \mathcal{F})$ and for any $\mathfrak{P} \subseteq \Omega$ define

$$\begin{aligned}
\mathcal{M}(\mathbb{F}) &:= \quad \{\mathbb{Q} \in \Pi \mid S \text{ is an } \mathbb{F}\text{-martingale under } \mathbb{Q}\}, \\
\mathcal{M}_{\mathfrak{P}}(\mathbb{F}) &:= \quad \{\mathbb{Q} \in \mathcal{M}(\mathbb{F}) \mid \mathbb{Q}(\mathfrak{P}) = 1\}, \\
\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}) &:= \quad \{\mathbb{Q} \in \mathcal{M}_{\mathfrak{P}}(\mathbb{F}) \mid \mathbb{E}_{\mathbb{Q}}[X] = \mathcal{P}(X) \; \forall X \in \mathcal{X}\}, \\
\Pi^f &:= \quad \{\mathbb{Q} \in \Pi \mid \text{supp}(\mathbb{Q}) \text{ is finite}\}, \\
\mathcal{M}^f(\mathbb{F}) &:= \quad \mathcal{M}(\mathbb{F}) \cap \Pi^f, \\
\mathcal{M}^f_{\mathfrak{P}}(\mathbb{F}) &:= \quad \mathcal{M}_{\mathfrak{P}}(\mathbb{F}) \cap \Pi^f, \\
\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}) &:= \quad \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}) \cap \Pi^f,
\end{aligned}$$

where the support of $\mathbb{P} \in \Pi$ is defined by $\text{supp}(\mathbb{P}) = \bigcap\{C \in \mathcal{F} \mid C \text{ closed}, \mathbb{P}(C) = 1\}$. For any $\Pi_0 \subset \Pi$, the family of Π_0-polar sets is given by $\mathcal{N}(\Pi_0) := \{N \subseteq A \in \mathcal{F} \mid \mathbb{Q}(A) = 0 \ \forall \mathbb{Q} \in \Pi_0\}$.

2.2 Arbitrage

We first present a series of notions of arbitrage which have been in use in the literature. They were introduced by various researchers in different settings. Many were actually written in settings where there are no options or static hedging in options is not considered, while here all of them are defined relative to the input $(\mathcal{X}, \mathcal{P}, \mathfrak{P}, \mathbb{F})$. From this perspective, our definitions could be seen as natural generalisations.

Definition 2.2.1.

1. $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ is a *robust uniformly strong arbitrage* (RUSA) if $\exists \epsilon > 0$ such that
$$\Psi_{X,H}(\omega) - \mathcal{P}(X) > \epsilon \quad \forall \omega \in \mathfrak{P}.$$
2. $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ is a *robust strong arbitrage* (RSA) or *model-independent arbitrage* if
$$\Psi_{X,H}(\omega) - \mathcal{P}(X) > 0 \quad \forall \omega \in \mathfrak{P}.$$
3. $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ is a *robust np-arbitrage* if
$$\Psi_{X,H}(\omega) - \mathcal{P}(X) \geq 0 \quad \forall \omega \in \mathfrak{P}$$
and $\psi^+_{X,H} \cap \mathfrak{P}$ has at least n elements, where $\psi^+_{X,H} = \{\omega \in \Omega \mid \ \Psi_{X,H}(\omega) > \mathcal{P}(X)\}$.
4. We say that there is a weak free lunch with vanishing risk (WFLVR) if there exist trading strategies $(X_k, H_k) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ and $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ such that $\Psi_{X_k,H_k} \to 0$ pointwise on $\mathfrak{P}$, $\lim_k \mathcal{P}(X_k)$ is well defined with $\lim_k \mathcal{P}(X_k) < 0$ and $\Psi_{X_k,H_k} \geq \Psi_{X,H}$.

Remark 2.2.2. When $\mathfrak{P} = \Omega$, A *robust strong arbitrage* is often called as a *model-independent arbitrage*; see Davis and Hobson [31], Cox and Obłój [25] and Burzoni et al. [20, 18]. Note that these papers did not consider a path restriction through a prediction set, as we do here. Hence we here refer the model-independent arbitrage to the robust strong arbitrage in the absence of prediction set. WFLVR is another notion of arbitrage that appeared in the literature before. It, alongside with various others weaker notions of arbitrage, was introduced mainly to study the pricing and hedging problem when there are infinitely many call or put options, see Cox and Obłój [25] and Cox et al. [24].

From the definition it is clear that RUSA is also a RSA. In general, robust uniformly strong arbitrage and robust strong arbitrage are two different notions. We illustrate this with two (motivating) examples, where we assume that Ω is the canonical space.

Example 2.2.3. Consider $n = 1$, $\mathfrak{P} = \{\mathbf{s} \in \Omega : \; s_0 = 1, \; s_1 > 1\}$, $\mathcal{F}_0$ is trivial and no options are traded.

We know for any $H_0 > 0$,

$$H_0(s_1 - s_0) > 0 \quad \text{ on } \mathfrak{P},$$

and this yields a robust strong arbitrage. However, it is also straightforward to see that there exist no robust uniformly strong arbitrages on $\mathfrak{P}$.

Example 2.2.4. Consider $n = 1$, $\mathfrak{P} = \Omega$, $\mathcal{F}_0$ is trivial and $\mathcal{X} = \{f(\mathbb{S}_1)\}$ with $\mathcal{P}(f(\mathbb{S}_1)) = 0$, where $f : \mathbb{R}_+ \to (0, \infty)$ is a given positive function such that $\lim_{x\to 0} f(x) = 0$ and $\lim_{x\to\infty} f(x) = 0$.

Then in this case, it is straightforward to see that $(f(\mathbb{S}_1), 0) \in \mathcal{A}_{\mathcal{X}}$ is a RSA. However, there is no RUSA as for any $H_0 \in \mathbb{R}_+$ and $a \in \mathbb{R}$,

$$\lim_{s_1\to 0} \big(af(s_1) + H_0(s_1 - s_0)\big) \leq \lim_{s_1\to 0} af(s_1) = 0.$$

2.2.1 Arbitrage *de la classe $\mathcal{S}$.*

Recall from above that

$$\psi^+_{X,H} = \{\omega \in \Omega \mid \; \Psi_{X,H}(\omega) > 0\}.$$

Let $\mathcal{S}$ be a class of measurable subsets of Ω such that $\varnothing \notin \mathcal{S}$.

Since, loosely speaking, an arbitrage opportunity is a a riskless portfolio which yields a positive profit in *some* circumstances (denoted by $\psi^+_{X,H}$), without a countervailing treat of loss in other circumstances. In order to formally describe this economic principle we need to specify the meaning of a “riskless portfolio”and that of a “true gain”. While, in a robust setup, the former can be naturally considered as a strategy whose returns are nonnegative in any state of the world (i.e., $\forall \omega \in \mathfrak{P}$), the meaning of a true gain can be more subjective and depend on agents' personal feelings and opinions. This is exactly the role that we attribute to $\mathcal{S}$.

Definition 2.2.5 (Burzoni et al. [20]). A trading strategy $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F})$ is an Arbitrage de la classe $\mathcal{S}$ if

- $\Psi_{X,H}(\omega) \geq 0 \ \forall \omega \in \Omega$ (i.e. never loses);
- $\psi^+_{X,H}$ contains a set in $\mathcal{S}$ (i.e. on a significant set).

2.2.2 Fundamental theorem of asset pricing without static hedging

Recall that $\mathbb{F}^{\mathbb{S}} := \{\mathcal{F}^{\mathbb{S}}_t\}_{t=0}^n$ is the natural filtration generated by $\mathbb{S}$.

Through the rest of the chapter, instead of considering general $\mathbb{F}$*, we will only focus on the case that* $\mathbb{F} = \mathbb{F}^{\mathbb{S}}$.

For ease of notation, without specifying filtration explicitly, objects which depend on the filtration are all with respect to $\mathbb{F}$, for example we write $\mathcal{M} = \mathcal{M}(\mathbb{F})$, $\mathcal{M}_{\mathfrak{P}} = \mathcal{M}_{\mathfrak{P}}(\mathbb{F})$ and $\mathcal{A} = \mathcal{A}(\mathbb{F})$.

We now present the connection between the absence of arbitrage de la classe $\mathcal{S}$ and the existence of a martingale measure with nice properties. It is clear from the definition that the absence of 1p-arbitrage is the strongest condition that one can assume in this robust framework. In fact, it has been shown in Proposition 17 of Burzoni et al. [20] that this no arbitrage condition not only implies there exists a martingale measure but also guarantee the existence of a martingale measure with full support. On the other hand we are interested in characterising those markets which can exhibit 1p-arbitrages but nevertheless admit a martingale measure.

In a one-period setting, 1p-arbitrage has been studied in Riedel [90] and connection between the absence of 1p-arbitrage and the existence of martingale measures has been established. In a multi-period setting, while it is not difficult to show that the absence of 1p-arbitrage implies $\mathcal{M} \neq \varnothing$, as we shall see from the following example, the converse is not true in general:

Example 2.2.6. Consider a one-period market with three assets and let $\mathbb{S}$ be the canonical process on Ω, where

$$\Omega = \{\omega \in (\mathbb{R}^3)^2 \ : \omega_0 = (1,0,0), \\ \omega_1^{(1)} \geq 0,\ \omega_1^{(2)} = \mathbb{1}_{\left\{\omega_1^{(1)} \geq 3\right\}},\ \omega_1^{(3)} = (\omega^{(1)} - 3)_+ - \mathbb{1}_{\left\{\omega^{(1)} \geq 2\right\}}\}.$$

It described the situation where there are three assets: The first one is a stock. The second one is a digital option which pays off when the stock exceeds 3 and the third one has the same payoff as a portfolio which consists of a long position in a call option with strike 3 and a short position in a digital option which pays off when the stock exceeds 2.

One immediately sees that in this example $\mathcal{M} \neq \varnothing$. For one who believes that $\mathbb{S}_1^{(1)} \geq 3$, Long one unit of the second asset (i.e., $H^1 = (0,1,0)$) is the arbitrage strategy that can realise a profit if their belief proves to be true. In fact $H^1 \cdot (\mathbb{S}_1 - \mathbb{S}_0) \geq 0$ on Ω, and therefore there exists a 1p-arbitrage for this market. It is also interesting to notice that for those who believe that $2 \leq \mathbb{S}_2^{(1)} < 3$, they just need to short one unit of the third asset and long any amount of the second asset (i.e., $H^2 = (0, a, -1)$). However, for whichever a they choose, $H^2 \cdot (\mathbb{S}_1 - \mathbb{S}_0)$ becomes negative when $\mathbb{S}_1^{(1)}$ is sufficiently large.

Therefore, with the canonical filtration $\mathbb{F}^{\mathbb{S}}$, H^1 and H^2 cannot be aggregated into one. This reflects on the disagreement of the arbitrage strategies among agents, which is similar in spirit to the example [4] studied in Davis and Hobson [31] that lead to what was called weak arbitrage.

These situations are managed in Burzoni et al. [20] with the enlargement of the original filtration. Generally it is shown in Burzoni et al. [20] that it is both necessary and convenient to adopt a filtration enlargement of $\mathbb{F}$, with respect to which the class of martingale measures remained"unchanged". Such filtration enlargement has properties defined below.

Definition 2.2.7. We say that a filtration $\widetilde{\mathbb{F}}$ is an arbitrage aggregating enlargement of $\mathbb{F}$, in which case we write $\mathcal{M}(\mathbb{F}) \leftrightarrows \mathcal{M}(\widetilde{\mathbb{F}})$, if $\mathbb{F} \subseteq \widetilde{\mathbb{F}}$ and the following are satisfied

- the restriction of any $\widetilde{\mathbb{Q}} \in \mathcal{M}(\widetilde{\mathbb{F}})$ to $\mathcal{F}_n$ belongs to $\mathcal{M}(\mathbb{F})$;

[4] The example there is in addition to a stock there are two call options with different strikes but traded at the same initial price which is strictly positive. In this example, robust strong arbitrage is shown to be impossible but model dependent arbitrage always exists.

- any $\mathbb{Q} \in \mathcal{M}(\mathbb{F})$ can be uniquely extended to an element of $\mathcal{M}(\widetilde{\mathbb{F}})$.

The way to progressively enlarge the filtration of $\mathbb{F}$ by adding a special process H was considered in Burzoni et al. [20], ie. the enlarged filtration $\widetilde{\mathbb{F}} = \{\widetilde{\mathcal{F}}_t\}_{t=0}^n$ is given by

$$\begin{aligned} \widetilde{\mathcal{F}}_t &:= \mathcal{F}_t \vee \sigma(H_0, \ldots, H_t),\ t \in \{0, \ldots, n-1\} \\ \widetilde{\mathcal{F}}_n &:= \mathcal{F}_n \vee \sigma(H_1, \ldots, H_{n-1}). \end{aligned} \tag{2.2.1}$$

The process H that was introduced and constructed there was called the Universal Arbitrage Aggregator, which is a process of multifunctions that exploit all the inefficiencies of the market with the following property: $(H \cdot \mathbb{S}) \geq 0$ with

$$\psi_H^+ = (\Omega^*)^{\complement},$$

where $(\Omega^*)^{\complement}$ is the largest $\mathcal{M}(\mathbb{F})$-polar set in the sense that every $\omega \in (\Omega^*)^{\complement}$ can not be charged by a finitely supported martingale measure, i.e.

$$(\Omega^*)^{\complement} := \{\omega^* \in \Omega\,:\,\mathbb{P}(\{\omega^*\}) = 0 \quad \forall \mathbb{P} \in \mathcal{M}^f\},$$

where $\mathcal{M}^f := \{\mathbb{P} \in \mathcal{M}\,:\,\text{supp}(\mathbb{P}) \text{ is finite}\}$. In particular, it has been shown in Burzoni et al. [20] that $\widetilde{\mathbb{F}}$ is an arbitrage aggregating filtration enlargement (see Definition 2.2.7). Moreover, it is clear that H is adapted to the enlarged filtration $\widetilde{\mathbb{F}}$. Finally, we have the framework for the formulation of the Fundamental Theorem of Asset Pricing in Burzoni et al. [20].

Theorem 2.2.8 (Theorem 1..2 in Burzoni et al. [20])**.** No Arbitrage de la classe $\mathcal{S}$ in $\mathcal{A}(\widetilde{\mathbb{F}}) \Leftrightarrow \mathcal{M}(\mathbb{F}) \neq \varnothing$ and $\mathcal{N}(\mathcal{M}(\mathbb{F}))$ does not contain sets of $\mathcal{S}$.

In this context an arbitrage opportunity is not anymore a strategy which is directly implementable but it is, more properly, a situation where different agents (represented by different probabilities $\mathbb{P}$) agree that there is an opportunity for a risk-free profit but, nevertheless, may disagree on how to achieve it.

Note that Theorem 2.2.8 shows that all the possible choices for $\mathcal{S}$ can be studied simultaneously and this is reflected in the properties of the class of $\mathcal{M}(\mathbb{F})$-polar sets $\mathcal{N}(\mathcal{M}(\mathbb{F}))$. The maximal $\mathcal{M}$-polar set have been studied in [20] and denoted by Ω^*. This set is very important for Superhedging Duality Theory that we present in the next section.

2.3 Robust FTAP and pricing–hedging duality without static hedging in the presence of prediction set

In the previous section, we discussed the use of an arbitrage aggregating filtration enlargement $\widetilde{\mathbb{F}}$ and the so-called Universal Arbitrage Aggregator which is a set-valued $\widetilde{\mathcal{F}}$-adapted stochastic process which allows to characterise those markets which can exhibit 1p-arbitrages but nevertheless admit martingale measures. In particular, for the extreme case of $\mathcal{S} = \{\Omega\}$, Theorem 2.2.8 asserts that the absence of model-independent arbitrage is equivalent to Ω not being an $\mathcal{M}$-polar set, which simply reduces to $\mathcal{M} \neq \varnothing$, i.e.,

$$\text{No } \mathcal{A}(\widetilde{\mathbb{F}})\text{-robust strong arbitrage} \Leftrightarrow \mathcal{M}(\mathbb{F}) \neq \varnothing.$$

Analogously, we might consider the cases of $\mathcal{S} = \{\mathfrak{P}\}$, where $\mathfrak{P}$ is a subset of Ω. The motivation of such consideration is to add agents' beliefs into the framework via $\mathfrak{P}$, which is a selection of paths that is deemed feasible paths. In this section, we are interested in generalising the aforementioned characterisation to $\mathfrak{P}$ and establish the connection between robust strong arbitrage on $\mathfrak{P}$ and the existence of martingale measures supported on $\mathfrak{P}$. Moreover, by giving a geometric decomposition of $\mathfrak{P}$ into $\mathfrak{P}^*$ and $(\mathfrak{P}^*)^{\complement}$ with $\mathfrak{P}^*$ being the maximal support of $\mathcal{M}_{\mathfrak{P}}(\mathbb{F})$, we provide a pricing–hedging duality with superhedging required on $\mathfrak{P}^*$.

In this section and the following, we will make use of $\mathcal{F}^{\mathcal{A}}$, the sigma algebra generated by analytic sets of $(\Omega, \mathcal{B}_\Omega)$ (see [36] Chapter III). Also define $\mathbb{F}^M := \{\mathcal{F}_t^M\}_{t\in I}$ as

$$\mathcal{F}_t^M := \bigcap_{P\in\mathcal{M}_{\Omega,\Phi}(\mathbb{F}^S)} \mathcal{F}_t^S \vee \mathcal{N}^P(\mathcal{F}_n^S) \ . \tag{2.3.1}$$

We also want to extend the definition of arbitrage aggregating filtration enlargement in Definition 2.2.7 to depend on input $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$.

Definition 2.3.1. We say that a filtration $\mathbb{G}$ is a $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$-arbitrage aggregating enlargement of $\mathbb{F}$, in which case we write $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{G}) \leftrightarrows \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})$, if $\mathbb{F} \subseteq \mathbb{G}$ and the following are satisfied

- the restriction of any $\widetilde{\mathbb{Q}} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{G})$ to $\mathcal{F}_n$ belongs to $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})$;
- any $\mathbb{Q} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})$ can be uniquely extended to an element of $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{G})$.

Remark 2.3.2. Consider two filtrations $\mathbb{F}^1$ and $\mathbb{F}^2$ satisfying $\mathbb{F}^{\mathbb{S}} \subseteq \mathbb{F}^1 \subseteq \mathbb{F}^2 \subseteq \mathbb{F}^M$. Then any $\mathbb{Q} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}^1)$ can be uniquely extended to a measure $\hat{\mathbb{Q}} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}^2)$; viceversa, for any $\hat{\mathbb{Q}} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}^2)$, the restriction $\hat{\mathbb{Q}}_{|\mathbb{F}^1}$ belongs to $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}^1)$. Since in this Chapter we only consider such filtrations, with a slight abuse of notation, we will write $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}^1) = \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}^2) = \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$, and analogously for $\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}$, and $\mathcal{M}_{\mathfrak{P}}$ and $\mathcal{M}^f_{\mathfrak{P}}$ (when there are no options to statically trade).

2.3.1 Efficient Belief $\mathfrak{P}^*$ and Universal Arbitrage Aggregator

The main results in Burzoni et al. [20] concerns the Polish space Ω rather than a subset of it, but nevertheless the main technical tools there are for a generic analytically measurable set $\mathfrak{P}$. Therefore, to decompose $\mathfrak{P}$ into $\mathfrak{P}^*$ and $(\mathfrak{P}^*)^{\complement}$ and construct a Universal Arbitrage Aggregator, it is recently shown that the same tools and procedures could be applied with only small modifications[5]. In the following, we shall give a review of the techniques used to construct $\mathfrak{P}^*$ and a Universal Arbitrage Aggregator which has positive gains on $(\mathfrak{P}^*)^{\complement}$.

For technical reasons, we will make use of the filtration $\mathbb{F}^{\mathcal{A}} := (\mathcal{F}^{\mathcal{A}}_t)_{t=0}^n$ where $\mathcal{F}^{\mathcal{A}}_t$ is the sigma algebra generated by the analytic sets of $(\Omega, \mathcal{F}^{\mathbb{S}}_t)$ (see Dellacherie and Meyer [36] Chapter III). Clearly, $\mathbb{F}^S \subset \mathbb{F}^{\mathcal{A}}$.

Lemma 2.3.3 (Lemma 3.1 in Burzoni et al. [17])**.** Fix any $t \in \{1, \dots, n\}$ and $\Gamma \in \mathcal{F}^{\mathcal{A}}$. There exist an index $\beta \in \{0, \dots, d\}$, random vectors $H^1, \dots, H^\beta \in \mathcal{L}(\Omega, \mathcal{F}^{\mathcal{A}}_{t-1}; \mathbb{R}^d)$, $\mathcal{F}^{\mathcal{A}}_t$-measurable sets $E^0, ..., E^\beta$ such that the sets $B^i := E^i \cap \Gamma$, $i = 0, \dots, \beta$, form a partition of Γ satisfying:

1. if $\beta > 0$ and $i = 1, \dots, \beta$ then: $B^i \neq \varnothing$; $H^i \cdot \Delta S_t(\omega) > 0$ for all $\omega \in B^i$ and $H^i \cdot \Delta S_t(\omega) \geq 0$ for all $\omega \in \cup_{j=i}^{\beta} B^j \cup B^0$.
2. $\forall H \in \mathcal{L}(\Omega, \mathcal{F}^{\mathcal{A}}_{t-1}; \mathbb{R}^d)$ s.t. $H \cdot \Delta S_t \geq 0$ on B^0 we have $H \cdot \Delta S_t = 0$ on B^0.

Remark 2.3.4. Clearly if $\beta = 0$ then $B^0 = \Gamma$ (which include the trivial case $\Gamma = \varnothing$). Notice also that for any $\Gamma \in \mathcal{F}^{\mathcal{A}}$ and $t = \{1, \dots, n\}$ we have that $H^i = H^{i,\Gamma}_t, B^i = B^{i,\Gamma}_t, \beta = \beta^\Gamma_t$ depend explicitly on t and Γ.

[5] The generalisation of the FTAP and pricing–hedging duality result from Burzoni et al. [20] Burzoni et al. [18] to include a prediction set is recently discussed in Burzoni et al. [19]. The results will be published as part of a joint work in Burzoni et al. [17].

Define

$$\mathfrak{P}_n := \quad \mathfrak{P}$$
$$\mathfrak{P}_{t-1} := \quad \mathfrak{P}_t \setminus \bigcup_{i=1}^{\beta_t} B_t^i, \quad t \in \{1, \dots, n\},$$

where $B_t^i := B_t^{i,\Gamma}$, $\beta_t := \beta_t^\Gamma$ are the sets and index constructed in Lemma 2.3.3 with $\Gamma = \mathfrak{P}_t$, for $1 \leq t \leq n$. Note that we can iteratively apply Lemma 2.3.3 at time $t-1$ since $\Gamma = \mathfrak{P}_t$ is $\mathcal{F}^{\mathcal{A}}$-measurable.

Corollary 2.3.5 (Corollary in Burzoni et al. [17])**.** For any $t \in \{1, \dots, n\}$, $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$ and $Q \in \mathcal{M}_{\mathfrak{P}}$ we have $\cup_{i=1}^{\beta_t} B_t^i$ is a subset of a Q-nullset. In particular $\cup_{i=1}^{\beta_t} B_t^i$ is an $\mathcal{M}_{\mathfrak{P}}$ polar set.

We set

$$\mathfrak{P}^* := \mathfrak{P}_0. \tag{2.3.2}$$

The proposition below provides a geometric decomposition of $\mathfrak{P}$ in two parts, $\mathfrak{P} = \mathfrak{P}^* \cup (\mathfrak{P}^*)^{\complement}$, which is an generalisation of $\Omega = \Omega^* \cup (\Omega^*)^{\complement}$. The set $\mathfrak{P}^*$ contains those events ω supported by martingale measures, namely, for any of those ω it is possible to construct a martingale measure (even with finite support) that assigns a positive probability to it. Therefore, it could be interpreted as the efficient version of $\mathfrak{P}$ and here we call it the efficient beliefs. Observe that such a decomposition is induced by S and it is determined prior to arbitrage considerations.

Proposition 2.3.6 (Lemma 3.8 in Burzoni et al. [17])**.** We have the following

$$\mathcal{M}_{\mathfrak{P}} \neq \varnothing \iff \mathfrak{P}^* \neq \varnothing \iff \mathcal{M}_{\mathfrak{P}} \cap \Pi^f =: \mathcal{M}_{\mathfrak{P}}^f \neq \varnothing.$$

If $\mathcal{M}_{\mathfrak{P}} \neq \varnothing$ then for any $\omega_* \in \mathfrak{P}^*$ there exists $\mathbb{Q} \in \mathcal{M}_{\mathfrak{P}}^f$ such that $\mathbb{Q}(\{\omega_*\}) > 0$.

A version of the Universal Arbitrage Aggregator can also be constructed from $B_t^{i,\mathfrak{P}_t}$ and $H_t^{i,\mathfrak{P}_t}$ provided by Remark 2.3.4.

Definition 2.3.7. We call (a selection of) the Universal Arbitrage Aggregator the process

$$H_{t-1}(\omega) := \sum_{i=1}^{\beta_t} H_t^{i,\mathfrak{P}_t}(\omega) \mathbf{1}_{B_t^{i,\mathfrak{P}_t}}(\omega) \tag{2.3.3}$$

for $t \in \{1, \dots, T\}$, where $H_t^{i,\mathfrak{P}_t}, B_t^{i,\mathfrak{P}_t}, \beta_t$ are provided by Remark 2.3.4.

Remark 2.3.8. Observe that from Lemma 2.3.3 item 1, $\Psi_H(\omega) \geq 0$ for all $\omega \in \mathfrak{P}$ and $\Psi_H(\omega) > 0$ for all $\omega \in (\mathfrak{P}^*)^{\complement}$. Also note that by construction we have that H_t is $\mathcal{F}^{\mathcal{A}}$-measurable for every $t \in \{0, \ldots, n-1\}$

The strategy $H := (H_t)_{t=0}^{n-1}$ is adapted to the enlarged filtration $\widetilde{\mathbb{F}} = \{\widetilde{\mathcal{F}}_t\}_{t=0}^{n}$ given by

$$\begin{aligned} \widetilde{\mathcal{F}}_t &:= \mathcal{F}_t^{\mathcal{A}} \vee \sigma(H_0, \ldots, H_t),\ t \in \{0, \ldots, n-1\}, \\ \widetilde{\mathcal{F}}_T &:= \mathcal{F}_n^{\mathcal{A}} \vee \sigma(H_0, \ldots, H_{n-1}). \end{aligned} \tag{2.3.4}$$

It is straight forward to see that it follows from Corollary 2.3.5 that $\widetilde{\mathbb{F}}$ is an arbitrage aggregating enlargement of $\mathbb{F}$. (See Definition 2.2.7 for the definition of arbitrage aggregating enlargement.)

We now provide the version of the fundamental theorem of asset pricing in this framework with no static hedging.

Theorem 2.3.9 (General FTAP)**.** Let $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$. Then, there exists a filtration $\widetilde{\mathbb{F}}$ such that $\mathbb{F}^{\mathbb{S}} \subseteq \widetilde{\mathbb{F}} \subseteq \mathbb{F}^{M}$ and

$$\text{No } \mathcal{A}(\widetilde{\mathbb{F}})\text{-robust strong arbitrage} \Leftrightarrow \mathcal{M}_{\mathfrak{P}} \neq \varnothing \Leftrightarrow \mathcal{M}_{\mathfrak{P}}^{f} \neq \varnothing.$$

Furthermore, for any $\omega \in \mathfrak{P}$, there exists $\mathbb{Q} \in \mathcal{M}_{\mathfrak{P}}^{f}$ such that $\mathbb{Q}(\{\omega\}) > 0$ if and only if $\omega \in \mathfrak{P}^*$. In particular,

$$\text{No 1p-Arbitrage on } \Omega \text{ in } \mathcal{A}(\mathbb{F}) \Leftrightarrow \mathfrak{P}^* = \mathfrak{P}.$$

2.3.2 Pricing–hedging duality without static hedging

Let g be a contingent claim, e.g., an Asian option or a lookback option. We say that a particular strategy H superreplicates g on $\mathfrak{P}$ with initial capital x if

$$x + \Psi_H(\omega) \geq g(\omega) \qquad \forall \omega \in \mathfrak{P}. \tag{2.3.5}$$

If superhedging is required for every $\omega \in \mathfrak{P}$, then there might be a duality gap for some measurable function g, meaning the infimum x over all $\mathfrak{P}$-superreplicating strategies is strictly larger than the supremum of expectation of g over all $\mathbb{Q} \in \mathcal{M}_{\mathfrak{P}}$. Examples could be found in section 4 of Burzoni et al. [18], where the authors studied model-independent pricing and hedging without beliefs (i.e., $\mathfrak{P} = \Omega$) and argued that to recover exact pricing–hedging duality for any measurable function g one needs to relax the requirement of (2.3.5) to hold on the smaller set $\Omega^* \subseteq \Omega$. To demonstrate the issue of superhedging everywhere, we give another simple example below:

Example 2.3.10. Consider a canonical setup. Let $\Omega = \{\omega \in \mathbb{R}^3_+ : \omega_0 = 1\}$ be the canonical space, $\mathbb{S} = (\mathbb{S}_0, \mathbb{S}_1, \mathbb{S}_2)$ be the canonical map and $\mathbb{F}$ be the canonical filtration. Suppose $g : \Omega \to \mathbb{R}$ is a contingent claim defined by

$$g(\omega) = \mathbb{1}_A(\omega),$$

where $A = \{\omega_1 = 0 \text{ and } \omega_2 > 0\}$.

Note that A is an $\mathcal{M}$-polar set. In fact, $(\Omega^*)^{\complement} = A$. Hence $P_\Omega = \sup_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_{\mathbb{Q}}[g] = 0$. However, to be able to superreplicate g, we need $H_0 \geq 0$ otherwise the loss can be unbounded since the path space is unbounded from above. Also H_1 is a constant on the $\mathcal{F}_1$-atom $\{\omega \in \Omega : \omega = 0\}$. Hence no matter how large H_1 is, by choosing an $\omega \in A$ such that $\omega_2 > 0$ is small enough, we show that it is necessary to require an initial capital equal to 1. Therefore, we conclude that $V_\Omega(g) > P_\Omega(g) = 0$.

The causes of the duality gap are not a single factor but multiple. On one hand, superhedging is required everywhere while we know the option pays off only on impossible events (i.e., $\omega \in (\Omega^*)^{\complement}$). On the other hand, the payoff function g here has no desirable continuity property around the boundary between Ω and Ω^*, which makes the cost to hedge those impossible events considerable. We also notice that the filtration can also play a part here. For example, the duality gap could disappear if instead we use a technical filtration even larger than the arbitrage aggregating enlargement of $\mathbb{F}$ given by (2.2.1). For example we can build an H by hand: $H_0 = 0$ and

$$H_1(\omega) = \sum_{n\in\mathbb{Z}} 2^{-n} \mathbb{1}_{A_n}(\omega),$$

where $A_n = \{\omega \in \Omega : \omega_1 = 0,\ \omega_2 \in (2^n, 2^{n+1}]\}$. Note that $\Psi_H(\omega) \geq g(\omega)$. In a non-canonical setup, there are even more factors that potentially affect the appearance of duality gap. But nevertheless, a version of pricing–hedging duality result also requires a delicate tradeoff among many factors including those we already mentioned above.

In Burzoni et al. [18], a version of the pricing–hedging duality theorem is established where the authors relax the superhedging requirement to be on Ω_* instead of Ω and consider the universal filtration. In this formulation, they demonstrate that there is always a pricing–hedging duality for any measurable claim and the value of the superhedging problem could be attained by some strategy.

The same procedure could be applied to $\mathfrak{P}$ and a generalised pricing–hedging duality theorem was then shown in Burzoni et al. [17].

Theorem 2.3.11. Let $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$ be an analytic subset of $(\Omega, \mathcal{B}_\Omega)$. We have that $\mathfrak{P}^*$, defined in (2.3.2) belongs to $\mathcal{F}^{\mathcal{A}}$, and for any $g \in \mathcal{L}(\Omega, \mathcal{F}^{\mathcal{A}}; \mathbb{R})$

$$V_{\mathfrak{P}^*}(g) = P_{\mathfrak{P}}(g) = \sup_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g]. \tag{2.3.6}$$

with $V_{\mathfrak{P}^*}(g) = \inf \left\{ x \in \mathbb{R} \mid \exists H \in \mathcal{A}(\mathbb{F}^{\mathcal{A}}) \text{ such that } x + (H \circ \mathbb{S})_n(\omega) \geq g(\omega) \ \forall \omega \in \mathfrak{P}^* \right\}$. In particular, the LHS is attained by some strategy $H \in \mathcal{A}(\mathbb{F}^{\mathcal{A}})$.

Remark 2.3.12. Another by product of the result is a dynamic programming principle for the pricing and hedging problems.

2.4 Robust FTAP and pricing–hedging duality with semi-static strategies for finitely many options

In this section[6], we present our contribution to the literature, that is to extend the results in the last section to allow for static trading in a finite number of options. We fix an arbitrary analytic set of paths $\mathfrak{P} \subseteq \Omega$, where Ω is a Polish space. Note that $\mathfrak{P}^*$ defined in (2.3.2) is analytic (by Theorem 2.3.11). We now add to the previous market k options $\mathcal{X} = (X_1, ..., X_k)$ and assume without loss of generality that they have zero initial cost, i.e., $\mathcal{P}(X_i) = 0$ for any $i \leq k$. We also assume that each X_j is $\mathcal{F}^{\mathcal{A}}$-measurable, superreplicable and subreplicable on $\mathfrak{P}^*$, i.e., $V_{\mathfrak{P}^*}(X_j) < \infty$ and $V_{\mathfrak{P}^*}(-X_j) < \infty$. Define $h \cdot \mathcal{X} := \sum_{j=1}^k h_j X_j$, $h \in \mathbb{R}^k$. Recall from Definition 2.1.2 that

$$\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}) := \{ \mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}(\mathbb{F}) \mid \mathbb{E}_{\mathbb{Q}}[X_j] = 0 \ \forall j = 1, ..., k \},$$

which are the market calibrated models with finite support. Define

$$\mathfrak{P}^*_{\mathcal{X}} := \left\{ \omega \in \mathfrak{P} \mid \exists \mathbb{Q} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F}) \text{ s.t. } \mathbb{Q}(\omega) > 0 \right\} \subseteq \mathfrak{P}^*. \tag{2.4.1}$$

We have by definition that for every $\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\mathbb{F})$ the support satisfies $supp(\mathbb{Q}) \subseteq \mathfrak{P}_{\mathcal{X}}$. Recall from Definition 2.1.1 that $V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$, the superhedging price of g on $\mathfrak{P}^*_{\mathcal{X}}$, is defined by

$$\inf \left\{ \mathcal{P}(X) \mid \exists (X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F}^{\mathcal{A}}) \text{ s.t. } \Psi_{X,H}(\omega) \geq g(\omega) \ \forall \omega \in \mathfrak{P}^*_{\mathcal{X}} \right\},$$

[6] The original research presented in Sections 2.4, 2.5 and 2.6 will be published as part of a joint work in Burzoni et al. [17].

which is also equivalent to

$$\inf\{x \in \mathbb{R} \mid \exists (h, H) \in \mathbb{R}^k \times \mathcal{A}(\mathbb{F}^{\mathcal{A}}) \text{ s.t. } x + (H \circ \mathbb{S})_n(\omega) + h \cdot \mathcal{X}(\omega) \geq g(\omega) \, \forall \omega \in \mathfrak{P}^*_{\mathcal{X}}\}. \tag{2.4.2}$$

We first present the robust FTAP on $\mathfrak{P} \subseteq \Omega$ with semi-static strategies. The technical details needed to establish it will be deferred to Section 2.5.

Theorem 2.4.1 (Robust FTAP on $\mathfrak{P} \subseteq \Omega$, Theorem 2.3 in Burzoni et al. [17]). Fix $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$ and $\mathcal{X}$ a finite set of statically traded options. Then, there exists a filtration $\widetilde{\mathbb{F}}$ such that $\mathbb{F}^{\mathbb{S}} \subseteq \widetilde{\mathbb{F}} \subseteq \mathbb{F}^M$ and

$$\text{No } \mathcal{A}_{\mathcal{X}}\text{-robust strong arbitrage on } \mathfrak{P} \iff \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \emptyset \iff \mathfrak{P}^*_{\mathcal{X}} \neq \emptyset.$$

Further, $\mathfrak{P}^*_{\mathcal{X}} \in \mathcal{F}^{\mathcal{A}}$ and there exists an Arbitrage Aggregator $(\alpha^*, H^*) \in A_{\mathcal{X}}(\widetilde{\mathbb{F}})$ such that $\alpha^* \cdot \mathcal{X} + (H^* \circ S)_T \geq 0$ on $\mathfrak{P}$.

Proof. It follows from Theorem 2.5.9 and Remark 2.5.8 below. □

Theorem 2.4.2. Let $\mathfrak{P} \subseteq \Omega$, $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$ and $\mathcal{X}$ be the vector of statically traded options. Suppose $\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ and

$$\max\{V_{\mathfrak{P}^*_{\mathcal{X}}}(X_j), V_{\mathfrak{P}^*_{\mathcal{X}}}(-X_j)\} < \infty, \quad \forall j \leq k. \tag{2.4.3}$$

Then, for any $\mathcal{F}^{\mathcal{A}}$-measurable g,

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(g) = \sup_{\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g] \tag{2.4.4}$$

and, if finite, the left hand side is attained by some strategy $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F}^{\mathcal{A}})$.

Moreover, if g satisfies $V_{\Omega^*_{\Phi}}(g) < \infty$ then the result holds also without assuming (2.4.3).

Proof. See Section 2.7.1. □

Remark 2.4.3. In fact, when (2.4.3) is not satisfied, if g is $\mathcal{F}^{\mathcal{A}}$–measurable and satisfies that $V_{\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$, then the pricing–hedging duality in (2.3.6) would still hold, and moreover the superhedging cost $V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ is attained by some strategy $(X, H) \in \mathcal{A}_{\mathcal{X}}(\mathbb{F}^{\mathcal{A}})$.

Remark 2.4.4 (On the structure of the proofs and analyticity of $\mathfrak{P}^*_{\mathcal{X}}$)**.** We provide a detailed description of the logical flow of the proofs in Section 2.7. We point out only that showing $\mathfrak{P}^*_{\mathcal{X}} \in \mathcal{F}^{\mathcal{A}}$ is involved. The case without options follows from the construction and is given in Lemma 2.3.6. However, in the case with options, we first establish Theorem 2.4.2 assuming this property which allows us then to deduce Proposition 2.5.1 in which $\mathfrak{P}^*_{\mathcal{X}} = \mathfrak{P}$ which belongs by assumption to $\mathcal{F}^{\mathcal{A}}$. Only this in turn allows us to establish in Lemma 2.5.5 that in all cases $\mathfrak{P}^*_{\mathcal{X}} \in \mathcal{F}^{\mathcal{A}}$. This then completes the proof of Theorem 2.4.2 in the general setting.

2.5 FTAP with semi-static strategies for finitely many options

In this section, we focus on the case where

$$\max\{V_{\mathfrak{P}^*}(X_j), V_{\mathfrak{P}^*}(-X_j)\} < 0, \quad \forall j \leq k.$$

Proposition 2.5.1. Suppose $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$. Then there is no 1p-arbitrage on $\mathfrak{P}$ with respect to $\mathbb{F}^{\mathcal{A}}$ if and only if $\mathfrak{P} = \mathfrak{P}^*_{\mathcal{X}}$.

Proposition 2.5.2. Suppose $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$ and $\mathcal{X}$ is not perfectly replicable on $\mathfrak{P}$. Then absence of 1p-arbitrage on $\mathfrak{P}$ with respect to $\mathbb{F}^{\mathcal{A}}$ is equivalent to the existence of an $\epsilon > 0$ such that $\mathcal{M}^f_{\mathcal{X},\mathcal{P}+\vec{x},\mathfrak{P}} \neq \varnothing$ for any $\vec{x} \in B_\epsilon(0)$, where $B_\epsilon(0)$ is a uniform ball around 0 with radius ϵ and $\mathcal{P} + \vec{x}$ denotes the pricing system on $\mathrm{Lin}(\mathcal{X})$ such that $\mathcal{P}(X_i) = x_i \ \forall i \leq k$.

Proof of Propositions 2.5.1 and 2.5.2.

We first prove Proposition 2.5.1.

"$\Leftarrow$" is clear since, by definition in (2.4.1), we can take a martingale measure which assigns a positive probability to ω on which the 1p-arbitrage has strictly positive payoff leading to a contradiction. We prove "$\Rightarrow$" by iteration on number of options used for static trading. No 1p-arbitrage using dynamic trading and $\mathcal{X}$ in particular means that there is no 1p-arbitrage using only dynamic trading which, from Theorem 2.3.9 implies that $\mathfrak{P} = \mathfrak{P}^*$ and hence for any $\omega \in \mathfrak{P}$ there exists $\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}$ such that $\mathbb{Q}(\{\omega\}) > 0$. Note also that if, for some $j \leq k$, X_j is replicable on $\mathfrak{P}^*$ by dynamic trading in S then $\mathbb{E}_{\mathbb{Q}}[X_j] = 0$ for every $\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}$. With no loss of generality we assume that $(X_1,\ldots,X_{k_1})$ is a vector of non-replicable options on $\mathfrak{P}^*$ with $k_1 \le k$. We now apply Theorem 2.3.11 to X_1 and argue that

$$m_1 := \min\{V_{\mathfrak{P}^*}(X_1), V_{\mathfrak{P}^*}(-X_1)\} > 0.$$

Indeed, if $m_1 < 0$ then we would have a robust strong arbitrage and if $m_1 = 0$, since superhedging price is attained, in order to avoid 1p-arbitrage we have to have $X_1 = (H \circ S)_n$ on $\mathfrak{P}$ for some $H \in \mathcal{A}(\mathbb{F}^{\mathcal{A}})$, which is a contradiction since X_1 is not replicable. This shows that $m_1 > 0$ which in turn implies there exist $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}^f_{\mathfrak{P}}$ such that $\mathbb{E}_{\mathbb{Q}_1}[X_1] > 0$ and $\mathbb{E}_{\mathbb{Q}_2}[X_1] < 0$. Then, for any $\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}$, there exist $\alpha, \beta, \gamma \in [0,1]$, $\alpha+\beta+\gamma = 1$ and $\mathbb{E}_{\alpha\mathbb{Q}_1+\beta\mathbb{Q}_2+\gamma\mathbb{Q}}[X_1] = 0$. Thus, for any $\omega \in \mathfrak{P}^*$ there exists $\mathbb{Q} \in \mathcal{M}^f_{\{X_1\},\mathcal{P},\mathfrak{P}}$ such that $\mathbb{Q}(\{\omega\}) > 0$. In particular, $\mathfrak{P}^*_{\{X_1\}} = \mathfrak{P}$ and we may apply Theorem 2.4.2 to $\mathfrak{P}$ and X_1. Define now

$$m_{1,j} := \min\{V_{\{X_1\},\mathcal{P},\mathfrak{P}^*}(X_j), V_{\{X_1\},\mathcal{P},\mathfrak{P}^*}(-X_j)\} \quad \forall j = 2,\ldots,k_1.$$

By absence of robust strong arbitrage we necessarily have $m_{1,j} \ge 0$ for every $j = 2,\ldots,k_1$. Let $j \in I_2 = \{j = 2,\ldots,k_1 \mid m_{1,j} = 0\}$, by no 1p-arbitrage, we have perfect replication of X_j using semi-static strategies with X_1 on $\mathfrak{P}$ and in consequence for any $\mathbb{Q} \in \mathcal{M}^f_{\{X_1\},\mathcal{P},\mathfrak{P}}$ we have $\mathbb{E}_{\mathbb{Q}}[X_j] = 0$ for all $j \in I_2$. We may discard these options and, up to re-numbering, assume that $(X_2,\ldots,X_{k_2})$ is a vector of the remaining options, non-replicable on $\mathfrak{P}$ with semi-static trading in X_1, with $k_2 \le k_1$. If $k_2 \ge 2$, $m_{1,2} > 0$ by Theorem 2.4.2 and absence of 1p-arbitrage using arguments as above. Hence, there exist $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}^f_{\{X_1\},\mathcal{P},\mathfrak{P}}$ such that $\mathbb{E}_{\mathbb{Q}_1}[X_2] > 0$ and $\mathbb{E}_{\mathbb{Q}_2}[X_2] < 0$. As above, this implies that $\mathfrak{P}^*_{\{X_1,X_2\}} = \mathfrak{P}^*_{\{X_1\}} = \mathfrak{P}$. We can iterate the above arguments and the procedure ends after at most k steps showing $\mathfrak{P}^*_{\mathcal{X}} = \mathfrak{P}$ as required.

We now prove Proposition 2.5.2.

Write $\mathcal{X}_{-i} = (X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_k)$. We now suppose there is no 1p-arbitrage on $\mathfrak{P}$ with respect to $\mathbb{F}^{\mathcal{A}}$. In this case, we know from above that $\mathfrak{P} = \mathfrak{P}^* = \mathfrak{P}^*_{\mathcal{X}}$. Indeed, when $\mathcal{X}$ is not perfectly replicable on $\mathfrak{P}$, we have

$$\epsilon_j := \min\{V_{\mathcal{X}_{-j},\mathcal{P},\mathfrak{P}^*}(X_j), V_{\mathcal{X}_{-j},\mathcal{P},\mathfrak{P}^*}(-X_j)\} > 0 \quad \forall j = 1,\ldots,k,$$

which by Theorem 2.4.2 implies for any $j = 1,\ldots,k$, there exist $\mathbb{Q}^{(j)}_+$ and $\mathbb{Q}^{(j)}_- \in \mathcal{M}^f_{\mathcal{X}_{-j},\mathcal{P},\mathfrak{P}^*}$ such that $\mathbb{E}_{\mathbb{Q}^{(j)}_+}[X_j] = \epsilon/2$ and $\mathbb{E}_{\mathbb{Q}^{(j)}_-}[X_j] = -\epsilon/2$. Take $\epsilon = \min_{1\le i\le k}\{\epsilon_j\}/2n$ and fix any $\vec{x} \in B_\epsilon(0)$. Let $\mathbb{Q}^{(j)} = \mathbb{Q}^{(j)}_+$ if $x_j \ge 0$ and $\mathbb{Q}^{(j)}_-$ otherwise. It is clear that

$$\sum_{j=1}^n |x_j|\mathbb{Q}^{(j)} + (1 - \sum_{j=1}^n |x_j|)\mathbb{Q}^{(0)} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P}+\vec{x},\mathfrak{P}},$$

where $\mathbb{Q}^{(0)}$ is any element in $\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}$. Therefore $\mathcal{M}^f_{\mathcal{X},\mathcal{P}+\vec{x},\mathfrak{P}} \neq \varnothing$ for any $\vec{x} \in B_\epsilon(0)$. The other direction is obvious. Therefore we have the equivalence as required.

□

Definition 2.5.3. A pathspace partition scheme $\mathcal{R}(\alpha^\bullet, H^\bullet)$ of $\mathfrak{P}$ is a collection of trading strategies $H^1, \ldots, H^\beta \in \mathcal{A}(\mathbb{F}^\mathcal{A})$, Universal Arbitrage Aggregator $\tilde{H}^0, \ldots, \tilde{H}^\beta$, $\alpha^1, \ldots, \alpha^\beta \in \mathbb{R}^k$, for some $\beta = 0, \ldots, k$, such that

(i) α^i, $1 \leq i \leq \beta$, are linearly independent,

(ii) for any $i \leq \beta$,

$$(H_i \circ \mathbb{S})_n + \alpha_i \cdot \mathcal{X} \geq 0 \text{ on } A^*_{i-1},$$

where $A_0 = \mathfrak{P}$, $A_i := \{(H^i \circ \mathbb{S})_n + \alpha^i \cdot \mathcal{X} = 0\} \cap A_{i-1}$ and A^*_i is the set $\mathfrak{P}^*$ in (2.3.2) with $\mathfrak{P} = A_i$ for $1 \leq i \leq \beta$,

(iii) for any $i = 0, \ldots, \beta$, $\tilde{H}^i$ is the Universal Arbitrage Aggregator, as defined in (2.3.3) substituting $\mathfrak{P}$ with A_i,

(iv) if $\beta < k$, then either $A^*_\beta = \varnothing$, or for any $\alpha \in \mathbb{R}^k$ orthonormal to the plane spanned by $\alpha^1, \ldots, \alpha^\beta$, there do not exist H such that

$$(H \circ \mathbb{S})_n + \alpha \cdot \mathcal{X} \geq 0 \text{ on } A^*_\beta.$$

We note that as defined in *(ii)* above, each $A_i \in \mathcal{F}^\mathcal{A}$ so that A^*_i in *(iii)* is well defined. The purpose of a pathspace partition scheme is to iteratively split subsets of pathspace on which a robust strong arbitarge strategy can be identified. For existence of a calibrated martingale measure it will be crucial to see whether this procedure exhausts the pathspace or not.

Definition 2.5.4. A pathspace partition scheme $\mathcal{R}(\alpha^\bullet, H^\bullet)$ is successful if $A^*_\beta \neq \varnothing$.

Note that if a partition scheme is successful then there are no 1p-arbitrage on A^*_β. When $\beta < k$ this follows from *(iv)* in Definition 2.5.3 and in the case of $\beta = k$ this follows from Theorem 2.4.1.

Lemma 2.5.5. For any $\mathcal{R}(\alpha^\bullet, H^\bullet)$, $A^*_i = \mathfrak{P}^*_{\{\alpha^j \cdot \mathcal{X} \,:\, j \leq i\}}$ for any $i \leq \beta$. Moreover, if $\mathcal{R}(\alpha^\bullet, H^\bullet)$ is successful, then $A^*_\beta = \mathfrak{P}^*_\mathcal{X}$.

Proof. If $\mathfrak{P}^* = \varnothing$, then the claim holds trivially. We now assume $\mathfrak{P}^* \neq \varnothing$, fix $\mathcal{R}(\alpha^\bullet, H^\bullet)$ and prove the claim by induction on i. For simplicity of notation, let $\mathfrak{P}^*_i := \mathfrak{P}^*_{\{\alpha^j\cdot\mathcal{X}\,:\, j\leq i\}}$. We have $A^*_0 = \mathfrak{P}^*_0 = \mathfrak{P}^*$ by definition. Suppose now $A^*_{i-1} = \mathfrak{P}^*_{i-1}$ for some $i \leq \beta$. Then, by definition, $\mathfrak{P}^*_i \subset \mathfrak{P}^*_{i-1} = A^*_{i-1}$. Further, since $(H_i \circ S)_n + \alpha^i \cdot \mathcal{X} \geq 0$ on A^*_{i-1} with strict inequality outside of A_i, it follows that $\mathfrak{P}^*_i \subset A_i$. Finally, from the properties of Universal Arbitrage Aggregator, we also have $\mathfrak{P}^*_i \subset A^*_i$. For the reverse inclusion consider $\omega \in A^*_i \setminus \mathfrak{P}^*_i$. By definition of A^*_i and Theorem 2.3.9 , there exists $\mathbb{Q} \in \mathcal{M}^f_{A^*_i}$ with $\mathbb{Q}(\{\omega\}) > 0$. Since on A^*_i, all options $\alpha^j \cdot \Phi$, $1 \leq j \leq i$, are perfectly replicated, it follows that $\mathbb{Q} \in \mathcal{M}^f_{\{\alpha^j\cdot\mathcal{X}\,:\,j\leq i\},\mathcal{P},\mathfrak{P}}$ so that $\omega \in \mathfrak{P}^*_i$.

Suppose now $\mathcal{R}(\alpha^\bullet, H^\bullet)$ is successful. In the case $\beta = k$, $\mathfrak{P}^*_{\mathcal{X}} = A^*_\beta$ by the above. Suppose $\beta < k$ so that the above shows $\mathfrak{P}^*_{\mathcal{X}} \subset \mathfrak{P}^*_\beta = A^*_\beta$. Observe that $A^*_\beta \in \mathcal{F}^{\mathcal{A}}$ by construction and Theorem 2.3.11, and there is no 1p-arbitrage on A^*_β. Then, by Proposition 2.5.1, we have $\mathfrak{P}^*_{\mathcal{X}} = A^*_\beta$ as required. □

Remark 2.5.6. The above lemma implies relative uniqueness of $\mathcal{R}(\alpha^\bullet, H^\bullet)$ in the sense that either every $\mathcal{R}(\alpha^\bullet, H^\bullet)$ is not successful or all $\mathcal{R}(\alpha^\bullet, H^\bullet)$ are successful and then $A^*_\beta = \mathfrak{P}^*_{\mathcal{X}}$. It follows also from Theorem 2.3.11 that $A^*_i \equiv \mathfrak{P}^*_{\{\alpha^j\cdot\Phi\,:\,j\leq i\}}$ belongs to $\mathcal{F}^{\mathcal{A}}$ for any $i \leq \beta$. In particular, if $\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ then $\mathfrak{P}^*_{\mathcal{X}}$ in (2.4.1) is in $\mathcal{F}^{\mathcal{A}}$ and assumptions of Theorem 2.4.2 are satisfied.

Definition 2.5.7. Given a pathspace partition scheme we define the Universal Arbitrage Aggregator as

$$(\alpha^\bullet, H^\bullet) = \left(\sum_{i=1}^{\beta}(\alpha^i \cdot \mathcal{X})\mathbf{1}_{A^*_i}, \sum_{i=1}^{\beta} H^i \mathbf{1}_{A^*_i} + \sum_{0=1}^{\beta} \tilde{H}^i \mathbf{1}_{A_i \setminus A^*_i}\right),$$

with $(\alpha^\bullet, H^\bullet) = (0, \tilde{H}^0 \mathbf{1}_{\mathfrak{P}\setminus\mathfrak{P}^*})$ if $\beta = 0$. To make the above Universal Arbitrage Aggregator adapted we need to enlarge the filtration. We therefore introduce the arbitrage aggregating filtration $\widetilde{\mathbb{F}}$ given by

$$\begin{aligned} \widetilde{\mathcal{F}}_t &= \mathcal{F}^{\mathcal{A}}_t \vee \{A_0, A^*_0, \ldots, A_\beta, A^*_\beta\} \vee \sigma(\tilde{H}^0_0, \ldots, \tilde{H}^\beta_0, \ldots, \tilde{H}^0_t, \ldots, \tilde{H}^\beta_t) \quad t < n, \\ \widetilde{\mathcal{F}}_n &= \mathcal{F}^{\mathcal{A}}_n \vee \{A_0, A^*_0, \ldots, A_\beta, A^*_\beta\} \vee \sigma(\tilde{H}^0_0, \ldots, \tilde{H}^\beta_0, \ldots, \tilde{H}^0_{n-1}, \ldots, \tilde{H}^\beta_{n-1}). \end{aligned} \tag{2.5.1}$$

Remark 2.5.8. It follows trivially from Lemma 2.5.5 that $\widetilde{\mathbb{F}}$ defined in (2.5.1) is a $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$-arbitrage aggregating filtration enlargement of $\mathbb{F}$.

Theorem 2.5.9 (Robust FTAP on $\mathfrak{P} \subseteq \Omega$)**.** Let $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$, $\mathcal{R}(\alpha^\bullet, H^\bullet)$ be a pathspace partition scheme and $\widetilde{\mathbb{F}}$ be given by (2.5.1). Then the following are equivalent:

1. $\mathcal{R}(\alpha^\bullet, H^\bullet)$ is successful.

2. $\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$.

3. $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$.

4. No $\mathcal{A}_{\mathcal{X}}(\widetilde{\mathbb{F}})$-robust strong arbitrage on $\mathfrak{P}$.

If the above hold then in particular the assumptions of Theorem 2.4.2 are satisfied and the pricing–hedging duality (2.3.6) holds.

Proof. By remark 2.5.6, $A^*_\beta = \mathfrak{P}^*_{\mathcal{X}}$ and hence (1) $\Leftrightarrow$ (2) follows by definition of Ω^*_Φ in (2.4.1). (2) $\Rightarrow$ (3) is obvious. To show (3) $\Rightarrow$ (4), observe that under $\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}$, expectation of $\Psi_{X,H}$ for any semi-static trading strategy $(X, H) \in \mathcal{A}_{\mathcal{X}}(\widetilde{\mathbb{F}})$ is zero which excludes the possibility of existence of a robust strong arbitrage. Finally, for (4) $\Rightarrow$ (1), note that since, for any $1 \leq i \leq \beta$, $(H^i \circ \mathbb{S})_n + \alpha^i \cdot \mathcal{X} > 0$ on $A_{i-1} \setminus A_i$ the selection of the Universal Arbitrage Aggregator, defined in Definition 2.3.7, (cf. Remark 2.3.8,) satisfies

$$\sum_{i=1}^{\beta} (H^i \circ \mathbb{S})_n + \alpha^i \cdot \mathcal{X} + (\tilde{H} \circ \mathbb{S})_n > 0 \quad \text{on } (A^*_\beta)^{\complement}.$$

The hypothesis (4) implies therefore that A^*_β is non-empty and hence the pathspace partition scheme is successful. We note that by Lemma 2.5.5, $\mathfrak{P}^*_{\mathcal{X}} = A^*_\beta \in \mathcal{F}^{\mathcal{A}}$ and the assumptions of Theorem 2.4.2 are satisfied.

□

2.6 Superhedging and FTAP for arbitrarily many options

In this section, we want to recover and extend the main result in Acciaio et al. [1] and possibly extend it to a non-canonical setting. Fix an arbitrary set of paths $\mathfrak{P} \subseteq \Omega$, $\mathfrak{P} \in \mathcal{F}^A$, where Ω is the enveloping Polish space. Recall that, by Theorem 2.3.11, $\mathfrak{P}^*$ defined in (2.3.2) belongs to $\mathcal{F}^{\mathcal{A}}$ and we assume it is non trivial. Here we allow for the possibility of static trading in options whose payoffs X_i, $i \in I$, are $\mathcal{F}^{\mathcal{A}}$-measurable functions, where I here is an index set (which can be uncountable). These options are traded at time $t = 0$ and, since the payoffs can be shifted by a constant, without loss of generality, we may assume that all these options have zero initial prices.

Given any $\widetilde{\mathcal{X}} \subseteq \mathcal{X} := \{X_i\}_{i\in I}$ and a σ-algebra $\mathcal{G}$, we let $\mathrm{Lin}(\widetilde{\mathcal{X}})(\mathcal{G})$ denote the set of finite linear combinations of elements of $\widetilde{\mathcal{X}}$, i.e.

$$\mathrm{Lin}(\widetilde{\mathcal{X}})(\mathcal{G}) = \Big\{ \sum_{i=1}^{m} a_i X_i \, : \, m \in \mathbb{N}, \;\; a_i \text{ is } \mathcal{G}\text{-measurable}, X_i \in \widetilde{\mathcal{X}} \Big\}.$$

Therefore, an admissible (semi-static) trading strategy is a pair (X, H) where

$$X \in \mathrm{Lin}(\Phi)(\mathcal{F}_0^{\mathcal{A}}) \;\; \text{and} \;\; H \in \mathcal{A}(\mathbb{F}^{\mathcal{A}}).$$

The total payoff associated to (X, H) is given by $X + (H \circ \mathbb{S})_n$, and the cost of following such a trading strategy is equal to the cost of setting up its static part, which is zero. We denote the class of admissible (semi-static) trading strategies by $\mathcal{A}_{\mathcal{X}}(\mathcal{F}^{\mathcal{A}})$.

Assumption 2.6.1. In this section, we assume all the options X_i are continuous derivatives on the underlying assets S, more precisely

$$X_i = g_i \circ \mathbb{S} \quad \text{for some continuous } g_i : \mathbb{R}_+^{d\times(n+1)} \to \mathbb{R}, \quad i \in I.$$

In addition, we assume $0 \in I$ and $X_0 = g_0(\mathbb{S}_n)$ for a convex super-linear function g_0 on $\mathbb{R}$, such that the other options have a slower growth at infinity:

$$\lim_{|x|\to\infty} \frac{g_i(x)}{m(x)} = 0, \quad \forall i \in I/\{0\}, \quad \text{where } m(x_0, \ldots, x_n) = \sum_{t=0}^{n} g_0(x_t).$$

The presence of X_0 has the effect of restricting non-trivial considerations to a compact set of values for S and then then continuity of X_i allows to aggregate different arbitrages without enlarging the filtration. This results in the following special case of the Robust Fundamental Theorem of Asset Pricing. Denote by $\widetilde{\mathcal{M}}_{\mathcal{X},\mathcal{P},\mathfrak{P}} := \{\mathbb{Q} \in \overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mid \mathbb{E}_{\mathbb{Q}}[X_i] = 0 \text{ for } i \in I \setminus \{0\}\}$, where we write $\mathcal{X}_0 = \{X_0\}$ and the closure is in the weak sense. Consider $\widetilde{\mathbb{F}} = \{\widetilde{\mathcal{F}}_t\}_{t\in I}$, the enlarged filtration given in (2.3.4). We have the following series of results concerning Robust Fundamental Theorem of Asset Pricing. First, we look at the case that only the option with the super-linear payoff is traded for static hedging.

Proposition 2.6.2. Consider $\mathfrak{P} \in \mathcal{F}^{\mathcal{A}}$ is such that $\mathfrak{P} = \mathfrak{P}^*$, $V_{\mathfrak{P}^*}(X_0) > 0$ and there exists an $\omega^* \in \mathfrak{P}$ such that $\mathbb{S}_0(\omega^*) = \mathbb{S}_1(\omega^*) = \ldots = \mathbb{S}_n(\omega^*)$. Then the following are equivalent:

(1) There is no robust uniformly strong arbitrage on $\mathfrak{P}$ in $\mathcal{A}_{\mathcal{X}_0}(\widetilde{\mathbb{F}})$;

(2) There is no robust strong arbitrage on $\mathfrak{P}$ in $\mathcal{A}_{\mathcal{X}_0}(\widetilde{\mathbb{F}})$;

(3) $\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}} \neq \varnothing$.

where we write $\mathcal{X}_0 = \{X_0\}$. Moreover, when any of these holds, for any upper semi-continuous $g : \mathbb{R}_+^{d\times(n+1)} \to \mathbb{R}$ that satisfies

$$\lim_{|x|\to\infty} \frac{g_+(x)}{m(x)} = 0, \tag{2.6.1}$$

where $m(x_0, ..., x_n) := \sum_{t=0}^{n} g_0(x_t)$, the following pricing–hedging duality holds:

$$V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S})) = \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})] = \sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})]. \tag{2.6.2}$$

We now turn to the general case.

Theorem 2.6.3. Under Assumption 2.6.1 and conditions of Proposition 2.6.2, if in addition we assume that the set of laws of $\mathbb{S}$ under measures $\mathbb{Q} \in \overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}$ is a closed subset of the space of probability measures on $\mathbb{R}_+^{d\times(n+1)}$ in the weak topology, then the following are equivalent:

(1) There is no robust uniformly strong arbitrage on $\mathfrak{P}$ in $\mathcal{A}_{\mathcal{X}}(\widetilde{\mathbb{F}})$;

(2) There is no robust strong arbitrage on $\mathfrak{P}$ in $\mathcal{A}_{\mathcal{X}}(\widetilde{\mathbb{F}})$;

(3) $\widetilde{\mathcal{M}}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$.

Moreover, when any of these holds, for any upper semi-continuous $g : \mathbb{R}_+^{d\times(n+1)} \to \mathbb{R}$ that satisfies (2.6.1), the following pricing duality holds:

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S})) = \sup_{\mathbb{Q}\in\widetilde{\mathcal{M}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})]. \tag{2.6.3}$$

Remark 2.6.4. The set of laws of $\mathbb{S}$ under measure $\mathbb{Q} \in \widetilde{\mathcal{M}}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}$ is a closed subset of the space of probability measures on $\mathbb{R}_+^{d\times(n+1)}$ in the weak topology, as soon as we assume that $\mathbb{S}$ is continuous with respect to ω.

Proof of Proposition 2.6.2.

(3) $\Rightarrow$ (2) and (2) $\Rightarrow$ (1) are obvious.

Step 1. To show (1) implies (3). We suppose there is no robust uniformly strong arbitrage on $\mathfrak{P}$ in $\mathcal{A}_{\mathcal{X}_0}(\widetilde{\mathbb{F}})$. We now apply Theorem 2.3.11 to $\mathcal{X}_0 = \{X_0\}$ and argue that $V_{\mathfrak{P}^*}(-X_0) \geq 0$. Indeed, if $V_{\mathfrak{P}^*}(-X_0) < 0$, since the superhedging price is attained and $\mathfrak{P} = \mathfrak{P}^*$, there exist $H \in \mathbb{F}^A$ and $x < 0$ such that

$$x + (H \circ \mathbb{S})_n(\omega) \geq -X_0, \quad \forall \omega \in \mathfrak{P}$$

which is clearly a robust uniformly strong arbitrage on $\mathfrak{P}$.

Step 2. We first consider the case that that $V_{\mathfrak{P}^*}(-X_0) > 0$. By Theorem 2.3.11, $\sup_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[-X_0] = V_{\mathfrak{P}^*}(-X_0) > 0$. Similarly, due to the assumption that $V_{\mathfrak{P}^*}(X_0) > 0$, $\inf_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[-X_0] < 0$. Hence, we have 0 is in the interior of the price bound formed by $\sup_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[X_0]$ and $\inf_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[X_0]$, i.e.,

$$\sup_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[X_0] > 0 \quad \text{and} \quad \inf_{\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[X_0] < 0. \tag{2.6.4}$$

Therefore we know $\mathcal{M}^f_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}} \neq \varnothing$. Moreover, for any $\omega \in \mathfrak{P}^*$ and $\mathbb{Q} \in \mathcal{M}^f_{\mathfrak{P}}$ such that $\mathbb{Q}(\{\omega\}) > 0$, we can find $\tilde{\mathbb{Q}} \in \mathcal{M}^f_{\mathfrak{P}}$ and $\alpha \in [0, 1]$ such that $\mathbb{E}_{\alpha\mathbb{Q}+(1-\alpha)\tilde{\mathbb{Q}}}[X_0] = 0$. Hence it is straightforward to see that in this case $\mathfrak{P}^*_{\mathcal{X}_0} = \mathfrak{P}^*$.

The remaining case is $V_{\mathfrak{P}^*}(-X_0) = 0$. In this case, by considering the ω^* such that $s_0 = \mathbb{S}_0(\omega^*) = \mathbb{S}_1(\omega^*) = \ldots = \mathbb{S}_n(\omega^*)$, we can rule out the possibility that $g_0(s_0) < 0$ as super-relicating $-X_0$ would require an initial capital of at least $-g_0(s_0)$ and it contradicts $V_{\mathfrak{P}^*}(-X_0) = 0$. We now look at the two sub-cases: $g_0(s_0) = 0$ or $g_0(s_0) > 0$. In the former case, it is obvious that the Dirac measure $\delta_{\omega^*} \in \mathcal{M}^f_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}$ and hence $\mathcal{M}^f_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}} \neq \varnothing$. For the later, we are going to argue that it is impossible. Indeed, if $g_0(s_0) > 0$, note that by convexity of g_0, for any $l \in 0, 1, \ldots, n-1$, $g_0(\mathbb{S}_n(\omega)) - \sum_{i=l+1}^n g_0'(\mathbb{S}_{i-1}(\omega))\Big(\mathbb{S}_i(\omega) - \mathbb{S}_{i-1}(\omega)\Big) > g_0(\mathbb{S}_l(\omega))$ for any $\omega \in \mathfrak{P}$, where g_0' here is the right derivative of g_0. In particular, when $l = 0$,

$$-g_0(s_0) - \sum_{i=1}^n g_0'(\mathbb{S}_{i-1}(\omega))\Big(\mathbb{S}_i(\omega) - \mathbb{S}_{i-1}(\omega)\Big) \geq -g_0(\mathbb{S}_n(\omega))) \quad \forall \omega \in \mathfrak{P},$$

which is a robust uniformly strong arbitrage on $\mathfrak{P}$. It is a clear contradiction to our assumption.

Step 3. We now show the pricing–hedging duality (2.6.2) for any g that is measurable and bounded. Again, we need to consider the two cases: $V_{\mathfrak{P}^*}(-X_0) > 0$ or $V_{\mathfrak{P}^*}(-X_0) = 0$. For the former, we know from Step 1 that $\mathfrak{P}^*_{X_0} = \mathfrak{P}^*$. Therefore, it follows from Remark 2.4.3 that for any g that is measurable and bounded from above,

$$V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S})) = \sup_{\mathbb{Q} \in \mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})].$$

In the later case, note that it follows from Step 1 that $g_0(s_0) = 0$, where $s_0 = \mathbb{S}_0(\omega^*) = \mathbb{S}_1(\omega^*) = \ldots = \mathbb{S}_n(\omega^*)$ for some $\omega^* \in \mathfrak{P}$. In fact, due to the strict convexity of g_0, for any $\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}$, the law of $\mathbb{S}_n$ under any $\mathbb{Q}$ is just the Dirac measure $\delta_{(s_0,\ldots,s_0)}$ on $\mathbb{R}^{d\times(n+1)}_+$. Otherwise, by Jensen's inequality, $\mathbb{E}_\mathbb{Q}[-g_0(\mathbb{S}_n)] > -g_0(\mathbb{E}_\mathbb{Q}[\mathbb{S}_n]) = 0$, and it would contradict $V_{\mathfrak{P}^*}(-X_0) = 0$ as $V_{\mathfrak{P}^*}(-X_0) \geq \sup_{\mathbb{Q}\in\mathcal{M}_\mathfrak{P}} \mathbb{E}_\mathbb{Q}[-X_0] > 0$. Therefore, we have

$$\sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_\mathbb{Q}[g(\mathbb{S})] = g(s_0,\ldots,s_0).$$

On the other hand, it is possible to super-replicate g with any initial capital larger than $g(s_0,\ldots,s_0)$. To see this, note that by convexity of g_0,

$$g_0(s_n) - \sum_{i=1}^n g_0'(s_{i-1})(s_i - s_{i-1}) \geq g_0(s_0) = 0,$$

for any $s = (s_0,\ldots,s_n) \in \mathbb{R}^{d\times(n+1)}_+$, with a strict inequality for any $s \in \mathbb{R}^{d\times(n+1)}_+$ such that s is not a constant path, i.e., $s_i \neq s_0$ for some $i \in \{0,\ldots,n\}$. In fact, it is bounded away from 0 outside any small ball of $(s_0,\ldots,s_0)$. Hence, due to the upper semi-continuity and boundedness of g, for any $\epsilon > 0$, there exists a sufficiently large K such that

$$g(s_0,\ldots,s_0)+\epsilon+K\Big\{g_0(\mathbb{S}_n(\omega))-\sum_{i=1}^n g_0'(\mathbb{S}_{i-1}(\omega))\big(\mathbb{S}_i(\omega)-\mathbb{S}_{i-1}(\omega)\big)\Big\} \geq g(\mathbb{S}(\omega)) \quad \forall\omega \in \mathfrak{P}^*$$

Therefore, $V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g) \leq g(s_0,\ldots,s_0) = \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_\mathbb{Q}[g(\mathbb{S})]$. The reverse inequality is easy and hence we obtain $V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g) = \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_\mathbb{Q}[g]$ as required.

Step 4. It remains to argue that the duality still holds true for any g that is upper semi-continuous and satisfies (2.6.1). We first argue that any upper semi-continuous $g : \mathbb{R}^{d\times(n+1)}_+ \to \mathbb{R}$, satisfying (2.6.1), can be super-replicated on $\mathfrak{P}^*$ by a strategy involving dynamic trading in S, static hedging in g_0 and cash. Define a synthetic option with payoff $\widetilde{m} : \mathbb{R}^{d\times(n+1)}_+ \to \mathbb{R}$ by

$$\widetilde{m}(x_0,\ldots,x_n) = \sum_{l=0}^n \Big\{g_0(x_n) - \sum_{i=l+1}^n g_0'(x_{i-1})(x_i - x_{i-1})\Big\}. \tag{2.6.5}$$

By convexity of g, we know that

$$\widetilde{m}(x_0,\ldots,x_n) = \sum_{l=0}^n \Big\{g_0(x_n)-\sum_{i=l+1}^n g_0'(x_{i-1})(x_i-x_{i-1})\Big\} \geq \sum_{l=0}^n g_0(x_l) = m(x_0,\ldots,x_n).$$

Since we assume there is no robust strong arbitrage, it is clear that $V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(\widetilde{m}(\mathbb{S})) = 0$.

For any $K > 0$, it follows from (2.6.1) that $\{x \in \mathbb{R}_+^{d\times(n+1)} : K\widetilde{m}(x) \leq g(x)\}$ is bounded. Hence we know from the semi-continuity of $\widetilde{m}$ and g that

$$\exists\, x_0 \in \mathbb{R}_+ \text{ such that } x_0 + K\widetilde{m}(\mathbb{S}(\omega)) \geq g(\mathbb{S}(\omega)) \quad \forall \omega \in \mathfrak{P}^*. \tag{2.6.6}$$

Hence by sublinearity of $V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(\cdot)$, we have

$$\begin{aligned} V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g) &\leq V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S}) - K\widetilde{m}(\mathbb{S})) + V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(K\widetilde{m}(\mathbb{S})) \\ &\leq \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - K\widetilde{m}(\mathbb{S})] + 0 \\ &\leq \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})], \end{aligned}$$

where the second inequality follows from the pricing–hedging duality for bounded claims that we established above and the fact that $V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(K\widetilde{m}(\mathbb{S})) = KV_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(\widetilde{m}(\mathbb{S})) = 0$.

The reverse inequality is always true and hence we obtain $V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g) = \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g]$ as required.

It remains to show that

$$\sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g] = \sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \mathbb{E}_{\mathbb{Q}}[g].$$

It is obvious that

$$\sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \mathbb{E}_{\mathbb{Q}}[g] \geq \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g].$$

For the inverse, note that from (2.6.6) $g - K\widetilde{m}$ is bounded from above. Hence for any sequence $\{\mathbb{Q}_n\}$ in $\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}$ such that $\mathbb{Q}_n \to \mathbb{Q}$ as $n \to \infty$,

$$\liminf_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[g(\mathbb{S})] = \liminf_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[g(\mathbb{S}) - K\widetilde{m}(\mathbb{S})] \geq \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - K\widetilde{m}(\mathbb{S})]. \tag{2.6.7}$$

On the other hand, as $\widetilde{m}$ is bounded from below,

$$\mathbb{E}_{\mathbb{Q}}[\widetilde{m}(\mathbb{S})] \leq \liminf_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[\widetilde{m}(\mathbb{S})] = 0.$$

Combining this with (2.6.7), we have

$$\liminf_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[g(\mathbb{S})] \geq \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})]$$

Therefore

$$\sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \mathbb{E}_{\mathbb{Q}}[g] \leq \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g].$$

and we have

$$\sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g] = \sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \mathbb{E}_{\mathbb{Q}}[g].$$

□

Proof of Theorem 2.6.3.

(3) ⇒ (2) and (2) ⇒ (1) are obvious.

Step 1. To show that (1) implies (3). Suppose there is no robust uniformly strong arbitrage on $\mathfrak{P}$ in $\mathcal{A}_{\mathcal{X}}(\widetilde{\mathbb{F}})$. We first know from Proposition 2.6.2 that $\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}} \neq \varnothing$ and hence, in particular, $\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \neq \varnothing$. We can use a variational argument to deduce the following equalities:

$$\begin{aligned} V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S})) &= \inf_{X\in\mathrm{Lin}(\mathcal{X}/\{X_0\})} V_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S})-X) \\ &= \inf_{X\in\mathrm{Lin}(\mathcal{X}/\{X_0\})} \sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})-X] \qquad (2.6.8) \\ &= \sup_{\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}} \inf_{X\in\mathrm{Lin}(\mathcal{X}/\{X_0\})} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})-X]. \qquad (2.6.9) \end{aligned}$$

where the equality between (2.6.8) and (2.6.9) is an application of min–max theorem (see Corollary 2 in Terkelsen [102]). To justify this, denote $\mathcal{Q}$ by the set of laws of $\mathbb{S}$ under measures $\mathbb{Q}\in\overline{\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}}$ and write $\mathrm{Lin}(\{g_i\}_{i\in I/\{0\}})$ to be the set of finite linear combinations of elements in $\{g_i\}_{i\in I/\{0\}}$. By assumption, $\mathcal{Q}$ is closed. Observe that $\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}$ is a subset of $\{\mathbb{Q}\in\mathcal{M}_{\mathfrak{P}} \mid \mathbb{E}_{\mathbb{Q}}[X_0]\le 0\}$. From Step 1 of the proof of Theorem 1.3 in Acciaio et al. [1], $\mathcal{Q}$ is weakly compact.

Notice that

$$\inf_{X\in\mathrm{Lin}(\mathcal{X}/\{X_0\})} \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})-X] = \inf_{G\in\mathrm{Lin}(\{g_i\}_{i\in I/\{0\}})} \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})-G(\mathbb{S})],$$

where $\mathbb{S} := (\mathbb{S}_t)_{t=0}^n$ is the canonical process on $\mathbb{R}_+^{d\times(n+1)}$. We apply min–max theorem (see Corollary 2 in Terkelsen [102]) to the compact convex set $\mathcal{Q}$, the convex set $\mathrm{Lin}(\{g_i\}_{i\in I/\{0\}})$, and the function

$$f(\mathbb{Q},G) = \int_{\mathbb{R}_+^{d\times(n+1)}} \Big(g(s_0,\dots,s_n) - G(s_0,\dots,s_n)\Big) d\mathbb{Q}(s_0,\dots,s_n).$$

Clearly f is affine in each of the variables. Furthermore, $f(\cdot,G)$ is upper semi-continuous on $\mathcal{Q}$. To see this, fix $K>0$ and $G\in\mathrm{Lin}(\{g_i\}_{i\in I/\{0\}})$, and let $\widetilde{g} =$

$g - G - K\widetilde{m}$, where $\widetilde{m} : \mathbb{R}_+^{d\times(n+1)} \to \mathbb{R}$ is defined in (2.6.5). It follows from the definition of $\widetilde{m}$ that for every $\mathbb{Q} \in \mathcal{Q}$,

$$f(\mathbb{Q}, G) = \mathbb{E}_{\mathbb{Q}}[\widetilde{g}(\mathbb{S})]. \tag{2.6.10}$$

It follows from (2.6.6) that $\widetilde{g}$ is bounded from above. Hence, for every sequence of $\{\mathbb{Q}_n\}_n \in \mathcal{Q}$ with $\mathbb{Q}_n \to \mathbb{Q}$ as $n \to \infty$ for some $\mathbb{Q}$ weakly, we have

$$\lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[\widetilde{g}_+(\mathbb{S})] = \mathbb{E}_{\mathbb{Q}}[\widetilde{g}_+(\mathbb{S})]$$

by Portmanteau theorem, and

$$\liminf_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[\widetilde{g}_-(\mathbb{S})] \geq \mathbb{E}_{\mathbb{Q}}[\widetilde{g}_+(\mathbb{S})]$$

by Fatou's lemma, where $\widetilde{g} := \widetilde{g}_+ - \widetilde{g}_-$ with $\widetilde{g}_+ = \widetilde{g} \vee 0$. Then

$$\limsup_{n\to\infty} f(\mathbb{Q}_n, G) = \limsup_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n}[\widetilde{g}(\mathbb{S})] \leq \mathbb{E}_{\mathbb{Q}}[\widetilde{g}(\mathbb{S})] = f(\mathbb{Q}, G).$$

Therefore, the assumptions of Corollary 2 in Terkelsen [102] are satisfied and we have

$$\begin{aligned}
\inf_{X\in \mathrm{Lin}(\mathcal{X}/\{X_0\})} \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - X] &= \inf_{G\in \mathrm{Lin}(\{g_i\}_{i\in I/\{0\}})} \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - G(\mathbb{S})] \\
&= \sup_{\mathbb{Q}\in\mathcal{Q}} \inf_{G\in \mathrm{Lin}(\{g_i\}_{i\in I/\{0\}})} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - G(\mathbb{S})] \\
&= \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \inf_{X\in \mathrm{Lin}(\mathcal{X}/\{X_0\})} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - X].
\end{aligned}$$

Take $g = 0$. If $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} = \varnothing$, then

$$\sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X}_0,\mathcal{P},\mathfrak{P}}} \inf_{X\in \mathrm{Lin}(\mathcal{X}/\{X_0\})} \mathbb{E}_{\mathbb{Q}}[-X] = -\infty,$$

and hence $V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*}(0) = -\infty$, which contradicts the no arbitrage assumption. Therefore we have (1) implies (3).

Step 2. To show the pricing–hedging duality, suppose now $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$. Then, from Proposition 2.6.2, we know $V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S}))$ is finite. If $\mathbb{Q} \notin \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$, then

$$\inf_{X\in \mathrm{Lin}(\mathcal{X}/\{X_0\})} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}) - X] = -\infty.$$

Therefore, in (2.6.9) it suffices to look at measure in $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ only, and hence we obtain

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}^*}(g(\mathbb{S})) = \sup_{\mathbb{Q}\in\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S})]$$

as required. $\square$

2.7 Complementary

In this section, we give the proof of Theorem 2.4.2.

2.7.1 Proof of Theorem 2.4.2

Write $\mathcal{X}_m = (X_1, \ldots, X_m)$, $1 \leq m \leq k$ with $\mathcal{X}_k = \mathcal{X}$. We prove the statement by induction on the number of static options used for superhedging. For this we consider the superhedging problem with additional options $\mathcal{X}_m$ on $\mathfrak{P}^*_{\mathcal{X}}$ and denote its superhedging cost by $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ which is defined as in (2.4.2) but with $\mathcal{X}_m$ replacing $\mathcal{X}$.

We first assume that $\mathfrak{P}^*_{\mathcal{X}_m} \in \mathcal{F}^{\mathcal{A}}$ *for all* $m \leq k$. *This will be justified at the end.*

The case $m = 0$ corresponds to super-hedging on $\mathfrak{P}^*_{\mathcal{X}}$ without options. Since by assumption $\mathfrak{P}^*_{\mathcal{X}} \in \mathcal{F}^{\mathcal{A}}$, the pricing–hedging duality follows from Theorem 2.3.11.

Now assume that for some $m < k$, for any $\mathcal{F}^{\mathcal{A}}$-measurable g, we have the following pricing–hedging duality

$$V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = \sup_{Q \in \mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g] \tag{2.7.1}$$

when either $V_{\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$ or (2.4.3) is satisfied, and $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ is attained by some strategy $(\alpha, H) \in \mathbb{R}^m \times \mathcal{A}(\mathbb{F}^{\mathcal{A}})$ when $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ is finite. We show that the same statement holds for $m + 1$. The proof proceeds in four steps.

Step 1. First observe that if X_{m+1} is replicable on $\mathfrak{P}^*_{\mathcal{X}}$ by semi-static portfolios with the static hedging part restricted to $\mathcal{X}_m$, i.e. $x + h \cdot \mathcal{X}_m(\omega) + (H \circ \mathbb{S})_n(\omega) = X_{m+1}(\omega)$, for any $\omega \in \mathfrak{P}^*_{\mathcal{X}}$, then necessarily $x = 0$ (otherwise $\mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}} = \varnothing$). Moreover since any such portfolio has zero expectation under measures in $\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ we have that $\mathbb{E}_{\mathbb{Q}}[X_{m+1}] = 0 \ \forall \mathbb{Q} \in \mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$. In particular $\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}} = \mathcal{M}^f_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ and (2.7.1) holds for $m + 1$.

Conversely, if the latter holds than $\mathcal{X}_{m+1}$ is replicable. Indeed, by the inductive hypothesis it follows from (2.7.1) and the fact that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ can be attained that there exists some $(H, h) \in \mathcal{A}(\mathbb{F}^{\mathcal{A}}) \times \mathbb{R}^m$ such that

$$h \cdot \mathcal{X}_m(\omega) + (H \circ \mathbb{S})_n(\omega) \geq X_{m+1}(\omega) \quad \forall \omega \in \mathfrak{P}^*_{\mathcal{X}}.$$

If the above inequality is strict for some $\tilde{\omega} \in \mathfrak{P}^*_{\mathcal{X}}$ then by taking expectation under $\tilde{\mathbb{Q}} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ such that $\tilde{\mathbb{Q}}(\tilde{\omega}) > 0$ we obtain

$$\mathbb{E}_{\tilde{\mathbb{Q}}}[h \cdot \mathcal{X}_m + (H \circ \mathbb{S})_n] > \mathbb{E}_{\tilde{\mathbb{Q}}}[\mathcal{X}_{m+1}] = 0. \tag{2.7.2}$$

But $\mathbb{E}_{\tilde{\mathbb{Q}}}[h \cdot \mathcal{X}_m + (H \circ \mathbb{S})_n] = 0$, which gives a contradiction.
As a consequence of the equivalence result we just verified, it is straightforward to see the claim holds for $m+1$ if $\mathcal{X}_{m+1}$ is replicable.

Step 2. We now look at the more interesting case, that is X_{m+1} is not replicable. In this case, by a similar reasoning as above, we have

$$\sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[X_{m+1}] > 0 \quad \text{and} \quad \inf_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[X_{m+1}] < 0. \tag{2.7.3}$$

Indeed, in the case that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) < \infty$, using the induction hypothesis that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1})$ is attained implies that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) > 0$ as otherwise we would obtain a contradiction as in (2.7.2) above. Hence $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1})$ is always strictly positive. Similarly $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1})$ is always strictly positive. Therefore, (2.7.3) follows from (2.7.1).

Given (2.7.3), we now show that also in the case that X^{m+1} is not replicable, for any g measurable

$$V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g],$$

when either $V_{\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$ or (2.4.3) is satisfied

We first use a variational argument to deduce the following equalities:

$$\begin{aligned} V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) &= \inf_{l\in\mathbb{R}} V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g - lX_{m+1}) \\ &= \inf_{l\in\mathbb{R}} \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g - lX_{m+1}] \\ &= \inf_{N} \inf_{|l|\le N} \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g - lX_{m+1}] \\ &= \inf_{N} \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \inf_{|l|\le N} \mathbb{E}_{\mathbb{Q}}[g - lX_{m+1}], \end{aligned} \tag{2.7.4}$$

The first equality follows from Theorem 2.3.11, the second from the inductive hypothesis and the last one is obtained with an application of min–max theorem (see Corollary 2 in Terkelsen [102]).

Note that for any $\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ we have $V_{\mathcal{X}_n,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) \geq \mathbb{E}_{\mathbb{Q}}[g] > -\infty$. Similarly, $V_{\mathcal{X}_{n+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) > -\infty$. Suppose $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$. Then in this case $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) \leq V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$. When N is sufficiently large,

$$\sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \inf_{|l|\leq N} \mathbb{E}_{\mathbb{Q}}[g - lX_{m+1}] - V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) < 1.$$

By taking $c = 1 + |V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)| + |V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)|$, we can argue that it suffices for us to consider $\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ such that $|\mathbb{E}_{\mathbb{Q}}[X_{m+1}]| \leq c/N$, a set we denote $\mathcal{M}^{f,c/N}_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ below. Indeed, for any $\mathbb{Q} \in \mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ such that $|\mathbb{E}_{\mathbb{Q}}[X_{m+1}]| > c/N$,

$$\inf_{|l|\leq N} E_Q[g - lX_{m+1}] < \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g] - c < V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g).$$

It follows that

$$\begin{aligned} V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) &= \inf_N \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \inf_{|l|\leq N} \mathbb{E}_{\mathbb{Q}}[g - lX_{m+1}] \\ &= \lim_{N\to\infty} \sup_{\mathbb{Q}\in\mathcal{M}^{f,c/N}_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \inf_{|l|\leq N} \mathbb{E}_{\mathbb{Q}}[g - lX_{m+1}] \\ &\leq \lim_{N\to\infty} \sup_{\mathbb{Q}\in\mathcal{M}^{f,c/N}_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g]. \end{aligned}$$

Now take a sequence $\{\mathbb{Q}_N\}_N \in \mathcal{M}^{f,c/N}_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ such that $\mathbb{E}_{\mathbb{Q}_N}[g] \geq \sup_{\mathbb{Q}\in\mathcal{M}^{f,c/N}_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g] - 1/N \ \forall N$. From (2.7.3), we know 0 is in the interior of the price bound formed by $\sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[X_{n+1}]$ and $\inf_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[X_{m+1}]$. Take $\mathbb{Q}_{\text{sup}} \in \mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ and $\mathbb{Q}_{\text{inf}} \in \mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ such that

$$\mathbb{E}_{\mathbb{Q}_{\text{sup}}}[X_{m+1}] \geq \frac{1}{2} \sup_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[X_{m+1}] \quad \text{and} \quad \mathbb{E}_{\mathbb{Q}_{\text{inf}}}[X_{m+1}] \leq \frac{1}{2} \inf_{\mathbb{Q}\in\mathcal{M}^f_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[X_{m+1}].$$

For every N, set $\tilde{\mathbb{Q}}_N = \mathbb{Q}_{\text{inf}}$ if $\mathbb{E}_{\mathbb{Q}}[X_{m+1}] \geq 0$, and $\mathbb{Q}_{\text{sup}}$ otherwise. Then for N sufficiently large we can find $\lambda_N \in [0,1]$ such that $\hat{\mathbb{Q}}_N = (1-\lambda_N)\mathbb{Q}_N + \lambda_N\tilde{\mathbb{Q}}_N \in \mathcal{M}^f_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$. Moreover, these λ_N satisfy $\lambda_N \to 0$, which leads to $\mathbb{E}_{\hat{\mathbb{Q}}_N}[g] - \mathbb{E}_{\mathbb{Q}_N}[g] \to 0$ as $N \to \infty$.

Hence

$$\begin{aligned}\lim_{N\to\infty} \sup_{\mathbb{Q}\in\mathcal{M}^{f,c/N}_{\mathcal{X}_n,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g] &= \lim_{N\to\infty} \mathbb{E}_{\mathbb{Q}_N}[g] \\ &= \lim_{N\to\infty} \mathbb{E}_{\hat{Q}_N}[g] \\ &\leq \sup_{\mathbb{Q}\in\mathcal{M}^{f}_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g].\end{aligned}$$

Therefore, $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) \leq \sup_{\mathbb{Q}\in\mathcal{M}^{f}_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g]$. The reverse inequality is easy and hence we obtain $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = \sup_{\mathbb{Q}\in\mathcal{M}^{f}_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g]$ as required.

Suppose $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = \infty$. In this case, $V_{\mathfrak{P}^*_{\mathcal{X}}}(g) \geq V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = \infty$. Hence, to show the pricing–hedging duality, we only need to consider the case that (2.4.3) is satisfied. First, it is clear that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) \leq V_{\mathfrak{P}^*}(X_{m+1}) < \infty$ and $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1}) \leq V_{\mathfrak{P}^*}(-X_{m+1}) < \infty$. Then, by super-linearity of $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$, we have

$$\begin{aligned}&V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g - lX_{m+1}) \\ \geq\ &V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) - V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(lX_{m+1}) \\ \geq\ &V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) - |l| \max\{V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}),\, V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1})\} = \infty.\end{aligned}$$

Therefore, it follows from (2.7.4) that $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = \infty$. On the other hand, we can take a sequence of measures $\{\mathbb{Q}_N\}_N \in \mathcal{M}^{f}_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ such that $\mathbb{E}_{\mathbb{Q}_N}[g] > N$. Again, using the same argument as above, we can construct a sequence of measures $\{\hat{\mathbb{Q}}_N\}_N \in \mathcal{M}^{f}_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}$ such that

$$\mathbb{E}_{\hat{\mathbb{Q}}_N}[X_{m+1}] = \lambda_N \mathbb{E}_{\mathbb{Q}_N}[X_{m+1}] + (1-\lambda_N)\mathbb{E}_{\tilde{\mathbb{Q}}_N}[X_{m+1}] = 0$$

and therefore

$$\begin{aligned}\lambda_N &= \frac{\mathbb{E}_{\tilde{\mathbb{Q}}_N}[X_{m+1}]}{\mathbb{E}_{\tilde{\mathbb{Q}}_N}[X_{m+1}] - \mathbb{E}_{\mathbb{Q}_N}[X_{m+1}]} \\ &\geq \frac{1}{2}\frac{\min\{V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}),\, V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1})\}}{V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) + V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1})} =: a > 0.\end{aligned}$$

Since we can only have $\mathbb{E}_{\tilde{\mathbb{Q}}_N}[g] = \mathbb{E}_{\mathbb{Q}_{\sup}}[g]$ or $\mathbb{E}_{\tilde{\mathbb{Q}}_N}[g] = \mathbb{E}_{\mathbb{Q}_{\inf}}[g]$ then for N sufficiently large $\mathbb{E}_{\mathbb{Q}_N}[g] - \mathbb{E}_{\tilde{\mathbb{Q}}_N}[g] > 0$. From

$$\mathbb{E}_{\hat{\mathbb{Q}}_N}[g] = \lambda_N(\mathbb{E}_{\mathbb{Q}_N}[g] - \mathbb{E}_{\tilde{\mathbb{Q}}_N}[g]) + \mathbb{E}_{\tilde{\mathbb{Q}}_N}[g] > a(\mathbb{E}_{\mathbb{Q}_N}[g] - \mathbb{E}_{\tilde{\mathbb{Q}}_N}[g]) + \mathbb{E}_{\tilde{\mathbb{Q}}_N}[g]$$

we deduce $\mathbb{E}_{\hat{\mathbb{Q}}_N}[g] \to +\infty$ as $N \to \infty$. Therefore, $\sup_{\mathbb{Q}\in\mathcal{M}^{f}_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}} \mathbb{E}_{\mathbb{Q}}[g] = \infty$.

Step 4. It remains to show that if either (2.4.3) or $V_{\mathfrak{P}^*}(g) < \infty$ is satisfied, $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ is attained by some strategy $(\alpha, H) \in \mathbb{R}^{m+1} \times \mathcal{A}_{\mathcal{X}_{m+1}}(\mathbb{F}^{\mathcal{A}})$ when $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ is finite. Consider $F : \mathbb{R} \to \mathbb{R} \cup \{-\infty, \infty\}$ defined by

$$F(l) = V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g - lX_{m+1}).$$

We first consider the case that (2.4.3) is satisfied. In this case, it is clear from above that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$. It is straightforward to see that F is continuous. Indeed, by super-linearity of $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(\cdot)$, we know that $\forall l_1, l_2 \in \mathbb{R}$,

$$F(l_1) - F(l_2) \leq V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}\big(-(l_1 - l_2)X_{m+1}\big), \tag{2.7.5}$$

which can lead to

$$|F(l_1) - F(l_2)| \leq |l_1 - l_2| \max\{V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}), V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1})\} \quad \forall l_1, l_2 \in \mathbb{R}.$$

Moreover, since $\mathcal{X}^{m+1}$ is not replicable, $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) > 0$ and $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1}) > 0$. Hence, by (2.7.5) that $F(l) \to \infty$ as $|l| \to \infty$. Therefore, F achieves its minimum in some $l^* \in \mathbb{R}$ for which $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) = V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g - l^* X^{m+1})$.

In the case that (2.4.3) is not satisfied, we consider g such that $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g) < \infty$. When $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) = V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1}) = \infty$, $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g - lX_{m+1}) = +\infty$ for any $l \neq 0$, and hence F is minimised at 0. When $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) = \infty$ and $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1}) < \infty$, $F(l) = \infty$ for any $l < 0$, while by similar argument as above we can show that F is continuous on $[0, \infty)$ such that $F(l) \to \infty$ as $l \to \infty$, and hence F achieves its minimum in some $l^* \in \mathbb{R}_+$. Similarly, when $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(X_{m+1}) < \infty$ and $V_{\mathcal{X}_m,\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(-X_{m+1}) = \infty$, F is continuous on $(-\infty, 0]$ such that $F(l) \to \infty$ as $l \to -\infty$, and hence F achieves its minimum in some $l^* \in (-\infty, 0]$.

It then follows from the inductive hypothesis that $V_{\mathcal{X}_{m+1},\mathcal{P},\mathfrak{P}^*_{\mathcal{X}}}(g)$ is attained.

Finally we come to our initial assumption that $\mathfrak{P}^*_{\mathcal{X}_m} \in \mathcal{F}^{\mathcal{A}}$ for all $m \leq k$. Observe that in the setup of Proposition 2.5.1 this assumption is automatically satisfied since we show that $\mathfrak{P} = \mathfrak{P}^*_{\mathcal{X}_m}$. Proposition 2.5.1 in turn allows us to establish Lemma 2.5.5 which, see Remark 2.5.6, then gives the required property $\mathfrak{P}^*_{\mathcal{X}_m} \in \mathcal{F}^{\mathcal{A}}$ for all $m \leq k$. The proof is complete.

Chapter 3

Robust pricing and hedging under trading restrictions and the emergence of local martingale models

In this chapter[1], we combine trading constraints with concepts from robust derivative pricing in discrete time. In the discrete-time setting, our results can be summarised as follows: we suppose that we are given a sequence of call price functions at maturity dates $T_1 < T_2 < \cdots < T_n$. We show that these prices are consistent with absence of (suitably defined) arbitrage opportunities if and only if they give rise to a sequence of probability measures $\mu_1, \ldots, \mu_n$ on $\mathbb{R}^n_+$ which satisfy natural ordering properties. These then, correspond to the implied marginal distributions of the asset under feasible risk-neutral measures. (Note that here and throughout, we assume that all assets are denominated in units of some numeraire, for example discounted by the money market account.) Classically, the measures would be in convex order. However, in the absence of the ability to short sell the asset, it is not possible to generate an arbitrage when $m_k = \int x\mu_k(\mathrm{d}x) > \int x\mu_{k+1}(\mathrm{d}x) = m_{k+1}$, and so the expected value of the asset according to the (implied) risk-neutral measure may be smaller at later maturities. We then show that the minimal price of a portfolio involving call options and long positions in the asset, and which superhedges a derivative for every path in $\mathfrak{P}$, is equal to the supremum of the expected value of the derivative's payoff, where the supremum is taken over all supermartingale measures which have full support on $\mathfrak{P}$, and under which the law of the asset at T_k is equal to μ_k. This result generalises

[1]This chapter has been published as Cox et al. [24]. The version here, except some minor adjustments to make it a coherent part of the thesis, is identical with the published work.

Corollary 1.1 in Beiglböck et al. [6] by including a restriction to a certain set of paths $\mathfrak{P}$ and a short selling constraint. Observe also that in the case where the measures μ_k all have the same mean, which is equal to the initial stock price s_0, the class of supermartingale measures is simply the class of martingale measures.

We also consider the case where the set of call options is replaced by put options with the same maturities. Since short selling of the asset is not permitted, one cannot immediately compare to the case where the call options are available to trade, even if the set of possible implied marginal laws remains the same. In this case, we show that a duality gap arises when the initial asset price s_0 is strictly larger than the implied mean m_k for some maturity T_k. In particular, there is no longer equality between the cheapest superhedge and the largest model-consistent price — rather, we see a difference which can be characterised in terms of the limit behaviour of the put prices as the strike goes to infinity.

The easiest example of this duality gap arises in considering the difference between the implied price of a forward contract written on the asset — if we take the forward to be a contract which pays the holder the value of the asset at some future date T_k, then the forward contract will have a model-implied price $m_k = \int x\mu_k(\mathrm{d}x)$, which in the cases of interest will be strictly smaller than the initial price of the asset s_0. In the case where call options are traded, the forward may be superhedged for m_k using call options (the call option with strike 0 has the same payoff as the forward). In the case where put options are traded, this is not the case — instead, the cheapest superreplicating strategy will simply be to purchase the asset at time 0, which has cost s_0.

Embedding the results into a continuous-time framework, we show that the duality gap may be interpreted as a financial bubble and link it to strict local martingales. This provides an intrinsic justification of strict local martingales as models for financial bubbles arising from a combination of trading restrictions and current market prices.

Finally, we note that in parallel to our research Fahim and Huang [46] and Bayraktar and Zhou [5] considered discrete-time robust pricing and hedging with trading restrictions. Fahim and Huang [46] use concepts from optimal martingale transport but assume a market input in the form of distributions μ_i already satisfying a set of assumptions which in our paper are characterised in terms of arbitrage opportunities. Bayraktar and Zhou [5] adopt the quasi-sure analysis of Bouchard and Nutz [12] with finitely many traded options. As a result, in both cases the pricing–hedging duality holds and no links are made to modelling of financial bubbles in discrete or continuous time. The focus of both papers is on general convex portfolio constraints.

This chapter is organised as follows. Section 3.2 discusses the robust modelling framework in discrete time. Sections 3.3 and 3.4 specialise respectively to the case when call or put options are traded. The latter in particular explores when a duality gap arises. Subsequently Section 3.5 focuses on the continuous-time setup. Several proofs are relegated to the Section 3.6.

3.1 Introduction and Preliminaries

3.1.1 Monge–Kantorovich duality and martingale optimal transport

The original setup of Monge–Kantorovich optimal transport problem is a two dimensional problem with probability spaces (X_1, μ_1), (X_2, μ_2). The problem is to find a "cheap" way of transporting μ_1 to μ_2 among all stochastic transport plans, where following Kantorovich, a stochastic transport plan can be formalised as probability measure π on $X_1 \times X_2$ which has X_1-marginal μ_1 and X_2-marginal μ_2.

The transport problem can be naturally extended to a multi-dimensional setting. Let $\Pi_{\vec{\mu}}$ be the set of all Borel probability measures on $\mathbb{R}_+^{n+1}$ with marginals $\delta_{s_0}, \mu_1, \ldots, \mu_n$, where $\mu_1, \cdots, \mu_n$ are probability measures on $\mathbb{R}_+$ which have a finite first moment. Given a measurable cost function $G : \mathbb{R}_+^n \to (-\infty, \infty]$, the primal objective is to minimise

$$I_\pi(G) := \int_{\mathbb{R}_+^n} G(x_1, \ldots, x_n)\pi(\mathrm{d}x_1, \ldots, \mathrm{d}x_n)$$

over $\pi \in \Pi_{\vec{\mu}}$. On the other hand, provided that there exist μ_i-integrable functions u_i, $i = 1, \cdots, n$ such that

$$u_1(s_1) + \ldots + u_n(s_n) \leq G(s_1, \ldots, s_n) \quad \forall s_1, \ldots, s_n \in \mathbb{R}_+, \tag{3.1.1}$$

a dual problem is to maximise

$$\int_{\mathbb{R}_+} u_1 \mathrm{d}\mu_1 + \cdots + \int_{\mathbb{R}_+} u_n \mathrm{d}\mu_n$$

over all μ_i-integrable functions $u_i : \mathbb{R}_+ \to \mathbb{R}$ subject to the constraint (3.1.1).

In the literature, starting already with Kantorovich, a few results have been established in different settings to show that the optimal values of the primal and dual problems agree with each other, including Theorem 2.14 in Kellerer [70] which states that for every Borel measurable $G : \mathbb{R}_+^n \to [0, \infty]$

$$\inf_{\pi \in \Pi_{\vec{\mu}}} I_\pi(G) = \sup\{\sum_{i=1}^{n} \int u_i \mathrm{d}\mu_i : u_1 \oplus \cdots \oplus u_n \leq G \,,\, u_i \text{ is } \mu_i\text{-integrable}\}. \qquad (3.1.2)$$

Another duality result of this type is Proposition 2.1 in Beiglböck et al. [6], which essentially says that if the cost function G is taken to be lower semi-continuous and $G > -\infty$, then the value of the dual problem is still preserved if dual maximisers u_i's are restricted to piece-wise linear functions. For the purpose of application in the financial context, this result has certain advantages as piece-wise linear functions u_i's here have a natural interpretation as a static hedging portfolio of call options. For a more complete account of the history of the problem and theory of optimal transport, see Villani [103, 104].

In the financial context, as already pointed out in Section 1.1, Breeden and Litzenberger [13] first made the observation that if many (all) European options for a given maturity trade then this is equivalent to fixing the implied marginal distributions of the assets under any risk neutral measure. Then, given $\mu_1, \dots, \mu_n$ as the implied marginal distributions and G as the payoff of an option, a modified primal problem, known as the martingale optimal transport problem, is to maximise

$$\mathbb{E}_{\mathbb{P}}[G(\mathbb{S}_1, \cdots, \mathbb{S}_n)]$$

over measures $\mathbb{P}$ such that $\mathbb{S}$, the canonical process on $\mathbb{R}_+^n$, is a martingale with respect to the canonical filtration, starts in a given point and has fixed marginal distributions as $\vec{\mu}$. We denote the set of all such martingale measures $\mathbb{P}$ by $\mathcal{M}_{\vec{\mu}}$. The dual problem of this is also a modification of the original one above. This can be formalised to minimise the setup cost of strategies among all super-hedging strategies of the option payoff, i.e., a static position (a portfolio of call options) $(u_i)_{1 \leq i \leq n}$ plus a dynamic hedging strategy $H = (H_i)_{i \leq n}$ with bounded measurable functions $H_j : \mathbb{R}_+^j \to \mathbb{R}_+$, $j = 0, \dots, n-1$ such that

$$\sum_{i=1}^{n} u_i(\mathbb{S}_i) + \sum_{j=0}^{n-1} H_j(\mathbb{S})(\mathbb{S}_{j+1} - \mathbb{S}_j) \geq G(\mathbb{S}).$$

Here and throughout, $\mathbb{R}_+^0 := \{0\}$, which simply means that H_0 is a nonnegative constant.

Following the consideration above, the natural question arises under which conditions $\mathcal{M}_{\vec{\mu}} \neq \varnothing$. It turns out this is a classical problem and was solved in Strassen [101]. It is now well understood that $\mathcal{M}_{\vec{\mu}} \neq \varnothing$ if and only if $\mu_1, \ldots, \mu_n$ have a finite first moment, mean s_0 and increase in *convex order*. Taking this as a starting point, Beiglböck et al. [6] showed that the pricing–hedging duality is an analogue of the Kantorovich duality in the optimal transport: for each upper semi-continuous payoff function G subject to linear growth

$$\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] = \inf\Big\{\sum_{i=1}^{n}\int u_i \mathrm{d}\mu_i \,:\, \Psi_{X,H} \geq G \text{ on } \Omega \text{ for some admissible } (X,H)\Big\}.$$

3.1.2 Supermartingale optimal transport

In this chapter, different from Beiglböck et al. [6], we start with no arbitrage consideration and show that the prices of call options are consistent with absence of (suitably defined) arbitrage opportunities if and only if they give rise to a sequence of probability measures $\mu_1, \ldots, \mu_n$ on $\mathbb{R}_+^n$, satisfying natural ordering properties and nonincreasing means. In the absence of the ability to short sell the asset, the dual problem becomes to find the minimal price of a portfolio involving call options and long positions in the asset. Accordingly, the primal elements considered here are supermartingale measures under which the law of the asset at T_k is equal to μ_k and hence the primal problem is a supermartingale optimal transport problem.

Historically, there has been relatively little study of asset prices which are strict supermartingales[2] under the risk-neutral measure in the literature. Their main appearance has been as models for the study of financial bubbles, where strict local-martingales are considered. We believe that our results, both in discrete time and in continuous time, contribute to and provide a novel perspective on the existing literature on financial bubbles.

3.1.3 Financial bubbles and local-martingale models

Bubbles, often defined as a deviation of the market price from the asset's fundamental value, were introduced by academic researchers as a concept to explain the failure of

[2] A strict supermartingale is a supermartingale which is not a martingale. Since we only consider nonnegative processes over finite time-horizons, a strict supermartingale is therefore a supermartingale which has a non-constant expected value. Similarly, a strict nonnegative local martingale is a local martingale which is also a strict supermartingale.

classical asset valuation models, e.g. present-value models. Perhaps the most famous and earliest known example was the tulip bubble in Holland that started in 1634. At the peak of the bubble, a single tulip bulb sold for an equivalent of 60,000 US Dollars today. More recent and financially relevant examples include the internet bubble of the 1990s – a speculative growth in stock prices of companies in the internet sector and related fields amid a surging supply of new internet IPOs, and the put warrant bubbles observed in China's financial market. The latter was concerned with put warrants with long maturities traded in the 2005-2008 period, and was documented and studied in Xiong and Yu [107]. These warrants were deeply out-of-money so almost certain to expire worthless, yet they traded in high volumes and at inflated prices throughout the contract lives. See Malkiel [76] for more examples.

In mathematical finance, the modelling of financial bubbles using local-martingale models can be traced back to Heston et al. [58], with subsequent contributions including Cox and Hobson [26]; Jarrow et al. [66, 67]. Our work also has interesting parallels to the recent work of Herdegen and Schweizer [57]. Before Heston et al. [58], a number of authors observed that in certain circumstances, models which were only strict local martingales arise naturally and/or are interesting of their own right (and can be attributed some financial interpretation); see Lewis [73]; Delbaen and Schachermayer [35]; Loewenstein and Willard [74]; Sin [97]. One of the most common examples of a naturally occurring class of local-martingale models is the class of CEV models, $dS_t = S_t^\alpha dB_t, S_0 = s_0$, where $\alpha > 1$. In the case where $\alpha = 2$, one recovers the inverse of a 3-dimensional Bessel process, which was studied in Delbaen and Schachermayer [35]. More recently, quadratic normal volatility (QNV) models have also been studied, which are mostly strict local-martingales, but typically calibrate well to market data; see Carr et al. [21].

We build our contribution to this literature by embedding the discrete-time results into a continuous-time framework. Consider a continuous-time framework with dynamic trading in the asset and call or put options traded initially for certain fixed maturities. Then the discrete setup is naturally included by considering trading strategies which only rebalance at the maturity dates of the options. Discrete-time supermartingale measures are obtained as projections of local-martingale measures which meet the given marginals. The duality gap is preserved when put options trade, and this gap has a possible interpretation as a financial bubble. To make this generalisation, it is necessary to introduce a pathwise superhedging requirement which enforces a collateral requirement. A similar requirement has already been considered in Cox and Hobson [26]. We therefore believe that an important consequence of this paper is the following interpretation of local-martingale models in financial applications: *local-martingale models naturally arise due to trading constraints.*

This has an impact on the existing literature on financial bubbles: intrinsically, we believe that models where asset prices are strict local martingales (under a risk-neutral measure) are models which arise due to constraints on possible trading strategies. They thus correspond exactly to *rational* or *speculative bubbles* in the asset pricing and economics literature. These are usually driven by short selling constraints and/or disagreement between the agents on the fundamental values due to heterogenous beliefs or overconfidence, see Hugonnier [63], Harrison and Kreps [52] and Scheinkman and Xiong [95]. Strict local martingale models are a very natural class of models for bubbles, since there is a natural notion of a 'fundamental' price which diverges from the traded price. However, as we show, this divergence is 'rational' and driven by the absence of arbitrage combined with trading restrictions, as in speculative bubbles. This is different from the case of an 'irrational bubble' when divergence between the market price of an asset and its fundamental price is driven by some behavioural aspect of market participants, rather than specific market features. In this sense, an important contribution of this article for the literature on bubbles is to divorce any notions of 'irrationality' from the financial study of strict local martingale models.

We also make the observation that although we present results on local martingale models in continuous time, our approach is firmly rooted in a discrete-time setup, and all pricing results in continuous time follow essentially from the corresponding discrete-time results. One interpretation is that these models therefore are the natural discrete-time analogues of local martingale models (in this sense, our results provide an alternative response to the first criticism that local martingale models for bubbles are only a feature of continuous-time models, as discussed in Protter [88]; see also Jarrow and Protter [65]). However, it seems to us that the implication more naturally runs in the other direction: in discrete time, our models are very natural and easily specified. In continuous time, however, local martingales are very subtle processes, and the difference between a local martingale and a martingale is not easy to detect — our paper provides a clear specification of a discrete-time setup which could be interpreted in continuous time as a local martingale model. As a result, in our setup local martingale phenomena arise naturally, and reflect specific market conditions. This contrasts with the arguments of e.g. Guasoni and Rásonyi [50], who argue against local martingale models on the basis that they can always be approximated by martingale models.

3.1.4 Short selling bans

Short selling bans as a regulatory tool to discourage speculation and stabilise markets have proved to be popular among emerging markets and during times of financial crisis. During the US subprime mortgage crisis, short selling of 797 financial stocks in US markets was banned by the SEC between September 19, 2008 and October 8, 2008. Around the same time, the South Korean Financial Supervisory Commission imposed an outright prohibition of short selling of any listed stocks in an attempt to curb the spread of malignant rumours in the market. The ban was lifted for non-financial stocks about a year later, while the constraints on financial stocks remained until November 2013. Interestingly, the US and South Korea both have very active derivatives markets and, in both examples, the bans on short selling did not extend to derivative markets. This allowed market makers and investors to use options to hedge portfolios and express pessimistic views. In light of a series of short selling bans across the globe, the question of their impact on stocks and derivatives markets is once again a matter of concern to academics and policy makers, see e.g. Battalio and Schultz [4], Hendershott et al. [55]. The current paper thus represents a theoretical contribution to this literature. Battalio and Schultz [4] study the US short selling ban in September 2008 and find that synthetic share prices for banned stocks, computed separately for puts and for calls, become significantly lower than the actual share prices, accompanied by increases in bid–ask spreads. The findings correspond to the setting of our paper with $m_k < s_0$, making it particularly interesting.

3.2 Robust framework for pricing and hedging

We consider a financial market with two assets: a risky asset S and a numeraire (e.g. the money market account). All prices are denominated in the units of the numeraire. In particular, the numeraire's price is thus normalised and equal to one. We assume initially that S is traded discretely in time at maturities $0 = T_0 < T_1 < T_2 < \ldots < T_n = T$. This is extended to a continuous-time setup in Section 3.5. The asset starts at $S_0 = s_0 > 0$ and is assumed to be nonnegative. We work on the canonical space with a fixed starting point $\Omega = \{(\omega_0, \ldots, \omega_n) \in \mathbb{R}_+^{n+1} : \omega_0 = s_0\}$. The coordinate process on Ω is denoted by $\mathbb{S} = (\mathbb{S}_i)_{i=0}^n$ i.e.,

$$\mathbb{S}_i : \Omega \to \mathbb{R}_+, \; \mathbb{S}_i(\omega_0, \omega_1, \dots, \omega_n) = \omega_i, \; i = 0, \dots, n,$$

and $\mathbb{F} = (\mathcal{F}_i)_{i=1}^n$ is its natural filtration, $\mathcal{F}_i = \sigma(\mathbb{S}_0, \dots, \mathbb{S}_i)$ for any $i = 0, \dots, n$.

We pursue here a robust approach and do not postulate any probability measure which would specify the dynamics for S. Instead we assume that there is a set $\mathcal{X}$ of market traded options with prices $\mathcal{P}(X)$, $X \in \mathcal{X}$, known at time zero. The trading is frictionless and options in $\mathcal{X}$ may be bought or sold at time zero at their known prices. Hence we extend $\mathcal{P}$ to a linear operator defined on

$$\mathrm{Lin}(\mathcal{X}) = \left\{ a_0 + \sum_{i=1}^{m} a_i X_i : m \in \mathbb{N}, a_0, a_i \in \mathbb{R}, X_i \in \mathcal{X} \text{ for all } i = 1, \dots, m \right\}.$$

As explained above, the numeraire has a constant price equal to one. Finally, the risky asset S may be traded at any T_i, $i = 0, \dots, n$; *however, short selling is not allowed.*

We consider two cases, namely when $\mathcal{X}$ is composed of call options or of put options:

$$\mathcal{X}_c = \{(\mathbb{S}_i - K)^+ : i = 1, \dots, n, \; K \in \mathbb{R}_+\},$$
$$\mathcal{X}_p = \{(K - \mathbb{S}_i)^+ : i = 1, \dots, n, \; K \in \mathbb{R}_+\}.$$

An admissible (semi-static) trading strategy is a pair (X, H) where $X \in \mathrm{Lin}(\mathcal{X})$ and $H = (H_j)$ are bounded nonnegative measurable functions $H_j : \mathbb{R}_+^j \to \mathbb{R}_+$, $j = 0, \dots, n-1$. Here and throughout, $\mathbb{R}_+^0 := \{0\}$, which simply means that H_0 is a nonnegative constant. The total payoff associated to (X, H) is given by

$$\Psi_{X,H}(\mathbb{S}) := X(\mathbb{S}) + \sum_{j=0}^{n-1} H_j(\mathbb{S}_1, \dots, \mathbb{S}_j)(\mathbb{S}_{j+1} - \mathbb{S}_j).$$

The cost of following such a trading strategy is equal to the cost of setting up its static part, i.e., of buying the options at time zero, and is equal to $\mathcal{P}(X)$. We denote the class of admissible (semi-static) trading strategies by $\mathcal{A}_{\mathcal{X}}$. We write $\mathcal{A}_c$ (resp., $\mathcal{A}_p$) for the case $\mathcal{X} = \mathcal{X}_c$ (resp., $\mathcal{X} = \mathcal{X}_p$). Note that since short selling is not allowed, these are genuinely different and, as we shall see, will give very different results. Indeed, note that in the former, the short selling of call options is allowed, including the strike zero i.e., the forward, providing a superreplication of the asset S, possibly at a strictly cheaper price than s_0. This feature is not present when dealing with put options.

We are interested in characterising and computing superhedging prices. All the quantities we have introduced are defined pathwise and the superhedging property is also required to hold pathwise. We have also made only mild assumptions on the market mechanisms (e.g. no frictions), but no specific modelling assumptions on the dynamics of the assets. A natural way to incorporate beliefs into the robust framework is through specifying the set $\mathfrak{P} \subset \Omega$ of "possible paths", i.e., paths we deem feasible and for which the hedging strategies are required to work. This can be thought of as specifying the maximal support of the plausible models. In this way, with the support ranging from all paths to e.g. paths in a binomial model, the robust framework can interpolate between model-independent and model-specific setups. The set $\mathfrak{P}$ might be obtained through time series analysis of the past data combined with modelling and a given agent's idiosyncratic views, and is referred to as the *prediction set*. Note that since there is no probability measure specified and hence no distinction between the *real* and the *risk-neutral* measure, it is very natural to combine two streams of information: time-series of past data and forward-looking option prices. This idea goes back to Mykland [78] and we refer to Nadtochiy and Obłój [79], Hou and Obłój [61] and Spoida [100] for more details and extended discussion.

We call the triplet $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$ of market traded options $\mathcal{X}$, their prices and prediction set the *robust modelling inputs*. The fundamental financial notions defined below, e.g. the arbitrage or the superreplication price, are implicitly relative to these inputs.

Definition 3.2.1. The *superreplication cost* of a derivative given by a payoff function $G : \Omega \to \mathbb{R}$, denoted by $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, is the smallest initial capital required to finance an admissible semi-static trading strategy which superreplicates G for every path in $\mathfrak{P}$, i.e.,

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \inf \Big\{ \mathcal{P}(X) \,:\, \exists (X,H) \in \mathcal{A}_{\mathcal{X}} \text{ s.t } \Psi_{X,H} \geq G \text{ on } \mathfrak{P} \Big\}.$$

Note that since $\omega_0 = s_0$ for all $\omega \in \Omega$, it is equivalent to see G as a function from Ω or from $\mathbb{R}^n_+$. We shall be tacitly switching between these viewpoints; the former is used when writing $G(\mathbb{S})$, the latter when imposing conditions on G, see e.g., (3.3.1) below.

Our aim is to understand when a pricing–hedging duality holds, i.e., when the superreplication price can be computed through the supremum of expectations of the payoff over a suitable class of probabilistic models.

Definition 3.2.2. A *market-calibrated model* is a probability measure $\mathbb{P}$ on $(\Omega, \mathbb{F})$ satisfying $\mathbb{P}(\mathfrak{P}) = 1$ and for any $(X, H) \in \mathcal{A}_{\mathcal{X}}$

$$\mathbb{E}_{\mathbb{P}}[\Psi_{X,H}(\mathbb{S})] \leq \mathcal{P}(X), \tag{3.2.1}$$

where here and throughout, we make the convention that $\infty - \infty = -\infty$ so that the LHS in (3.2.1) is always well defined. The set of market-calibrated models is denoted by $\mathcal{M}^{-}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$.

Remark 3.2.3. It follows from the definition that if $\mathcal{M}^{-}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ then for any Borel function $G : \Omega \to \mathbb{R}$

$$P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \sup_{\mathbb{P} \in \mathcal{M}^{-}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] \leq V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G). \tag{3.2.2}$$

We sometimes refer to the LHS of the above inequality as the *primal value* and to the RHS as the *dual value*. The convention is borrowed from the literature on the martingale optimal transport problem. Both sides of (3.2.2) could be interpreted as a notion of "price" of the asset. The superhedging price $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ arises from (efficient) trading. The primal value $P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ arises from modelling. Indeed, it represents the worst model price, among models consistent with the prices observed in the market, can be thought of as the *fundamental price* of G. We also emphasise that either or both of these may be different from the *market prices*, which are observable and equal to s_0 for the dynamically traded asset and to $\mathcal{P}(X)$ for $X \in \mathcal{X}$. The case of strict inequality in (3.2.2) admits an interpretation as a *financial bubble* — an issue we consider in detail further below.

Remark 3.2.4. It follows from the definition that under any market-calibrated model $\mathbb{P}$, the canonical process $\mathbb{S} = (\mathbb{S}_i)_{i=1}^n$ is a supermartingale. Such a measure is called a *supermartingale measure*. Furthermore, for any $X \in \mathcal{X}$, (3.2.1) holds for both $(X, 0)$ and $(-X, 0)$ so that $\mathbb{P}$ is calibrated to options in $\mathcal{X}$, i.e., $\mathbb{P}$ satisfies $\mathbb{E}_{\mathbb{P}}[X] = \mathcal{P}(X)$ for any $X \in \mathcal{X}$.

Definition 3.2.5. We say there is a *robust uniformly strong arbitrage* if there exists a trading strategy $(X, H) \in \mathcal{A}_{\mathcal{X}}$ with a negative price $\mathcal{P}(X) < 0$ and a nonnegative payoff $\Psi_{X,H}(\mathbf{s}) \geq 0$ for all $\mathbf{s} \in \mathfrak{P}$.

Remark 3.2.6. By definition, it is clear that the market admits a robust uniformly strong arbitrage if and only if $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(0) < 0$. When $\mathfrak{P} = \Omega$ the notion of *robust uniformly strong arbitrage* corresponds to a *model-independent arbitrage*; see Davis and Hobson [31] and Cox and Obłój [25]. In a general robust setting, the existence of an arbitrage in the above sense may depend on the modelling assumptions, expressed through $\mathfrak{P} \subset \Omega$, which justifies the terminology. Equally, we stress that the arbitrage is required to be uniform in outcomes $\omega \in \mathfrak{P}$ to distinguish from a slightly weaker notion, used in Acciaio et al. [1], of a strategy $(X, H) \in \mathcal{A}_{\mathcal{X}}$ with $\mathcal{P}(X) \leq 0$ and a positive payoff $\Psi_{X,H}(\mathbf{s}) > 0$. We refer to the latter as *robust strong arbitrage.* See Definition 2.2.1. The two notions are not equivalent in general. However, we can show that they are equivalent in our setup when $\mathcal{X} = \mathcal{X}_p$. Likewise, when $\mathcal{X} = \mathcal{X}_c$, we can show they are equivalent in our setup when either $\mathfrak{P} = \Omega$ or property (iii) in Condition 3.3.1 below holds[3].

The three papers mentioned in the above remark also showed that typically absence of robust uniformly strong arbitrage is not sufficient to guarantee a (robust) fundamental theorem of asset pricing holds, and introduced various weaker notions or additional assumptions. Here we follow Cox and Obłój [25]:

Definition 3.2.7. We say that there is a *weak free lunch with vanishing risk* (WFLVR) if there exist admissible trading strategies $(X_k, H_k) \in \mathcal{A}_{\mathcal{X}}$ and $(X, H) \in \mathcal{A}_{\mathcal{X}}$ such that $\Psi_{X_k,H_k} \to 0$ pointwise on $\mathfrak{P}$, $\lim_k \mathcal{P}(X_k)$ is well defined with $\lim_k \mathcal{P}(X_k) < 0$ and $\Psi_{X_k,H_k} \geq \Psi_{X,H}$.

Note that the requirement that $\lim_k \mathcal{P}(X_k)$ exists is made with no loss of generality as we could always pass to a subsequence of strategies. Note also that a robust uniformly strong arbitrage is by definition also a WFLVR. A version of the fundamental theorem of asset pricing in our context, given below in Proposition 3.3.2 for the case $\mathcal{X} = \mathcal{X}_c$ and in Proposition 3.4.2 for the case $\mathcal{X} = \mathcal{X}_p$, states that absence of WFLVR is equivalent to existence of a market-calibrated model. Further, as in Davis and Hobson [31] and Cox and Obłój [25], we can characterise absence of WFLVR through the properties of $\mathcal{P}$. This is insightful and is one of the reasons why we prefer to keep the option payoffs fixed, e.g. call options in $\mathcal{X}_c$, and discuss their prices $\mathcal{P}$, as opposed to considering shifted payoffs $X - \mathcal{P}(X)$ and eliminating $\mathcal{P}$ from the discussion as in e.g. Acciaio et al. [1].

[3] We chose not to present these argument here as they are lengthy and deviate from the main focus of this Chapter. Instead they will be part of a separate note on arbitrage in a robust setting, available on request or through the authors' webpages.

3.3 Robust pricing–hedging duality when call options trade

In this section we consider the market in which call options are traded, i.e., $\mathcal{X} = \mathcal{X}_c$. Our main result states that we recover the duality known from the case when short selling restrictions are not present. *Throughout we assume that $\mathfrak{P}$ is a closed subset of Ω.*

3.3.1 Market input and no arbitrage

We start by establishing a robust fundamental theorem of asset pricing for our setting which links absence of arbitrage, properties of call prices and existence of a market-calibrated model.

Condition 3.3.1. Let $\mathcal{X} = \mathcal{X}_c$ and $c_i(K) := \mathcal{P}((\mathbb{S}_i - K)^+)$, $i = 1, \ldots, n$, $K \geq 0$. Then

(i) $c_i(x)$ is a nonnegative, convex, decreasing function of x on $\mathbb{R}_+$,

(ii) $s_0 \geq c_1(0) \geq \ldots \geq c_n(0) \geq 0$ and $c_i'(0+) \geq -1$, where c_i' is the right derivative,

(iii) $c_i(K) \to 0$ as $K \to \infty$,

(iv) for any $x \in \mathbb{R}_+$, $c_i(0) - c_i(x)$ is nonincreasing in i.

A robust fundamental theorem of asset pricing in our setup reads as follows.

Proposition 3.3.2. Suppose $\mathfrak{P}$ is a closed subset of Ω and $\mathcal{X} = \mathcal{X}_c$. Then there is no WFLVR if and only if there exists a market-calibrated model, which then implies Condition 3.3.1 holds. Furthermore, if $\mathfrak{P} = \Omega$, then Condition 3.3.1 implies absence of WFLVR.

Proposition 3.3.3. Suppose $\mathfrak{P} = \Omega$ and $\mathcal{X} = \mathcal{X}_c$. Then Condition 3.3.1 (i),(ii) and (iv) are necessary and sufficient for absence of robust uniformly strong arbitrage. In consequence, when these conditions hold but Condition 3.3.1 (iii) fails, there is no robust uniformly strong arbitrage, but a market-calibrated model does not exist.

We defer the proofs of Propositions 3.3.2 and 3.3.3 to Sections 3.6.2 and 3.6.4.

Remark 3.3.4. If we assume that there is no robust uniformly strong arbitrage, then we can immediately deduce that $c_i(0)$ is nonincreasing in i. Indeed, if there exists some i such that $c_i(0) < c_{i+1}(0)$, then by taking $H_i = 1$, $H_j = 0$ for $j \neq i$ and $X = (\mathbb{S}_i - 0)^+ - (\mathbb{S}_{i+1} - 0)^+$, we have $\mathcal{P}(X) = c_i(0) - c_{i+1}(0) < 0$ but $\Psi_{X,H} = 0$ which shows that (X, H) is a robust uniformly strong arbitrage.

3.3.2 Robust pricing–hedging duality and (super-)martingale optimal transport

Our main theorem in the section states that the pricing–hedging duality is preserved under a ban on short selling restrictions when call options are traded.

Theorem 3.3.5. Suppose the market input $(\mathcal{X}_c, \mathcal{P}, \mathfrak{P})$ admits no WFLVR. Let $G : \mathbb{R}^n_+ \to [-\infty, \infty)$ be an upper semi-continuous function such that

$$G(s_1, \ldots, s_n) \leq K(1 + s_1 + \ldots + s_n) \tag{3.3.1}$$

on $\mathbb{R}^n_+$ for some constant K. Then the pricing–hedging duality holds, i.e.,

$$P_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G) = V_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G). \tag{3.3.2}$$

Remark 3.3.6. Our proof of this result follows closely Beiglböck et al. [6] and is an application of the duality theory from optimal transport, which allows us to express the dual problem as a min–max calculus of variations where the infimum is taken over functions corresponding to the delta hedging terms and marginal constraint, and the supremum is taken over all market-calibrated models. The proof will be given in Sections 3.6.3 and 3.6.5.

Remark 3.3.7. Recall from Remark 3.2.3 that the case of strict inequality in (3.2.2) may be thought of as a natural model for a financial bubble. From (3.3.2) we see that this never happens when call options are traded, $\mathcal{X} = \mathcal{X}_c$. It is still possible that

$$s_0 > c_n(0) = P_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(\mathbb{S}_n) = V_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(\mathbb{S}_n)$$

so that the *market price* for the asset S, which is s_0 is strictly greater than its fundamental price $c_n(0)$. However it is not clear if this could be seen as a bubble. In this case the market does not satisfy the *no dominance* principle of Merton [77]: the asset $\mathbb{S}$ is strictly dominated by a call with zero strike and one could argue that $c_n(0)$ is in fact the correct market price for $\mathbb{S}$. This situation is akin to the case of bubbles in complete markets described in Jarrow et al. [66]. We shall see in Section 3.4.2 below that bubbles appear in a meaningful way when put options and not call options are traded.

Note that, by Proposition 3.3.2, no WFLVR is equivalent to $\mathcal{M}^-_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}} \neq \varnothing$ and it implies Condition 3.3.1. Following classical arguments going back to Breeden and Litzenberger [13], we can then define probability measures μ_i on $\mathbb{R}_+$ by

$$\mu_i([0,K]) = 1 + c_i'(K) \quad \text{ for } K \in \mathbb{R}_+. \tag{3.3.3}$$

Naturally the market prices $\mathcal{P}$, or $c_i(K)$, are uniquely encoded via μ_i with

$$c_i(K) = \mathcal{P}((\mathbb{S}_i - K)^+) = \int (s-K)^+ \mu_i(\mathrm{d}s).$$

To make the link with the (super-)martingale transport explicit, we may think of (μ_i) as the inputs. Note that by Remark 3.2.4 the set of market-calibrated models $\mathcal{M}^-_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}$ is simply the set of probability measures $\mathbb{P}$ on $\mathbb{R}^n_+$ such that $\mathbb{S}_0 = s_0$, $\mathbb{S}$ is a supermartingale and $\mathbb{S}_i$ is distributed according to the measure μ_i. Accordingly we use the notation $\mathcal{M}^-_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}} = \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$ and $P_{\vec{\mu},\mathfrak{P}}(G) := \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G]$. Likewise we write $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = V_{\vec{\mu},\mathfrak{P}}(G)$. Note that we have dropped the explicit reference to call options. This is justified since in fact we can allow any μ_i-integrable functions for the static part of trading strategies. To state this as a corollary, we first rewrite Condition 3.3.1 in terms of $\mu_1,\dots,\mu_n$ as follows:

Assumption 3.3.8. The probability measures $\mu_1,\dots,\mu_n$ on $\mathbb{R}_+$ satisfy

1. $s_0 \geq \int_{\mathbb{R}_+} x\mu_1(\mathrm{d}x) \geq \ldots \geq \int_{\mathbb{R}_+} x\mu_n(\mathrm{d}x)$;
2. the sequence $(\int \phi \,\mathrm{d}\mu_i)_{1\le i\le n}$ is nonincreasing for any concave and nondecreasing function $\phi : \mathbb{R} \to \mathbb{R}_+$.

Then Theorem 3.3.5 may be restated as follows.

Corollary 3.3.9. Assume $\mu_1,\dots,\mu_n$ satisfy Assumption 3.3.8 and $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}} \neq \varnothing$. Let the function $G : \mathbb{R}^n_+ \to [-\infty,\infty)$ be upper semi-continuous and satisfies (3.3.1). Then

$$\begin{aligned} P_{\vec{\mu},\mathfrak{P}}(G) =& \inf\Big\{ \sum_{i=0}^n \int u_i(s)\mu_i(\mathrm{d}s) \;:\; u_i : \mathbb{R}_+ \to \mathbb{R} \text{ with linear growth and} \\ & \qquad H_i : \mathbb{R}^i_+ \to \mathbb{R}_+ \text{ bounded s.t. } \Psi_{(u_i),(H_i)} \geq G \text{ on } \mathfrak{P} \Big\} \\ =& V_{\vec{\mu},\mathfrak{P}}(G), \end{aligned} \tag{3.3.4}$$

where $\mu_0 := \delta_{s_0}$. Further, if $\int x\mu_i(\mathrm{d}x) = s_0$ for $i = 1,\dots,n$, then (3.3.4) also holds with $H_i : \mathbb{R}^i_+ \to \mathbb{R}$.

Proof. By taking expectations in the pathwise superhedging inequality, we have for any $\mathbb{P} \in \mathcal{M}^{-}_{\vec{\mu},\mathfrak{P}}$ and any superhedging strategy $((u_i),(H_i))$ as in (3.3.4) that

$$\mathbb{E}_{\mathbb{P}}[G] \leq \sum_{i=1}^{n} \int u_i(s)\mu_i(\mathrm{d}s)$$

and hence an inequality "$\leq$" follows in the first equality in (3.3.4), see also Remark 3.2.3. The inequality "$\leq$" in the second equality in (3.3.4) is obvious because we take the infimum over a smaller set of superhedging strategies. The result (3.3.4) then follows by Theorem 3.3.5. The last statement is clear since in the special case that $\mu_1, \ldots, \mu_n$ have the same mean, $\mathcal{M}^{-}_{\vec{\mu},\mathfrak{P}}$ is the set of martingale measures with marginals μ_i. □

Remark 3.3.10. (3.3.4) is a general statement of Corollary 1.1 in Beiglböck et al. [6] in the presence of a prediction set $\mathfrak{P}$, while $P_{\vec{\mu},\mathfrak{P}}(G) = V_{\vec{\mu},\mathfrak{P}}(G)$ follows directly from Theorem 3.3.5.

The implication of Corollary 3.3.9 is that in a market without bubbles, a short-selling ban does not make any difference to the robust superhedging prices, and the martingale transport cost of G with prediction set $\mathfrak{P}$ is equal to the robust ($\mathfrak{P}$)-superhedging price of G.

3.4 Put options as hedging instruments

We specify now to the case $\mathcal{X} = \mathcal{X}_p$ when put options are traded. The set of semi-static trading strategies (X, H) is denoted by $\mathcal{A}_p$. In this case, the options can not be used to superreplicate the asset. This, as we shall see, has important consequences for pricing and hedging.

3.4.1 Pricing–hedging duality for options with bounded payoffs

We start with a brief discussion of the market input, no arbitrage and existence of market-calibrated models.

Condition 3.4.1. Let $\mathcal{X} = \mathcal{X}_p$ and $p_i(K) := \mathcal{P}((K - \mathbb{S}_i)^+)$, $i = 1, \ldots, n$, $K \geq 0$. Then

(i) $p_i(x)$ is a nonnegative, convex, increasing function of x on $\mathbb{R}_+$,

(ii) $s_0 \geq \lim_{x\to\infty}(x - p_1(x)) \geq \ldots \geq \lim_{x\to\infty}(x - p_n(x)) \geq 0$ and for every i $0 \leq p_i'(0+) \leq 1$,

(iii) $p_i(K) \to 0$ as $K \to 0$,

(iv) for any $x \in \mathbb{R}_+$, $p_i(x)$ is nondecreasing in i.

A robust fundamental theorem of asset pricing analogous to the one in Proposition 3.3.2 holds also in this setup.

Proposition 3.4.2. Suppose $\mathfrak{P}$ is a closed subset of Ω and $\mathcal{X} = \mathcal{X}_p$. Then there is no WFLVR if and only if there exists a market-calibrated model, which then implies Condition 3.4.1. Furthermore, if $\mathfrak{P} = \Omega$, then Condition 3.4.1 implies the absence of WFLVR.

A direct analogue of Proposition 3.3.3 holds also in this setup. Further, if Condition 3.4.1 is satisfied, similarly to (3.3.3), we can define probability measures μ_i on $\mathbb{R}_+$ by

$$\mu_i([0, K]) = p_i'(K) \quad \text{for } K \in \mathbb{R}_+, \tag{3.4.1}$$

which satisfy the same properties as before, namely Assumption 3.3.8. The set of market-calibrated models is simply $\mathcal{M}^-_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}} = \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$ and only depends on the marginals μ_i and not on whether these were derived from put or from call prices. In consequence, we have $P_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) = P_{\vec{\mu},\mathfrak{P}}(G)$.

The situation on the dual side — the superhedging problem — is different. Indeed, we have seen in Corollary 3.3.9 that in the case of call options, we could relax the static part of the portfolio from combinations of call options to combinations of any functions with linear growth without affecting the superhedging price. In contrast, when put options are traded, their combinations are always bounded and such a relaxation is not possible. We stress this in the notation and write $V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) := V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G)$. Our first result shows that when G is bounded, then trading puts instead of calls has no impact on the superhedging price, as one would expect.

Theorem 3.4.3. Suppose the market input $(\mathcal{X}_p, \mathcal{P}, \mathfrak{P})$ admits no WFLVR or equivalently that $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}} \neq \varnothing$. In particular, Condition 3.4.1 is satisfied and (3.4.1) defines measures which satisfy Assumption 3.3.8. Let $G : \mathbb{R}^n_+ \to [-\infty, \infty)$ be an upper semi-continuous function bounded from above. Then

$$P_{\vec{\mu},\mathfrak{P}}(G) = P_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) = V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) = V^{(p)}_{\vec{\mu},\mathfrak{P}}(G). \tag{3.4.2}$$

The proof is given in Sections 3.6.3 and 3.6.5 and is similar to the proof of Theorem 3.3.5.

The above result may be extended to functions G which are not necessarily bounded, but have sub-linear growth. We state one such extension which is used later. In contrast, the duality in (3.4.2) will fail for G which has linear growth – a theme we explore in the subsequent sections.

Corollary 3.4.4. In the setup of Theorem 3.4.3, assume that G is an upper semi-continuous function such that $G_M(s_1,\dots,s_n) := G(s_1,\dots,s_n)-(\sum_{i=1}^n s_i)/M$ is bounded from above for any $M > 1$. Then (3.4.2) holds for G.

Proof. We have

$$\begin{aligned} V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) \leq & V^{(p)}_{\vec{\mu},\mathfrak{P}}(G_M) + V^{(p)}_{\vec{\mu},\mathfrak{P}}((\mathbb{S}_1+\cdots+\mathbb{S}_n)/M) \\ \leq & P_{\vec{\mu},\mathfrak{P}}(G_M) + V^{(p)}_{\vec{\mu},\mathfrak{P}}((\mathbb{S}_1+\cdots+\mathbb{S}_n)/M) \\ \leq & P_{\vec{\mu},\mathfrak{P}}(G) + \frac{ns_0}{M}, \end{aligned}$$

where we used the obvious inequality $0 \leq V^{(p)}_{\vec{\mu},\mathfrak{P}}(\mathbb{S}_i) \leq s_0$, $i = 1,\dots,n$. By letting $M \to \infty$, we get $V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) \leq P_{\vec{\mu},\mathfrak{P}}(G)$. The other inequality $V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) \geq P_{\vec{\mu},\mathfrak{P}}(G)$ is true in all generality (see Remark 3.2.3), and we conclude that $V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = P_{\vec{\mu},\mathfrak{P}}(G)$. □

3.4.2 Duality gap and bubbles

We come back to the topic of financial bubbles considered in Remarks 3.2.3 and 3.3.7. We start with a motivating example of a simple one-period model, $n = 1$. The prediction set is of the form $\mathfrak{P} = \{s_0\} \times \mathfrak{P}_1$ for some $\mathfrak{P}_1 \subset \mathbb{R}_+$. We assume the market admits no WFLVR which is equivalent to saying that μ defined via (3.4.1) is a probability measure supported on $\mathfrak{P}_1$ and satisfies $\int x\mu(\mathrm{d}x) \leq s_0$. We assume the prediction set $\mathfrak{P}_1$ is unbounded and consider an option with an upper semi-continuous payoff function $G : \mathbb{R}_+ \to [-\infty,\infty)$ such that $|G(x)| \leq K|x|$ for some K and let

$$\limsup_{x\to\infty,\, x\in\mathfrak{P}_1} \frac{G(x)}{x} =: \beta \in [-\infty,\infty).$$

A semi-static trading strategy $(X,H) \in \mathcal{A}_p$ here is a pair with $X \in \mathcal{X}_p$ and $H \geq 0$. If it superreplicates G, i.e.,

$$\Psi_{X,H}(s_1) := X(s_1) + H(s_1 - s_0) \geq G(s_1), \quad s_1 \in \mathfrak{P}_1,$$

then necessarily $H \geq \beta^+$ since X is bounded, where $\beta^+ = \beta \vee 0$. Therefore, we find

$$\begin{aligned}
V^{(p)}_{\mu,\mathfrak{P}}(G) =& \inf\Big\{ \int X(s_1)\mu(\mathrm{d}s_1) : (X,H) \in \mathcal{A}_p \text{ s.t } \Psi_{X,H} \geq G \text{ on } \mathfrak{P}\Big\} \\
=& \inf_{H_0 \geq \beta_+} \Bigg(H_0 s_0 + \inf\Big\{ \int X(s_1)\mu(\mathrm{d}s_1) : (X,H) \in \mathcal{A}_p \text{ s.t} \\
& \qquad \Psi_{X,H}(s_1) \geq G(s_1) - H_0 s_1 \,\forall\, s_1 \in \mathfrak{P}_1 \Big\}\Bigg) \\
=& \inf_{H_0 \geq \beta^+} \Big\{ H_0 s_0 + P_{\mu,\mathfrak{P}}(G(\mathbb{S}) - H_0 \mathbb{S}_1)\Big\} \\
=& \inf_{H_0 \geq \beta^+} \Big(H_0 s_0 + \int_{\mathbb{R}_+} (G(s_1) - H_0 s_1)\mu(\mathrm{d}s_1)\Big) \qquad (3.4.3) \\
=& \int_{\mathbb{R}_+} G(s_1)\mu(\mathrm{d}s_1) + \inf_{H_0 \geq \beta_+} \Big(H_0 (s_0 - \int_{\mathbb{R}_+} s_1 \mu(\mathrm{d}s_1))\Big) \\
=& \int_{\mathbb{R}_+} G(s_1)\mu(\mathrm{d}s_1) + \beta^+ (s_0 - \int_{\mathbb{R}_+} s_1 \mu(\mathrm{d}s_1)) \\
=& P_{\mu,\mathfrak{P}}(G) + \beta^+ (s_0 - \int_{\mathbb{R}_+} s_1 \mu(\mathrm{d}s_1)).
\end{aligned}$$

It follows that if the mean of μ is strictly smaller than s_0, then we have a duality gap for G with linear growth. The intuitive reason is clear: buying the asset directly is implicitly more expensive than constructing a position using put options. If G has a bounded payoff then the latter is feasible as seen in Theorem 3.4.3. However, for G with a linear growth, any superhedging portfolio has to include the asset $\mathbb{S}$ and is hence more expensive, as seen above. When $G(s_1) = (s_1 - K)^+$, $\limsup_{x\to\infty,\, x\in\mathfrak{P}_1} G(x)/x = 1$ and we obtain

$$V^{(p)}_{\mu,\mathfrak{P}}(G) = P_{\mu,\mathfrak{P}}(G) + \left(s_0 - \int_{\mathbb{R}_+} x\mu(\mathrm{d}x)\right).$$

Likewise, taking $G(s_1) = s_1$, we have

$$s_0 = V^{(p)}_{\mu}(\mathbb{S}_1) \geq P_{\mu,\mathfrak{P}}(\mathbb{S}_1) = f_0 := \int_{\mathbb{R}_+} x\mu(\mathrm{d}x).$$

The market has a bubble — a misalignment of market and fundamental prices — if the forward price f_0 implied by the put options is strictly smaller than the spot price s_0. This should be contrasted with the situation in Remark 3.3.7, where the bubble arose due to dominated assets.

The difference between these situations can be summarised as follows: in order to have a financially meaningful market, we must always have the following inequalities for assets with quoted prices:

$$\begin{aligned} \begin{array}{c}\text{market}\\ \text{price}\end{array} &\geq \text{cheapest superreplication price} \\ &\geq \sup\{\text{model-implied prices}\} \\ &= \begin{array}{c}\text{fundamental}\\ \text{price.}\end{array} \end{aligned}$$

The first inequality here follows from the fact that we can superreplicate an asset by purchasing it, and we may have a strict inequality without a simple arbitrage if it is not possible (due to portfolio constraints) to short-sell the asset. However, in the case where there is a strict inequality here, the market contains a dominating portfolio — that is, the super-replicating strategy strictly dominates the purchase of the asset at the market price, and so Merton's *no-dominance* principle fails. In general, one would not expect such markets to exist — even if arbitrage were not possible, one would expect equilibrium to close the gap, since no (rational) market participants would purchase the asset at its market price. On the other hand, the second inequality here is rational — there is no a priori need for the superreplication price and the model-implied prices to agree.

As a result, markets where the fundamental price and the market price differ for some assets, but where the fundamental price and the superreplication price always agree, are mathematically possible, even if they are not economically plausible. This corresponds to the setup described in Section 3.3. A second case is also possible, and economically more plausible, where the superreplication price of an asset coincides with its market price but is different from the fundamental price. This corresponds to the setup described in Section 3.4. One of the main contributions of this paper is that we provide specific characterisations of markets where both behaviours are possible.

In a more classical framework, the two cases described above are encapsulated in the difference between the complete setting of Cox and Hobson [26] and Jarrow et al. [66], and the incomplete models of Jarrow et al. [67]. In the former, completeness of the markets implies that equality always holds between the cost of the cheapest super-replicating strategy for an option and the model-implied price of the option. In contrast, in the latter, Merton's no-dominance condition implies that the inequality between the market price and the superreplication price is an equality. However, in Jarrow et al. [67], the existence of a bubble depends on the choice of some pricing measure to determine the "market-price". In the current (robust) setting, we are able to define the fundamental price in a concrete manner (dependent only on the market prices and the prediction set), leading to a possibly clearer characterisation of a bubble, which does not need some external "selection" procedure.

We now extend the above discussion to a general n-marginal setting. In the one-period case above, G was a European option and the size of the gap between its market and fundamental prices was given simply as a product of its linear growth coefficient and the bubble size $s_0 - f_0$. In the general setting, we can not compute explicitly the duality gap for an arbitrary payoff G. We give below a characterisation which then allows us to obtain explicit expressions for most of the typically traded exotic options.

Theorem 3.4.5. Assume $\mu_1, \ldots, \mu_n$ satisfy Assumption 3.3.8. Suppose the payoff function $G : \mathbb{R}_+^n \to [-\infty, \infty)$ is upper semi-continuous and satisfies (3.3.1) on $\mathbb{R}_+^n$ for some K. Define $\beta_i : \mathbb{R}_+^i \to \mathbb{R}$ recursively by setting $\beta_n = 0$ and

$$\begin{aligned} &\beta_i(s_1, \ldots, s_i) \\ &= \sup_{s_{i+2}, \ldots, s_n \in \mathbb{R}_+} \limsup_{x \to \infty} \Bigg(\Big(\beta_{i+1}(s_1, \ldots, s_i, x) \\ &\quad + \frac{G(s_1, \ldots, s_i, x, s_{i+2}, \ldots, s_n)}{x} \Big) \mathbb{1}_{\mathfrak{P}}(s_1, \ldots, s_i, x, s_{i+2}, \ldots, s_n) \Bigg) \vee 0, \end{aligned} \tag{3.4.4}$$

for $i = 0, \ldots, n-1$. If $G_\beta(\mathbb{S}) := G(\mathbb{S}) - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)$ is upper semi-continuous and bounded from above then

$$V^{(p)}_{\vec{\mu}, \mathfrak{P}}(G) = \sup_{\mathbb{P} \in \mathcal{M}^-_{\vec{\mu}, \mathfrak{P}}} \mathbb{E}_{\mathbb{P}} \Big[G - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i) \Big] = P_{\vec{\mu}, \mathfrak{P}}(G_\beta).$$

More generally, the result remains true if there exists a sequence $(0, \beta^{(N)}) \in \mathcal{A}_p$ such that $G_{\beta^{(N)}}(\mathbb{S})$ is upper semi-continuous, bounded from above on $\mathfrak{P}$ for every N and $G_{\beta^{(N)}}(\mathbb{S}) \to G_\beta(\mathbb{S})$ pointwise as $N \to \infty$.

The proof is reported in Section 3.6.6. Here we show how the above result applies in the case of an Asian or a lookback option when $\mathfrak{P} = \Omega$.

Remark 3.4.6. This result stands in stark contrast with the existing literature on pricing under constraints on short-selling. For example, in a general (classical) setting, where prices are assumed to be locally bounded semimartingales under some probability measure $\mathbb{P}$, and under a restriction on short-selling, Pulido [89, Theorem 4.1] shows that there is no duality gap.

Example 1 An Asian option has payoff function $G : \mathbb{R}_+^n \to \mathbb{R}_+$ defined by

$$G(s_1, \dots, s_n) = \Big(\frac{\sum_{i=1}^n s_i}{n} - K \Big)^+.$$

In this case, as for any $i = 1, \dots, n$ and $s_1, \dots, s_i, s_{i+2}, \dots, s_n \in \mathbb{R}_+$, we have

$$\lim_{x \to \infty} \frac{G(s_1, \dots, s_i, x, s_{i+2}, \dots, s_n)}{x} = \frac{1}{n},$$

(3.4.4) can be simplified to

$$\beta_i(s_1, \dots, s_i) = \sup_{s_{i+2}, \dots, s_n \in \mathbb{R}_+} \limsup_{x \to \infty} \beta_{i+1}(s_1, \dots, s_i, x) + \frac{1}{n}.$$

This yields $\beta_i = (n-i)/n$ for $i = 0, \dots, n-1$. It is clear that

$$G_\beta(\mathbb{S}) = G(\mathbb{S}) - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_{i+1} - \mathbb{S}_i) = \Big(\frac{\sum_{i=1}^n \mathbb{S}_i}{n} - K \Big)^+ - \frac{\sum_{i=1}^n \mathbb{S}_i}{n} + s_0$$

is continuous and bounded from above. Therefore, by Theorem 3.4.5

$$\begin{aligned} V_{\vec{\mu}}^{(p)}(G) &= \sup_{\mathbb{P} \in \mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}} \Big[G - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_{i+1} - \mathbb{S}_i) \Big] \\ &= \sup_{\mathbb{P} \in \mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G] + \frac{1}{n} \sum_{i=1}^n \Big(s_0 - \int_{\mathbb{R}_+} x \mu_i(\mathrm{d}x) \Big). \end{aligned}$$

Example 2 The second example we consider is a *lookback option with a knock-in feature*, whose payoff function $G : \mathbb{R}_+^n \to \mathbb{R}_+$ is given by

$$G(s_1, \dots, s_n) = \big(\max_{0 \le i \le n} s_i - K \big)^+ \mathbb{1}_{\{\min_{0 \le i \le n} s_i \le B\}}.$$

In particular, when $B = \infty$, it is just a lookback call option with strike K. By (3.4.4),

$$\beta_{n-1}(s_1, \dots, s_{n-1}) = \sup_{s_n \in \mathbb{R}_+} \limsup_{x \to \infty} \frac{G(s_1, \dots, s_{n-1}, x)}{x} = \mathbb{1}_{\{\min_{0 \le i \le n-1} s_i \le B\}}.$$

Since for $i = 0, \dots, n-2$ and $s_1, \dots, s_i \in \mathbb{R}_+$, we have

$$\alpha_i(s_1, \dots, s_i) := \sup_{s_{i+2}, \dots, s_n \in \mathbb{R}_+} \lim_{x \to \infty} \frac{G(s_1, \dots, s_i, x, s_{i+2}, \dots, s_n)}{x} = 1.$$

(3.4.4) can then be simplified to

$$\beta_i(s_1, \dots, s_i) = \sup_{s_{i+2}, \dots, s_n \in \mathbb{R}_+} \limsup_{x \to \infty} \beta_{i+1}(s_1, \dots, s_i, x) + 1, \quad i = 0, \dots, n-2,$$

from which we can derive

$$\beta_i(s_1,\dots,s_i)=(n-1-i)+\mathbb{1}_{\{\min\limits_{0\le j\le i} s_j\le B\}},\quad i=0,\dots,n-1.$$

It is not hard to see that

$$\begin{aligned}
G_\beta(\mathbb{S}) =& G(\mathbb{S})-\beta_0(\mathbb{S}_1-s_0)-\sum_{i=1}^{n-1}\beta_i(\mathbb{S}_1,\dots,\mathbb{S}_i)(\mathbb{S}_{i+1}-\mathbb{S}_i)\\
=& G(\mathbb{S})+\beta_0 s_0-\sum_{i=0}^{n-2}\big(\beta_i(\mathbb{S}_1,\dots,\mathbb{S}_i)-\beta_{i+1}(\mathbb{S}_1,\dots,\mathbb{S}_{i+1})\big)\mathbb{S}_{i+1}\\
&-\beta_{n-1}(\mathbb{S}_1,\dots,\mathbb{S}_{n-1})\mathbb{S}_n\\
=& ns_0+\Big(\max_{0\le i\le n}\mathbb{S}_i-K\Big)^+\mathbb{1}_{\{\min\limits_{0\le i\le n}\mathbb{S}_i\le B\}}-\sum_{i=1}^{n}\mathbb{S}_i\\
&+\sum_{i=1}^{n}\mathbb{1}_{\{\min_{0\le j\le i-1}\mathbb{S}_j>B\}}\mathbb{1}_{\{\mathbb{S}_i\le B\}}\mathbb{S}_i
\end{aligned}$$

is bounded from above. Now define continuous functions $\beta_i^{(N)}:\mathbb{R}_+^i\to\mathbb{R}_+$ by

$$\beta_i^{(N)}(s_1,\dots,s_i)=\begin{cases} n-i & \text{if } \min\limits_{0\le j\le i} s_j\le B\\ n-i-1+N\Big(B+\frac{1}{N}-\min_{0\le j\le i} s_j\Big)^+ & \text{otherwise.}\end{cases}$$

Similarly, we can show $G_{\beta^{(N)}}(\mathbb{S})=G(\mathbb{S})-\sum_{i=0}^{n-1}\beta_i^{(N)}(\mathbb{S}_1,\dots,\mathbb{S}_i)(\mathbb{S}_{i+1}-\mathbb{S}_i)$ is bounded from above. Also $\beta_i^{(N)}\to\beta_i$ as $N\to\infty$ for any $i=0,\dots,n-1$. Then $G_{\beta^{(N)}}\to G_\beta$ pointwise as $N\to\infty$ and hence by Theorem 3.4.5

$$V_{\vec{\mu}}^{(p)}(G)=\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-}\mathbb{E}_{\mathbb{P}}\Big[G-\sum_{i=0}^{n-1}\beta_i(\mathbb{S}_1,\dots,\mathbb{S}_i)(\mathbb{S}_{i+1}-\mathbb{S}_i)\Big].$$

As shown in the lookback option example above, the duality gap is not only dependent on G and the marginal distributions μ_i, but also on how the μ_i are (optimally) transported. In the case that $\mathfrak{P}$ is a strict subset of Ω, it may become increasingly hard to calculate β and check the assumption of Theorem 3.4.5. We develop now an argument which connects asymptotically the duality gap of G in the presence of a prediction set and the duality gaps of penalised functions of G in the absence of a prediction set. In particular, it provides an alternative way to compute the duality gap when $\mathfrak{P}$ is an arbitrary closed set.

Assume the market input $(\mathcal{X}_p, \mathcal{P}, \mathfrak{P})$ admits no WFLVR and G is upper-semi continuous subject to

$$G(s_1, \ldots, s_n) \leq K(1 + s_1 + \ldots + s_n), \quad (s_1, \ldots, s_n) \in \mathbb{R}_+.$$

Under this assumption, we argue first that if $P_{\vec{\mu},\mathfrak{P}}(G) = -\infty$, then $V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = -\infty$. By Proposition 3.4.2, absence of WFLVR is equivalent to $\mathcal{M}^-_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}} \neq \varnothing$. It follows from the sublinearity of $V^{(p)}_{\vec{\mu},\mathfrak{P}}$ and Theorem 3.4.3 that

$$\begin{aligned} V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) &\leq V^{(p)}_{\vec{\mu},\mathfrak{P}}\big(G - K(1 + \mathbb{S}_1 + \cdots + \mathbb{S}_n)\big) + V^{(p)}_{\vec{\mu},\mathfrak{P}}\big(K(1 + \mathbb{S}_1 + \cdots + \mathbb{S}_n)\big) \\ &\leq P_{\vec{\mu},\mathfrak{P}}(G - K(1 + \mathbb{S}_1 + \cdots + \mathbb{S}_n)) + nKs_0 + K = -\infty. \end{aligned}$$

From now on we make an additional assumption that $P_{\vec{\mu},\mathfrak{P}}(G) > -\infty$. Define $G^{(N)} : \mathbb{R}^n_+ \to [-\infty, \infty)$ by

$$G^{(N)}(s_1, \ldots, s_n) = G(s_1, \ldots, s_n) - N\lambda_{\mathfrak{P}}(s_1, \ldots, s_n),$$

where $\lambda_{\mathfrak{P}}(\mathbb{S}) := (1 + \mathbb{S}_1 + \ldots + \mathbb{S}_n)\mathbb{1}_{\{(\mathbb{S}_1,\ldots,\mathbb{S}_n)\notin\mathfrak{P}\}}$ is as defined in (3.6.9). Then note that

$$V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = \inf_{N\geq 1} V^{(p)}_{\vec{\mu}}(G^{(N)}). \tag{3.4.5}$$

The inequality "$\leq$" is clear. On the other hand, "$\geq$" follows from the fact that $G^{(N)}$ is decreasing in N and given any $(X, H) \in \mathcal{A}_p$, $\Psi_{X,H} \geq -N(1 + \sum_{i=1}^N \mathbb{S}_i)$ for N sufficiently large.

Since $\mathfrak{P}$ is closed, $-\mathbb{1}_{\{(\mathbb{S}_1,\ldots,\mathbb{S}_n)\notin\mathfrak{P}\}}$ is an upper semi-continuous function and hence $G^{(N)}$ is upper semi-continuous. Then the problem is reduced to the case that $\mathfrak{P} = \Omega$, for which we have a formula to calculate the duality gap if the contingent claim satisfies all the assumptions in Theorem 3.4.5. Now let

$$\gamma_N := V^{(p)}_{\vec{\mu}}(G^{(N)}) - \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G^{(N)}].$$

It follows by (3.4.5) that

$$V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = \inf_{N\geq 1} \Big\{ \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G^{(N)}] + \gamma_N \Big\} = \lim_{N\to\infty} \Big\{ \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G^{(N)}] + \gamma_N \Big\}.$$

In addition, we can deduce that

$$\inf_{N\geq 1} \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G^{(N)}] = \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu}}} \inf_{N\geq 1} \mathbb{E}_{\mathbb{P}}[G^{(N)}] = \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G],$$

where the first equality is achieved by using the min–max theorem (Corollary 2 in Terkelsen [102]) and the second equality holds as $\inf_{N\geq 1} \mathbb{E}_{\mathbb{P}}[G - N\lambda_{\mathfrak{P}}] = -\infty$ for any $\mathbb{P} \in (\mathcal{M}^-_{\vec{\mu}} \setminus \mathcal{M}^-_{\vec{\mu},\mathfrak{P}})$ but

$$\inf_{N\geq 1} \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu}}} \mathbb{E}_{\mathbb{P}}[G^{(N)}] \geq \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G] > -\infty.$$

Hence, the limit of γ_N exists and by writing $\gamma = \lim_{N\to\infty} \gamma_N$, we have

$$V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G] + \gamma.$$

3.5 Continuous time: local martingales, bubbles and pricing

We now turn to continuous-time models to explore the link between options' prices, trading constraints, speculative bubbles and strict local martingales. Let $\Omega = \mathbb{D}([0,T], \mathbb{R}_+)$ be the space of nonnegative right-continuous functions with left limits on $[0,T]$ and $\mathbb{S} = (\mathbb{S}_t : t \leq T)$ be the canonical process on Ω with $(\mathcal{F}_t)$ denoting its natural filtration.

Now consider the case when put options are traded for $n \geq 1$ maturities $0 < T_1 < \ldots < T_n = T$,

$$\mathcal{X}_p := \{(K - \mathbb{S}_{T_i})^+ : 1 \leq i \leq n, K \geq 0\}, \quad p_i(K) := \mathcal{P}((K - \mathbb{S}_{T_i})^+).$$

We need to impose some assumptions on the prediction set $\mathfrak{P}$.

Assumption 3.5.1. The prediction set $\mathfrak{P} \subset \Omega$ satisfies $\omega(0) = s_0$ for every $\omega \in \mathfrak{P}$ and

$$\text{for any } \omega \in \mathfrak{P} \text{ and any stopping time } \tau, \quad \omega^\tau = (\omega(t \wedge \tau(\omega)) : t \leq T) \in \mathfrak{P}.$$

Further the set $\mathfrak{P}_{\vec{T}} := \{(\omega_0, \omega_{T_1}, \ldots, \omega_{T_n}) : \omega \in \mathfrak{P}\}$ is closed.

The first condition corresponds to $\mathfrak{P}$ being closed under stopping and will imply that any superhedging strategy in fact satisfies a collateral requirement, see Remark 3.5.2 below. The second point is technical and will enable us to compare the continuous-time setting to the discrete-time setting.

There are several possible choices for the class of admissible dynamic trading strategies. They will typically lead to the same superhedging price, provided the admissible class is large enough, but to different sets of market-calibrated models. Here, to make the connection with discrete-time setup clearer, we consider dynamic trading strategies H which are predictable piecewise constant processes with finitely many jumps. More precisely, we consider $H : [0,T] \times \Omega \to \mathbb{R}$ such that for any $\omega \in \Omega$, $H(\omega) : [0,T] \to \mathbb{R}_+$ is a simple nonnegative function (piecewise constant with finitely many jumps) and for any $t \in [0,T]$ and for any $\omega_1, \omega_2 \in \Omega$ such that $\omega_1(s) = \omega_2(s)$ for $s \in [0,t)$ we have $H_t(\omega_1) = H_t(\omega_2)$. We call such H is admissible and write $H \in \mathcal{A}$. Note that for $H \in \mathcal{A}$, the stochastic integral $\int_0^t H_{u-} \mathrm{d}\mathbb{S}_u$ is a sum and hence defined pathwise.

An admissible semi-static trading strategy is a pair (X,H) with a linear combination of put options $X(\omega) = a_0 + \sum_{i=1}^m a_i X_i(\omega)$, $m \geq 0$, $a_i \in \mathbb{R}$, $X_i \in \mathcal{X}_p$ and $H \in \mathcal{A}$. Its payoff at time T is given by

$$\Psi_{X,H}(\mathbb{S}) = X(\mathbb{S}) + \int_0^T H_{u-} \mathrm{d}\mathbb{S}_u.$$

We recall that the family of admissible semi-static trading strategies is denoted by $\mathcal{A}_p = \mathcal{A}_{\mathcal{X}_p}$ and the superreplication price $V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}$ was given in Definition 3.2.1.

Remark 3.5.2. Note that because $\mathfrak{P}$ is closed under stopping (cf. Assumption 3.5.1), it follows that if $(X,H) \in \mathcal{A}_p$ superhedges G on $\mathfrak{P}$, then in fact

$$\Psi_{X,H}(\mathbb{S}^t) = X(\mathbb{S}^t) + \int_0^t H_{u-} \mathrm{d}\mathbb{S}_u \geq G(\mathbb{S}^t), \quad t \leq T, \text{ on } \mathfrak{P}, \tag{3.5.1}$$

where $\mathbb{S}^t = (\mathbb{S}_{u\wedge t} : u \leq T)$. In other words, (X,H) satisfies a collateral requirement. As we shall see below, this feature will contribute towards the emergence of bubbles.

Static trading arguments, as in the proof of Propositions 3.3.2, show that absence of WFLVR implies that $p_i(K)$ satisfy the properties listed in Condition 3.4.1 and hence we can use (3.4.1) to define probability measures $\vec{\mu} = (\mu_i)_{i=1}^n$ which satisfy Assumption 3.3.8. The set of market-calibrated models $\mathcal{M}^-_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ is as given in Definition 3.2.2. Note that Remark 3.2.3 is in force, with the convention $\infty - \infty = -\infty$. Finally, let $\mathcal{M}^{\mathrm{loc}}_{\vec{\mu},\mathfrak{P}}$ denote the set of all calibrated local martingale measures on $(\Omega, \mathbb{F}_T)$, i.e., all $\mathbb{P}$ such that $\mathbb{S}$ is a $\mathbb{P}$-local martingale and $\mathbb{E}_{\mathbb{P}}[(K-\mathbb{S}_{T_i})^+] = p_i(K)$, $K \geq 0$, or equivalently $\mathbb{S}_{T_i} \sim \mu_i$, $1 \leq i \leq n$. It is easy to see that $\mathcal{M}^-_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ is the set of measures $\mathbb{P}$ under which $\mathbb{S}_{T_i} \sim \mu$ and $\mathbb{S}$ is a $\mathbb{P}$-supermartingale and in particular $\mathcal{M}^{\mathrm{loc}}_{\vec{\mu},\mathfrak{P}} \subset \mathcal{M}^-_{\mathcal{X},\mathcal{P},\mathfrak{P}}$.

Consider now a European option with payoff $G(\omega) = G(\omega_{T_n})$, or more generally an upper semi-continuous $G(\omega) = G(\omega_{T_1}, \dots, \omega_{T_n})$. We can then compare the present setting to that of a discrete n-period model with traded put options at prices $\mathcal{P}$, where short-selling is prohibited, and with a prediction set $\mathfrak{P}_{\vec{T}}$, as considered in Section 3.4. Denote the corresponding primal and dual values by $P^d_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}_{\vec{T}}}$ and $V^d_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}_{\vec{T}}}$. Note that the discrete superhedging problem naturally embeds into the continuous-time one. A discrete-time trading strategy corresponds to a nonnegative (H_t) constant on every $[T_i, T_{i+1})$ for $i = 0, \dots, n-1$, which is in $\mathcal{A}$. On the primal side, for any $\mathbb{P} \in \mathcal{M}^{\text{loc}}_{\vec{\mu},\mathfrak{P}}$ the vector $(\mathbb{S}_0, \mathbb{S}_{T_1}, \dots, \mathbb{S}_{T_n})$ is a market-calibrated discrete-time model; so for G as above, continuous-time calibrated models are embedded in discrete ones. In summary,

$$P_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) \le P^d_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}_{\vec{T}}}(G) \quad \text{and} \quad V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) \le V^d_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}_{\vec{T}}}(G).$$

In some cases we can establish an equality in the first inequality. For example, when

$$\mathfrak{P} = \{\omega \in \Omega : \omega(0) = s_0\}$$

then any discrete-time market model may be seen as a continuous-time one with the asset being constant on any $[T_i, T_{i+1})$. We can then conclude that there is no duality gap in the continuous-time setting from the results Theorem 3.4.3 and Corollary 3.4.4 in discrete time.

However, our prime interest is in the case when the pricing–hedging duality fails. We can use the results of Section 3.4.2 to understand the case of European options.

Proposition 3.5.3. Suppose that the prediction set $\mathfrak{P}$ satisfies Assumption 3.5.1 and $\mathfrak{P}_T := \{\omega(T) : \omega \in \mathfrak{P}\}$ is unbounded, and that market prices are such that $\mathcal{M}^{\text{loc}}_{\vec{\mu},\mathfrak{P}} \neq \varnothing$. Consider a continuous function $G : \mathbb{R}_+ \to [-\infty, \infty)$ with linear growth (i.e., (3.3.1) holds) such that the limit $\beta := \lim_{s\to\infty, s\in\mathfrak{P}_T} \frac{G(s)}{s}$ is well defined and nonnegative. Then,

$$\begin{aligned} V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) &= \sup_{\mathbb{P}\in\mathcal{M}^{\text{loc}}_{\vec{\mu},\mathfrak{P}}} \lim_{n\to\infty} \mathbb{E}_{\mathbb{P}}[G_+(\mathbb{S}_{T\wedge\tau_n^{\mathbb{P}}}) - G_-(\mathbb{S}_T)] \\ &= \int G(s)\mu_n(\mathrm{d}s) + \beta\Big(s_0 - \int_{\mathbb{R}_+} s\mu_n(\mathrm{d}s)\Big), \end{aligned}$$

where we implicitly set $G(\omega) = G(\omega(T))$, $(\tau_n^{\mathbb{P}})$ is a localising sequence for $\mathbb{S}$ under $\mathbb{P}$, and $G_+ = G \vee 0$, $G_- = -(G \wedge 0)$.

We first give two remarks before proving the above result.

Remark 3.5.4. If the forward price implicit in the put options,

$$f_0 = \int s\mu_n(\mathrm{d}s) = \lim_{K\to\infty}(K - p_n(K)),$$

is cheaper than the spot, $s_0 > f_0$, then the market has a bubble. The *market price* s_0 is strictly greater than the *fundamental price*, given by

$$\sup_{\mathbb{P}\in\mathcal{M}^{\mathrm{loc}}_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}_T)] = \int G(s)\mu_n(\mathrm{d}s).$$

The correction is equal to $\beta^+(s_0 - f_0)$. This is the same correction as exhibited in Theorem 5.2 in Cox and Hobson [26], see also Section 6.1 in Jarrow et al. [67]. In Cox and Hobson [26], the bubble was driven by a collateral requirement and a strict local martingale property. While the former is a natural trading restriction the latter appears artificial. In our robust framework a bubble is triggered by trading restrictions and properties of market prices of options. The difference is that we take market prices as given and adopt a robust framework. A bubble arises when these prices are misaligned with the asset price, $s_0 > f_0$, while an arbitrage does not arise because of the trading restrictions. In our setup, the trading restrictions take the form of a short selling ban and, as highlighted in Remark 3.5.2 above, a collateral requirement.

Remark 3.5.5. The assumption $\mathcal{M}^{\mathrm{loc}}_{\vec{\mu},\mathfrak{P}} \neq \varnothing$ is an implicit assumption on $\mathfrak{P}$ and market prices. It is satisfied e.g. when $\mathfrak{P}$ is equal to all paths, or all continuous paths, which start in s_0, and put prices $p_i(K)$ satisfy the properties listed in Condition 3.4.1. The latter is equivalent to $\vec{\mu} = (\mu_i)_{i=1}^n$, defined via (3.4.1), satisfying Assumption 3.3.8.

Proof of Proposition 3.5.3. As explained above, we can directly compare the continuous-time setting with a discrete-time setting from Section 3.4 with the same put prices and prediction set

$$\mathfrak{P}_{\vec{T}} = \{(\omega_0, \omega_{T_1}, \ldots, \omega_{T_n}) : \omega \in \mathfrak{P}\}.$$

Using (3.4.3), which is a one-marginal result, we immediately have

$$V^d_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}_{\vec{T}}}(G) \leq V^p_{\mu_n,\mathfrak{P}_T}(G) = \int G(s)\mu_n(\mathrm{d}s) + \beta^+\Big(s_0 - \int_{\mathbb{R}_+} s\mu_n(\mathrm{d}s)\Big)$$

and hence we conclude that

$$V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G) \leq \int G(s)\mu_n(\mathrm{d}s) + \beta^+\Big(s_0 - \int_{\mathbb{R}_+} s\mu_n(\mathrm{d}s)\Big). \tag{3.5.2}$$

Consider a superhedging strategy (X, H) and $\mathbb{P} \in \mathcal{M}^{\text{loc}}_{\bar{\mu},\mathfrak{P}}$ with a reducing sequence (τ_n) for $\mathbb{S}$ under $\mathbb{P}$. If $G(\mathbb{S}_{T\wedge\tau_n}) > 0$, it follows from (3.5.1) and $G_- \geq 0$ that

$$X(\mathbb{S}_{T\wedge\tau_n}) + \int_0^{T\wedge\tau_n} H_{u-}\mathrm{d}\mathbb{S}_u \geq G_+(\mathbb{S}_{T\wedge\tau_n}) - G_-(\mathbb{S}_T).$$

Otherwise, $G(\mathbb{S}_{T\wedge\tau_n}) \leq 0$ and then

$$X(\mathbb{S}_T) + \int_0^T H_{u-}\mathrm{d}\mathbb{S}_u \geq G(\mathbb{S}_T) \geq G_+(\mathbb{S}_{T\wedge\tau_n}) - G_-(\mathbb{S}_T)$$

Therefore

$$\begin{aligned} G_+(\mathbb{S}_{T\wedge\tau_n}) - G_-(\mathbb{S}_T) \leq X(\mathbb{S}_T) &+ \big(X(\mathbb{S}_{T\wedge\tau_n}) - X(\mathbb{S}_T)\big)\mathbb{1}_{\{G(\mathbb{S}_{T\wedge\tau_n})>0\}} \\ &+ \int_0^T \tilde{H}_{u-}\mathrm{d}\mathbb{S}_u, \end{aligned}$$

where $\tilde{H}_u = H_u\mathbb{1}_{\{u\leq\tau_n\wedge T\}} + H_u\mathbb{1}_{\{u>\tau_n\wedge T\}}\mathbb{1}_{\{G(\mathbb{S}_{T\wedge\tau_n})\leq 0\}}$. We note that $\tilde{H} \in \mathcal{A}$ and hence the expectation of the integral is nonpositive under $\mathbb{P}$. Further, $\tau_n \wedge T = T$ for n large enough (which may depend on the path) and X is bounded so we may apply the dominated convergence theorem to conclude that

$$\limsup_{n\to\infty} \mathbb{E}_{\mathbb{P}}[G_+(\mathbb{S}_{T\wedge\tau_n}) - G_-(\mathbb{S}_T)] \leq \mathbb{E}_{\mathbb{P}}[X(\mathbb{S}_T)] = \mathcal{P}(X) \tag{3.5.3}$$

and hence the LHS is a lower bound on $V_{\mathcal{X}_p,\mathcal{P},\mathfrak{P}}(G)$. Finally we compute the LHS. Note that for any $\epsilon > 0$, $G(s) - (\beta - \epsilon)s$ is bounded from below on $\mathfrak{P}_T$. It follows, applying Fatou's Lemma and noting that $\mathbb{P} \in \mathcal{M}^{\text{loc}}_{\bar{\mu},\mathfrak{P}}$ implies $\mathbb{S}_T$ and $\mathbb{S}_{T\wedge\tau_n}$ are almost surely in $\mathfrak{P}_T$, that

$$\begin{aligned} \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}[G_+(\mathbb{S}_{T\wedge\tau_n})] &\geq \liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}}[G_+(\mathbb{S}_{T\wedge\tau_n}) - (\beta-\epsilon)\mathbb{S}_{T\wedge\tau_n}] + (\beta-\epsilon)s_0 \\ &\geq \mathbb{E}_{\mathbb{P}}[G_+(\mathbb{S}_T) - (\beta-\epsilon)\mathbb{S}_T] + (\beta-\epsilon)s_0 \\ &= \int G_+(s)\mu_n(\mathrm{d}s) + (\beta-\epsilon)\left(s_0 - \int s\mu_n(\mathrm{d}s)\right). \end{aligned}$$

We conclude that the upper bound in (3.5.2) coincides with the lower bound obtained by taking $\epsilon \searrow 0$ and the infimum over superhedging strategies in (3.5.3), as required.

□

The above statement may be extended to G which depends on the values of the asset at the intermediate maturities, i.e., $G(\mathbb{S}) = G(\mathbb{S}_{T_1}, \ldots, \mathbb{S}_{T_n})$, by using Theorem 3.4.5. We do not pursue this here.

3.6 Proofs

3.6.1 Preliminary Results

In this section and in Sections 3.6.3 and 3.6.5, we assume that $\mu_1, \dots, \mu_n$ are probability measures on $\mathbb{R}_+$ which have a finite first moment. Let $\Pi_{\vec{\mu}}$ be the set of all Borel probability measures on Ω with marginals $\delta_{s_0}, \mu_1, \dots, \mu_n$ and denote by $\mathcal{M}^-_{\vec{\mu}}$ the set of probability measures $\mathbb{P}$ on Ω such that $\mathbb{S}$ is a supermartingale and $\mathbb{S}_i$ is distributed according to μ_i. We also write $C_b(\mathbb{R}^j_+, \mathbb{R}_+)$ to denote the set of continuous, bounded and nonnegative functions f on $\mathbb{R}^j_+$ and $C_c(\mathbb{R}^j_+, \mathbb{R}_+)$ for the subset of continuous nonnegative and compactly supported functions.

Lemma 3.6.1. Let $\pi \in \Pi_{\vec{\mu}}$. Then the following are equivalent:

1. $\pi \in \mathcal{M}^-_{\vec{\mu}}$.
2. For $0 \le j \le n-1$ and for every $H \in C_c(\mathbb{R}^j_+, \mathbb{R}_+)$, we have

$$\int_\Omega H(x_1, \dots, x_j)(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \dots, x_n) \le 0 \tag{3.6.1}$$

Proof of Lemma 3.6.1. Claim (1) asserts that whenever $A \subseteq \mathbb{R}^j_+$, $j \le n-1$, is Borel measurable, then

$$\int_{\Omega_{s_0}} \mathbb{1}_A(x_1, \dots, x_j)(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \dots, x_n) \le 0. \tag{3.6.2}$$

To see (3.6.2) $\Rightarrow$ (3.6.1), we fix any $j = 0, \dots, n-1$ and $H \in C_c(\mathbb{R}^j_+, \mathbb{R}_+)$ and define simple functions $f_k : \mathbb{R}^j_+ \to \mathbb{R}$ by $f_k = 2^{-k}\lfloor 2^k H \rfloor$. Then $0 \le f_k \uparrow H$ and it follows from the dominated convergence theorem and (3.6.2) that (3.6.1) is satisfied.

To show (3.6.1) $\Rightarrow$ (3.6.2), first consider $A \in \mathbb{R}^j_+$ such that A is open and bounded. Since $\mathbb{1}_A$ is lower semi-continuous, there exists a sequence $(f_k)_{k\ge1} \subseteq C_c(\mathbb{R}^j_+, \mathbb{R}_+)$ such that $0 \le f_k \le \mathbb{1}_A$ and $f_k \uparrow \mathbb{1}_A$. Therefore, the dominated convergence theorem implies that

$$\int_\Omega \mathbb{1}_A(x_1, \dots, x_j)(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \dots, x_n) \le 0.$$

Now consider an arbitrary open set $A \in \mathbb{R}^j_+$. We can write $A = \cup_{n\ge1} A^{(n)}$ with $A^{(n)} := A \cap \{S \in \mathbb{R}^j_+ : \|S\| < n\}$ being open and bounded. Then by the dominated convergence theorem

$$\int_\Omega \mathbb{1}_A(x_1, \dots, x_j)(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \dots, x_n) \le 0.$$

Finally, if $A \subseteq \mathbb{R}_+^j$ is a Borel set, then by Corollary 3.12 in Bruckner et al. [15], for every $N > 0$, there is an open set $A_N \subseteq \mathbb{R}_+^j$ such that $A \subseteq A_N$ with $\pi(A_N) \leq \pi(A) + \frac{1}{N}$. It follows that

$$\begin{aligned}
&\int_\Omega \mathbb{1}_A(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \ldots, x_n) \\
&= \int_\Omega \mathbb{1}_{A_N}(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \ldots, x_n) - \int_\Omega \mathbb{1}_{A_N \setminus A}(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \ldots, x_n) \\
&\leq \int_\Omega \mathbb{1}_{A_N \setminus A} x_j \mathrm{d}\pi(x_1, \ldots, x_n) \\
&= \int_{A_N \setminus A} \mathbb{1}_{\left\{x_j \geq \sqrt{N}\right\}} x_j \mathrm{d}\pi(x_1, \ldots, x_n) + \int_{A_N \setminus A} \mathbb{1}_{\left\{x_j < \sqrt{N}\right\}} x_j \mathrm{d}\pi(x_1, \ldots, x_n) \\
&\leq \int_{\mathbb{R}_+} \mathbb{1}_{\left\{x_j \geq \sqrt{N}\right\}} x_j \mu_j(\mathrm{d}x) + \frac{\sqrt{N}}{N} \quad \to 0 \text{ as } N \to \infty.
\end{aligned}$$

□

Lemma 3.6.2. For a closed $\mathfrak{P} \subseteq \Omega$ the set $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$ is compact in the weak topology.

Proof. Since $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$ is a subset of the compact set $\Pi_{\vec{\mu}}$, it suffices to prove that $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$ is a closed subset of $\Pi_{\vec{\mu}}$. By Lemma 3.6.1,

$$\mathcal{M}^-_{\vec{\mu}} = \bigcap_{j=0}^{n-1} \bigcap_{H \in C_b(\mathbb{R}_+^j, \mathbb{R}_+)} \left\{ \pi \in \Pi_{\vec{\mu}} : \int_{\mathbb{R}_+^n} H(x_1, \ldots, x_j)(x_{j+1} - x_j)\mathrm{d}\pi(x_1, \ldots, x_n) \leq 0 \right\}.$$

Therefore, by Lemma 2.2 in Beiglböck et al. [6], $\mathcal{M}^-_{\vec{\mu}}$ is a closed subset of $\Pi_{\vec{\mu}}$ in the weak topology.

To show $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$ is a closed subset of $\mathcal{M}^-_{\vec{\mu}}$, we take any sequence $(\mathbb{Q}_n) \in \mathcal{M}_{\vec{\mu}}$ such that $\mathbb{Q}_n(\mathfrak{P}) = 1$ and $\mathbb{Q}_n \to \mathbb{Q}$ for some $\mathbb{Q} \in \mathcal{M}_{\vec{\mu}}$ as $n \to \infty$. Then by weak convergence of measures, for $\mathfrak{P} \subseteq \Omega$ closed, $\mathbb{Q}(\mathfrak{P}) \geq \limsup_{n\to\infty} \mathbb{Q}_n(\mathfrak{P}) = 1$. It follows that $\mathcal{M}_{\vec{\mu},\mathfrak{P}}$ is a closed subset of $\mathcal{M}_{\vec{\mu}}$ and hence closed in $\Pi_{\vec{\mu}}$ in the weak topology. □

To prove Theorems 3.3.5 and 3.4.3, we use the following version of Monge-Kantorovich duality theorem, which is essentially proposition 2.1 in Beiglböck et al. [6]. The proposition is rewritten here to suit the notation and purpose of this paper.

Lemma 3.6.3. For any G that is upper semi-continuous and bounded from above we have

$$\sup_{\pi \in \Pi_{\vec{\mu}}} \mathbb{E}_\pi[G] = \inf\{\mathcal{P}(X) : X \in \mathcal{A}_o, \text{ s.t. } X \geq G \text{ on } \Omega\}, \quad \text{where } o \in \{c, p\}.$$

Further, the result remains true with $o = c$ for any upper-semi continuous G that satisfies 3.3.1.

The call option case is just Proposition 2.1 in Beiglböck et al. [6]. The put option case follows from Equation (A.1) in the proof of Proposition 2.1 in Beiglböck et al. [6], which in our notation simply states that for any G that is upper semi-continuous and bounded from above

$$\sup_{\pi\in\Pi_{\vec{\mu}}} \mathbb{E}_\pi[G] = \inf\Big\{ \sum_{i=0}^{n} \int u_i \mathrm{d}\mu_i : u_i \in C_b(\mathbb{R}_+, \mathbb{R}) \text{ s.t. } \sum_{i=1}^{n} u_i(\mathbb{S}_i) \geq G(\mathbb{S}) \text{ on } \Omega \Big\}. \tag{3.6.3}$$

Note that given any $f \in C_b(\mathbb{R}_+, \mathbb{R})$, $\epsilon > 0$ and $i = 1, \ldots, n$, there is some $u : \mathbb{R}_+ \to \mathbb{R}$ of the form $a_0 + \sum_{j=1}^{m} a_j(K_j - s_j)^+$ with $u \geq f$ and $\int (u - f)\mathrm{d}\mu_i < \epsilon$. Therefore we may change the class of admissible functions in (3.6.3) from $C_b(\mathbb{R}_+, \mathbb{R})$ to $u : \mathbb{R}_+ \to \mathbb{R}$ taking the form $a_0 + \sum_{j=1}^{m} a_j(K_j - s)^+$.

3.6.2 Proofs of FTAP in the setting with traded calls or puts: case where $\mathfrak{P} = \Omega$

The proofs of Propositions 3.3.2 and 3.4.2 are virtually identical; so we only give the proof of Proposition 3.3.2. We include with it the proof of Proposition 3.3.3. In this section, we only give the proof in the case where $\mathfrak{P} = \Omega$. This allows us then to prove Theorem 3.3.5 when $\mathfrak{P} = \Omega$ which in turn is used to establish the results when $\mathfrak{P} \subsetneq \Omega$.

Step 1: "$\exists$ MCM (market-Calibrated Model) $\Longrightarrow$ no WFLVR"
First we show that the existence of a market-calibrated model implies no WFLVR. Fix a market-calibrated model $\mathbb{P}$ and any $(X_k, H_k) \in \mathcal{A}_{\mathcal{X}}$ and $(X, H) \in \mathcal{A}_{\mathcal{X}}$ such that $\Psi_{X_k,H_k} \to 0$ pointwise on $\mathfrak{P}$, $\lim_k \mathcal{P}(X_k)$ is well defined and $\Psi_{X_k,H_k} \geq \Psi_{X,H}$. Then by Fatou's lemma, we get

$$\liminf_k \mathbb{E}_{\mathbb{P}}[\Psi_{X_k,H_k}] \geq \mathbb{E}_{\mathbb{P}}[\liminf_k \Psi_{X_k,H_k}] = 0$$

and hence

$$\lim_k \mathcal{P}(X_k) = \lim_k \mathbb{E}_{\mathbb{P}}[X_k] \geq \liminf_k \mathbb{E}_{\mathbb{P}}[\Psi_{X_k,H_k}] \geq \mathbb{E}_{\mathbb{P}}[\liminf_k \Psi_{X_k,H_k}] = 0.$$

Step 2: "no WFLVR $\Longrightarrow$ Condition 3.3.1"
It is straightforward and classical that the absence of a robust uniformly strong arbitrage implies Condition 3.3.1 (i)–(ii). Note that since $c_i(\cdot)$ are convex $c_i(0+)$ is well defined. Let $\alpha_i := \lim_{K\to\infty} c_i(K)$ which is well-defined by Condition 3.3.1 (i) with $\alpha_i \geq 0$ for any $i = 1, \ldots, n$. If $\alpha_i > 0$ for some $i = 1, \ldots, n$ then $(X_k, (0))$ with $X_k = -(\mathbb{S}_i - k)^+$ is a WFLVR since $X_k \to 0$ pointwise as $k \to \infty$ and $\mathcal{P}(X_k) = -c_i(k) \to -\alpha_i < 0$. If Condition 3.3.1 (iv) is violated, then for some $K \in \mathbb{R}_+$ and i, $c_i(0) - c_i(K) < c_{i+1}(0) - c_{i+1}(K)$. Consider

$$X = (\mathbb{S}_i - 0)^+ - (\mathbb{S}_i - K)^+ - (\mathbb{S}_{i+1} - 0)^+ + (\mathbb{S}_{i+1} - K)^+ \in \mathcal{X}_c,$$
$$H_i = \mathbb{1}_{\{\mathbb{S}_i < K\}} \text{ and } H_j = 0 \text{ for } j \neq i.$$

Then (X, H) is a robust uniformly strong arbitrage since $\Psi_{X,H} \geq 0$, but $\mathcal{P}(X) < 0$. We conclude that no WFLVR implies Condition 3.3.1. Moreover, absence of a robust uniformly strong arbitrage implies Condition 3.3.1 (i), (ii) and (iv).
Step 3: "If $\mathfrak{P} = \Omega$ then Condition 3.3.1 $\Longrightarrow \exists$ MCM"
Next we show that Condition 3.3.1 implies the existence of a market-calibrated model when $\mathfrak{P} = \Omega$. It follows from Condition 3.3.1 (i), (ii) and (iii) that we can derive from the observed prices of call options probability measures $\vec{\mu} = (\mu_1, \ldots, \mu_n)$ with μ_i on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ such that for any $i = 1, \ldots, n$ and $K \in \mathbb{R}_+$,

$$c_i(K) := \int (x - K)^+ \mu_i(\mathrm{d}x) \quad \text{and} \quad c_i(0) = \int x \mu_i(\mathrm{d}x),$$

where $\mathcal{B}(\mathbb{R}_+)$ is the Borel σ-algebra of $\mathbb{R}_+$. In fact, due to Breeden and Litzenberger [13], μ_i can be defined via

$$\mu_i([0, K]) = 1 + c_i'(K) \qquad \text{for } K \in \mathbb{R}_+.$$

In addition, given $\mu_1, \ldots, \mu_n$ derived from the observed market prices of call options, Strassen's Theorem ([101]) states that Condition 3.3.1 (iv) holds if and only if for any convex nonincreasing function $\phi : \mathbb{R}_+ \to \mathbb{R}$, the sequence $(\int \phi \, d\mu_i)_{i\geq 1}$ is nondecreasing, which is the necessary and sufficient condition for the existence of a nonnegative supermartingale having marginals $\mu_1, \ldots, \mu_n$. Therefore, when $\mathfrak{P} = \Omega$, the absence of WFLVR implies the existence of a market-calibrated model which happens if and only if Condition 3.3.1 is satisfied.
Step 4: "Additional arguments for Proposition 3.3.3"
Given the above steps, to show Proposition 3.3.3, it remains to argue that Condition 3.3.1 (i), (ii) and (iv) imply that there is no robust uniformly strong arbitrage when

$\mathfrak{P} = \Omega$. Suppose to the contrary that there exists a semi-static strategy (X, H) such that $\Psi_{X,H} \geq 0$ and $\mathcal{P}(X) < \epsilon < 0$. As X is a finite linear combination of elements of $\mathcal{X}_c$, we let $K_{\max}$ be the largest among the strikes of call options present in X. Then, for any $\delta > 0$ small enough, there exists a sequence of functions $(c_i^{(\delta)})_{i=1}^n$ satisfying Condition 3.3.1 (i)–(iv) and such that $|c_i(K) - c_i^{(\delta)}(K)| \leq \delta$, for any $i = 1, \ldots, n$ and $K \leq K_{\max}$. In fact, we can construct functions $(c_i^{(\delta)})_{i=1}^n$ in the following way. For any $i = 1, \ldots, n$, we can first define a function $\tilde{c}_i$ by $\tilde{c}_i(K) = c_i(0) - (1 - \frac{\delta i}{2ns_0})(c_i(0) - c_i(K))$ if $c_i'(0+) < 0$, $\tilde{c}_i(K) = (c_i(0) - \frac{\delta(n+1-i)}{2nK_{\max}}K) \vee 0$ otherwise. Note that if $c_i'(0+) = 0$, then $c_j \equiv c_j(0)$ for any $j \geq i$. Then $|\tilde{c}_i(K) - c_i(K)| \leq \delta/2$ for $K \leq K_{\max}$, and for δ sufficiently small, $(\tilde{c}_i)_{i=1}^n$ satisfies Condition 3.3.1 (i)–(iv) and $\tilde{c}_i(K) - \tilde{c}_i(0)$ is strictly decreasing in i for $K \in (0, K_{\max}]$. Then, for any $i \leq n$, we can find a convex, decreasing function $c_i^{(\delta)}$ which approximates $\tilde{c}_i$ arbitrarily closely on $[0, K_{\max}]$ and satisfies $c_i^{(\delta)}(0) = \tilde{c}_i(0)$, $\tilde{c}_{i+1}(0) - \tilde{c}_{i+1} \geq c_i^{(\delta)}(0) - c_i^{(\delta)} \geq \tilde{c}_i(0) - \tilde{c}_i$ and $c_i^{(\delta)}(K) \to 0$ as $K \to \infty$. By the arguments above, with $\mathcal{P}^{(\delta)}$ corresponding to prices $(c_i^{(\delta)})$, $\mathcal{P}^{(\delta)}$ and $(c_i^{(\delta)})$ satisfy no WFLVR and hence there is no robust uniformly strong arbitrage, so $\mathcal{P}^{(\delta)}(X) \geq 0$. However, we can take δ small enough so that $|\mathcal{P}(X) - \mathcal{P}^{(\delta)}(X)| < \epsilon/2$ which gives the desired contradiction and completes the proof of Proposition 3.3.3.

3.6.3 Proof of Theorems 3.3.5 and 3.4.3: case where $\mathfrak{P} = \Omega$ and G is bounded

We now give the proof of Theorems 3.3.5 and 3.4.3 for bounded and upper semi-continuous G in the case where $\mathfrak{P} = \Omega$. In this case, we can apply Proposition 3.3.2 or 3.4.2. The proof of the general case will be postponed. Since the proof of Theorem 3.4.3 is virtually identical to that of Theorem 3.3.5, we only give the proof of Theorem 3.3.5.

Proof of Theorems 3.3.5. We first prove Theorem 3.3.5 for bounded and upper semi-continuous G in the case where $\mathfrak{P} = \Omega$.

By Proposition 3.3.2 in the case where $\mathfrak{P} = \Omega$, absence of WFLVR is equivalent to $\mathcal{M}^{-}_{\mathcal{X}_c,\mathcal{P},\Omega} \neq 0$, for which to hold Condition 3.3.1 is both necessary and sufficient. Following the classical arguments in Breeden and Litzenberger [13], by defining probability measures μ_i on $\mathbb{R}_+$ via $\mu_i([0, K]) = 1 + c_i'(K)$ for $K \in \mathbb{R}_+$, we can encode the market prices $\mathcal{P}$, or $c_i(K)$, via (μ_i) with $c_i(K) = \mathcal{P}((\mathbb{S}_i - K)^+) = \int (s - K)^+ \mu_i(\mathrm{d}s)$. Hence $\mathcal{M}^{-}_{\mathcal{X}_c,\mathcal{P},\Omega} = \mathcal{M}^{-}_{\vec{\mu}}$.

By Remark 3.2.3, to show (3.3.2), it suffices to show $V_{\mathcal{X}_c,\mathcal{P},\Omega}(G) \leq P_{\mathcal{X}_c,\mathcal{P},\Omega}(G)$.

Define $G_H : \mathbb{R}_+^n \to [-\infty, \infty)$ by

$$G_H(\mathbb{S}) := G(\mathbb{S}) - \sum_{j=0}^{n-1} H_j(\mathbb{S}_1, \ldots, \mathbb{S}_j)(\mathbb{S}_{j+1} - \mathbb{S}_j).$$

It is clear that if $H_j \in C_c(\mathbb{R}_+^j, \mathbb{R}_+)$ for every j, then $G_H(\mathbb{S})$ is upper semi-continuous and bounded, and hence satisfies (3.3.1). We can deduce that

$$\begin{aligned}
V_{\mathcal{X}_c,\mathcal{P},\Omega}(G) &= \inf_{(X,H)\in\mathcal{A}_c \text{ s.t. } \Psi_{X,H}\geq G} \mathcal{P}(X) && (3.6.4)\\
&\leq \inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \inf \Big\{\mathcal{P}(X) : X \in \mathcal{A}_c, \text{ s.t. } X \geq G_H \text{ on } \Omega\Big\} && (3.6.5)\\
&= \inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \sup_{\pi\in\Pi_{\vec{\mu}}} \Big\{\int_{\mathbb{R}_+^n} G_H(s_1, \ldots, s_n)\mathrm{d}\pi(s_1, \ldots, s_n)\Big\} && (3.6.6)\\
&= \sup_{\pi\in\Pi_{\vec{\mu}}} \inf_{H_j\in C_c(\mathbb{R}_+^n,\mathbb{R}_+)} \Big\{\int_{\mathbb{R}_+^n} G_{,H}(s_1, \ldots, s_n)\mathrm{d}\pi(s_1, \ldots, s_n)\Big\}, && (3.6.7)
\end{aligned}$$

where the equality between (3.6.5) and (3.6.6) is guaranteed by Lemma 3.6.3. In order to justify the equality between (3.6.6) and (3.6.7) we apply min–max theorem (see Corollary 2 in Terkelsen [102]) to the compact convex set $\Pi_{\vec{\mu}}$, the convex set $\mathbb{R}_+ \times C_c(\mathbb{R}_+, \mathbb{R}_+) \times \ldots \times C_c(\mathbb{R}_+^{n-1}, \mathbb{R}_+)$, and the function

$$f(\pi, (H_j)) = \int_{\mathbb{R}_+^n} G_H(s_1, \ldots, s_n)\mathrm{d}\pi(s_1, \ldots, s_n).$$

Clearly f is affine in each of the variables. Furthermore, by the Portmanteau theorem, $f(\cdot, (H_j))$ is upper semi-continuous on $\Pi_{\vec{\mu}}$. Therefore, the assumptions of Corollary 2 in Terkelsen [102] are satisfied.

The last step is to establish the following equality

$$\sup_{\pi\in\Pi_{\vec{\mu}}} \inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \Big\{\int_{\mathbb{R}_+^n} G_H(s_1, \ldots, s_n)\mathrm{d}\pi(s_1, \ldots, s_n)\Big\} = \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]. \tag{3.6.8}$$

If $\pi \notin \mathcal{M}_{\vec{\mu}}^-$, then by Lemma 3.6.1, there is a $H_j \in C_c(\mathbb{R}_+^j, \mathbb{R}_+)$ for some j such that

$$B = \int_{\mathbb{R}_+^n} H_j(s_1, \ldots, s_j)(s_{j+1} - s_j)\mathrm{d}\pi(s_1, \ldots, s_n) > 0.$$

By scaling, B can be arbitrarily large. Hence, if $\pi \notin \mathcal{M}_{\vec{\mu}}^-$, then

$$\inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \Big\{\int_{\mathbb{R}_+^n} G_H(s_1, \ldots, s_n)\mathrm{d}\pi(s_1, \ldots, s_n)\Big\} = -\infty.$$

Since G is bounded and $\mathcal{M}_{\vec{\mu}}^- \neq \varnothing$, $V_{\mathcal{X}_c,\mathcal{P}}(G) \geq \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G] > -\infty$. Therefore, in the LHS of (3.6.8), it suffices to consider $\pi \in \mathcal{M}_{\vec{\mu}}^-$ and then

$$\inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \sum_{j=0}^{n-1} \int H_j(s_1,\ldots,s_j)(s_j - s_{j+1})\mathrm{d}\pi = 0.$$

Hence

$$\begin{aligned}
&\sup_{\pi\in\mathcal{M}_{\vec{\mu}}^-} \inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \Big\{ \int_{\mathbb{R}_+^n} G_H(s_1,\ldots,s_n)\mathrm{d}\pi(s_1,\ldots,s_n)\Big\} \\
&\leq \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G] + \sup_{\pi\in\mathcal{M}_{\vec{\mu}}^-} \inf_{H_j\in C_c(\mathbb{R}_+^j,\mathbb{R}_+)} \sum_{j=0}^{n-1} \int H_j(s_1,\ldots,s_j)(s_j - s_{j+1})\mathrm{d}\pi \\
&= \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G].
\end{aligned}$$

□

3.6.4 Completing the proof of Proposition 3.3.2

In this section, we complete the proof of Proposition 3.3.2 in the case where $\mathfrak{P} \subsetneq \Omega$.

Step 5: "no WFLVR $\Longrightarrow \exists$ MCM"

It remains to argue that when $\mathfrak{P}$ is a closed subset of Ω such that $\mathfrak{P} \subsetneq \Omega$ and Condition 3.3.1 is satisfied, the non-existence of a market-calibrated model concentrated on $\mathfrak{P}$ implies the existence of a WFLVR. In fact, in this case, it is a robust uniformly strong arbitrage. Define a lower semi-continuous function $\lambda_{\mathfrak{P}} : \mathbb{R}_+^n \to \mathbb{R}$ by

$$\lambda_{\mathfrak{P}}(s_1,\ldots,s_n) = \mathbb{1}_{\{(s_1,\ldots,s_n)\notin\mathfrak{P}\}}. \tag{3.6.9}$$

Then we apply Theorem 3.3.5 to the prediction set Ω and $-\lambda_{\mathfrak{P}}$ and find that

$$V_{\mathcal{X}_c,\mathcal{P},\Omega}(-\lambda_{\mathfrak{P}}) = \sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{X}_c,\mathcal{P},\Omega}^-} \mathbb{E}_{\mathbb{P}}[-\lambda_{\mathfrak{P}}] := \alpha.$$

If $\alpha = 0$, then there exists a sequence $(\mathbb{P}_k) \in \mathcal{M}_{\mathcal{X}_c,\mathcal{P},\Omega}^-$ such that $\mathbb{P}_k(\mathfrak{P}^\complement) \to 0$. By Lemma 3.6.2, $\mathcal{M}_{\mathcal{X}_c,\mathcal{P},\Omega}^-$ is compact and closed. Hence $(\mathbb{P}_k)_{k\in\mathbb{N}}$ has a subsequence converging to some $\mathbb{P} \in \mathcal{M}_{\mathcal{X}_c,\mathcal{P},\Omega}^-$. In fact, by weak convergence of measures, $\mathbb{P}(\mathfrak{P}^\complement) = 0$ and hence $\mathbb{P} \in \mathcal{M}_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}^-$. This shows that in the absence of a market-calibrated model, we have $\alpha < 0$. Therefore we can conclude that no market-calibrated model concentrated on $\mathfrak{P}$ implies the existence of a robust uniformly strong arbitrage (and hence WFLVR). Together with the results of Section 3.6.2, this completes the proof of Proposition 3.3.2.

With the complete proof of Proposition 3.3.2, we are now able to give a proof of Theorem 3.3.5 in the general case where $\mathfrak{P} \neq \Omega$.

3.6.5 Completing the proof of Theorems 3.3.5 and 3.4.3: case where $\mathfrak{P} \subseteq \Omega$

We now complete the proof of Theorems 3.3.5 and 3.4.3 for G satisfying (3.3.1) in the case where $\mathfrak{P} \subseteq \Omega$. Again, since they are virtually identical, we only give the proof of Theorem 3.3.5 here.

If (3.3.2) holds for G, then (3.3.2) is still true for any function $\tilde{G} = G + X$ with X of the form $a_0 + \sum_{i=1}^n a_i(S_i - K_i)^+$. Therefore, without loss of generality, we may and do assume that G is bounded from above.

Recall from (3.6.9) that $\lambda_{\mathfrak{P}}(s_1, \ldots, s_n) = \mathbb{1}_{\{(s_1,\ldots,s_n)\notin\mathfrak{P}\}}$ is bounded and lower semi-continuous and hence $G \vee (-N) - N\lambda_{\mathfrak{P}}$ is bounded and upper semi-continuous for each $N \in \mathbb{N}$. We also notice that

$$V_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G) \leq V_{\mathcal{X}_c,\mathcal{P},\Omega}\big(G \vee (-N) - N\lambda_{\mathfrak{P}}\big)$$

for each $N \in \mathbb{N}$, since any super-replicating portfolio of $G \vee (-N) - N\lambda_{\mathfrak{P}}$ on Ω naturally superreplicates G on $\mathfrak{P}$. Thus

$$\begin{aligned} V_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G) &\leq \inf_{N\geq 0} V_{\mathcal{X}_c,\mathcal{P},\Omega}(G \vee (-N) - N\lambda_{\mathfrak{P}}) \\ &= \inf_{N\geq 0} P_{\mathcal{X}_c,\mathcal{P},\Omega}(G \vee (-N) - N\lambda_{\mathfrak{P}}) \\ &= \inf_{N\geq 0} \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G \vee (-N) - N\lambda_{\mathfrak{P}}]. \end{aligned}$$

Define $f_N : \mathcal{M}_{\vec{\mu}}^- \to (-\infty, \infty)$ by $f_N(\mathbb{P}) = \mathbb{E}_{\mathbb{P}}[G \vee (-N) - N\lambda_{\mathfrak{P}}]$. Note that f_N is upper semi-continuous on $\mathcal{M}_{\vec{\mu}}^-$ and $f_N \geq f_{N+1}$ for every $N \in \mathbb{N}$. Hence, applying the min–max theorem (see Corollary 1 in Terkelsen [102]) to the compact convex set $\mathcal{M}_{\vec{\mu}}^-$ and $(f_N)_{N\in\mathbb{N}}$, we have

$$\inf_{N\geq 0} \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G \vee (-N) - N\lambda_{\mathfrak{P}}] = \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \inf_{N\geq 0} \mathbb{E}_{\mathbb{P}}[G \vee (-N) - N\lambda_{\mathfrak{P}}].$$

Define $G_{\mathfrak{P}}$ by $G_{\mathfrak{P}} = G$ on $\mathfrak{P}$ and $-\infty$ elsewhere. Note that $G_{\mathfrak{P}}$ is the pointwise limit of $G \vee (-N) - N\lambda_{\mathfrak{P}}$ as $N \to \infty$. Then by Fatou's lemma,

$$\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \inf_{N\geq 0} \mathbb{E}_{\mathbb{P}}[G \vee (-N) - N\lambda_{\mathfrak{P}}] \leq \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu}}^-} \mathbb{E}_{\mathbb{P}}[G_{\mathfrak{P}}] = \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P}}^-} \mathbb{E}_{\mathbb{P}}[G].$$

Therefore, we have $V_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G) \leq P_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G)$, which together with Remark 3.2.3 leads us to conclude that

$$V_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G) = P_{\mathcal{X}_c,\mathcal{P},\mathfrak{P}}(G).$$

3.6.6 Proof of Theorem 3.4.5

Proof. Given a semi-static superreplicating strategy (X, H) we have by definition for any $(s_1, \ldots, s_n) \in \mathfrak{P}$

$$X(s_1, \ldots, s_n) + \sum_{i=0}^{n-1} H_i(s_1, \ldots, s_i)(s_{i+1} - s_i) \geq G(s_1, \ldots, s_n). \tag{3.6.10}$$

We start with the following

Claim: *If (X, H) is a semi-static super-replicating strategy of G on the prediction set $\mathfrak{P}$, then $H_j \geq \beta_j$ for any $i = 0, \ldots, n-1$.*

We prove the claim by mathematical induction. When $j = n-1$, we fix $\vec{s}_{n-1} := (s_1, \ldots, s_{n-1})$. Letting $s_n \in \mathfrak{P}(\vec{s}_{n-1}, n) := \{x : (s_1, \ldots, s_{n-1}, x) \in \mathfrak{P}\}$ go to infinity, it follows from (3.6.10) that

$$H_{n-1}(s_1, \ldots, s_{n-1}) \geq \limsup_{x\to\infty,\, x\in\mathfrak{P}(\vec{s}_{n-1},n)} \frac{G(s_1, \ldots, s_{n-1}, x)}{x}.$$

This, together with $H_{n-1} \geq 0$, yields

$$\begin{aligned} H_{n-1}(s_1, \ldots, s_{n-1}) &\geq \limsup_{x\to\infty} \left\{ \frac{G(s_1, \ldots, s_{n-1}, x)}{x} \mathbb{1}_{\mathfrak{P}}(s_1, \ldots, s_{n-1}, x) \right\} \vee 0 \\ &= \beta_{n-1}(s_1, \ldots, s_{n-1}). \end{aligned}$$

Now suppose the claim holds for $j = i+1$ with $i \leq n-2$. We fix a vector $\vec{s}_{n-1} := (s_1, \ldots, s_i, s_{i+2}, \ldots, s_n)$ and denote

$$\mathfrak{P}(\vec{s}_{n-1}, i+1) := \{x : (s_1, \ldots, s_i, x, s_{i+2}, \ldots, s_n) \in \mathfrak{P}\}.$$

If $\mathfrak{P}(\vec{s}_{n-1}, i+1)$ is unbounded, then by taking $x \in \mathfrak{P}(\vec{s}_{n-1}, i+1)$ to infinity, (3.6.10) implies

$$\begin{aligned} H_i(s_1, \ldots, s_i) \geq \limsup_{x\to\infty,\, x\in\mathfrak{P}(\vec{s}_{n-1},i+1)} \Big(& H_{i+1}(s_1, \ldots, s_i, x) \\ & + \frac{G(s_1, \ldots, s_i, x, s_{i+2}, \ldots, s_n)}{x} \Big). \end{aligned}$$

If the right hand side is nonnegative, then

$$\begin{aligned}
&\limsup_{x\to\infty,\, x\in\mathfrak{P}(\vec{s}_{n-1},i+1)} \left(H_{i+1}(s_1,\dots,s_i,x) + \frac{G(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)}{x}\right) \\
=&\limsup_{x\to\infty} \Bigg(\mathbb{1}_{\mathfrak{P}}(s_1,\dots,s_i,x,s_{i+2},\dots,s_n) \\
&\qquad \times \left(H_{i+1}(s_1,\dots,s_i,x) + \frac{G(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)}{x}\right)\Bigg).
\end{aligned}$$

Hence, as $H_i \geq 0$, when $\mathfrak{P}(\vec{s}_{n-1}, i+1)$ is unbounded, we have

$$\begin{aligned}
&H_i(s_1,\dots,s_i) \\
\geq&\limsup_{x\to\infty} \Bigg(\Bigg(H_{i+1}(s_1,\dots,s_i,x) \\
&+ \frac{G(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)}{x}\Bigg)\mathbb{1}_{\mathfrak{P}}(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)\Bigg) \vee 0 \\
\geq&\limsup_{x\to\infty} \Bigg(\Bigg(\beta_{i+1}(s_1,\dots,s_i,x) \\
&+ \frac{G(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)}{x}\Bigg)\mathbb{1}_{\mathfrak{P}}(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)\Bigg) \vee 0. \qquad (3.6.11)
\end{aligned}$$

On the other hand, when $\mathfrak{P}(\vec{s}_{n-1}, i+1)$ is bounded, we notice that

$$\begin{aligned}
&\limsup_{x\to\infty} \Bigg(\left(\beta_{i+1}(s_1,\dots,s_i,x) + \frac{G(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)}{x}\right) \\
&\qquad \times \mathbb{1}_{\mathfrak{P}}(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)\Bigg) = 0.
\end{aligned}$$

Hence, the inequality in (3.6.11) is true in either case. In addition, as it holds for all $s_1,\dots,s_i,s_{i+2},\dots,s_n \in \mathbb{R}_+$, we can conclude that

$$\begin{aligned}
&H_i(s_1,\dots,s_i) \\
\geq& \sup_{s_{i+2},\dots,s_n\in\mathbb{R}_+} \limsup_{x\to\infty} \Bigg(\left(H_{i+1}(s_1,\dots,s_i,x) + \frac{G(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)}{x}\right) \\
&\qquad \times \mathbb{1}_{\mathfrak{P}}(s_1,\dots,s_i,x,s_{i+2},\dots,s_n)\Bigg) \vee 0 \\
=&\beta_i(s_1,\dots,s_i), \qquad \text{for any } s_1,\dots,s_i \in \mathbb{R}_+.
\end{aligned}$$

This ends the induction and the proof of the claim.

It follows from the claim above that for any $(X, H) \in \mathcal{A}_p$ that superreplicates G on

$\mathfrak{P}$ and any $\mathbb{P} \in \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$, we have

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[G] \leq & \mathbb{E}_{\mathbb{P}}\Big[X(\mathbb{S}) + \sum_{i=0}^{n-1} H_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)\Big] \\ \leq & \mathbb{E}_{\mathbb{P}}\Big[X(\mathbb{S}) + \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)\Big],\end{aligned}$$

which implies that

$$V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) \geq \sup_{\mathbb{P} \in \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}\Big[G - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)\Big].$$

For the converse inequality, we let $\mathcal{B}_b(\mathbb{R}^d_+, \mathbb{R}_+)$ be the set of all bounded and measurable functions $f : \mathbb{R}^d_+ \to \mathbb{R}_+$ and denote by $\mathcal{Z}$ the collection of all tuples of functions $(H_j)_{j=0}^{n-1} \in \mathbb{R}_+ \times \mathcal{B}_b(\mathbb{R}_+, \mathbb{R}_+) \times \ldots \times \mathcal{B}_b(\mathbb{R}^{n-1}_+, \mathbb{R}_+)$ such that

$$G_H(\mathbb{S}) := G(\mathbb{S}) - \sum_{i=0}^{n-1} H_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)$$

is upper semi-continuous and bounded from above on $\mathfrak{P}$. Note that $\mathcal{Z}$ is a convex subset of $\mathbb{R}_+ \times \mathcal{B}_b(\mathbb{R}_+, \mathbb{R}_+) \times \ldots \times \mathcal{B}_b(\mathbb{R}^{n-1}_+, \mathbb{R}_+)$. Then we can apply min–max theorem (Corollary 2 in Terkelsen [102]) to the compact convex set $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$, $\mathcal{Z}$ and the function

$$f(\pi, (H_j)) = \int \Big(G(s_1, \ldots, s_n) - \sum_{i=0}^{n-1} H_i(s_1, \ldots, s_i)(s_{i+1} - s_i)\Big) \mathrm{d}\pi(s_1, \ldots, s_n).$$

Clearly f is affine in each of the variables and by the Portmanteau theorem $f(\cdot\,, (H_j))$ is upper semi-continuous on $\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$. So the assumptions of Corollary 2 in Terkelsen [102] are satisfied, and we find that

$$\begin{aligned}V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) =& \inf \Big\{\mathcal{P}(X) : (X, H) \in \mathcal{A}_p \text{ s.t. } \Psi_{X,H} \geq G \text{ on } \mathfrak{P}\Big\} && (3.6.12)\\ \leq & \inf_{H \in \mathcal{Z}} \inf \Big\{\mathcal{P}(X) \,:\, (X, \tilde{H}) \in \mathcal{A}_p \text{ s.t. } \Psi_{X,H+\tilde{H}} \geq G \text{ on } \mathfrak{P}\Big\} && (3.6.13)\\ =& \inf_{H \in \mathcal{Z}} \inf \Big\{\mathcal{P}(X) \,:\, (X, \tilde{H}) \in \mathcal{A}_p \text{ s.t. } \Psi_{X,\tilde{H}} \geq G_H \text{ on } \mathfrak{P}\Big\} && (3.6.14)\\ =& \inf_{H \in \mathcal{Z}} \sup_{\mathbb{P} \in \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}\big[G_H(\mathbb{S})\big] = \sup_{\mathbb{P} \in \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \inf_{H \in \mathcal{Z}} \mathbb{E}_{\mathbb{P}}\big[G_H(\mathbb{S})\big]\end{aligned}$$

where the inequality between (3.6.12) and (3.6.13) is by restricting the delta-hedging terms to a smaller set, and the equality between (3.6.13) and (3.6.14) follows from Theorem 3.4.3.

To conclude, from the assumption we know there exists a sequence $\{\beta^{(N)}\}$ in $\mathcal{Z}$ such that $G_{\beta^{(N)}}(\mathbb{S}) = G(\mathbb{S}) - \sum_{i=0}^{n-1} \beta_i^{(N)}(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)$ is upper semi-continuous, bounded from above on $\mathfrak{P}$ and $G_{\beta^{(N)}} \to G_\beta$ pointwise as $N \to \infty$. Hence, by Fatou's Lemma

$$\limsup_{N\to\infty} \mathbb{E}_{\mathbb{P}}[G_{\beta^{(N)}}(\mathbb{S})] \leq \mathbb{E}_{\mathbb{P}}\Big[G - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)\Big]$$

holds for any $\mathbb{P} \in \mathcal{M}^-_{\vec{\mu},\mathfrak{P}}$, and therefore we have

$$\sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \inf_{H\in\mathcal{Z}} \mathbb{E}_{\mathbb{P}}\big[G_H(\mathbb{S})\big] \leq \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}\Big[G - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)\Big],$$

which leads us to conclude that

$$V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) \leq \sup_{\mathbb{P}\in\mathcal{M}^-_{\vec{\mu},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}\Big[G - \sum_{i=0}^{n-1} \beta_i(\mathbb{S}_1, \ldots, \mathbb{S}_i)(\mathbb{S}_{i+1} - \mathbb{S}_i)\Big].$$

□

Chapter 4

Robust pricing–hedging duality without options and dynamic programming principle

4.1 Introduction

The robust approach to pricing and hedging in continuous time has been an active field of research in mathematical finance over the recent years. However, compared to discrete time, it is relatively less well understood due to all the technical issues concerning measurability, admissibility etc. involved in building up a well defined stochastic integral to represent the capital gain from trading, which could arise in the classical approach to the pricing and hedging problems once one extends the analysis from discrete time to continuous time.

So far in the literature, there have been three main approaches to resolve the issues of defining a pathwise integral. The first one is the closest to the classical approach. It is considered in the quasi-sure setting where a family of probability measures is considered simultaneously and the integral constructed therein needs to be well defined and consistent with the classical stochastic integral under any probability measure, see e.g. [37; 99; 12; 8; 82; 80; 87; 80; 85; 81] and the references therein. As shown in these papers, among many others, the main difficulty usually lies in aggregating the integrand. In contrast, the second and third ones are to construct an integral pathwise without any direct reference to a probability measure. The difference between these two is that the former focus on building a pathwise integral for a more general class of integrators which, instead of being well-defined for every path, could work well for paths with desirable regularity, for example paths that admit pathwise quadratic variations, see e.g. [105; 69; 9; 47], while the later puts more restrictions on the integrands, usually concerning a certain degree of smoothness of the paths of integrands, and as a result usually the regularity of path is not needed and the pathwise integral therein can be defined for any path, see e.g. [39; 41].

In this chapter and the following, we follow the third approach to define a pathwise integral that represents the gain from continuous trading, and then a continuous-time superhedging problem can be defined. One of our contributions here is to establish an "unconstrained" pricing–hedging duality, which states that for any derivative with bounded and uniformly continuous payoff function G, the minimal initial set-up cost of a portfolio consisting of cash and dynamic trading in the risky assets which superhedges the payoff G for every nonnegative continuous path, is equal to the supremum of the expected value of G over all nonnegative continuous martingale measures[1]. This result is shown through an elaborate discretisation procedure built on ideas in [39; 41].

The duality result involves neither static trading nor dynamic trading in options, but can serve as a critical building block to establish a general pricing–hedging duality via a variational approach as in Galichon et al. [49], as will be seen in the next chapter.

Our second contribution is to provide a dynamic programming principle (DPP), which relates the time-0 superhedging cost (resp model price) to any later time-t supherhedging cost (resp model price). We show that looking at the pricing and the hedging problems on $[0, T]$ is the same as first looking at the pricing and the hedging problems on $[T_1, T]$ and then on $[0, T_1]$.

Note that the pricing and the hedging problems on $[T_1, T]$ shall naturally depend on the information given at T_1 or in our case the price path up to time T_1 as we choose the filtration to be canonical. Hence, a properly defined superhedging cost and model price on $[T_1, T]$ shall reflect the path dependence, and that requires us to extend the robust framework in Dolinsky and Soner [39] to have the flexibility to incorporate the case that the filtration has a non-trivial initial σ-algebra. However, as we shall emphasise here, the case of a filtration with a non-trivial initial σ-algebra is not the main subject studied here. Neither is the case of a non-canonical filtration. Our focus here and in the next chapter is on the case when trading strategies are adapted to the natural filtration $\mathbb{F}$ of the price process $\mathbb{S}$. For consideration beyond this, we refer the reader to Aksamit et al. [2].

[1] Note that here and throughout, we assume that all assets are discounted or, more generally, are expressed in terms of some numeraire.

The chapter is organised as follows. Section 4.2 introduces our robust framework for pricing and hedging and defines the primal (pricing) and the dual (hedging) problems. In Section 4.3, we present the unconstrained pricing–hedging duality in Theorem 4.3.1. Then in Sections 4.4 we establish the dynamic programming principle for pricing and hedging problems with proofs. Theorem 4.3.1 is proved in Sections 4.5 and 4.6. The proof proceeds via discretisation: of the primal problem in Section 4.5 and of the dual problem in Section 4.6. Proofs of two auxiliary results are relegated to the Appendix.

4.2 Robust Modelling Framework

4.2.1 Traded assets

We consider a financial market with $d+1$ assets: a numeraire (e.g. the money market account) and d underlying assets $S^{(1)}, \ldots, S^{(d)}$, which may be traded at any time $t \leq T$. All prices are denominated in the units of the numeraire. In particular, the numeraire's price is thus normalised and equal to one. We assume that the price path $S_t^{(i)}$ of each risky asset is continuous. The assets start at $S_0 = (1, \ldots, 1)$ and are assumed to be nonnegative. We work on the canonical space $\mathcal{C}([0,T], \mathbb{R}_+^d)$, the set of all $\mathbb{R}_+^d$-valued continuous functions on $[0,T]$.

We pursue here a robust approach and do not postulate any probability measure which would specify the dynamics for S.

4.2.2 Trading strategies

We now discuss the notion of atoms of a σ-field and introduce the right class of trading strategies with respect to the canonical filtration $\mathbb{F}$. We refer to Dellacherie and Meyer [36] Chapter 1 Section 9-12 (pages 14-16) for useful details.

For a measurable space $(\Omega, \mathcal{F}_T)$ and a sub σ-field $\mathcal{G} \subset \mathcal{F}_T$ we introduce the following equivalence relation.

Definition 4.2.1. Let ω and $\widetilde{\omega}$ be two elements of Ω, and $\mathcal{G} \subset \mathcal{F}_T$ be a σ-field. Then we say that ω and $\widetilde{\omega}$ are $\mathcal{G}$-equivalent, and write $\omega \sim_{\mathcal{G}} \widetilde{\omega}$, if for each $G \in \mathcal{G}$ we have $\mathbb{1}_G(\omega) = \mathbb{1}_G(\widetilde{\omega})$.

We call $\mathcal{G}$-atoms the equivalence classes in Ω with respect to this relation. We denote by A^ω the atom which contains ω: $A^\omega = \bigcap\{A : A \in \mathcal{G}, \omega \in A\}$.

Note that if $\mathcal{G}$ is countably generated, $\mathcal{G} = \sigma(B_n : n \geq 1)$, then each atom is an element of $\mathcal{G}$ as $A^\omega = \bigcap_n C_n$ is a countable intersection, where $C_n = B_n$ if $\omega \in B_n$ and $C_n = B_n^c$ if $\omega \in B_n^c$. Also, it is then enough to check the relation from Definition 4.2.1 on the generators of $\mathcal{G}$. Finally, we note that in our setting, $\omega \sim_{\mathcal{F}_t} \widetilde{\omega}$ if and only if $\omega_u = \widetilde{\omega}_u$ for each $u \leq t$.

We only consider trading strategies γ which are of finite variation. This allows us, similarly to Dolinsky and Soner [39], to define stochastic integrals pathwise simply via the integration by parts formula:

$$\int_0^t \gamma_u \mathrm{d}S_u := \gamma_t \cdot S_t - \gamma_0 \cdot S_0 - \int_0^t S_u \mathrm{d}\gamma_u,$$

where we write $a \cdot b$ to denote the usual scalar product for any $a, b \in \mathbb{R}^d$ and the last term on the right hand side is a Stieltjes integral.

Further, γ is required to be progressively measurable with respect to a filtration which, in our context, is the natural filtration generated by the canonical process. More precisely, we have:

Definition 4.2.2. (i) We say that a map $\gamma : \Omega \to \mathcal{D}([0,T], \mathbb{R}^d)$ is *measurable*, if the mapping $(t, \omega) \to \gamma_t(\omega)$ is measurable on $[0,T] \times \Omega$ with respect to the product σ-algebra $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$.
(ii) γ is an $\mathbb{F}$-*adapted* process if $\omega \to \gamma_t(\omega)$ is $\mathcal{F}_t$-measurable for every $t \leq T$.
(iii) Finally, γ is an $\mathbb{F}$-*progressive measurable* process if for every $t \leq T$ the mapping $(s, \omega) \to \gamma_s(\omega)$ is measurable on $[0,t] \times \Omega$ with respect to the product σ-algebra $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$.

Remark 4.2.3. Galmarino's test states that for a measurable τ, it is an $\mathbb{F}$-stopping time if and only if for every t, the properties $\tau \leq t$ and $\omega_s = \upsilon_s$ for any $s \leq t$, imply that $\tau(\omega) = \tau(\upsilon)$. Moreover, given an $\mathcal{F}_T$-measurable random variable $\phi : \Omega \to \mathbb{R}^d$ and an $\mathbb{F}$-stopping time $\tau \leq T$, ϕ is $\mathcal{F}_\tau$-measurable if and only if

$$\forall \upsilon, \omega \in \Omega, \;\; \upsilon_u = \omega_u, \; \forall u \in [0, \tau(\upsilon)] \quad \Rightarrow \quad \phi(\upsilon) = \phi(\omega).$$

For more details, see Dellacherie and Meyer [36] Chapter IV. Then it is clear that for a measurable map γ, it is $\mathbb{F}$-*adapted* if and only if

$$\forall \upsilon, \omega \in \Omega, \; t \leq T, \;\; \upsilon_u = \omega_u, \; \forall u \in [0, t] \quad \Rightarrow \quad \gamma_t(\upsilon) = \gamma_t(\omega). \tag{4.2.1}$$

In particular, every $\mathbb{F}$-adapted measurable process is progressively measurable and even optional. (For detailed explanation of the properties above and more properties of the stopping times and processes on the canonical space paired with the canonical filtration, we refer the reader to Chapter 4 of Dellacherie and Meyer [36].)

We say γ is admissible if $\gamma : \Omega \to \mathcal{D}([0,T], \mathbb{R}^d_+)$ is progressively measurable and of finite variation, satisfying

$$\int_0^t \gamma_u(S)\mathrm{d}S_u \geq -M, \quad \forall\, S \in \Omega,\, t \in [0,T], \text{ for some } M > 0. \tag{4.2.2}$$

Let $\mathcal{A}$ be the set of such integrands. The set of simple integrands, i.e., $\gamma \in \mathcal{A}$ such that $\gamma(\omega)$ is a simple function $\forall \omega \in \Omega$, is denoted $\mathcal{A}^{sp}$. Later, we consider hedging only on some time interval $[T_1, T]$ and hence we need to extend the admissibility to filtration with non-trivial initial σ-algebra: γ is admissible on $[T_1, T]$ if $\gamma : \Omega \to \mathcal{D}([0,T], \mathbb{R}^d_+)$ is progressively measurable on $[T_1, T]$ and of finite variation, satisfying

$$\int_{T_1}^t \gamma_u(S)\mathrm{d}S_u \geq -M(S), \quad \forall\, S \in \Omega,\, t \in [T_1,T], \text{ for some } M \in \mathrm{L}^0(\Omega, \mathcal{F}_{T_1}). \tag{4.2.3}$$

When the admissibility (4.2.3) is only required on some time interval $[T_1, T_2]$ we write $\mathcal{A}([T_1, T_2])$ for the respective sets of trading strategies.

We are now in a position to define the main quantities of interest: the robust pricing and hedging prices of an option.

We start with superhedging problem.

Definition 4.2.4.

1. $\gamma \in \mathcal{A}$ is said to *superreplicate* G on Ω if
$$\int_0^T \gamma_u(S)\mathrm{d}S_u \geq G(S), \;\; \forall S \in \Omega. \tag{4.2.4}$$
2. The (minimal) superhedging cost of G is defined as
$$\mathbf{V}(G) := \inf \Big\{ x \,:\, \exists \gamma \in \mathcal{A} \text{ s.t. } \gamma \text{ superreplicates } G - x \text{ on } \Omega \Big\} \tag{4.2.5}$$
3. Finally, we let $\mathbf{V}^{sp}(G)$ denote the super-replicating cost of G in (5.2.3) using $\gamma \in \mathcal{A}^{sp}$.

4.2.3 Market models

We now turn to the pricing problem. In the classical approach markets with no arbitrage are modelled using martingale measures. We denote by $\mathcal{M}$ the set of probability measures $\mathbb{P}$ on $(\Omega, \mathcal{F}_T, \mathbb{F})$ such that $\mathbb{S}$ is an $\mathbb{F}$-martingale under $\mathbb{P}$.

Whenever we have $\mathbb{P} \in \mathcal{M}$ it provides us with a feasible no-arbitrage price $\mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$ for a derivative with payoff G. The robust price for G is given as

$$\mathbf{P}(G) := \sup_{\mathbb{P} \in \mathcal{M}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})],$$

where throughout the expectation is defined with the convention that $\infty - \infty = -\infty$.

In some instances, for technical reasons, it will be convenient to consider only $\mathbb{P}$ arising within a Brownian setup. We denote by $\underline{\mathcal{M}}$ the collection of $\mathbb{P} \in \mathcal{M}$ such that $\mathbb{P} = \mathbb{P}^W \circ M^{-1}$ for some continuous martingale M defined on some probability space satisfying the usual assumptions $(\Omega^W, \mathcal{F}_T^W, \mathbb{F}^W, P^W)$ with a finite dimensional Brownian motion $\{W_t\}_{t \geq 0}$ which generates the filtration $\mathbb{F}^W$. We write $\underline{\mathbf{P}} := \sup_{\mathbb{P} \in \underline{\mathcal{M}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$.

4.3 Pricing–hedging duality

Our prime interest here is in establishing a robust pricing–hedging duality in the unconstrained case, that is when there are no options available for hedging purpose. Given a non-traded derivative with payoff G we have two candidate robust prices for it. The first one, $\mathbf{V}(G)$, is obtained through pricing-by-hedging arguments. The second one, $\mathbf{P}(G)$, is obtained by pricing-via-expectation arguments. In a classical setting, the analoguous two prices are equal. This is trivially true in a complete market and is a fundamental result for incomplete markets, see Theorem 5.7 in Delbaen and Schachermayer [34].

Theorem 4.3.1. For any bounded and uniformly continuous $G : \Omega \to \mathbb{R}$ we have

$$\mathbf{V}^{sp}(G) = \mathbf{V}(G) = \mathbf{P}(G) = \underline{\mathbf{P}}(G).$$

An analogous duality in a quasi-sure setting was obtained in Possamaï et al. [87] and earlier papers, as discussed therein. However, while similar in spirit, there is no immediate link between our results or proofs and these in [87]. Here, we consider a comparatively smaller set of admissible trading strategies and we require a pathwise superhedging property. Consequently, we also need to impose stronger regularity constraints on G. The inequality

$$\mathbf{V}(G) \geq \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$$

is straightforward. Indeed, since every $\gamma \in \mathcal{A}$ is progressively measurable, the integral $\int_0^{\cdot} \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u$, defined pathwise via integration by parts, agrees a.s. with the stochastic integral under $\mathbb{P}$. Then, by (4.2.2), the stochastic integral is a $\mathbb{P}$-super-martingale and hence $\mathbb{E}_{\mathbb{P}}\left[\int_0^T \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u\right] \leq 0$. The result follows since γ and $\mathbb{P}$ were arbitrary.

Sections 4.5 and 4.6 are mainly devoted to the proof of the much harder reverse inequality

$$\mathbf{V}(G) \leq \sup_{\mathbb{P}\in\underline{\mathcal{M}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})], \tag{4.3.1}$$

which then implies Theorem 4.3.1. The proof proceeds through discretisation of both the primal and the dual problems.

4.4 Dynamic programming principle

In optimisation, dynamic programming is an approach that transforms a complex global optimisation problem into a sequence of sub-problems. In our context, to analyse the pricing and the hedging problems, we provide a dynamic programming principle (DPP), which relates the time-0 superhedging cost $\mathbf{V}$ (resp $\mathbf{P}$) to any later time-t supherhedging cost (resp model price), which will be defined later.

More formally, we divide our problems into two time intervals: $[0, T_1]$ and $[T_1, T]$. First we look at the pricing and the hedging problems on $[T_1, T]$ and then on $[0, T_1]$, and we will show that it is the same as looking at the pricing and the hedging problems on $[0, T]$.

We begin with two propositions where we develop the dynamic programming principle for two cases: superhedging cost and market model price. We will first show the time-t supherhedging cost (resp model price) is uniformly continuous with respect to $\|\cdot\|_{[0,t]}$, where $||\cdot||_{[0,t]}$ is the sup norm on Ω such that $||v - \widetilde{v}||_{[0,t]} = \sup_{0\leq t\leq T_1} |v_t - \widetilde{v}_t|$, for any $v, \widetilde{v} \in \Omega$. We also write $\|\cdot\|$ to denote $\|\cdot\|_{[0,T]}$.

Proposition 4.4.1. Let B^ω denote the $\mathcal{F}_{T_1}$-atom containing ω. Then for a bounded uniformly continuous ξ the following hold.

(i) The mapping $\mathbf{V}^{[T_1,T]}(\xi):\Omega\to\mathbb{R}$ defined as

$$\mathbf{V}^{[T_1,T]}(\xi)(\omega):=\inf\{x\in\mathbb{R}:\exists\,\gamma\in\mathcal{A}([T_1,T])\ \text{ such that } x+\int_{T_1}^{T}\gamma_t \mathrm{d}S_t\geq\xi\ \forall S\in B^{\omega}\}$$

is uniformly continuous with respect to $\|\cdot\|_{[0,T_1]}$, and $\mathcal{F}_{T_1}$-measurable.
(ii) The dynamic programming principle holds in the form:

$$\mathbf{V}^{[0,T]}(\xi)=\mathbf{V}^{[0,T_1]}\left(\mathbf{V}^{[T_1,T]}(\xi)\right).$$

Proof. In the proof we denote $\widetilde{\xi}:=\mathbf{V}^{[T_1,T]}(\xi)$.
(i) Let $v:=(v^1,...,v^d)\in\Omega$ and $\widetilde{v}:=(\widetilde{v}^1,...,\widetilde{v}^d)\in\Omega$. Note that

$$\left|\widetilde{\xi}((v^1,..,v^d))-\widetilde{\xi}((\widetilde{v}^1,..,\widetilde{v}^d))\right|\leq\sum_{k=1}^{d}\left|\widetilde{\xi}((\widetilde{v}^1,..,\widetilde{v}^{k-1},v^k,..,v^d))-\widetilde{\xi}((\widetilde{v}^1,..,\widetilde{v}^k,v^{k+1},..,v^d))\right|.$$

Thus, to establish uniform continuity of $\widetilde{\xi}$, it is enough to consider v and $\widetilde{v}$ which differ on one coordinate only and, without loss of generality, we may assume that $d=1$.

Consider a small $\delta>0$. Suppose that $||v-\widetilde{v}||_{[0,T_1]}\leq\delta$, $|v_{T_1}-\widetilde{v}_{T_1}|=D\geq 0$, $v_{T_1}>0$ and $\widetilde{v}_{T_1}>0$.
In the first step we show that $\widetilde{\xi}(\widetilde{v})\leq\widetilde{\xi}(v)+\varepsilon$ for an appropriately chosen ε, depending only on ξ and δ. For each $\eta>0$ there exists an admissible strategy γ such that

$$\widetilde{\xi}(v)+\int_{T_1}^{T}\gamma_t\mathrm{d}S_t\geq\xi-\eta \text{ on } B^v.$$

Let $\lambda=v_{T_1}/\widetilde{v}_{T_1}\in(0,\infty)$ and define the path modification mapping $\alpha^{v,\widetilde{v}}$ by

$$\alpha(\omega):=\alpha^{v,\widetilde{v}}(\omega):=\begin{cases} v|_{[0,T_1]}\otimes\frac{v_{T_1}}{\widetilde{v}_{T_1}}\omega|_{[T_1,T]} & \omega\in B^{\widetilde{v}}\\ \widetilde{v}|_{[0,T_1]}\otimes\frac{\widetilde{v}_{T_1}}{v_{T_1}}\omega|_{[T_1,T]} & \omega\in B^{v}\\ \omega & \omega\notin B^v\cup B^{\widetilde{v}}\end{cases}\qquad(4.4.1)$$

where $\lambda\omega$ is a multiplicative modification of ω by λ in Ω and $v|_{[0,T_1]}\otimes\lambda\omega|_{[T_1,T]}$ means that the path is equal to v on $[0,T_1]$ and to $\lambda\omega|$ on $[T_1,T]$. Note that α is a bijection satisfying $\alpha=\alpha^{-1}$. Introduce a stopping time

$$\widetilde{\tau}(\omega):=\tau^{v,\widetilde{v}}(\omega):=\begin{cases}\inf\{t>T_1:\omega_t-\widetilde{v}_{T_1}\geq\widetilde{v}_{T_1}D^{-\frac{1}{2}}\}\wedge T & \omega\in B^{\widetilde{v}}\\ \inf\{t>T_1:\omega_t-v_{T_1}\geq v_{T_1}D^{-\frac{1}{2}}\}\wedge T & \omega\in B^{v}\\ T_1 & \omega\notin B^v\cup B^{\widetilde{v}}\end{cases}.\qquad(4.4.2)$$

To show that $\widetilde{\xi}(\widetilde{v}) \leq \widetilde{\xi}(v) + \varepsilon$ we will consider a strategy $\lambda\gamma \circ \alpha + \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}} \mathbb{1}_{[T_1,\widetilde{\tau})}$ on $B^{\widetilde{v}}$. Both terms of this strategy are clearly $\mathbb{F}$-adapted.

Then, we obtain

$$\begin{aligned}
&\widetilde{\xi}(v) + \lambda \int_{T_1}^{T} \gamma(\alpha(\omega))_t \mathrm{d}S_t(\omega) + \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\widetilde{\tau}} - \widetilde{v}_{T_1}) \\
&\quad = \widetilde{\xi}(v) + \int_{T_1}^{T} \gamma(\alpha(\omega))_t \mathrm{d}(S\alpha(\omega))_t + \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\widetilde{\tau}} - \widetilde{v}_{T_1}) \\
&\quad \geq \xi \circ \alpha(\omega) - \eta + \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\widetilde{\tau}} - \widetilde{v}_{T_1})
\end{aligned}$$

where the first equality is due to our definition of integration. In the case that $\widetilde{\tau}(\omega) = T$ one has

$$\frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\widetilde{\tau}} - \widetilde{v}_{T_1}) \geq -\frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}\widetilde{v}_{T_1} = -D^{1/4}$$

and

$$\|\alpha(\omega) - \omega\| \leq \delta \vee (\lambda - 1)\widetilde{v}_{T_1}(D^{-\frac{1}{2}} + 1) \leq 2\delta^{1/2}. \tag{4.4.3}$$

Thus, for $\widetilde{\tau}(\omega) = T$, it follows that

$$\xi \circ \alpha(\omega) - \eta + \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\tau} - \widetilde{v}_{T_1}) \geq \xi(\omega) - e_{\xi}(2\delta^{1/2}) - \eta - D^{1/4}$$

where e_ξ is modulus of continuity of ξ. Hence, for $\widetilde{\tau}(\omega) = T$, we deduce $\widetilde{\xi}(\widetilde{v}) \leq \widetilde{\xi}(v) + e_{\xi}(2\delta^{1/2}) + D^{1/4}$. In the case that $\widetilde{\tau}(\omega) < T$ one has

$$\frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\widetilde{\tau}} - \widetilde{v}_{T_1}) = \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}\widetilde{v}_{T_1} D^{-\frac{1}{2}} = D^{-1/4}$$

and

$$\xi \circ \alpha(\omega) - \eta + \frac{D^{\frac{1}{4}}}{\widetilde{v}_{T_1}}(\omega_{\tau} - \widetilde{v}_{T_1}) \geq -||\xi|| - \eta + D^{-1/4}$$

which, for D small enough ($D \leq (2||\xi||)^{-4}$), is larger than $\xi(\omega)$. We deduce that $\widetilde{\xi}(\widetilde{v}) \leq \widetilde{\xi}(v) + e_{\xi}(2\delta^{1/2}) + D^{1/4}$.
Therefore, $\widetilde{\xi}$ is uniformly continuous on $\{\omega \in \Omega : \|\omega\| > 0\}$.
To complete the proof, we now consider the case where $\widetilde{v}_{T_1} = 0$. Let, for some small $\delta > 0$, $||v - \widetilde{v}||_{[0,T_1]} \leq \delta$ and $v_{T_1} = D > 0$. Firstly notice that $\widetilde{v}$ must satisfy $\widetilde{\xi}(\widetilde{v}) = \xi(\widetilde{v}|_{[0,T_1]} \otimes 0|_{[T_1,T]})$ since we can buy any amount of stock at price 0 at time T_1 thus only constant path is relevant, and therefore $\widetilde{\xi}(\widetilde{v}) \leq \widetilde{\xi}(v) + e_{\xi}(\delta)$. Now consider the strategy γ for $\omega \in B^v$ defined as

$$\gamma(\omega) := \delta^{-1/2}\mathbb{1}_{[T_1,\sigma(\omega))} \quad \text{where} \quad \sigma(\omega) := \inf\{t > T_1 : \omega_t - v_{T_1} \geq \delta^{1/4}\}.$$

Then, whenever $\sigma(\omega) < T$,

$$\widetilde{\xi}(\widetilde{v}) + \int_{T_1}^{T} \gamma(\omega)_t \mathrm{d}S_t(\omega) = \widetilde{\xi}(\widetilde{v}) + \delta^{-1/2}(\omega_\sigma - v_{T_1}) = \widetilde{\xi}(\widetilde{v}) + \delta^{-1/4}$$

which, for d small enough, is larger than $\xi(\omega)$. Otherwise, if $\sigma(\omega) = T$

$$\widetilde{\xi}(\widetilde{v}) + \int_{T_1}^{T} \gamma(\omega)_t \mathrm{d}S_t(\omega) \geq \widetilde{\xi}(\widetilde{v}) - \delta^{1/2} \geq \xi(\omega) - e_\xi(2\delta^{1/4}) - \delta^{1/2}$$

since $||\widetilde{v} - \omega|| \leq 2\delta^{1/4}$. Therefore, $\widetilde{\xi}(v) \leq \widetilde{\xi}(\widetilde{v}) + e_\xi(2\delta^{1/4}) + \delta^{1/2}$.

(ii) Let $V^1 := \mathbf{V}^{[0,T]}(\xi)$ and $V^2 := \mathbf{V}^{[0,T_1]}(\widetilde{\xi})$. For each $\eta > 0$ there exists a strategy $\gamma \in \mathcal{A}([0,T])$ such that

$$V^1 + \int_0^{T_1} \gamma_t \mathrm{d}S_t + \int_{T_1}^{T} \gamma_t \mathrm{d}S_t \geq \xi - \eta \quad \text{on} \quad \Omega.$$

Let $\tau(S) := \inf\{t > 0 \,:\, V^1 + \int_0^t \gamma_u \mathrm{d}S_u \geq \sup_{\omega\in\Omega} \xi(\omega) - \eta\} \wedge T$. It is a stopping time. Hence $\widetilde{\gamma} := \gamma \mathbb{1}_{[0,\tau]} \in \mathcal{A}([0,T])$ and satisfies that

$$V^1 + \int_0^{T_1} \widetilde{\gamma}_t \mathrm{d}S_t + \int_{T_1}^{T} \widetilde{\gamma}_t \mathrm{d}S_t \geq \xi - \eta \quad \text{on} \quad \Omega.$$

Moreover, for any $t \geq T_1$,

$$\int_{T_1}^{t} \widetilde{\gamma}_u \mathrm{d}S_u \geq \xi - \sup_{\omega\in\Omega} \xi(\omega) \quad \text{on} \quad \Omega,$$

and therefore $\widetilde{\gamma} \in \mathcal{A}([T_1, T])$.

In particular, for a fixed $\omega \in \Omega$, the superhedging holds on B^ω. Since $V^1 + \int_0^{T_1} \widetilde{\gamma}_t \mathrm{d}S_t$ is constant on B^ω, we deduce that $V^1 + \int_0^{T_1} \widetilde{\gamma}_t \mathrm{d}S_t \geq \widetilde{\xi}$ on $\Omega|_{[0,T_1]}$ and therefore $V^1 \geq V^2$.

To prove the reverse inequality take $z > V^2$. First, there exists $\gamma^{(1)} \in \mathcal{A}([0,T_1])$ such that $z + \int_0^{T_1} \gamma_t^{(1)} \mathrm{d}S_t \geq \widetilde{\xi}$ on Ω. If, for each $\delta > 0$, there exists a strategy $\gamma^{(2)} \in \mathcal{A}([T_1,T])$ such that $z + \int_0^{T_1} \gamma_t^{(1)} \mathrm{d}S_t + \int_{T_1}^{T} \gamma_t^{(2)} \mathrm{d}S_t \geq \xi - \delta$, then clearly $z \geq V^1$.

We now show the existence of such $\gamma^{(2)}$ for every $\eta > 0$. Fix $\delta > 0$. Let $\{\omega^n\}_{n\in\mathbb{N}}$ be a countable dense subset of $\Omega|_{[0,T_1]}$ and $B^n := B^{\omega^n}$, and denote the closed ball around ω^n of radius δ by $\widetilde{B}^n(\delta) := \{\omega : \sup_{t\in[0,T_1]} |\omega_t - \omega_t^n| \leq \delta\}$. Define the path modification mapping $\alpha^{n,\omega}$ by $\alpha^{n,\omega} := \alpha^{\omega^n,\omega}$ where $\alpha^{\omega^n,\omega}$ is given in (4.4.1). Note that

$\alpha^{n,\omega}$ is a bijection satisfying $\alpha^{n,\omega} = (\alpha^{n,\omega})^{-1}$. We now take $\{\gamma^n\}_n$, a set of strategies in $\mathcal{A}([T_1, T])$, such that

$$\widetilde{\xi}(\omega^n) + \int_{T_1}^{T} \gamma_u^n(S)\mathrm{d}S_u \geq \xi - \delta \quad \text{ on } B^n. \tag{4.4.4}$$

Let us consider $\widetilde{\gamma}^n : \Omega \to$ defined by $\widetilde{\gamma}_n(\omega) = 0$ if $\omega \notin \widetilde{B}^n(D)$ and for $\omega \in \widetilde{B}^n(D)$

$$\widetilde{\gamma}^n(\omega) := \begin{cases} \frac{\omega^n_{T_1}}{\omega_{T_1}}\gamma^n \circ \alpha^{n,\omega} + \frac{D^{\frac{1}{4}}}{\omega_{T_1}}\mathbb{1}_{[T_1, \tau^{\omega^n,\omega})} & \text{if } \omega^n_{T_1} \geq \delta, \\ \delta^{-1/2}\mathbb{1}_{[T_1,\sigma(\omega))} & \text{if } \omega^n_{T_1} < \delta, \end{cases}$$

where $\sigma(\omega) := \inf\{t > T_1 : \omega_t - \omega_{T_1} \geq \delta^{1/4}\}$. It follows from above that there exists a constant $\epsilon(D, \delta)$ which depends on D and δ with $\epsilon(D, \delta) \to 0$ as $D, \delta \to 0$, such that

$$\widetilde{\xi}(S) + \int_{T_1}^{T} \widetilde{\gamma}_u^n(S)\mathrm{d}S_u \geq \xi(S) - \epsilon(D, \delta).$$

The strategy $\widetilde{\gamma}^n$ is clearly $\mathcal{F}_T$-measurable. We also notice that it is adapted to $\mathbb{F}$ on $[T_1, T]$ since it is straightforward to see that for any $\omega, \upsilon \in \Omega$ such that $\omega_u = \upsilon_u$ for any $u \leq [t, T]$ with $t \geq T_1$, $\widetilde{\gamma}^n_u(\omega) = \widetilde{\gamma}^n_u(\upsilon)$ on $[T_1, t]$. Hence, $\widetilde{\gamma}^n \in \mathcal{A}([T_1, T])$. In addition, we know that for any n, $t \in [T_1, T]$ and $S \in B^n$, there exists $\widetilde{S}$ such that $\widetilde{S}_u = S_u$ for any $u \leq t$ and $\widetilde{S}_u = S_t$ for any $u \geq t$, and therefore

$$\int_{T_1}^{t} \widetilde{\gamma}_u^n(S)\mathrm{d}S_u = \int_{T_1}^{T} \widetilde{\gamma}_u^n(\tilde{S})\mathrm{d}\tilde{S}_u \geq \xi(\widetilde{S}) - \delta - \widetilde{\xi}(\omega^n) \geq 2\inf_{\omega\in\Omega}\xi(\omega) - 1. \tag{4.4.5}$$

Let us now define $\gamma^{(2)}$ by

$$\gamma^{(2)}(\omega) := \sum_n \mathbb{1}_{C^n}(\omega)\widetilde{\gamma}^n(\omega) \quad \text{where} \quad C^n := \widetilde{B}^n \backslash \bigcup_{k=1}^{n-1} \widetilde{B}^k.$$

It is then straightforward to see that $\gamma^{(2)}$ is progressively measurable and satisfies the admissibility condition in (4.2.2).

□

Define the set of measures $\mathcal{M}_A^{[T_1,T]}$ concentrated on $A \in \mathcal{F}_T$ as follows:

$$\mathcal{M}_A^{[T_1,T]} := \{\mathbb{P} : \mathbb{S} \text{ is an } \mathbb{F}\text{-martingale on } [T_1, T] \text{ and } \mathbb{P}(A) = 1\}.$$

Also write $\mathcal{M}^{[T_1,T]} = \mathcal{M}_\Omega^{[T_1,T]}$.

Proposition 4.4.2. Let B^ω denote the $\mathcal{F}_{T_1}$-atom containing ω and assume that $\mathcal{M}_{B^\omega}^{[T_1,T]} \neq \varnothing$ for each ω. Then for a bounded uniformly continuous ξ the following hold.
(i) The mapping $\mathbf{P}^{[T_1,T]}(\xi)$ defined as $\mathbf{P}^{[T_1,T]}(\xi)(\omega) := \sup_{\mathbb{P}\in\mathcal{M}_{B^\omega}^{[T_1,T]}} \mathbb{E}_\mathbb{P}[\xi]$ is uniformly continuous and $\mathcal{F}_{T_1}$-measurable.
(ii) The dynamic programming principle holds in the form:

$$\mathbf{P}^{[0,T]}(\xi) = \mathbf{P}^{[0,T_1]}(\mathbf{P}^{[T_1,T]}(\xi)).$$

Proof. In the proof we denote $\widehat{\xi} := \mathbf{P}^{[T_1,T]}(\xi)$.
(i) Let $v := (v^1, ..., v^d) \in \Omega$ and $\widetilde{v} := (\widetilde{v}^1, ..., \widetilde{v}^d) \in \Omega$. Note that

$$\left|\widehat{\xi}((v^1,..,v^d)) - \widehat{\xi}((\widetilde{v}^1,..,\widetilde{v}^d))\right| \leq \sum_{k=1}^{d} \left|\widehat{\xi}((\widetilde{v}^1,..,\widetilde{v}^{k-1},v^k,..,v^d)) - \widehat{\xi}((\widetilde{v}^1,..,\widetilde{v}^k,v^{k+1},..,v^d))\right|.$$

Thus, to prove uniform continuity of $\widetilde{\xi}$, it is enough to consider v and $\widetilde{v}$ which differ on one coordinate only and, without loss of generality, we may assume that $d = 1$. Suppose that $||v - \widetilde{v}||_{[0,T_1]} \leq \delta$ and $|v_{T_1} - \widetilde{v}_{T_1}| = D \geq 0$. It is enough to show that $\widehat{\xi}(\widetilde{v}) \leq \widehat{\xi}(v) + \varepsilon$ for an appropriately chosen ε depending only on ξ and δ. Take $\mathbb{P} \in \mathcal{M}_{B^{\widetilde{v}}}^{[T_1,T]}$, i.e., $\mathbb{P}(B^{\widetilde{v}}) = 1$, where $B^{\widetilde{v}} := \{\omega : \omega_t = \widetilde{v}_t \text{ for } t \in [0,T_1]\}$, $\mathbb{P}(\mathbb{S}_{T_1} = \widetilde{v}_{T_1}) = 1$ and $\mathbb{E}_\mathbb{P}[\mathbb{S}_t \mathbb{1}_F] = \mathbb{E}_\mathbb{P}[\mathbb{S}_s \mathbb{1}_F]$ for each $T_1 \leq s \leq t \leq T$ and $F \in \mathcal{F}_s$. Define measure $\bar{\mathbb{P}}$ as $\bar{\mathbb{P}} = \mathbb{P}\circ\alpha$ with path modification α given in (4.4.1). Then $\bar{\mathbb{P}}$ is an element of $\mathcal{M}_{B^v}^{[T_1,T]}$ since $\bar{\mathbb{P}}(B^v) = \mathbb{P}(\alpha(B^v)) = \mathbb{P}(B^{\widetilde{v}}) = 1$, it is a martingale measure on $[T_1, T]$ as for $T_1 \leq s \leq t \leq T$ and $F \in \mathcal{F}_s$

$$\mathbb{E}_{\bar{\mathbb{P}}}[\mathbb{S}_t \mathbb{1}_F] = \mathbb{E}_\mathbb{P}[(\mathbb{S}\circ\alpha)_t \mathbb{1}_{\alpha(F)}] = \mathbb{E}_\mathbb{P}[(\mathbb{S}\circ\alpha)_s \mathbb{1}_{\alpha(F)}] = \mathbb{E}_{\bar{\mathbb{P}}}[\mathbb{S}_s \mathbb{1}_F],$$

where the second equality follow by $\alpha(F) \in \mathcal{F}_s$. The latter is true since, for any Borel set B, one has $\alpha(\{Z \in B\} \cap B^v) = B^{\widetilde{v}} \cap \{Z \in A\}$; the σ-field $\mathcal{F}_s$ coincides with trivial σ-field up to $\mathbb{P}$-null sets and up to $\bar{\mathbb{P}}$-null sets; the general case follows from the monotone class argument. Hence, with $\widetilde{\tau}$ defined in (4.4.2),

$$\begin{aligned} |\mathbb{E}_\mathbb{P}[\xi] - \mathbb{E}_{\bar{\mathbb{P}}}[\xi]| &= |\mathbb{E}_\mathbb{P}[\xi] - \mathbb{E}_\mathbb{P}[\xi\circ\alpha]| \\ &= \mathbb{E}_\mathbb{P}[(\xi - \xi\circ\alpha)\mathbb{1}_{\{\widetilde{\tau}=T\}}] + \mathbb{E}_\mathbb{P}[(\xi - \xi\circ\alpha)\mathbb{1}_{\{\widetilde{\tau}<T\}}] \\ &\leq e_\xi(2\delta^{1/2}) + 2||\xi||\frac{D}{D + D^{1/2}}, \end{aligned} \tag{4.4.6}$$

where in the last inequality we used (4.4.3), Doob's inequality and the fact that

$$\mathbb{P}(\widetilde{\tau} < T) = \mathbb{P}\left(\sup_{t\in[T_1,T]} \mathbb{S}_t \geq \widetilde{v}_{T_1}(1 + D^{-1/2})\right) \leq \frac{\widetilde{v}_{T_1}}{\widetilde{v}_{T_1}(1 + D^{-1/2})} = \frac{D}{D + D^{1/2}}.$$

(ii) To prove that $\mathbf{P}^{[0,T]}(\xi) \leq \mathbf{P}^{[0,T_1]}(\mathbf{P}^{[T_1,T]}(\xi))$ it is enough to note that:

$$\sup_{\mathbb{P}\in\mathcal{M}^{[0,T]}} \mathbb{E}_{\mathbb{P}}[\xi] = \sup_{\mathbb{P}\in\mathcal{M}^{[0,T]}} \mathbb{E}_{\mathbb{P}}[\mathbb{E}_{\mathbb{P}_\omega}[\xi]] \leq \sup_{\mathbb{P}\in\mathcal{M}^{[0,T]}} \mathbb{E}_{\mathbb{P}}\left[\sup_{\bar{\mathbb{P}}\in\mathcal{M}^{[T_1,T]}_{B^\omega}} \mathbb{E}_{\bar{\mathbb{P}}}[\xi]\right]$$

where $\{\mathbb{P}_\omega\}$ is regular conditional distribution with respect to $\mathcal{F}_{T_1}$ and where in the last step we used measurability implied by assertion (i).

Now we will show the remaining inequality. Let $\{\omega^n\}_n$ be a countable dense subset of $\Omega|_{[0,T_1]}$ and $B^n := B^{\omega^n}$. Define the path modification mapping $\alpha^{n,\omega}$ by $\alpha^{n,\omega} := \alpha^{\omega^n,\omega}$ where $\alpha^{\omega^n,\omega}$ is given in (4.4.1). Note that $\alpha^{n,\omega}$ is a bijection satisfying $\alpha^{n,\omega} = (\alpha^{n,\omega})^{-1}$. For any $\mathbb{P}_n \in \mathcal{M}^{[T_1,T]}_{B^n}$ the measure $\mathbb{P}_n \circ \alpha^{n,\omega}$ belongs to $\mathcal{M}^{[T_1,T]}_{B^\omega}$. Moreover, similarly to (4.4.6), we obtain that

$$|\mathbb{E}_{\mathbb{P}_n}[\xi] - \mathbb{E}_{\mathbb{P}_n\circ\alpha^{n,\omega}}[\xi]| \leq e_\xi(2\delta^{1/2}) + 2||\xi||\frac{\delta}{\delta+\delta^{1/2}}$$

whenever $||\omega^n - \omega||_{[0,T_1]} \leq \delta$. Let us consider probability kernel $N_n : \Omega \to \mathcal{M}^{[T_1,T]}$ defined by $N_n(\omega) := \mathbb{P}_n \circ \alpha^{n,\omega}$. The kernel N_n is $\mathcal{F}_T$-measurable, i.e., $N_n(\omega, F) = \mathbb{P}_n \circ \alpha^{n,\omega}(F)$ is $\mathcal{F}_T$-measurable for any $F \in \mathcal{F}_T$, since $(\omega,\widetilde{\omega}) \to \mathbb{1}_F \circ \alpha^{n,\omega}(\widetilde{\omega})$ is $\mathcal{F}_T \otimes \mathcal{F}_T$-measurable and bounded thus $\mathbb{E}_{\mathbb{P}_n}[\mathbb{1}_F \circ \alpha^{n,\omega}]$ is an $\mathcal{F}_T$-measurable (see [11, Section 3.3]). Then, since N_n is constant on atoms of $\mathcal{F}_{T_1}$, we deduce from Blackwell's Theorem (see Theorem 8.6.7 in Cohn [23] or Chapter III §26 pages 80-81 in Dellacherie and Meyer [36]) that N_n is $\mathcal{F}_{T_1}$-measurable probability kernel.

Denoting the closed ball around ω^n of radius δ by $\widetilde{B}^n(\delta) := \{\omega : \sup_{t\in[0,T_1]} |\omega_t - \omega^n_t| \leq \delta\}$, we observe that

$$\sup_{\mathbb{P}\in\mathcal{M}^{[T_1,T]}_{\widetilde{B}^n(\delta)}} \mathbb{E}_{\mathbb{P}}[\xi] = \sup_{\omega\in\widetilde{B}^n(\delta)} \sup_{\mathbb{P}\in\mathcal{M}^{[T_1,T]}_{B^\omega}} \mathbb{E}_{\mathbb{P}}[\xi] = \sup_{\omega\in\widetilde{B}^n(\delta)} \widehat{\xi}(\omega) \leq \widehat{\xi}(\omega^n) + \varepsilon(\delta)$$

where the inequality follows from uniform continuity of $\widehat{\xi}$.

Fix $\varepsilon > 0$. Then we can chose $\delta > 0$ and family of measures $\mathbb{P}^\varepsilon_n \in \mathcal{M}^{[T_1,T]}_{B^n}$ for each n such that

$$\varepsilon/2 + \mathbb{E}_{\mathbb{P}^\varepsilon_n}[\xi] \geq \widehat{\xi}(\omega) \quad \forall\omega \in \widetilde{B}^n(\delta) \quad \text{and} \quad e_\xi(2\delta^{1/2}) + 2||\xi||\delta/(\delta+\delta^{1/2}) \leq \varepsilon/2.$$

Let us now define the $\mathcal{F}_{T_1}$-measurable probability kernel N^ε as

$$N^\varepsilon(\omega) := \mathbb{1}_{C^n}(\omega)\, \mathbb{P}^\varepsilon_n \circ \alpha^{n,\omega} \quad \text{where} \quad C^n := \widetilde{B}^n \setminus \bigcup_{k=1}^{n-1} \widetilde{B}^k.$$

The probability kernel N^ε is constructed such that it satisfies

$$\varepsilon + \mathbb{E}_{N^\varepsilon(\omega)}(\xi) \geq \widehat{\xi}(\omega) \quad \forall \omega \in \Omega \quad \text{and} \quad N^\varepsilon(\omega) \in \mathcal{M}^{[0,T_1]}_{B^\omega}.$$

There as well exists a measure $\mathbb{P}^\varepsilon \in \mathcal{M}^{[0,T_1]}$ such that $\varepsilon + \mathbb{E}_{\mathbb{P}^\varepsilon}[\widehat{\xi}] \geq \sup_{\mathbb{P}\in\mathcal{M}^{[0,T_1]}} \mathbb{E}_{\mathbb{P}}[\widehat{\xi}]$. The concatenation of measures $\bar{\mathbb{P}}^\varepsilon := \mathbb{P}^\varepsilon \otimes N^\varepsilon$ (see Section 3.1 in El Karoui and Tan [45]), defined, for each $F \in \mathcal{F}_T$, as

$$\bar{\mathbb{P}}^\varepsilon(F) = \mathbb{E}_{\mathbb{P}^\varepsilon}\Big[\sum_n \mathbb{1}_{C^n} N^\varepsilon(F)\Big]$$

is a probability measure. Note that regular conditional probabilities of $\bar{\mathbb{P}}^\varepsilon$ w.r.t $\mathcal{F}_{T_1}$ are equal to N^ε and $\mathrm{d}\,\bar{\mathbb{P}}^\varepsilon|_{\mathcal{F}_{T_1}} = \mathrm{d}\,\mathbb{P}^\varepsilon|_{\mathcal{F}_{T_1}}$. Thus, for $s \leq t$ and $F_s \in \mathcal{F}_s$, we have

$$\begin{aligned}
\mathbb{E}_{\bar{\mathbb{P}}^\varepsilon}\big[(\mathbb{S}_t - \mathbb{S}_s)\mathbb{1}_{F_s}\big] &= \mathbb{E}_{\bar{\mathbb{P}}^\varepsilon}\big[\mathbb{E}_{N^\varepsilon}[(\mathbb{S}_t - \mathbb{S}_s)\mathbb{1}_{F_s}]\big] \\
&= \mathbb{E}_{\bar{\mathbb{P}}^\varepsilon}\big[\mathbb{E}_{N^\varepsilon}[(\mathbb{S}_{s\vee T_1} - \mathbb{S}_s)\mathbb{1}_{F_s}]\big] \\
&= \mathbb{E}_{\bar{\mathbb{P}}^\varepsilon}\big[(\mathbb{S}_{s\vee T_1} - \mathbb{S}_s)\mathbb{1}_{F_s}\big] \\
&= \mathbb{E}_{\mathbb{P}^\varepsilon}\big[(\mathbb{S}_{s\vee T_1} - \mathbb{S}_s)\mathbb{1}_{F_s}\big] \\
&= 0
\end{aligned}$$

which shows that $\mathbb{S}$ is a $\bar{\mathbb{P}}^\varepsilon$-martingale. Moreover $\bar{\mathbb{P}}^\varepsilon$ satisfies

$$\mathbb{E}_{\bar{\mathbb{P}}^\varepsilon}[\xi] = \mathbb{E}_{\mathbb{P}^\varepsilon}\big[\mathbb{E}_{N^\varepsilon}[\xi]\big] \geq \mathbb{E}_{\mathbb{P}^\varepsilon}[\widehat{\xi}] - \varepsilon \geq \sup_{\mathbb{P}\in\mathcal{M}^{[0,T_1]}} \mathbb{E}_{\mathbb{P}}[\widehat{\xi}] - 2\varepsilon.$$

The proof is completed. □

Remark 4.4.3. The dynamic programming principle for **P** stated in Proposition 4.4.2 (ii) is linked to conditional sublinear expectations studied in Nutz and van Handel [82] (see Theorem 2.3 in Nutz and van Handel [82]). Since there is more structure in our set-up we prove it relying on uniform continuity of ξ instead of a general analytic selection argument.

4.5 Discretisation of the dual

This and the subsequent section, are devoted to the proof of (4.3.1) which in turn implies Theorem 4.3.1. The strategy of the proof is inspired by Dolinsky and Soner [39] and proceeds via discretisation, of the dual side in this section and of the primal side in Section 4.6. The duality between discrete counterparts is obtained using classical probabilistic results of Föllmer and Kramkov [48].

We start by describing a discretisation of a continuous path, often referred to as the "Lebesgue discretisation" which will often used. In particular, it will be central to Section 4.5.

Definition 4.5.1. For a positive integer N and any $S \in \Omega$, we set $\tau_0^{(N)}(S) = 0$ and $m_0^{(N)}(S) = 0$, then define

$$\tau_k^{(N)}(S) = \inf\left\{t \geq \tau_{k-1}^{(N)}(S) : |S_t - S_{\tau_{k-1}^{(N)}(S)}| = \frac{1}{2^N}\right\} \wedge T$$

and let $m^{(N)}(S) = \min\{k \in \mathbb{N} : \tau_k^{(N)}(S) = T\}$.

Following the observation that $m^{(N)}(S) < \infty \; \forall S \in \Omega$, we say the sequence of stopping times $0 = \tau_0^{(N)} < \tau_1^{(N)} < \cdots < \tau_{m^{(N)}}^{(N)} = T$ forms a Lebesgue partition of $[0, T]$ on Ω. Similar partitions were studied previously, see e.g. Vovk [105]. Their main appearances have been as tools to build a pathwise version of the Itô's integral. They can also be interpreted, from a financial point of view, as candidate times for rebalancing portfolio holdings, see Whalley and Wilmott [106].

Remark 4.5.2. Note that $m^{(N-2)}(S) \leq m^{(N)}(\tilde{S})$ for any $S, \tilde{S} \in \Omega$ such that $\|S - \tilde{S}\| < 2^{-N}$. To justify this, notice that for each $i < m^{(N-2)}(S)$, $\{\tilde{S}_t \,:\, t \in (\tau_{i-1}^{(N-2)}(S), \tau_i^{(N-2)}(S)] \cap \{k/2^N \,:\, k \in \mathbb{N}_+\}\}$ has at least three elements, which implies that for each $i < m^{(N-2)}(S)$ there exist at least one $j < m^{(N)}(\tilde{S})$ such that $\tau_j^{(N)}(\tilde{S}) \in (\tau_{i-1}^{(N-2)}(S), \tau_i^{(N-2)}(S)]$. In consequence, for any sequence $(\mathbb{P}^{(k)})_{k \geq 1}$ converging to $\mathbb{P}$ weakly and bounded non-increasing function $\phi : \mathbb{N} \to \mathbb{R}$

$$\mathbb{E}_{\mathbb{P}}\Big[\phi\big(m^{(N)}(\mathbb{S})\big)\Big] \leq \liminf_{k \to \infty} \mathbb{E}_{\mathbb{P}^{(k)}}\Big[\phi\big(m^{(N-2)}(\mathbb{S})\big)\Big]. \tag{4.5.1}$$

4.5.1 A discrete-time approximation through simple strategies

Denote by $\mathcal{A}_N$ the set of $\gamma \in \mathcal{A}$ for which we only allow trading in the risky assets to take place at the moments $0 = \tau_0^{(N)}(S) < \tau_1^{(N)}(S) < \cdots < \tau_{m^{(N)}(S)}^{(N)}(S) = T$ and $|\gamma| \leq N$. Set

$$\mathbf{V}^{(N)}(G) := \inf\Big\{x \,:\, \exists \gamma \in \mathcal{A}_N \text{ s.t. } \gamma \text{ superreplicates } G - x\Big\}.$$

Then it is obvious from the definition of $\mathbf{V}^{(N)}$ that $\mathbf{V}^{(N_1)}(G) \geq \mathbf{V}^{(N_2)}(G) \geq \mathbf{V}(G)$ for any $N_2 \geq N_1$, and in fact, the following result states that $\mathbf{V}^{(N)}(G)$ converges to $\mathbf{V}(G)$ asymptotically.

Theorem 4.5.3. For uniformly continuous and bounded G,

$$\lim_{N\to\infty} \mathbf{V}^{(N)}(G) = \mathbf{V}(G).$$

Theorem 4.5.4. For any uniformly continuous and bounded G, $\alpha, \beta \geq 0$ and $D \in \mathbb{N}$

$$\mathbf{V}(G - \alpha \wedge (\beta\sqrt{m^{(D)}})) \leq \mathbf{P}(G - \alpha \wedge (\beta\sqrt{m^{(D-2)}})),$$

where $m^{(D)}$ is defined in Definition 4.5.1.

4.5.2 A countable class of piecewise constant functions

In this section, we construct a countable set of piecewise constant functions which can give approximations to any continuous function S to a certain degree. It will be achieved in three steps. The first step is to use the Lebesgue partition defined in the last section to discretise a continuous function into a piecewise constant function whose jump times are the stopping times. Due to the arbitrary nature of jump times and jump sizes, $F^{(N)}(S)$, the piecewise constant function generated through this procedure, will take values in an uncountable set. To overcome this, in the subsequent two steps, we restrict the jump times and the jump sizes to a countable set and hence define a class of approximating schemes. Note that the final approximation also involves a shift of $F^{(N)}(S)$ to the left, both in time and space. It is for the purpose of dealing with some delicate questions of measurability. Later, we use the final approximation to lift predictable maps defined on $\mathbb{D}([0,T],\mathbb{R}^d)$ to the initial space Ω, and the shifts can result in the lifted maps on Ω being progressively measurable. The lifting procedure is defined in Definition 4.5.7 and the measurability is proved in Lemma 4.5.8 below.

Step 1. Let $\tau_k^{(N)}(S)$ and $m^{(N)}(S)$ be defined as in Subsection 4.5.1. To simplify notations, in this section we often suppress their dependences on S and N and write

$$m = m^{(N)}(S), \quad \tau_k = \tau_k^{(N)}(S) \qquad \text{for any } k, N.$$

Our first naive approximation $F^{(N)} : \Omega \to \mathbb{D}([0,T],\mathbb{R}^d)$ is as follows:

$$F_t^{(N)}(S) = \sum_{k=0}^{m-1} S_{\tau_k} \mathbb{1}_{[\tau_k,\tau_{k+1})}(t) + S_T \mathbb{1}_{\{T\}}(t) \quad \text{for } t \in [0,T],\ S \in \Omega. \tag{4.5.2}$$

Note that $F^{(N)}(\mathbb{S})$ is piecewise constant and $\|F^{(N)}(\mathbb{S}) - \mathbb{S}\| \leq 1/2^N$.

Step 2. Define a map $\pi^{(N)} : \mathbb{R}^d_+ \to A^{(N)} := \{2^{-N}k \,:\, k = (k_1, \ldots, k_d) \in \mathbb{N}^d\}$ as

$$\pi^{(N)}(x)_i := 2^{-N}\lceil 2^N x_i \rceil, \quad i = 1, \ldots, d.$$

We then define our second approximation $\check{F}^{(N)} : \Omega \to \mathbb{D}([0,T], \mathbb{R}^d)$ by

$$\begin{aligned}\check{F}^{(N)}_t(S) =& (S_0 - \pi^{(N+1)}(S_{\tau_1})) + \sum_{k=0}^{m-2} \pi^{(N+k+1)}(S_{\tau_{k+1}}) \mathbb{1}_{[\tau_k, \tau_{k+1})}(t) \\ &+ \pi^{(N+m)}(S_{\tau_m}) \mathbb{1}_{[\tau_{m-1}, T]}(t) \qquad t \in [0,T].\end{aligned}$$

Step 3.

We now construct the shifted jump times $\hat{\tau}^{(N)}_k : \Omega \to \mathbb{Q}_+ \cup \{T\}$. Firstly, set $\hat{\tau}^{(N)}_0 = 0$. Then, for any $S \in \Omega$ and $k = 1, \cdots, m^{(N)}(S)$, let

$$\Delta\hat{\tau}^{(N)}_k = \frac{p_k}{q_k} \quad \text{with } (p_k, q_k) = \operatorname{argmin}\{p + q \,:\, (p,q) \in \mathbb{N}^2,\ \tau^{(N)}_{k-1} - \hat{\tau}^{(N)}_{k-1} < \frac{p}{q} \le \Delta\tau^{(N)}_k\}$$

if $k < m^{(N)}(S)$ and otherwise

$$\Delta\hat{\tau}^{(N)}_k = T - \hat{\tau}^{(N)}_{m^{(N)}-1},$$

where $\Delta\tau^{(N)}_k := \tau^{(N)}_k - \tau^{(N)}_{k-1}$. Lastly, define $\hat{\tau}^{(N)}_k := \sum_{i=1}^k \Delta\hat{\tau}^{(N)}_i$. Here we also suppress the dependences of these shifted jump times on S and N and write

$$\hat{\tau}_k = \hat{\tau}^{(N)}_k(S) \qquad \text{for any } k, N.$$

Clearly $0 = \hat{\tau}_0 < \hat{\tau}_1 < \hat{\tau}_2 \cdots < \hat{\tau}_m = T$, $\tau_{k-1} < \hat{\tau}_k \le \tau_k\ \forall\, k < m$ and $\hat{\tau}_m = \tau_m = T$. These $\hat{\tau}$'s are the shifted versions of τ's, and are uniquely defined for any S. We are going to use $\hat{\tau}$'s to define a class of approximating schemes.

We can define an approximation $\hat{F}^{(N)} : \Omega \to \mathbb{D}([0,T], \mathbb{R}^d)$ by

$$\begin{aligned}\hat{F}^{(N)}_t(S) =& (S_0 - \pi^{(N+1)}(S_{\tau_1})) + \sum_{k=0}^{m-2} \pi^{(N+k+1)}(S_{\tau_{k+1}}) \mathbb{1}_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t) \\ &+ \pi^{(N+m)}(S_{\tau_m}) \mathbb{1}_{[\hat{\tau}_{m-1}, T]}(t) \qquad t \in [0,T].\end{aligned}$$

Notice that $\hat{F}^{(N)}(\mathbb{S})$ is piecewise constant and

$$\begin{aligned}\|\hat{F}^{(N)}(\mathbb{S}) - \mathbb{S}\| \le& \|\hat{F}^{(N)}(\mathbb{S}) - \check{F}^{(N)}(\mathbb{S})\| + \|\check{F}^{(N)}(\mathbb{S}) - F^{(N)}(\mathbb{S})\| + \|F^{(N)}(\mathbb{S}) - \mathbb{S}\| \\ \le& \frac{2}{2^{N-1}} + \frac{2}{2^N} + \frac{1}{2^N} < \frac{1}{2^{N-3}}.\end{aligned} \tag{4.5.3}$$

Definition 4.5.5. Let $\hat{\mathbb{D}}^{(N)} \subset \mathbb{D}([0,T], \mathbb{R}^d)$ be the set of functions $f = (f^{(i)})_{i=1}^d$ which satisfy the following,

1. for any $i = 1, \dots, d$, $f^{(i)}(0) = 1$,
2. f is piecewise constant with jumps only at times $t_1, \cdots, t_{l-1} \in \mathbb{Q}_+$ for some $l < \infty$,
   where $t_0 = t_{l_0} = 0 < t_1 < t_2 < \cdots < t_{l-1} < T$,
3. for any $k = 1, \dots, l-1$ and $i = 1, \dots, d$, $f^{(i)}(t_k) - f^{(i)}(t_{k-1}) = j/2^{N+k}$, for $j \in \mathbb{Z}$ with $|j| \le 2^k$,
4. $\inf_{t \in [0,T],\, 1 \le i \le d} f^{(i)}(t) \ge -2^{-N+3}$,
5. if $f^{(i)}(t_k) = -2^{-N+3}$ for some $i \le d$ and $k \le l-1$, then $f(t_j) = f(t_k)\ \forall\, k < j < l$.

It is clear that $\hat{\mathbb{D}}^{(N)}$ is countable.

4.5.3 A countable probabilistic structure

Let $\hat{\Omega} := \mathbb{D}([0,T], \mathbb{R}^d)$ be the space of all right continuous functions $f : [0,T] \to \mathbb{R}^d$ with left-hand limits. Denote by $\hat{\mathbb{S}} = (\hat{\mathbb{S}}_t)_{0 \le t \le T}$ the canonical process on the space $\hat{\Omega}$.

The set $\hat{\mathbb{D}}^{(N)}$ is a countable subset of $\hat{\Omega}$. There exists a local martingale measure $\hat{\mathbb{P}}^{(N)}$ on $\hat{\Omega}$ which satisfies $\hat{\mathbb{P}}^{(N)}(\hat{\mathbb{D}}^{(N)}) = 1$ and $\hat{\mathbb{P}}^{(N)}(\{f\}) > 0$ for all $f \in \hat{\mathbb{D}}^{(N)}$. In fact, such a local martingale measure $\hat{\mathbb{P}}^{(N)}$ on $\hat{\mathbb{D}}^{(N)}$ can be constructed 'by hand'. Indeed, we can construct a continuous Markov chain that undergoes transitions in the finite number of allowed values in the way that the mean is preserved, with jump times decided via an exponential clock. Let $\hat{\mathbb{F}}^{(N)} := \{\hat{\mathcal{F}}_t^{(N)}\}_{0 \le t \le T}$ be the filtration generated by the process $\hat{\mathbb{S}}$ and satisfying the usual assumptions (right continuous and contains $\hat{\mathbb{P}}^{(N)}$-null sets).

In the last section, we saw definitions of $\hat{\tau}_k^{(N)}$ on Ω. Here we extend their definitions to $\bigcup_{N \in \mathbb{N}} \hat{\mathbb{D}}^{(N)}$. Define the jump times by setting $\hat{\tau}_0(\hat{\mathbb{S}}) = 0$ and for $k > 0$,

$$\hat{\tau}_k(\hat{\mathbb{S}}) = \inf\left\{t > \hat{\tau}_{k-1}(\hat{\mathbb{S}}) : \hat{\mathbb{S}}_t \ne \hat{\mathbb{S}}_{t-}\right\} \wedge T. \tag{4.5.4}$$

Next we introduce the random time before T

$$m(\hat{\mathbb{S}}) := \min\{k : \hat{\tau}_k(\hat{\mathbb{S}}) = T\}.$$

Observe that for $S \in \Omega$, $\hat{F}^{(N)}(S) \in \hat{\mathbb{D}}^{(N)}$, $\hat{\tau}_k(\hat{F}^{(N)}(S)) = \hat{\tau}_k(S)$ for all k and $m(\hat{F}^{(N)}(S)) = m^{(N)}(S)$. It follows that the definitions are consistent.

In this context, a trading strategy $(\hat{\gamma}_t)_{t=0}^T$ on the filtered probability space $(\hat{\Omega}, \hat{\mathbb{F}}^{(N)}, \hat{\mathbb{P}}^{(N)})$ is a predictable stochastic process with respect to $\hat{\mathbb{F}}^{(N)}$. Thus, $\hat{\gamma}$ is a map from $\mathbb{D}([0,T],\mathbb{R}^d)$ to $\mathcal{D}([0,T],\mathbb{R}^d)$. Choose $a \in \mathcal{D}([0,T],\mathbb{R}^d)$ such that $a \notin \hat{\gamma}(\hat{\mathbb{D}}^{(N)})$ and then define a map $\phi : \mathbb{D}([0,T],\mathbb{R}^d) \to \mathcal{D}([0,T],\mathbb{R}^d)$ by $\phi(\hat{S}) = \hat{\gamma}(\hat{S})$ if $\hat{S} \in \hat{\mathbb{D}}^{(N)}$, and equal to a otherwise. Since $\hat{\mathbb{P}}^{(N)}$ has full support on $\hat{\mathbb{D}}^{(N)}$, $\hat{\gamma} = \phi(\hat{\mathbb{S}})$ $\hat{\mathbb{P}}^{(N)}$-a.s.. In particular, for any $A \in \mathcal{B}(\mathbb{R})$, the symmetric difference of $\{\hat{\gamma}_t \in A\}$ and $\{\phi(\hat{\mathbb{S}})_t \in A\}$ is a null set for $\hat{\mathbb{P}}^{(N)}$. Thus ϕ is a predictable map. Furthermore, since $\hat{\mathbb{P}}^{(N)}$ charges all elements in $\hat{\mathbb{D}}^{(N)}$, for any $\upsilon, \tilde{\upsilon} \in \hat{\mathbb{D}}^{(N)}$ and $t \in [0,T]$.

$$\upsilon_u = \tilde{\upsilon}_u \quad \forall u \in [0,t) \quad \Rightarrow \quad \phi(\upsilon)_t = \phi(\tilde{\upsilon})_t. \tag{4.5.5}$$

Indeed, suppose there exist $t \in [0,T]$ and $\upsilon, \tilde{\upsilon} \in \hat{\mathbb{D}}^{(N)}$ such that $\upsilon_u = \tilde{\upsilon}_u$ for all $u \in [0,t)$ and $\phi(\upsilon)_t \neq \phi(\tilde{\upsilon})_t$. Since $\hat{\gamma}$ is predictable, we have

$$\hat{\mathcal{F}}_{t-}^{(N)} \ni \{\hat{\gamma}_t = \phi(\upsilon)_t\} \cap \{\mathbb{S}_u = \upsilon_u, u < t\} = \{\hat{\gamma}_t = \phi(\upsilon)_t\} \cap \{\mathbb{S}_u = \upsilon_u, u < t\} \cap \{\mathbb{S}_t \neq \tilde{\upsilon}_t\},$$

which is a contradiction since $\{\hat{\gamma}_t = \phi(\upsilon)_t\} \cap \{\mathbb{S}_u = \upsilon_u, u < t\}$ is not a null set and hence not in $\hat{\mathcal{F}}_{t-}^{(N)}$. We conclude that any predictable process $\hat{\gamma}$ has a version ϕ that is predictable in the sense of (4.5.5). In what follows we always take this version.

In this section, we formally define the probabilistic super-replicating problem and later build a connection between the probabilistic super-replication problem on the discretised space and the pathwise discretised robust hedging problem. For the rest of the section, we write $\int_{t_1}^{t_2}$ to mean $\int_{(t_1,t_2]}$.

As G is defined only on Ω, to consider paths in $\hat{\Omega}$, we need to extend the domain of G to $\hat{\Omega}$. For most of the financial contracts, the extension is natural. However, here we pursue a general approach. We first define a projection function $\dashv : \hat{\Omega} \to \mathcal{C}([0,T],\mathbb{R}^d)$ by

$$\dashv(\hat{S}) = \begin{cases} \hat{S} & \text{if } \hat{S} \text{ is continuous} \\ \sum_{k=0}^{m(\hat{S})-1} \left(\frac{\hat{S}_{\hat{\tau}_{k+1}} - \hat{S}_{\hat{\tau}_k}}{\hat{\tau}_{k+1} - \hat{\tau}_k} (t - \hat{\tau}_k) + \hat{S}_{\hat{\tau}_k} \right) \mathbb{1}_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t) & \text{if } \hat{S} \in \bigcup_{N \in \mathbb{N}} \hat{\mathbb{D}}^{(N)} \\ \omega^1 & \text{otherwise,} \end{cases}$$

where ω^1 is the constant path equal to 1. In fact, when $\hat{S} \in \bigcup_{N \in \mathbb{N}} \hat{\mathbb{D}}^{(N)}$, $\dashv(\hat{S})$ is the minimum of 0 and the linear interpolation function of

$$\left((\hat{\tau}_0(\hat{S}), \hat{S}_{\hat{\tau}_0(\hat{S})}), \ldots, (\hat{\tau}_{m(\hat{S})}(\hat{S}), \hat{S}_{\hat{\tau}_{m(\hat{S})}(\hat{S})}) \right).$$

We then can define $\hat{G} : \hat{\Omega} \to \Omega$ via this explicit projection $\daleth$ by $\hat{G}(\hat{S}) = G(\daleth(\hat{S}) \vee 0)$, where $\hat{S} \vee 0 := \big((\hat{S}_t^{(1)} \vee 0, \ldots, \hat{S}_t^{(d)} \vee 0)\big)_{0\le t\le T}$ for any $S \in \hat{\Omega}$.

Note that G and $\hat{G}$ are equal on Ω. In addition, for every $N \in \mathbb{N}$ and $\hat{S} \in \hat{\mathbb{D}}^{(N)}$, we have

$$\|\daleth(\hat{S}) - \hat{S}\| \le 2^{-N+1}. \tag{4.5.6}$$

Therefore, we can deduce that

$$\begin{aligned}\|\daleth(\hat{F}(S)) \vee 0 - S\| \le& \|\daleth(\hat{F}(S)) \vee 0 - \hat{F}(S) \vee 0\| + \|\hat{F}(S) \vee 0 - S\| \\ \le& 2^{-N+1} + 2^{-N+3} \quad \forall S \in \Omega.\end{aligned} \tag{4.5.7}$$

where the last inequality follows from (4.5.3) and (4.5.6).

Similarly, for each $D \in \mathbb{N}$, we define $\hat{m}^{(D)} : \hat{\Omega} \to \Omega$ by $\hat{m}^{(D)}(\hat{S}) = m^{(D)}(\daleth(\hat{S}) \vee 0)$. Then by Remark 4.5.2 and (4.5.7), when N is sufficiently large,

$$\hat{m}^{(D-2)}(\hat{F}^{(N)}(S))) \le m^{(D)}(S) \quad \forall\, S \in \Omega. \tag{4.5.8}$$

Definition 4.5.6.

1. $\hat{\gamma} : \hat{\Omega} \to \mathcal{D}([0,T], \mathbb{R}^d)$ is $\hat{\mathbb{P}}^{(N)}$-admissible if $\hat{\gamma}$ is predictable and bounded by N, and the stochastic integral $(\int_0^t \hat{\gamma}_u(\hat{\mathbb{S}})\mathrm{d}\hat{\mathbb{S}}_u)_{0\le t\le T}$ is well defined under $\hat{\mathbb{P}}^{(N)}$, satisfying that $\exists\, M > 0$ such that

$$\int_0^t \hat{\gamma}_u(\hat{\mathbb{S}})\mathrm{d}\hat{\mathbb{S}}_u \ge -M \quad \hat{\mathbb{P}}^{(N)} - \text{a.s.}, \quad t \in [0,T). \tag{4.5.9}$$

2. An admissible strategy $\hat{\gamma}$ is said to $\hat{\mathbb{P}}^{(N)}$-*superreplicate* $\hat{G}$ if

$$\int_0^T \hat{\gamma}_u(\hat{\mathbb{S}})\mathrm{d}\hat{\mathbb{S}}_u \ge \hat{G}(\hat{\mathbb{S}}), \quad \hat{\mathbb{P}}^{(N)} - \text{a.s.}. \tag{4.5.10}$$

3. The super-replicating cost of $\hat{G}$ is defined as

$$\hat{\mathbb{V}}^{(N)} := \inf\{x \,:\, \exists\hat{\gamma} \text{ s.t. } \hat{\gamma} \text{ is } \hat{\mathbb{P}}^{(N)}\text{-admissible and } \hat{\mathbb{P}}^{(N)}\text{-superreplicates } \hat{G} - x\}$$

For the rest of the section we will establish connections between probabilistic superhedging problems and discretised robust hedging problems. Our reasoning is close to the one in Dolinsky and Soner [39].

Definition 4.5.7. Given a predictable stochastic process $(\hat{\gamma}_t)_{t=0}^T$ on $(\hat{\Omega}, \hat{\mathbb{F}}^{(N)}, \hat{\mathbb{P}}^{(N)})$, we define $\gamma^{(N)} : \Omega \to \mathbb{D}([0,T], \mathbb{R}^d)$ by

$$\gamma_t^{(N)}(S) := \sum_{k=0}^{m-1} \hat{\gamma}_{\hat{\tau}_k}(\hat{F}^{(N)}(S)) \mathbb{1}_{(\tau_k, \tau_{k+1}]}(t), \tag{4.5.11}$$

where $\tau_k = \tau_k^{(N)}(S)$, $m = m^{(N)}(S)$ are given in Definition 4.5.1 and $\hat{\tau}_k = \hat{\tau}_k(\hat{F}^{(N)}(S))$ are given in (4.5.4).

Lemma 4.5.8. For any admissible process $\hat{\gamma}$ in the sense of Definition 4.5.6, $\gamma^{(N)}$ defined in (4.5.11) is $\mathbb{F}$-predictable.

Proof. We first show that equipping Ω and $\mathbb{D}([0,T], \mathbb{R}^d)$ with respective σ-algebras $\mathcal{F}_T$ and $\hat{\mathcal{F}}_T^{(N)}$, $\hat{F}^{(N)} : \Omega \to \mathbb{D}([0,T], \mathbb{R}^d)$ is measurable.

Since $\hat{\mathbb{P}}^{(N)}$ has full support on $\hat{\mathbb{D}}^{(N)}$, for any $A \in \hat{\mathcal{F}}_T^{(N)}$,

$$\{\hat{F}^{(N)} \in A\} = \cup_{\hat{S} \in \hat{\mathbb{D}}^{(N)}} \{S \in \Omega \, : \, \hat{F}^{(N)}(S) = \hat{S}\}.$$

It is clear from the construction that $\hat{\tau}_k^{(N)}$, $m^{(N)}$ and $\pi^{(N+k)}$ are all $\mathcal{F}_T$ measurable, and we notice that for any $\hat{\mathbb{S}} \in \hat{\mathbb{D}}^{(N)}$

$$\{S \in \Omega \, : \, \hat{F}^{(N)}(S) = \hat{S}\} = \{S \in \Omega \, : \, m^{(N)} = m, \hat{\tau}_k^{(N)} = t_k, \pi^{(N+i)} = s_k \forall k < m\}$$

for some m, t_k, s_k. Therefore, we can conclude that $\hat{F}^{(N)}$ has the desired measurability.

To prove $\gamma^{(N)}$ is $\mathbb{F}$-predictable, we need to show $\hat{\gamma}_{\hat{\tau}_k} \circ \hat{F}^{(N)}$ is $\mathcal{F}_{\tau_k}$ measurable, or equivalently (due to Remark 4.2.3) that for any $\omega, \upsilon \in \Omega$ such that $\omega_u = \upsilon_u$ for any $u \leq \theta := \tau_k^{(N)}(\omega) = \tau_k^{(N)}(\upsilon)$,

$$\hat{\gamma}_\theta\big(\hat{F}^{(N)}(\omega)\big) = \hat{\gamma}_\theta\big(\hat{F}^{(N)}(\upsilon)\big).$$

It follows from the definition of $\hat{F}^{(N)}$ that

$$\hat{F}_u^{(N)}(\omega) = \hat{F}_u^{(N)}(\upsilon) \quad \forall u \in [0, \theta).$$

Hence, by (4.5.5)

$$\hat{\gamma}_\theta\big(\hat{F}^{(N)}(\omega)\big) = \hat{\gamma}_\theta\big(\hat{F}^{(N)}(\upsilon)\big).$$

Therefore, we conclude that $\gamma^{(N)}$ defined in (4.5.11) is $\mathbb{F}$-predictable. □

The following theorem is crucial. It states that the probabilistic super-replicating value is asymptotically larger than the value of the discretised robust hedging problem.

Proposition 4.5.9. For uniformly continuous and bounded G, $\alpha, \beta \geq 0$ and $D \in \mathbb{N}$, we have

$$\liminf_{N\to\infty} \mathbf{V}^{(N)}\Big(G(\mathbb{S}) - \alpha \wedge (\beta\sqrt{m^{(D)}(\mathbb{S})})\Big) \leq \liminf_{N\to\infty} \hat{\mathbb{V}}^{(N)}\Big(\hat{G}(\hat{\mathbb{S}}) - \alpha \wedge (\beta\sqrt{\hat{m}^{(D-2)}(\hat{\mathbb{S}})})\Big). \tag{4.5.12}$$

Proof. See Section 4.7.1. □

4.5.4 Duality for the discretised problems

Definition 4.5.10.

1. Let $\hat{\Pi}^{(N)}$ be the set of all probability measures $\hat{\mathbb{Q}}$ which are equivalent to $\hat{\mathbb{P}}^{(N)}$.
2. For any $\kappa \geq 0$, denote $\hat{\mathbb{M}}^{(N)}(\kappa)$ by the set of all probability measures $\hat{\mathbb{Q}} \in \hat{\Pi}^{(N)}$ such that
$$\mathbb{E}_{\hat{\mathbb{Q}}}\Bigg[\sum_{k=1}^{m(\hat{\mathbb{S}})}\sum_{i=1}^{d}|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{\mathbb{S}}^{(i)}_{\hat{\tau}_k}|\hat{\mathcal{F}}^{(N)}_{\hat{\tau}_k-}] - \hat{\mathbb{S}}^{(i)}_{\hat{\tau}_{k-1}}|\Bigg] \leq \frac{\kappa}{N},$$
where $\hat{\tau}_k = \hat{\tau}_k(\hat{\mathbb{S}})$ and $m = m(\hat{\mathbb{S}})$ are as defined in (4.5.4).

Lemma 4.5.11. Suppose $\hat{G}$ is bounded by $\kappa - 1$. Then, there are at most finitely many $N \in \mathbb{N}$ such that $\hat{\mathbb{M}}^{(N)}(2\kappa) = \varnothing$ and

$$\liminf_{N\to\infty} \hat{\mathbb{V}}^{(N)}(\hat{G}(\hat{\mathbb{S}})) \leq \liminf_{N\to\infty} \sup_{\hat{\mathbb{Q}}\in\hat{\mathbb{M}}^{(N)}(2\kappa)} \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})]. \tag{4.5.13}$$

Proof. Since for any $\hat{\mathbb{Q}} \in \hat{\Pi}^{(N)}$ the support of $\hat{\mathbb{Q}}$ is $\hat{\mathbb{D}}^{(N)}$, of which elements are piece-wise constant, the canonical process $\hat{\mathbb{S}}$ is therefore a semi-martingale under $\hat{\mathbb{Q}}$. Moreover, it has the following decomposition, $\hat{\mathbb{S}} = \hat{M}^{\hat{\mathbb{Q}}} + \hat{A}^{\hat{\mathbb{Q}}}$ where

$$\begin{aligned}\hat{A}^{\hat{\mathbb{Q}}}_t &= \sum_{k=1}^{m(\hat{\mathbb{S}})}\Big[\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{\mathbb{S}}_{\hat{\tau}_k}|\hat{\mathcal{F}}^{(N)}_{\hat{\tau}_k-}] - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}}\Big]\mathbb{1}_{[\hat{\tau}_k,\hat{\tau}_{k+1})}(t), \quad t < T,\\ \hat{A}^{\hat{\mathbb{Q}}}_T &:= \lim_{t\uparrow T}\hat{A}^{\hat{\mathbb{Q}}}_t\end{aligned}$$

is a predictable process of bounded variation and $\hat{M}^{\hat{\mathbb{Q}}}$ is a martingale under $\hat{\mathbb{Q}}$. Then, similar to Dolinsky and Soner [41], it follows from Example 2.3 and Proposition 4.1 in Föllmer and Kramkov [48] that

$$\hat{\mathbb{V}}^{(N)}(\hat{G}(\hat{\mathbb{S}})) = \sup_{\hat{\mathbb{Q}}\in\hat{\Pi}^{(N)}} \mathbb{E}_{\hat{\mathbb{Q}}}\left[\hat{G}(\hat{\mathbb{S}}) - N\sum_{k=1}^{m(\hat{\mathbb{S}})}\sum_{i=1}^{d} |\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{\mathbb{S}}^{(i)}_{\hat{\tau}_k}|\hat{\mathcal{F}}^{(N)}_{\hat{\tau}_k-}] - \hat{\mathbb{S}}^{(i)}_{\hat{\tau}_{k-1}}|\right]. \tag{4.5.14}$$

By Proposition 4.5.9,

$$\liminf_{N\to\infty} \hat{\mathbb{V}}^{(N)}(\hat{G}(\hat{\mathbb{S}})) \geq \liminf_{N\to\infty} \mathbf{V}^{(N)}(G) \geq \mathbf{P}(G) > -\kappa.$$

Then, in (4.5.14), it suffices to consider the supremum over $\hat{\mathbb{M}}^{(N)}(2\kappa)$. In particular, $\hat{\mathbb{M}}^{(N)}(2\kappa) \neq \varnothing$ for N large enough.

□

4.6 Discretisation of the primal

4.6.1 Approximation of Martingale Measures

Next, we show that we can lift any discrete martingale measure in $\hat{\mathbb{M}}^{(N)}(c)$ to a continuous martingale measure in $\underline{\mathcal{M}}$ such that the difference of expected value of G under this continuous martingale measure and the expected value of $\hat{G}$ under the original discrete martingale measure is within a bounded error, which goes to zero as $N \to \infty$. Through this, we connect the primal problems on the discretised space to the approximation of the primal problems on the space of continuous functions asymptotically.

Proposition 4.6.1. If G is uniformly continuous and bounded by $\kappa - 1$ for some $\kappa \geq 1$, then for any $\alpha, \beta \geq 0$, $D \in \mathbb{N}$

$$\limsup_{N\to\infty} \sup_{\hat{\mathbb{Q}}\in\hat{\mathbb{M}}^{(N)}(2\kappa+\alpha)} \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}}) - \alpha\wedge(\beta\sqrt{\hat{m}^{(D)}(\hat{\mathbb{S}})})] \leq \sup_{\mathbb{P}\in\underline{\mathcal{M}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \alpha\wedge(\beta\sqrt{m^{(D-2)}(\mathbb{S})})]. \tag{4.6.1}$$

Proof. Let $f_e : \mathbb{R}^d_+ \to \mathbb{R}_+$ be the modulus of continuity of G, i.e.,

$$|G(\omega) - G(\upsilon)| \leq f_e(|\omega - \upsilon|) \quad \text{for any } \omega, \upsilon \in \Omega$$

and $\lim_{x\searrow 0} f_e(x) = 0$. Recall from Lemma 4.5.11 that $\hat{\mathbb{M}}^{(N)}(2\kappa+2\alpha) \neq \varnothing$ for N large enough. Hence, to show (4.6.1), it suffices to prove that for any $\hat{\mathbb{Q}} \in \hat{\mathbb{M}}^{(N)}(2\kappa+2\alpha)$

$$\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}}) - \alpha \wedge (\beta\sqrt{\hat{m}^{(D)}(\hat{\mathbb{S}})})] \leq \sup_{\mathbb{P}\in\underline{\mathcal{M}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \alpha \wedge (\beta\sqrt{m^{(D-2)}(\mathbb{S})})] + g(1/N), \tag{4.6.2}$$

for some $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{x\searrow 0} g(x) = 0$. We now fix N and $\hat{\mathbb{Q}} \in \hat{\mathbb{M}}^{(N)}(2\kappa+2\alpha)$ and prove (4.6.2) in four steps.

Step 1. We will first construct a semi-martingale $\hat{Z} = \hat{M} + \hat{A}$ on a Wiener space $(\Omega^W, \mathcal{F}^W, P^W)$ such that

$$\left|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})] - E^W[\hat{G}(\hat{Z})]\right| \leq \kappa 2^{-N+1}, \tag{4.6.3}$$

where $\hat{M}$ is constructed from a martingale and both have piece-wise constant paths.

Since the measure $\hat{\mathbb{Q}}$ is supported on $\hat{\mathbb{D}}^{(N)}$, the canonical process $\hat{\mathbb{S}}$ is a pure jump process under $\hat{\mathbb{Q}}$, with a finite number of jumps $\hat{\mathbb{Q}}$-a.s. Consequently there exists a deterministic positive integer m_0 (depending on N) such that

$$\hat{\mathbb{Q}}(\hat{m}(\hat{\mathbb{S}}) > m_0) < 2^{-N}. \tag{4.6.4}$$

It follows that

$$|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}}^{\hat{\tau}_{m_0}})]| \leq \kappa 2^{-N+1}. \tag{4.6.5}$$

Notice that by definition of $\hat{\mathbb{D}}^{(N)}$, the law of $\hat{\mathbb{S}}^{\hat{\tau}_{m_0}}$ under $\hat{\mathbb{Q}}$ is also supported on $\hat{\mathbb{D}}^{(N)}$.

Let $(\Omega^W, \mathcal{F}^W, P^W)$ be a complete probability space together with a standard m_0+2-dimensional Brownian motion $\left\{W_t = (W_t^{(1)}, \cdots, W_t^{(m_0+2)})\right\}_{t=0}^{\infty}$ and the natural filtration $\mathcal{F}_t^W = \sigma\{W_s | s \leq t\}$. With a small modification to Lemma 5.1 in Dolinsky and Soner [39], we can construct a sequence of stopping times (with respect to Brownian filtration) $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_{m_0}$ together with $\mathcal{F}_{\sigma_i}^W$-measurable random variable Y_i's such that

$$\begin{aligned}&\mathcal{L}_{P^W}\big((\sigma_1, \ldots, \sigma_{m_0}, Y_1, \ldots, Y_{m_0})\big)\\ =&\mathcal{L}_{\hat{\mathbb{Q}}}\big((\hat{\tau}_1, \ldots, \hat{\tau}_{m_0}, \hat{\mathbb{S}}_{\hat{\tau}_1} - \hat{\mathbb{S}}_{\hat{\tau}_0}, \ldots, \hat{\mathbb{S}}_{\hat{\tau}_{m_0}} - \hat{\mathbb{S}}_{\hat{\tau}_{m_0-1}})\big).\end{aligned} \tag{4.6.6}$$

(Detailed construction is provided in the Section 4.7.2.)

Define X_i as

$$X_i = E^W[Y_i | \mathcal{F}_{\sigma_{i-1}}^W \vee \sigma(\sigma_i)], \quad i = 1, \ldots, m_0.$$

Note that $|X_i| \leq 2^{-N}$. Also by construction of σ_i's and Y_i's, we have

$$E^W[Y_i|\mathcal{F}^W_{\sigma_{i-1}} \vee \sigma(\sigma_i)] = E^W[Y_i|\vec{\sigma}_i, \vec{Y}_{i-1}],$$

where $\vec{\sigma}_i := (\sigma_1, \ldots, \sigma_i)$, $\vec{Y}_i := (Y_1, \ldots, Y_i)$ and E^W is the expectation with respect to P^W.

From these, we can construct a jump process $(\hat{A}_t)_{t=0}^T$ by

$$\hat{A}_t = \sum_{j=1}^{m_0} X_j \mathbb{1}_{[\sigma_j, T]}.$$

In particular, for $k \leq m_0$

$$\hat{A}_{\sigma_k} = \sum_{j=1}^{k} X_j.$$

Set a martingale $(M_t)_{t=0}^T$ as

$$M_t = 1 + E^W\Big[\sum_{j=1}^{m_0}(Y_j - X_j)|\mathcal{F}^W_t\Big], \quad t \in [0,T]. \tag{4.6.7}$$

Since all Brownian martingales are continuous, so is M. Moreover, Brownian motion increments are independent and therefore,

$$M_{\sigma_k} = 1 + \sum_{j=1}^{k}(Y_j - X_j), \quad P^W - \text{a.s.}, \quad k \leq m_0.$$

We now introduce a stochastic process $(\hat{M}_t)_{t=0}^T$, on the Brownian probability space, by setting $\hat{M}_t = M_{\sigma_k}$ for $t \in [\sigma_k, \sigma_{k+1})$, $k < m_0$ and $\hat{M}_t = \hat{M}_{\sigma_{m_0}}$ for $t \in [\sigma_{m_0}, T]$. Note that as $|Y_i - X_i| \leq 2^{-N+1}$, for any $k \leq m_0$ and $t \leq T$

$$\begin{aligned}
&\big|\hat{M}_{t\wedge\sigma_{k+1}\vee\sigma_k} - M_{t\wedge\sigma_{k+1}\vee\sigma_k}\big| \\
=&\Big|\sum_{j=k+1}^{m_0} E^W[(Y_j - X_j)|\mathcal{F}^W_{t\wedge\sigma_{k+1}\vee\sigma_k}]\Big| \\
=&\Big|\sum_{j=k+2}^{m_0} E^W\big[E^W[(Y_j - X_j)|\mathcal{F}^W_{\sigma_{j-1}} \vee \sigma(\sigma_j)]\mathcal{F}^W_{t\wedge\sigma_{k+1}\vee\sigma_k}\big] + E^W[Y_{k+1} - X_{k+1}|\mathcal{F}^W_{t\wedge\sigma_{k+1}\vee\sigma_k}]\Big| \\
=&\Big|E^W[Y_{k+1} - X_{k+1}|\mathcal{F}^W_{t\wedge\sigma_{k+1}\vee\sigma_k}]\Big| \leq E^W[|Y_{k+1} - X_{k+1}||\mathcal{F}^W_{t\wedge\sigma_{k+1}\vee\sigma_k}] \leq 2^{-N+1}
\end{aligned}$$

and hence

$$\|\hat{M} - M\| < 2^{-N+2}. \tag{4.6.8}$$

We also notice that $\hat{Z} = \hat{M} + \hat{A}$ satisfies $\hat{Z}_0 = \hat{\mathbb{S}}_0$ and

$$\mathcal{L}_{P^W}\big((\sigma_1, \ldots, \sigma_{m_0}, Y_1, \ldots, Y_{m_0})\big) = \mathcal{L}_{\hat{\mathbb{Q}}}\big((\hat{\tau}_1, \ldots, \hat{\tau}_{m_0}, \hat{\mathbb{S}}_{\hat{\tau}_1} - \hat{\mathbb{S}}_{\hat{\tau}_0}, \ldots, \hat{\mathbb{S}}_{\hat{\tau}_{m_0}} - \hat{\mathbb{S}}_{\hat{\tau}_{m_0-1}})\big).$$

It follows that

$$E^W[\hat{G}(\hat{Z})] = \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}}^{\tau_{m_0}})]. \tag{4.6.9}$$

In particular, by (4.6.5) we see that (4.6.3) holds.

Step 2. We will shortly construct a continuous martingale M^{θ_0} from M such that M^{θ_0} is bounded below by $-2^{-N+2} - N^{-\frac{1}{2}}$ and

$$|E^W[\hat{G}(M^{\theta_0})] - \mathbb{E}_{\hat{\mathbb{Q}}}\big[\hat{G}(\hat{\mathbb{S}})\big]| \leq c^2 N^{-\frac{1}{2}} + 2f_e(N^{-\frac{1}{2}} + 2^{-N+2}) + 2^{-N}. \tag{4.6.10}$$

As the law of $\hat{Z}$ is the same as $\hat{\mathbb{S}}^{m_0}$ under $\hat{\mathbb{Q}}$, it follows from the fact that $\hat{\mathbb{Q}}$ is supported on $\hat{\mathbb{D}}^{(N)}$ and any $f \in \hat{\mathbb{D}}^{(N)}$ is above -2^{-N+3} that

$$\hat{Z} \geq -2^{N+3}, \qquad P^W\text{-a.s..} \tag{4.6.11}$$

Then, by combining this with (4.5.6) and (4.6.8), we can deduce that

$$\begin{aligned}
\|\exists(\hat{Z}) - M\| \leq & \|\exists(\hat{Z}) - \hat{Z} \vee 0\| + \|\hat{Z} \vee 0 - \hat{Z}\| + \|\hat{Z} - M\| \\
\leq & 2^{-N+1} + 2^{N+3} + \|\hat{M} - M\| + \|\hat{A}\| \\
\leq & 2^{-N+4} + N^{-\frac{1}{2}}, \qquad \text{whenever } \max_{1 \leq i \leq d} \sum_{k=1}^{m_0} |X_k^{(i)}| \leq N^{-\frac{1}{2}}.
\end{aligned}$$

It follows that

$$\begin{aligned}
|\hat{G}(M) - \hat{G}(\hat{Z})| = & |G(M \vee 0) - G(\exists(\hat{Z}) \vee 0)| \\
\leq & f_e(2^{-N+4} + N^{-\frac{1}{2}}), \qquad \text{whenever } \max_{1 \leq i \leq d} \sum_{k=1}^{m_0} |X_k^{(i)}| \leq N^{-\frac{1}{2}},
\end{aligned}$$

where we use the fact that $\|\exists(\hat{Z}) \vee 0 - M \vee 0\| \leq \|\exists(\hat{Z}) - M\|$. Hence, since $\hat{G}$ is bounded by κ

$$\big|E^W[\hat{G}(M)] - E^W[\hat{G}(\hat{Z})]\big| \leq f_e(2^{-N+4} + N^{-\frac{1}{2}}) + 2\kappa P^W\Big(\max_{1 \leq i \leq d} \sum_{k=1}^{m_0} |X_k^{(i)}| > N^{-\frac{1}{2}}\Big).$$

Note that

$$\begin{aligned}
X_k = E^W\big[Y_k | \vec{\sigma}_k, \vec{Y}_{k-1}\big] \overset{(d)}{=} & \mathbb{E}_{\hat{\mathbb{Q}}}\big[\hat{\mathbb{S}}_{\hat{\tau}_k} - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}} | \vec{\hat{\tau}}_k, \overrightarrow{\Delta\hat{\mathbb{S}}}_{\hat{\tau}_{k-1}}\big] \\
= & \mathbb{E}_{\hat{\mathbb{Q}}}\big[\hat{\mathbb{S}}_{\hat{\tau}_k} - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}} | \mathcal{F}^{(N)}_{\hat{\tau}_k -}\big]
\end{aligned}$$

where $\Delta\hat{\mathbb{S}}_k = \hat{\mathbb{S}}_{\tilde{\tau}_k} - \hat{\mathbb{S}}_{\tilde{\tau}_k-}$ for $k \leq m_0$ and hence

$$E^W\Big[\sum_{i=1}^{d}\sum_{k=1}^{m_0}|X_k^{(i)}|\Big] = \mathbb{E}_{\hat{\mathbb{Q}}}\Big[\sum_{k=1}^{m_0}\sum_{i=1}^{d}|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{\mathbb{S}}^{(i)}_{\hat{\tau}_k}|\hat{\mathcal{F}}^{(N)}_{\hat{\tau}_k-}] - \hat{\mathbb{S}}^{(i)}_{\hat{\tau}_{k-1}}|\Big].$$

By Markov inequality and definition of $\hat{\mathbb{M}}^{(N)}(2\kappa)$, we have

$$\begin{aligned}P^W\Big(\sum_{i=1}^{d}\sum_{k=1}^{m_0}|X_k^{(i)}| > N^{-\frac{1}{2}}\Big) \leq & \sqrt{N}E^W\Big[\sum_{i=1}^{d}\sum_{k=1}^{m_0}|X_k^{(i)}|\Big] \\ \leq & \sqrt{N}\mathbb{E}_{\hat{\mathbb{Q}}}\Big[\sum_{k=1}^{m}\sum_{i=1}^{d}|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{\mathbb{S}}^{(i)}_{\hat{\tau}_k}|\hat{\mathcal{F}}_{\hat{\tau}_k-}] - \hat{\mathbb{S}}^{(i)}_{\hat{\tau}_{k-1}}|\Big] \leq 2\kappa N^{-\frac{1}{2}}.\end{aligned} \tag{4.6.12}$$

Therefore, we have

$$|E^W[\hat{G}(M)] - E^W[\hat{G}(\hat{Z})]| \leq f_e(2^{-N+4} + N^{-\frac{1}{2}}) + 4\kappa^2 N^{-\frac{1}{2}}. \tag{4.6.13}$$

By (4.6.8), (4.6.11) and (4.6.12)

$$\begin{aligned}&P^W\big(\inf_{0\leq t\leq T}\min_{1\leq i\leq d} M_t^{(i)} > -2^{-N+4} - N^{-\frac{1}{2}} \text{ and } \max_{d\leq i\leq d}\|M^{(i)}\| < \kappa + 1 + 2^{-N+2} + N^{-\frac{1}{2}}\big) \\ \geq & 1 - 2\kappa N^{-\frac{1}{2}}.\end{aligned} \tag{4.6.14}$$

Hence the stopped process M^{θ_0}, with

$$\theta_0 := \inf\big\{t \geq 0 : \min_{1\leq i\leq d} M_t^{(i)} \leq -2^{-N+4} - N^{-\frac{1}{2}} \text{ or } \max_{d\leq i\leq d}\|M^{(i)}\| \geq \kappa+1+2^{-N+2}+N^{-\frac{1}{2}}\big\},$$

satisfies

$$|E^W[\hat{G}(M)] - E^W[\hat{G}(M^{\theta_0})]| \leq 4\kappa^2 N^{-\frac{1}{2}}. \tag{4.6.15}$$

By (4.6.5), (4.6.13) and (4.6.15), it follows that

$$\begin{aligned}&|E^W[\hat{G}(M^{\theta_0})] - \mathbb{E}_{\hat{\mathbb{Q}}}\big[\hat{G}(\hat{\mathbb{S}})\big]| \\ \leq & |E^W[\hat{G}(M^{\theta_0})] - E^W[\hat{G}(M)]| + |E^W[\hat{G}(M)] - E^W[\hat{G}(\hat{Z})]| + |\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}}^{\hat{\tau}_{m_0}})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})]| \\ \leq & 4\kappa^2 N^{-\frac{1}{2}} + 4\kappa^2 N^{-\frac{1}{2}} + f_e(2^{-N+4} + N^{-\frac{1}{2}}) + \kappa 2^{-N+1}.\end{aligned} \tag{4.6.16}$$

Similarly, by (4.6.8) and (4.6.12), we have

$$P^W(\|\hat{Z} - M^\theta\| \geq 2^{-N+2} + N^{-\frac{1}{2}}) \leq 2\kappa N^{-\frac{1}{2}}. \tag{4.6.17}$$

Step 3. The next step is to modify the martingale M^{θ_0} in such way that Γ, the new continuous martingale, is nonnegative.

Write $\epsilon_N = 2^{-N+4} + N^{-\frac{1}{2}}$ and define a nonnegative $\mathcal{F}_T^W$-measurable random variable Λ by $\Lambda = (M_{T\wedge\theta_0} + \epsilon_N)/(1+\epsilon_N)$. Then

$$|\Lambda - M_{T_n\wedge\theta_0}| = \left|\epsilon_N \frac{1 - M_{T\wedge\theta_0}}{1+\epsilon_N}\right| \le \epsilon_N(1 + |M_{T\wedge\theta_0}|).$$

Note that for any $i > d$ $\|\Lambda^{(i)}\| \le \kappa + 1 + 2^{-N+2} + N^{-\frac{1}{2}} + \epsilon_N \le \kappa + 2$ for N large enough. We now construct a continuous martingale from the Λ by taking conditional expectations:

$$\Gamma_t = E^W[\Lambda|\mathcal{F}_t^W], \quad t \in [0,T]$$

and $\Lambda \ge 0$ implies that Γ is nonnegative and $\Gamma_0^{(i)} = 1$ $\forall\, i \le d$. Hence $\mathbb{P}^{(N)} := P^W \circ (\Gamma_t)^{-1} \in \underline{\mathcal{M}}$.

We first notice that

$$E^W[|M^{(i)}_{T\wedge\theta_0}|] = \mathbb{E}^W[M^{(i)}_{T\wedge\theta_0} - 2(M^{(i)}_{T\wedge\theta_0})^-] \le \mathbb{E}^W[M^{(i)}_{T\wedge\theta_0} + 2] = 3 \quad \forall i = 1, \ldots, d.$$

Then by Doob's martingale inequality

$$P^W(\|\Gamma - M^{\theta_0}\| \ge \epsilon_N^{1/2}) \le \epsilon_N^{-1/2} \sum_{i=1}^d E^W[|\Lambda^{(i)} - M^{(i)}_{T\wedge\theta_0}|] \le \epsilon_N^{-1/2} 4d\epsilon_N = 4d\epsilon_N^{1/2}. \tag{4.6.18}$$

This together with (4.6.16) yields

$$\begin{aligned}
&\left|E^W[G(\Gamma)] - \mathbb{E}_{\hat{\mathbb{Q}}}[G(\hat{\mathbb{S}})]\right| \\
\le& E^W\left[|G(\Gamma) - \hat{G}(M^{\theta_0})|\right] + \left|E^W[\hat{G}(M^{\theta_0})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})]\right| \\
\le& E^W\left[|G(\Gamma) - G(M^{\theta_0} \vee 0)| \mathbb{1}_{\left\{\|\Gamma - M^{\theta_0}\| < \epsilon_N^{1/2}\right\}}\right] \\
&\quad + 8\kappa d\epsilon_N^{1/2} + 8\kappa^2 N^{-\frac{1}{2}} + f_e(2^{-N+4} + N^{-\frac{1}{2}}) + \kappa 2^{-N+1} \\
\le& f_e(\epsilon_N^{1/2}) + 8\kappa d\epsilon_N^{1/2} + 9\kappa^2\epsilon_N^{1/2} + f_e(\epsilon_N^{1/2}) \le 2f_e(\epsilon_N^{1/2}) + 17\kappa^2 d\epsilon_N^{1/2}.
\end{aligned}$$

Finally, we can deduce from (4.6.17) and (4.6.18) that

$$P^W(\|\hat{Z} - \Gamma\| \ge 2^{-N+2} + N^{-\frac{1}{2}} + \epsilon_N^{1/2}) \le 2\kappa N^{-\frac{1}{2}} + 4d\epsilon_N^{1/2}. \tag{4.6.19}$$

Notice that when N is sufficiently large such that $2^{-N+2} + N^{-\frac{1}{2}} + \epsilon_N^{1/2} < 2^{-D-1}$, on the event $\left\{\omega \in \Omega^W : \|\hat{Z}(\omega) - \Gamma(\omega)\| < 2^{-N+2} + N^{-\frac{1}{2}} + \epsilon_N^{1/2} \text{ and } \hat{Z}(\omega) \in \hat{\mathbb{D}}^{(N)}\right\}$, we can deduce from (4.5.6) and (4.6.11) that

$$\begin{aligned}
|\daleth(\hat{Z}) \vee 0 - \Gamma| &\le |\daleth(\hat{Z}) \vee 0 - \daleth(\hat{Z})| + |\daleth(\hat{Z}) - Z| + |Z - \Gamma| \\
&< 2^{-N+3} + 2^{-N+1} + 2^{-D+1} \le 2^{-D}.
\end{aligned}$$

and hence by Remark 4.5.2 the inequality $\hat{m}^{(D)}(\hat{Z}) \geq m^{(D-2)}(\Gamma)$ holds on $\{\omega \in \Omega^W : \|\hat{Z}(\omega) - \Gamma(\omega)\| < 2^{-N+2} + N^{-\frac{1}{2}} + \epsilon_N^{1/2}$ and $\hat{Z}(\omega) \in \hat{\mathbb{D}}^{(N)}\}$.

It follows that

$$\begin{aligned}\mathbb{E}_{\hat{\mathbb{Q}}}[\alpha \wedge (\beta\sqrt{\hat{m}^{(D)}(\hat{\mathbb{S}})})] \geq & \mathbb{E}_{\hat{\mathbb{Q}}}[\alpha \wedge (\beta\sqrt{\hat{m}^{(D)}(\hat{\mathbb{S}}^{m_0})})] \\ = & E^W[\alpha \wedge (\beta\sqrt{\hat{m}^{(D)}(\hat{Z})})] \\ \geq & E^W[\alpha \wedge (\beta\sqrt{m^{(D-2)}(\Gamma)})] - \alpha\Big(2\kappa N^{-\frac{1}{2}} + 4d\epsilon_N^{1/2}\Big).\end{aligned}$$

□

4.7 Proofs

4.7.1 Proof of Proposition 4.5.9

Proof. Fix $N \geq 6$. Choose $f_e : \mathbb{R}_+ \to \mathbb{R}_+$ such that $|G| \leq \kappa$, $|G(\omega) - G(\upsilon)| \leq f_e(|\omega - \upsilon|)$ for any $\omega, \upsilon \in \Omega$ and $\lim_{x\to 0} f_e(x) = 0$. Define $G^{(N)} : \Omega \to \mathbb{R}$ as

$$G^{(N)}(S) := \hat{G}(S) - f_e(2^{-N+4}) - \frac{6dN}{2^N}.$$

and

$$\mathbf{V}^{(N)}(G - \alpha \wedge (\beta\sqrt{m^{(D)}})) = \mathbf{V}^{(N)}(G^{(N)} - \alpha \wedge (\beta\sqrt{m^{(D)}})) + f_e(2^{-N+4}) + \frac{6dN}{2^N}.$$

Hence, to show (4.5.12), it suffices to show

$$\mathbf{V}^{(N)}(G^{(N)} - \alpha \wedge (\beta\sqrt{m^{(D)}})) \leq \hat{\mathbb{V}}^{(N)}\Big(\hat{G} - \alpha \wedge (\beta\sqrt{\hat{m}^{(D-2)}})\Big). \tag{4.7.1}$$

The rest of proof is structured to establish (4.7.1). Given a probabilistic semi-static portfolio $\hat{\gamma}$ which superreplicates $\hat{G} - \alpha \wedge (\beta\sqrt{\hat{m}^{(D-2)}}) - x$, we will argue that the lifted progressively measurable trading strategy $\gamma^{(N)}$ superreplicates $G^{(N)} - \alpha \wedge (\beta\sqrt{m^{(D)}}) - x$. To simplify notations, throughout the rest of the proof, we fix $S \in \Omega$ and write $\hat{F} := \hat{F}^{(N)}(S)$.

Super-replication:

We first notice that for any $j < m - 1$

$$\begin{aligned}&|(S_{\tau_{j+1}} - S_{\tau_j}) - (\hat{F}_{\hat{\tau}_j} - \hat{F}_{\hat{\tau}_{j-1}})| \\ \leq & |S_{\tau_{j+1}} - \hat{F}_{\hat{\tau}_j}| + |S_{\tau_j} - \hat{F}_{\hat{\tau}_{j-1}}| \leq \frac{1}{2^{N+j+1}} + \frac{1}{2^{N+j}} = \frac{3}{2^{N+j+1}}.\end{aligned}$$

It follows that for any $k < m$,

$$\begin{aligned}
&\Big|\int_0^{\tau_k} \gamma_u^{(N)}(S)\mathrm{d}S_u - \int_0^{\hat\tau_k} \hat\gamma_u(\hat F)\mathrm{d}\hat F_u| \\
\leq&\Big|\sum_{j=0}^{k-1} \hat\gamma_{\hat\tau_j}(\hat F)\cdot(S_{\tau_{j+1}} - S_{\tau_j}) - \sum_{j=0}^{k-1}\hat\gamma_{\hat\tau_{j+1}}(\hat F)\cdot(\hat F_{\hat\tau_{j+1}} - \hat F_{\hat\tau_j})\Big| \\
\leq&\sum_{j=0}^{k-2}\Big|\hat\gamma_{\hat\tau_{j+1}}(\hat F)\cdot\big((S_{\tau_{j+2}} - S_{\tau_{j+1}}) - (\hat F_{\hat\tau_{j+1}} - \hat F_{\hat\tau_j})\big)\Big| + \frac{2dN}{2^{N-1}} \\
\leq&\sum_{j=0}^{\infty}\frac{Nd}{2^{N+j+2}} + \frac{2dN}{2^{N-1}} \leq \frac{5dN}{2^N}.
\end{aligned} \tag{4.7.2}$$

In addition,

$$\begin{aligned}
&\Big|\int_{\tau_{m-1}}^{T} \gamma_u^{(N)}(S)\mathrm{d}S_u - \int_{\hat\tau_{m-1}}^{T} \hat\gamma_u(\hat F)\mathrm{d}\hat F_u\Big| \\
=&\Big|\hat\gamma_{\hat\tau_{m-1}}(\hat F)\cdot(S_T - S_{\tau_{m-1}}) - \hat\gamma_{\hat\tau_m}(\hat F)\cdot(\hat F_{\hat\tau_m} - \hat F_{\hat\tau_{m-1}})\Big| \leq \frac{Nd}{2^N}.
\end{aligned} \tag{4.7.3}$$

Hence,

$$x + \int_0^T \gamma_u^{(N)}(S)\mathrm{d}S_u \geq x + \int_0^T \hat\gamma_u(\hat F)\mathrm{d}\hat F_u - \frac{5dN}{2^N} - \frac{dN}{2^N} \tag{4.7.4}$$

$$\geq \hat G(\hat F) - \alpha\wedge(\beta\sqrt{\hat m^{(D-2)}(\hat F)}) - \frac{6dN}{2^N} \tag{4.7.5}$$

$$\geq G(S) - \alpha\wedge(\beta\sqrt{m^{(D)}(S)}) - f_e(2^{-N+4}) - \frac{6dN}{2^N} \tag{4.7.6}$$

$$= G^{(N)}(S)$$

where the inequality between (4.7.4) and (4.7.5) follows from the super-replicating property of $\hat\gamma$ and the fact that $\hat{\mathbb{P}}^{(N)}(f) > 0$, $\forall f \in \hat{\mathbb{D}}^{(N)}$, the inequality between (4.7.5) and (4.7.6) is justified by (4.5.7) and (4.5.8).

Admissibility:

Now, for a given $t < T$, let $k < m$ be the largest integer so that $\tau_k(S) \leq t$. It follows from (4.7.2) and (4.7.3) that

$$\int_0^t \gamma_u^{(N)}(S)\mathrm{d}S_u = \int_0^{\tau_k} \gamma_u^{(N)}(S)\mathrm{d}S_u + \int_{\tau_k}^t \gamma_u^{(N)}(S)\mathrm{d}S_u$$

$$\geq \int_0^{\hat\tau_k} \hat\gamma_u(\hat F)\mathrm{d}\hat F_u - \frac{5dN}{2^N} - dN\max_i|S_t^{(i)} - S_{\tau_k}^{(i)}| \tag{4.7.7}$$

$$\geq -M - \frac{6dN}{2^N}. \tag{4.7.8}$$

where the inequality between (4.7.7) and (4.7.8) follows from the admissibility of $\hat{\gamma}$ and the fact that $\hat{\mathbb{P}}^{(N)}(f) > 0, \forall f \in \hat{\mathbb{D}}^{(N)}$. Hence, $\gamma^{(N)}$ is admissible.

□

4.7.2 Construction of σ's and Y in Proposition 4.6.1

We first introduce the notation $\vec{x}_m := (x_1, \ldots, x_m)$ for any integer m and denote $\Pi(E)$ by the set of probability measures on E. In addition, set

$$\mathbb{T} := \mathbb{Q}_+ \cup \{a \geq 0 \,:\, a = T - b \text{ for some } b \in \mathbb{Q}_+\},$$
$$\mathcal{S}_k := \{\frac{1}{2^{N+k}}(a_1, \ldots, a_d) \,:\, a_j \in \mathbb{Z},\, |a_j| \leq 2^k,\, j = 1, \ldots, d\}.$$

For $k = 1, \ldots, m_0$, define the function $\Psi_k : \mathbb{T}^k \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_k \to \Pi(\mathbb{R})$ by

$$\Psi_k(\vec{\alpha}_k; \vec{\beta}_k) := \mathcal{L}_{\hat{\mathbb{Q}}}(\hat{\tau}_{k+1} - \hat{\tau}_k | \hat{\tau}_i - \hat{\tau}_{i-1} = \alpha_i, \hat{\mathbb{S}}_{\hat{\tau}_i} - \hat{\mathbb{S}}_{\hat{\tau}_{i-1}} = \beta_i, i \leq k) \tag{4.7.9}$$

and $\Phi_k : \mathbb{T}^k \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_{k-1} \to \Pi(\mathbb{R}^d)$ by

$$\Phi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}) := \mathcal{L}_{\hat{\mathbb{Q}}}(\hat{\mathbb{S}}_{\hat{\tau}_k} - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}} | \hat{\tau}_j - \hat{\tau}_{j-1} = \alpha_j, \hat{\mathbb{S}}_{\hat{\tau}_i} - \hat{\mathbb{S}}_{\hat{\tau}_{i-1}} = \beta_i, j \leq k, i \leq k-1), \tag{4.7.10}$$

where we set $\mathcal{L}_{\hat{\mathbb{Q}}}(\cdot|\varnothing) \equiv \delta_{\vec{0}}$. Next, let $\mathcal{B}$ be the set of barriers, where a barrier R is a closed subset of $[0, +\infty] \times [-\infty, +\infty]$ such that $(+\infty, x) \in R$ for all $x \in [-\infty, +\infty]$, $(t, \pm\infty) \in R$ for all $t \in [0, +\infty]$, and satisfies that if $(t, x) \in R$ then $(s, x) \in R$ whenever $s > t$. Then, for any $k \leq m_0$, we can find $\Upsilon_k : \mathbb{T}^k \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_k \to [-\infty, \infty]$ and $B_k : \mathbb{T}^k \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_k \to \mathcal{B}$ such that

$$\mathcal{L}_{\mathbb{P}^W}\Big(\sum_{l=0} s_{l+1} \mathbb{1}_{\big\{\Upsilon_k(\vec{\alpha}_k; \vec{\beta}_{k-1}, s_{k,l}) \leq W^{(1)}_{\alpha_k} < \Upsilon_k(\vec{\alpha}_k; \vec{\beta}_{k-1}, s_{k,l+1})\big\}}\Big) = \Phi_k(\vec{\alpha}_k; \vec{\beta}_k), \tag{4.7.11}$$

where $\{s_{k,1}, s_{k,2}, \ldots, s_{k,l}, \ldots\}$ is an enumeration of $\mathcal{S}_k$ and

$$\mathcal{L}_{P^W}\Big(\tau_{B_k(\vec{\alpha}_k; \vec{\beta}_k)}(W^{(1)})\Big) = \Psi_k(\vec{\alpha}_k; \vec{\beta}_k). \tag{4.7.12}$$

where $\tau_{B_k(\vec{\alpha}_k; \vec{\beta}_k)}(W^{(1)})$ is the first hitting time of $B_k(\vec{\alpha}_k; \vec{\beta}_k)$ by $W^{(1)}$. Here the existence of $B_k(\vec{\alpha}_k; \vec{\beta}_k)$ can be justified by Theorem 2.3 in Ekström and Janson [44].

Now we set $\sigma_0 \equiv 0$ and define the random variables $\sigma_1, \ldots, \sigma_{m_0}, Y_1, \ldots, Y_{m_0}$ recursively as follow

$$\begin{aligned}
&\Delta_i = \tau_{B_{i-1}(\Delta_{i-1}; Y_{i-1})}(\{W^{(1)}_{t+\sigma_{i-1}} - W^{(1)}_{\sigma_{i-1}}\}_{t\geq 0}) \\
&\sigma_i = \sigma_{i-1} + \Delta_i \\
&Y_i = \mathbb{1}_{\{\sigma_i < T\}} \sum_{j=1}^{\infty} s_{i,j} \mathbb{1}_{\big\{\Upsilon_i(\vec{\Delta}_i; \vec{Y}_{i-1}; s_{i,j-1}) \leq W^{(i+1)}_{\sigma_i} - W^{(i+1)}_{\sigma_{i-1}} < \Upsilon_i(\vec{\Delta}_i; \vec{Y}_{i-1}; s_{i,j})\big\}},
\end{aligned} \tag{4.7.13}$$

Note that σ_i are stopping times with respect to the Brownian filtration. Fix $k \leq m_0$ and $(\vec{\alpha}_k; \vec{\beta}_{k-1}) \in \mathbb{T}^k \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_{k-1}$. By the strong Markov property and the independence of the Brownian motion increments, it follows from (4.7.12) that

$$\mathcal{L}_{P^W}(\Delta_k | (\vec{\Delta}_{k-1}, \vec{Y}_{k-1}) = (\vec{\alpha}_{k-1}, \vec{\beta}_{k-1})) = \Psi_{k-1}(\vec{\alpha}_{k-1}; \vec{\beta}_{k-1}). \tag{4.7.14}$$

Similarly, from (4.7.11) and (4.7.14), we have

$$\mathcal{L}_{P^W}\Big(Y_k | \vec{\Delta}_k = \vec{\sigma}_k, \vec{Y}_{k-1} = \vec{\beta}_{k-1}\Big) = \Phi(\vec{\alpha}_k; \vec{\beta}_{k-1}). \tag{4.7.15}$$

Therefore, using (4.7.9)-(4.7.10) and (4.7.14)-(4.7.15), we conclude that

$$\mathcal{L}_{P^W}\big((\vec{\sigma}_{m_0}; \vec{Y}_{m_0})\big) = \mathcal{L}_{\hat{\mathbb{Q}}}\big((\vec{\hat{\tau}}_{m_0}; \vec{\Delta\hat{\mathbb{S}}}_{m_0})\big),$$

where $\Delta\hat{\mathbb{S}}_k = \hat{\mathbb{S}}_{\hat{\tau}_k} - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}}$, $k \leq m_0$.

Chapter 5

Robust pricing–hedging duality with options in continuous time

5.1 Introduction

In the previous chapter, we studied the continuous-time robust pricing and hedging problems in a market where only dynamic trading in stocks is allowed. In this chapter, we follow the same robust framework as before but add both static and dynamic hedging in options into the framework. The typical trading strategy therefore shall consist not only of dynamic trading in stocks, but also dynamic trading in some options with static hedging in others. Our main interest here is to establish general pricing–hedging duality results. In particular, we shall look at two important cases: The first one is when there are only finitely many options traded; the second one is when all European call options for multiple maturities are traded. The later one is closely related to the so called martingale optimal transport problem and the setup often assumed in many papers of the model-independent finance literature, including Dolinsky and Soner [39] which has inspired our studies of the robust pricing and hedging problems.

Our second main contribution is to propose a *robust approach* which subsumes the model-independent setting but allows to include assumptions and move gradually towards the model-specific setting. In this sense, we strive to provide a setup which connects and interpolates between the two ends of spectrum considered by Merton [77]. In contrast, all of the above works on model-independent approach stay within Merton [77]'s "universally accepted" setting and analyse the implications of incorporating the ability to trade some options at given market prices for the outputs:

prices and hedging strategies of other contingent claims. We amend this setup and allow one to express modelling beliefs. These are articulated in a pathwise manner. More precisely, we allow the modeller to deem certain paths *impossible* and exclude them from then analysis: the superhedging property is only required to hold on the remaining set of paths $\mathfrak{P}$. This is reflected in the form of the pricing–hedging duality we obtain.

Our approach to establishing the pricing–hedging duality involves both discretisation, as in Dolinsky and Soner [39], as well as a variational approach as in Galichon et al. [49]. We first prove an "unconstrained" duality result: Theorem 5.3.2 states that for any derivative with bounded and uniformly continuous payoff function G, the minimal initial set-up cost of a portfolio consisting of cash and dynamic trading in the risky assets (some of which could be options themselves) which superhedges the payoff G for every nonnegative continuous path, is equal to the supremum of the expected value of G over all nonnegative continuous martingale measures[1]. This result is shown through an elaborate discretisation procedure building on ideas in [39; 41]. Subsequently, we develop a variational formulation which allows us to add statically traded options, or specification of prediction set $\mathfrak{P}$, via Lagrange multipliers. In some cases this leads to "constrained" duality result, similar to ones obtained in works cited above, with superhedging portfolios allowed to trade statically the market options and martingale measures required to reprice these options. In particular Theorems 5.3.10 and 5.3.14 extend the duality obtained respectively in Davis et al. [32] and [39]. However in general we obtain an *asymptotic* duality result with the dual and primal problems defined through a limiting procedure. The primal value is the limit of superhedging prices on ϵ-neighbourhoods of $\mathfrak{P}$ and the dual value is the limit of supremum of expectation of the payoff over ϵ-(miss)calibrated models, see Definitions 5.2.1 and 5.3.16.

The chapter is organised as follows. Section 5.2 introduces our robust framework for pricing and hedging and defines the primal (pricing) and dual (hedging) problems. Section 5.3 contains all the main results. First, in Section 5.3.1, we present the unconstrained pricing–hedging duality in Theorem 5.3.2 and derive constrained (asymptotic) duality results under suitable compactness assumptions. This allows us in particular to treat the case of finitely many traded options. Then in Sections 5.3.2–5.3.3 we apply the previous results to the martingale optimal transport case.

[1]Note that here and throughout, we assume that all assets are discounted or, more generally, are expressed in terms of some numeraire.

All the results except Theorem 5.3.2 are proved in Section 5.4. Theorem 5.3.2 is proved in Section 5.5.

5.2 Robust Modelling Framework

5.2.1 Traded assets

We consider a financial market with $d+1$ assets: a numeraire (e.g. the money market account) and d underlying assets $S^{(i)}, \ldots, S^{(d)}$, which may be traded at any time $t \leq T_n$. All prices are denominated in the units of the numeraire. In particular, the numeraire's price is thus normalised and equal to one. We assume that the price path $S_t^{(i)}$ of each risky asset is continuous. The assets start at $S_0 = (1, \ldots, 1)$ and are assumed to be nonnegative. We work on the canonical space $\mathcal{C}([0, T_n], \mathbb{R}_+^d)$, the set of all $\mathbb{R}_+^d$-valued continuous functions on $[0, T_n]$.

We pursue here a robust approach and do not postulate any probability measure which would specify the dynamics for S. Instead we assume that there is a set $\mathcal{X}$ of market traded options with prices known at time zero, $\mathcal{P}(X)$, $X \in \mathcal{X}$. In all generality, an option $X \in \mathcal{X}$ is just a mapping $X : \mathcal{C}([0, T_n], \mathbb{R}_+^d) \to \mathbb{R}$, measurable with respect to the σ-field generated by coordinate process. However most often we will consider European options, i.e., $X(S) = f(S_{T_i})$ for some f and for maturities $0 < T_1 < \ldots < T_n = T$. The trading is frictionless so prices are linear and options in $\mathcal{X}$ may be bought or sold at time zero at their known prices.

Further, we allow some of the options to be traded continuously. We do this by augmenting the set of risky assets so that there are $d + K$ assets which may be traded at any time $t \leq T_n$: d underlying assets S and K options $X_1^{(c)}(S), \ldots, X_K^{(c)}(S)$. We assume that $X_1^{(c)}, \ldots, X_K^{(c)}$ are European options with maturity T_n and have continuous price paths. In addition, they have nonnegative payoffs and their prices today $\mathcal{P}(X_i^{(c)})$'s are strictly positive. Hence, by normalisation, we can assume without loss of generality the price of each option starts at 1 and never goes below 0. We now consider a natural extension of the path space

$$\Omega = \{f \in \mathcal{C}([0, T_n], \mathbb{R}_+^{d+K}) \, : \, f_0 = (1, \ldots, 1)\}.$$

The coordinate process on Ω is denoted $\mathbb{S} = (\mathbb{S}_t)_{0 \leq t \leq T_n}$ i.e.,

$$\mathbb{S} = (\mathbb{S}^{(1)}, \ldots, \mathbb{S}^{(d+K)}) \, : \, [0, T] \to \mathbb{R}_+^{d+K},$$

and $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T_n}$ is its natural filtration. However, not every ω in Ω is a good candidate for price path of these assets. It can be seen from the fact that the prices of option $X_i^{(c)}$ and S at time T_n should always respect the payoff function $X_i^{(c)}$. Therefore, for the purpose of pricing and hedging duality, we only need to consider the set of possible price paths of these $d+K$ assets, denoted $\mathcal{I}$, i.e.,

$$\mathcal{I} = \{\omega \in \Omega : \omega_{T_n}^{(d+i)} = X_i^{(c)}(\omega_{T_n}^{(1)}, \ldots, \omega_{T_n}^{(d)})/\mathcal{P}(X_i^{(c)})\ \forall i \le K\}.$$

$\mathcal{I}$, called the information space, encodes not only the prices of these underlying assets and options at time zero, but also future payoff constraints.

5.2.2 Trading strategies

A trading strategy consists of two parts. The first part is static hedging X, which is a linear combination of market traded options. In contrast, the other part, known as the dynamic trading, features a potentially continuous trading in the underlying asset and a few selected European options. Heuristically, the capital gain from this trading activity takes the integral form of $\int \gamma_u(S)\mathrm{d}S_u$. To define this integral properly, we follow Section 4.2.2 and consider $\gamma : \Omega \to \mathcal{D}([0,T_n], \mathbb{R}_+^{d+K})$ which is of finite variation and progressively measurable with respect to a filtration which, in our context, is the natural filtration generated by the canonical process. Moreover, γ is admissible if $\gamma : \Omega \to \mathcal{D}([0,T_n], \mathbb{R}_+^{d+K})$ is progressively measurable and of finite variation, satisfying

$$\int_0^t \gamma_u(S)\mathrm{d}S_u \ge -M, \quad \forall\, S \in \mathcal{I},\ t \in [0,T_n],\ \text{for some } M > 0. \tag{5.2.1}$$

Let $\mathcal{A}$ be the set of such integrands. The set of simple integrands, i.e., $\gamma \in \mathcal{A}$ such that $\gamma(\omega)$ is a simple function $\forall \omega \in \Omega$, is denoted $\mathcal{A}^{sp}$. For precise definitions and more details, we refer the reader to Section 4.2.2.

An admissible (semi-static) trading strategy, in our context, is a pair (X, γ) where $X = a_0 + \sum_{i=1}^m a_i X_i$, for some m, $X_i \in \mathcal{X}$ and $a_0, a_i \in \mathbb{R}$, $i = 1, \ldots, m$ and $\gamma \in \mathcal{A}$. The cost of following such a trading strategy is equal to the cost of setting up its static part, i.e., of buying the options at time zero, and is equal to

$$\mathcal{P}(X) := a_0 + \sum_{i=1}^m a_i \mathcal{P}(X_i).$$

We denote the class of admissible (semi-static) trading strategies by $\mathcal{A}_{\mathcal{X}}$ and $\mathcal{A}_{\mathcal{X}}^{sp}$ for $\gamma \in \mathcal{A}$ or $\mathcal{A}^{sp}$ respectively.

5.2.3 Beliefs

As argued in the Introduction, we allow our agents to express modelling beliefs. These are encoded as restrictions of the pathspace and may come from time series analysis of the past data, or idiosyncratic views about market in the future. Put differently, we are allowed to rule out paths which we deem *impossible*. The paths which remain are referred to as *prediction set* or *beliefs*. Note that such beliefs may also encode one agent's superior information about the market.

We will consider pathwise arguments and require that they work provided the price path S falls into the prediction set $\mathfrak{P} \subseteq \mathcal{I}$. Any path falling out of $\mathfrak{P}$ will be ignored in our considerations. This binary way of specifying beliefs is motivated by the fact that in the end we only see one path and hence we are interested in arguments which work pathwise. Nevertheless, the approach is very parsimonious and as $\mathfrak{P}$ changes from all paths in $\mathcal{I}$ to a support of a given model we essentially interpolate between model-independent and model-specific setups. It also allows to incorporate the information from time-series of data coherently into the option pricing setup, as no probability measure is fixed and hence no distinction between real world and risk neutral measures is made. The idea of such a prediction set first appeared in Mykland [78]; also see Nadtochiy and Obłój [79] and [24] for an extended discussion.

As the agent rejects more and more paths, i.e., takes $\mathfrak{P}$ smaller and smaller, the framework's outputs – the robust price bounds, should get tighter and tighter. This can be seen as a way to quantify the impact of making assumptions or acquiring additional insights or information.

5.2.4 Superreplication

Our prime interest is in understanding robust pricing and hedging of a derivative with payoff $G : \Omega \to \mathbb{R}$ whose price is not quoted in the market. Our main results will consider bounded payoffs G and, since the setup is frictionless and there are no trading restrictions, without any loss of generality we may consider only the superhedging price. The subhedging follows by considering $-G$.

Definition 5.2.1.

1. A portfolio $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$ is said to *superreplicate* G on $\mathfrak{P}$ if

$$X(S) + \int_0^{T_n} \gamma_u(S) \mathrm{d}S_u \geq G(S), \quad \forall S \in \mathfrak{P}. \tag{5.2.2}$$

2. The (minimal) super-replicating cost of G on $\mathfrak{P}$ is defined as

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \inf \Big\{ \mathcal{P}(X) \,:\, \exists (X,\gamma) \in \mathcal{A}_{\mathcal{X}} \text{ s.t. } (X,\gamma) \text{ superreplicates } G \text{ on } \mathfrak{P} \Big\}. \tag{5.2.3}$$

3. The approximate super-replicating cost of G on $\mathfrak{P}$ is defined as

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \inf \Big\{ \mathcal{P}(X) \,: \exists (X,\gamma) \in \mathcal{A}_{\mathcal{X}} \text{ s.t. } (X,\gamma) \text{ superreplicates } G \text{ on } \mathfrak{P}^{\epsilon} \text{ for some } \epsilon > 0 \Big\}, \tag{5.2.4}$$

where $\mathfrak{P}^{\epsilon} = \{\omega \in \mathcal{I} \,:\, \inf_{\upsilon \in \mathfrak{P}} \|\omega - \upsilon\| \leq \epsilon\}$.

4. Finally, we let $V^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, respectively $\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, denote the super-replicating cost of G in (5.2.3), respectively in (5.2.4), but with $(X,\gamma) \in \mathcal{A}^{sp}_{\mathcal{X}}$.

5.2.5 Market models

Our aim is to relate the robust (super)hedging cost, as introduced above, to the classical pricing-by-expectation arguments. To this end we look at all classical models which reprice market traded options.

Definition 5.2.2. We denote by $\mathcal{M}$ the set of probability measures $\mathbb{P}$ on $(\Omega, \mathcal{F}_{T_n}, \mathbb{F})$ such that $\mathbb{S}$ is a $\mathbb{P}$–martingale and let $\mathcal{M}_{\mathcal{I}}$ be the set of probability measures $\mathbb{P} \in \mathcal{M}$ such that $\mathbb{P}(\mathcal{I}) = 1$.
In this Chapter, a probability measure $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$ is called a $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$–market model, or simply a calibrated model, if $\mathbb{P}(\mathfrak{P}) = 1$ and $\mathbb{E}_{\mathbb{P}}[X] = \mathcal{P}(X)$ for all $X \in \mathcal{X}$. The set of such measures is denoted $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$.
More generally, a probability measure $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$ is called an $\eta - (\mathcal{X}, \mathcal{P}, \mathfrak{P})$–market model if $\mathbb{P}(\mathfrak{P}^{\eta}) > 1 - \eta$ and $|\mathbb{E}_{\mathbb{P}}[X] - \mathcal{P}(X)| < \eta$ for all $X \in \mathcal{X}$. The set of such measures is denoted $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$.

Whenever we have $\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ it provides us with a feasible no-arbitrage price $\mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$ for a derivative with payoff G. The robust price for G is given as

$$P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})],$$

where throughout the expectation is defined with the convention that $\infty - \infty = -\infty$. In the cases of particular interest, $(\mathcal{X}, \mathcal{P})$ will determine uniquely the marginal distributions of $\mathbb{S}$ at given maturities and $P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ is then the value of the corresponding martingale optimal transport problem. We will often use this terminology, even in the case of arbitrary $\mathcal{X}$.

In practice, the market prices $\mathcal{P}$ are an idealised concept and may be obtained from averaging of bid-ask spread or otherwise. It might not be natural to require a perfect calibration and the concept of η–market model allows for a controlled degree of mis-calibration. This leads to the approximate value given as

$$\widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \lim_{\eta\searrow 0} \sup_{\mathbb{P}\in\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})].$$

As we will show below, both the approximate superhedging cost and the approximate robust pricing cost, while being motivated by practical considerations, appear very naturally when considering abstract pricing–hedging duality.

In some instances, for technical reasons, it will be convenient to consider only $\mathbb{P}$ arising within a Brownian setup. We denote the collection of $\mathbb{P}\in\mathcal{M}$ such that $\mathbb{P}=\mathbb{P}^W\circ M^{-1}$ for some continuous martingale M defined on some probability space satisfying the usual assumptions $(\Omega^W, \mathcal{F}^W_{T_n}, \mathbb{F}^W, P^W)$ with a finite dimensional Brownian motion $\{W_t\}_{t\geq 0}$ which generates the filtration $\mathbb{F}^W$. We write $\underline{\mathcal{M}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ to denote $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\underline{\mathcal{M}}$, $\underline{\mathcal{M}}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ for $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\underline{\mathcal{M}}$ and $\underline{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}}\mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$, $\widetilde{\underline{P}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) := \lim_{\eta\searrow 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}}\mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$.

5.3 Main results

In Chapter 4, we established a robust pricing–hedging duality in a simple setting without beliefs and options. Our prime interest here, as discussed in the Introduction, is in extending the robust pricing–hedging duality to include beliefs and options. More precisely, in our context, given a non-traded derivative with payoff G, we have two candidate robust prices for it: The first one, $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, is obtained through pricing-by-hedging arguments; The second one, $P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, is obtained by pricing-via-expectation arguments. And we want to show that they agree with each other.

Within the present pathwise robust approach, the pricing–hedging duality was obtained for specific payoffs G in literature linking robust approach with the Skorokhod embedding problem, see Hobson [60] or Obłój [83] for discussion. Subsequently, an abstract result was established in Dolinsky and Soner [39], when $\Omega=\mathcal{I}=\mathfrak{P}$, $n=d=1$, $K=0$ and $\mathcal{X}$ is the set of all call and put options with $\mathcal{P}(X)=\int_0^\infty X(x)\mu(\mathrm{d}x)$ for all $X \in \mathcal{X}$, where μ is a probability measure on $\mathbb{R}_+$ with mean equal to 1:

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) \quad \text{for a 'strongly continuous' class of bounded } G\ .$$

The result was extended to unbounded claims by broadening the class of admissible strategies and imposing a technical assumption on μ. Below we extend this duality to a much more general setting of abstract $\mathcal{X}$, possibly involving options with multiple maturities, a multidimensional setting and with an arbitrary prediction set $\mathfrak{P}$.

Note that, for any Borel $G : \Omega \to \mathbb{R}$, the inequality

$$V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \geq P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \tag{5.3.1}$$

is true as long as there is at least one $\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ and at least one $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$ which superreplicates G on $\mathfrak{P}$. Indeed, since γ is progressively measurable, the integral $\int_0^{\cdot} \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u$, defined pathwise via integration by parts, agrees a.s. with the stochastic integral under $\mathbb{P}$. Then, by (5.2.1), the stochastic integral is a $\mathbb{P}$ super-martingale and hence $\mathbb{E}_{\mathbb{P}}\Big[\int_0^{T_n} \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u\Big] \leq 0$. This in turn implies that

$$\mathbb{E}_{\mathbb{P}}\Big[G(\mathbb{S})\Big] \leq \mathcal{P}(X).$$

The result follows since (X, γ) and $\mathbb{P}$ were arbitrary.

5.3.1 General duality

We first consider the case without constraints: $\mathcal{X} = \varnothing$ and $\mathfrak{P} = \mathcal{I}$. As this context will be our reference point we introduce notation to denote the super-hedging cost and the robust price. We let

$$\mathbf{V}_{\mathcal{I}}(G) := \inf\Big\{x\ :\ \exists \gamma \in \mathcal{A} \text{ s.t. } \gamma \text{ superreplicates } G - x \text{ on } \mathcal{I}\Big\}, \quad \mathbf{P}_{\mathcal{I}}(G) := \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]. \tag{5.3.2}$$

We also write $\underline{\mathbf{P}}_{\mathcal{I}}(G)$ for $\sup_{\mathbb{P} \in \underline{\mathcal{M}}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$ and $\mathbf{V}^{sp}_{\mathcal{I}}(G)$ for super-replicating cost of G using $\gamma \in \mathcal{A}^{sp}$.

Assumption 5.3.1. Either $K = 0$ or $X^{(c)}_1, \ldots, X^{(c)}_K$ are bounded and uniformly continuous with market prices $\mathcal{P}(X^{(c)}_1), \ldots, \mathcal{P}(X^{(c)}_K)$ satisfying that there exists an $\epsilon > 0$ such that for any $(p_k)_{1 \leq k \leq K}$ with $|\mathcal{P}(X^{(c)}_k) - p_k| \leq \epsilon$ for all $k \leq K$, $\mathcal{M}_{\tilde{\mathcal{I}}} \neq \varnothing$, where

$$\tilde{\mathcal{I}} := \{\omega \in \Omega\ :\ \omega^{(d+i)}_{T_n} = X^{(c)}_i(\omega^{(1)}_{T_n}, \ldots, \omega^{(d)}_{T_n})/p_i\ \forall i \leq K\}.$$

Theorem 5.3.2. Under Assumption 5.3.1, for any bounded and uniformly continuous $G : \Omega \to \mathbb{R}$ we have

$$\mathbf{V}^{sp}_{\mathcal{I}}(G) = \mathbf{V}_{\mathcal{I}}(G) = \mathbf{P}_{\mathcal{I}}(G) = \underline{\mathbf{P}}_{\mathcal{I}}(G).$$

The result is an extension of Theorem 4.3.1. It can be proven either directly through a discretisation of both the primal and the dual problems, similar to the that in the proof of Theorem 4.3.1, or using Theorem 4.3.1 and a variation approach, which is what we will present below and in Section 5.5. First we see that the inequality

$$\mathbf{V}_{\mathcal{I}}(G) \geq \sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})]$$

is a special case of (5.3.1). Section 5.5 is devoted to the proof of the much harder reverse inequality

$$\mathbf{V}_{\mathcal{I}}(G) \leq \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})], \tag{5.3.3}$$

which then implies Theorem 5.3.2.

We let $\mathrm{Lin}(\mathcal{X})$ denote the set of finite linear combinations of elements of $\mathcal{X}$ and

$$\mathrm{Lin}_N(\mathcal{X}) = \Big\{a_0 + \sum_{i=1}^m a_i X_i \,:\, m \in \mathbb{N},\, X_i \in \mathcal{X},\, \sum_{i=0}^m |a_i| \leq N\Big\}.$$

Then, similarly to e.g. Proposition 5.2 in Henry-Labordère et al. [56], a calculus of variations characterisation of $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ is a corollary of Theorem 5.3.2. From that we are able to deduce pricing–hedging duality between the approximate values.

Corollary 5.3.3. Under Assumption 5.3.1, let $\mathfrak{P}$ be a measurable subset of $\mathcal{I}$ and $\mathcal{X}$ such that all $X \in \mathcal{X}$ are uniformly continuous and bounded. Then for any uniformly continuous and bounded $G : \Omega \to \mathbb{R}$ we have:

$$\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \inf_{X\in\mathrm{Lin}_N(\mathcal{X}),\, N\geq 0} \Big\{\mathbf{P}_{\mathcal{I}}(G - X - N\lambda_{\mathfrak{P}}) + \mathcal{P}(X)\Big\}, \tag{5.3.4}$$

where $\lambda_{\mathfrak{P}}(\omega) := \inf_{\upsilon\in\mathfrak{P}} \|\omega - \upsilon\| \wedge 1$.

Remark 5.3.4. As a by-product of the proof of Corollary 5.3.3, we show that for any bounded G,

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \inf_{N\geq 0} \widetilde{V}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G - N\lambda_{\mathfrak{P}}) \quad \text{and} \quad \widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \inf_{N\geq 0} \widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G - N\lambda_{\mathfrak{P}}). \tag{5.3.5}$$

Assumption 5.3.5. $\mathrm{Lin}_1(\mathcal{X})$ is a compact subset of $\mathcal{C}(\Omega, \mathbb{R})$ and every $X \in \mathcal{X}$ is bounded and uniformly continuous.

Theorem 5.3.6. Given $\mathcal{I}$, $\mathfrak{P}$ and $\mathcal{X}$ satisfy conditions in Corollary 5.3.3, if $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ for any $\eta > 0$, then for any uniformly continuous and bounded $G : \Omega \to \mathbb{R}$ we have

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \geq \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G), \tag{5.3.6}$$

and if $\mathcal{X}$ satisfies Assumption 5.3.5, then $\underline{\mathcal{M}}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ for any $\eta > 0$ and equality holds:

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \underline{\widetilde{P}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G). \tag{5.3.7}$$

Example 5.3.7 (Finite $\mathcal{X}$)**.** Consider $\mathcal{X} = \{X_1, \ldots, X_m\}$, where X_i's are bounded and uniformly continuous. In this case, $\mathrm{Lin}_1(\mathcal{X})$ is a convex and compact subset of $\mathcal{C}(\Omega, \mathbb{R})$. Therefore, if $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ for any $\eta > 0$, we can apply Theorem 5.3.6 to conclude $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$.

We end this section with consideration if the approximate superhedging and robust prices, $\widetilde{V}, \widetilde{P}$, are close to the precise values V, P. First, we focus on the case of finitely many traded put options and no beliefs. We consider

$$\mathcal{X} = \{(K^{(i)}_{k,j} - \mathbb{S}^{(i)}_{T_j})^+,\ 1 \leq i \leq d,\ 1 \leq j \leq n,\ 1 \leq k \leq m(i,j)\}, \tag{5.3.8}$$

where $0 < K^{(i)}_{k,j} < K^{(i)}_{k',j}$ for any $k < k'$ and $m(i,j) \in \mathbb{N}$. To simplify the notation, we write

$$\mathcal{P}((K^{(i)}_{k,j} - \mathbb{S}^{(i)}_{T_j})^+) = p_{k,i,j} \quad \forall i, j, k.$$

Assumption 5.3.8. Market put prices are such that there exists an $\epsilon > 0$ such that for any $(\tilde{p}_{k,i,j})_{i,j,k}$ with $|\tilde{p}_{k,i,j} - p_{k,i,j}| \leq \epsilon$ for all i, j, k, there exists a $\tilde{\mathbb{P}} \in \mathcal{M}_{\mathcal{I}}$ such that

$$\tilde{p}_{k,i,j} = \mathbb{E}_{\tilde{\mathbb{P}}}[(K^{(i)}_{k,j} - \mathbb{S}^{(i)}_{T_j})^+] \quad \forall i, j, k.$$

Remark 5.3.9. Assumption 5.3.8 can be rephrased as saying that the market prices $(\mathcal{X}, \mathcal{P})$ are in the interior of the no-arbitrage region.

Theorem 5.3.10. Let $\mathcal{X}$ be given in (5.3.8), prices $\mathcal{P}$ be such that Assumption 5.3.8 holds and $\mathcal{I}$ satisfy Assumption 5.3.1. Then for any uniformly continuous and bounded $G : \Omega \to \mathbb{R}$, we have

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G).$$

The above result establishes a general robust pricing–hedging duality when finitely many put options are traded. It extends in many ways the duality obtained in Davis et al. [32] for the case of $d = n = 1$ and $K = 0$. Note that in general $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G)$ so it follows from Example 5.3.7 that in Theorem 5.3.10 we also have $\widetilde{P}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G)$. These equalities may still hold, but may also fail dramatically, when non-trivial beliefs are specified. We present two examples to highlight this.

Example 5.3.11. In this example we consider $\mathfrak{P}$ corresponding to Black-Scholes model. For simplicity, consider the case without any traded options $K = 0, \mathcal{X} = \varnothing, d = 1$ and let[2]

$$\mathfrak{P} = \{\omega \in \Omega \,:\, \omega \text{ admits quadratic variation and } d\langle\omega\rangle_t = \sigma^2\omega_t^2 dt, 0 \leq t \leq T\}.$$

Then $\mathcal{M}_{\mathfrak{P}} = \{\mathbb{P}_\sigma\}$, where $\mathbb{S}$ is a geometric Brownian motion with constant volatility σ under $\mathbb{P}_\sigma$. The duality in Theorem 5.3.6 then gives that for any bounded and uniformly continuous G

$$\begin{aligned}\widetilde{V}_{\mathfrak{P}}(G) &= \inf\{x \,:\, \exists\gamma \in \mathcal{A} \text{ s.t. } \gamma \text{ superreplicates } G - x \text{ on } \mathfrak{P}^\epsilon \text{ for some } \epsilon > 0\}\\ &= \lim_{\eta\searrow 0} \sup_{\mathbb{P}\in\mathcal{M}^\eta_{\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G].\end{aligned}$$

However in this case, $\mathfrak{P}^\epsilon = \Omega$ and $\mathcal{M}^\epsilon_{\mathfrak{P}} = \mathcal{M}$ for any $\epsilon > 0$. The above then boils down to the duality in Theorem 5.3.2 and we have

$$\widetilde{V}_{\mathfrak{P}}(G) = V_{\mathcal{I}}(G) = \sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}}[G] \geq \mathbb{E}_{\mathbb{P}_\sigma}[G] = P_{\mathfrak{P}}(G), \tag{5.3.9}$$

where for most G the inequality is strict.

Example 5.3.12. Consider again the case with no traded options, $K = 0$ and $\mathcal{X} = \varnothing$ and let

$$\mathfrak{P} = \{\omega \in \Omega \,:\, \|\omega\| \leq b\} \quad \text{ for some } b \geq 1.$$

Given a bounded and uniformly continuous payoff function G, consider the duality in Theorem 5.3.6. For each $N \in \mathbb{N}$, we pick $\mathbb{P}^{(N)} \in \mathcal{M}^{1/N}_{\mathfrak{P}}$ such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G] \geq \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G] - 1/N.$$

[2]See also Step 4 in the proof of Theorem 5.3.21 in Section 5.4.

Let τ be the first hitting time of $b + 1/N$ by $\mathbb{S}$ and define $\tilde{\mathbb{S}}^{(N)}$ by $\tilde{\mathbb{S}}^{(N)}_t = \mathbb{S}_0 + \frac{b}{b+1/N}(\mathbb{S}_{t\wedge\tau} - \mathbb{S}_0)$. By definition, $\mathbb{P}^{(N)} \circ (\tilde{\mathbb{S}}^{(N)})^{-1} \in \mathcal{M}_{\mathfrak{P}}$. Also note that $\mathbb{P}^{(N)}(\tau < T) \leq 1/N$. Hence by uniform continuity of G it is straightforward to see that

$$|\mathbb{E}_{\mathbb{P}^{(N)}}[G(\tilde{\mathbb{S}}^{(N)})] - \mathbb{E}_{\mathbb{P}^{(N)}}[G(S)]| \to 0 \quad \text{as } N \to \infty,$$

which leads to $\widetilde{P}_{\mathfrak{P}}(G) = P_{\mathfrak{P}}(G)$. As $V_{\mathfrak{P}}(G) \leq \widetilde{V}_{\mathfrak{P}}(G) = \widetilde{P}_{\mathfrak{P}}(G)$, we then conclude that

$$\widetilde{V}_{\mathfrak{P}}(G) = \widetilde{P}_{\mathfrak{P}}(G) = P_{\mathfrak{P}}(G) = V_{\mathfrak{P}}(G).$$

5.3.2 Martingale optimal transport duality for bounded claims

We focus now on the cases when $(\mathcal{X}, \mathcal{P})$ determines uniquely certain distributional properties of $\mathbb{S}$ under any $\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$. We start with the case when $\mathcal{X}$ is large enough so that the market prices $\mathcal{P}$ pin down the (joint) distribution of $(\mathbb{S}^{(1)}_T, \ldots, \mathbb{S}^{(d)}_T)$ under any calibrated model. Later we consider the case when only marginal distributions of $\mathbb{S}^{(i)}_T$ for $i \leq d$ are fixed. In the former case we limit ourselves to one maturity and $\mathfrak{P} = \mathcal{I}$ which simplifies the exposition. It is possible to extend these results along the lines of the latter case, when we consider prices at multiple maturities and a non-trivial prediction set, however this would increase the complexity of the proof significantly.

Let $n = 1$ and $T = T_n$. We assume market prices for a rich family of basket options are available. We consider

$$\mathcal{X} \text{ s.t. } \mathrm{Lin}(\mathcal{X}) \text{ is a dense subset of } \{f(\mathbb{S}^{(1)}_T, \ldots, \mathbb{S}^{(d)}_T) \,|\, f : \mathbb{R}^d_+ \to \mathbb{R} \text{ bounded, Lip. cont.}\}. \tag{5.3.10}$$

In particular, $\mathcal{X}$ is large enough to determine uniquely the distribution of $\mathbb{S}_T$ under any calibrated model, i.e., there exists a unique probability distribution π on $\mathbb{R}^d_+$ such that

$$\mathbb{E}_{\mathbb{P}}[X] = \mathcal{P}(X) = \int_{\mathbb{R}^d_+} X(s_1, \ldots, s_d)\pi(\mathrm{d}s_1, \ldots, \mathrm{d}s_d), \quad \forall X \in \mathcal{X}, \mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}. \tag{5.3.11}$$

As an example, we could take $\mathcal{X}$ equal to the RHS in (5.3.10). A martingale measure $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$ is a calibrated model if and only if the distribution of $\mathbb{S}^{(1)}_T, \ldots, \mathbb{S}^{(d)}_T$ under $\mathbb{P}$ is π. Accordingly we write $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}} = \mathcal{M}_{\pi,\mathcal{I}}$ with $\underline{\mathcal{M}}_{\pi,\mathcal{I}}$, $P_{\pi,\mathcal{I}}$ etc. defined analogously. Note that in a Brownian setting, we can always define a continuous martingale M valued in $\mathbb{R}^{d+K}_+$ with $M_0 = 1$, $(M^{(1)}_T, \ldots, M^{(d)}_T) \sim \pi$ and $M^{(d+i)}_T = X^{(c)}_i(M^{(1)}_T, \ldots, M^{(d)}_T)$

for every $i \leq K$ simply by taking conditional expectations of a suitably chosen random variable distributed according to π and satisfying payoff constraints. It follows that the following equivalence holds.

Lemma 5.3.13. For a probability measure π on $\mathbb{R}^d_+$, $\underline{\mathcal{M}}_{\pi,\mathcal{I}} \neq \varnothing$ if and only if $\mathcal{M}_{\pi,\mathcal{I}} \neq \varnothing$ if and only if

$$\int_{\mathbb{R}^d_+} s_i \pi(\mathrm{d}s_1, \ldots, \mathrm{d}s_d) = 1, \;\; i = 1, \ldots, d. \tag{5.3.12}$$

Note that if (5.3.12) fails, then one of the forwards is mispriced leading to arbitrage opportunities[3]. We exclude this situation from our setup. The following is then a multi-dimensional extension of the pricing–hedging duality in Dolinsky and Soner [39].

Theorem 5.3.14. Consider traded options $\mathcal{X}$ and information space $\mathcal{I}$ satisfying (5.3.10) and Assumption 5.3.1, with market prices $\mathcal{P}$ such that $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}} \neq \varnothing$. Then for any uniformly continuous and bounded G, we have

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = P_{\pi,\mathcal{I}}(G).$$

It is clear that the above result holds if instead of assuming every $X \in \mathcal{X}$ is bounded and Lipschitz continuous, we allow bounded and uniformly continuous European payoffs, as long as $\mathcal{X}$ contains a subset made of bounded and Lipschitz continuous payoffs, which is rich enough to guarantee uniqueness of π which satisfies (5.3.11).

We now turn to the case when $\mathcal{X}$ is much smaller and the market prices determine marginal distributions of $\mathbb{S}^{(i)}_T$ for $i \leq d$. For concreteness, let us consider the case when put options are traded

$$\mathcal{X} = \{(K - \mathbb{S}^{(i)}_{T_j})^+ : i = 1, \ldots, d, \; j = 1, \ldots, n, \; K \in \mathbb{R}_+\}. \tag{5.3.13}$$

Arbitrage considerations, see e.g. Cox and Obłój [25] and Chapter 3, show that absence of (weak type of) arbitrage is equivalent to $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$. Note that the latter is equivalent to market prices $\mathcal{P}$ being encoded by probability measures $(\mu^{(i)}_j)$ with

$$p_{i,j}(K) = \mathcal{P}((K - \mathbb{S}^{(i)}_{T_j})^+) = \int (K - s)^+ \mu^{(i)}_j(\mathrm{d}s), \tag{5.3.14}$$

[3]This may be, depending on the sign of mispricing and the admissibility criterion, a *strong arbitrage* in Cox and Obłój [25] or *model-independent arbitrage* in Davis and Hobson [31] and Acciaio et al. [1] or else a weaker type of approximate arbitrage, e.g. a *weak free lunch of vanishing risk*; see Cox and Obłój [25] and Chapter 3.

where, for each $i = 1, \dots, d$, $\mu_1^{(i)}, \dots, \mu_n^{(i)}$ have finite first moments, mean 1 and increase in *convex order* ($\mu_1^{(i)} \preceq \mu_2^{(i)} \preceq \cdots \preceq \mu_n^{(i)}$), i.e., $\int \phi(x)\mu_1^{(i)}(\mathrm{d}x) \leq \ldots \leq \int \phi(x)\mu_n^{(i)}(\mathrm{d}x)$ for any convex function $\phi : \mathbb{R}_+ \to \mathbb{R}$. In fact, as noted already by Breeden and Litzenberger [13], probability measures $\mu_j^{(i)}$ are defined by

$$\mu_j^{(i)}([0,K]) = p'_{i,j}(K+) \quad \text{for } K \in \mathbb{R}_+.$$

We may think of $(\mu_j^{(i)})$ and $\mathfrak{P}$ as the modelling inputs. The set of calibrated market models $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ is simply the set of probability measures $\mathbb{P} \in \mathcal{M}$ such that $\mathbb{S}_{T_j}^{(i)}$ is distributed according to $\mu_j^{(i)}$, and $\mathbb{P}(\mathfrak{P}) = 1$. Accordingly, we write $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} = \mathcal{M}_{\vec{\mu},\mathfrak{P}}$ and $P_{\vec{\mu},\mathfrak{P}}(G) = P_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$. Furthermore, since $\mu_j^{(i)}$'s all have means equal to 1, under any $\mathbb{P} \in \mathcal{M}_{\vec{\mu},\mathfrak{P}}$, $\mathbb{S}$ is a (true) martingale.

Remark 5.3.15. It follows, see Strassen [101], that $\mathcal{M}_{\vec{\mu},\mathcal{I}} \neq \varnothing$ if and only if $\mu_1^{(i)}, \dots, \mu_n^{(i)}$ have finite first moments, mean 1 and increase in *convex order*, for any $i = 1, \dots, d$. However, in general, the additional constraints associated with a non-trivial $\mathfrak{P} \subsetneq \mathcal{I}$ are much harder to understand.

In this context we can improve Theorem 5.3.6 and narrow down the class of approximate market models requiring that they match exactly the marginal distributions at the last maturity.

Definition 5.3.16. Let $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}$ be the set of all measure $\mathbb{P} \in \mathcal{M}$ such that $\mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_j}^{(i)})$, the law of $\mathbb{S}_{T_j}^{(i)}$ under $\mathbb{P}$ satisfies

$$\mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_n}^{(i)}) = \mu_n^{(i)} \text{ and } d_p(\mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_j}^{(i)}), \mu_j^{(i)}) \leq \eta, \text{ for } j = 1, \dots, n-1, \ i = 1, \dots, d,$$

and furthermore $\mathbb{P}(\mathfrak{P}^\eta) \geq 1 - \eta$. Finally, let

$$\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) := \lim_{\eta \searrow 0} \sup_{\mathbb{P} \in \mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})].$$

Note that $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta} \subset \mathcal{M}^{\epsilon(\eta)}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ for a suitable choice[4] of $\epsilon(\eta)$ which converges to zero as $\eta \to 0$. It follows that $\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) \leq \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$. The following result extends and sharpens the duality obtained in Theorem 5.3.6 to the current setting.

Theorem 5.3.17. Let $\mathfrak{P}$ be a measurable subset of $\mathcal{I}$, $\mathcal{X}$ be given by (5.3.13) and $\mathcal{P}$ be such that, for any $\eta > 0$, $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta} \neq \varnothing$, where $\vec{\mu}$ is defined via (5.3.14). Then for any uniformly continuous and bounded G the robust pricing–hedging duality holds between the approximate values:

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G).$$

[4]One can take $\epsilon(\eta) = \sqrt{\eta} + 2f(1/\sqrt{\eta})$ with $f(K) = \max_{1 \leq i \leq d} \{p_{i,n}(K) - K + 1\}$.

5.3.3 Martingale optimal transport duality with exact marginal matching

Theorem 5.3.14 extends the duality obtained in [39]. In general we obtain an *asymptotic* duality result with the dual and primal problems defined through a limiting procedure. In this section, we want to focus on establishing a duality result without any asymptotic approximation. As already seen from Theorem 5.3.14, in a setting where there is a single marginal and the prediction set is absent, this type of duality result can be obtained without imposing further conditions on the payoff function G other than uniform continuity. However, to achieve this goal in a more general setting, we will impose stricter conditions on the payoff function G and prediction set $\mathfrak{P}$.

Assumption 5.3.18. G is bounded, uniformly continuous w.r.t. the sup norm $\|\cdot\|$ and satisfies that there exists a constant $L > 0$ such that for all $\upsilon, \hat{\upsilon} \in \mathcal{D}([0, T_n], \mathbb{R}^d_+)$ be of the form

$$\upsilon_t = \sum_{i=1}^{n} \sum_{j=0}^{m_i - 1} \upsilon_{i,j} \mathbb{1}_{[t_{i,j}, t_{i,j+1})}(t) + \upsilon_{n,m_n-1} \mathbb{1}_{T_n}(t),$$
$$\hat{\upsilon}_t = \sum_{i=1}^{n} \sum_{j=1}^{m_i - 1} \upsilon_{i,j} \mathbb{1}_{[\hat{t}_{i,j}, \hat{t}_{i,j+1})}(t) + \upsilon_{n,m_n-1} \mathbb{1}_{T_n}(t)$$

where $t_{i,0} = \hat{t}_{i,0} = T_i \ \forall 0 \le i \le n-1$, $t_{i,m_i-1} = \hat{t}_{i,m_i-1} = T_i \ \forall 1 \le i \le n$,

$$|G(\upsilon) - G(\hat{\upsilon})| \le L \|\upsilon\|^p \sum_{i=1}^{n} \sum_{i=1}^{m_i} |\Delta t_{i,j} - \Delta \hat{t}_{i,j}| \tag{5.3.15}$$

where as usual $\Delta t_{i,j} := t_{i,j} - t_{i,j-1}$ and $\Delta \hat{t}_{i,j} := \hat{t}_{i,j} - \hat{t}_{i,j-1}$.

Note that Assumption 5.3.18 is close in spirit to Assumption 2.1 in [39]. Despite their proximity, our assumption here is strictly weaker, which can be seen from the fact that it includes European options having intermediate maturities, in contrast to Assumption 2.1 in [39].

Definition 5.3.19. We say $\mathfrak{P}$ is *time invariant* if for any nondecreasing continuous function $f : [0, T_n] \to [0, T_n]$ such that $f(0) = 0$ and $f(T_i) = T_i$ for any $i = 1, \ldots, n$, $S \in \mathfrak{P}$ implies $(S_{f(t)})_{t \in [0, T_n]} \in \mathfrak{P}$.

We also need to assume that μ's admit a finite p^{th} moment.

Assumption 5.3.20. Assume $\vec{\mu} = (\mu_j^{(i)} : i = 1, \ldots, d,\ j = 1, \ldots, n)$ are probability measures on $\mathbb{R}_+$, with mean 1, admitting finite p-th moment for some $p > 1$ and $\mu_1^{(i)} \preceq \mu_2^{(i)} \preceq \cdots \preceq \mu_n^{(i)}$, $i = 1, \ldots, d$.

Theorem 5.3.21. Let $\vec{\mu}$ satisfy Assumption 5.3.20 and $\mathfrak{P}$ be closed and time invariant. Then, under Assumption 5.3.1, for every G that satisfies Assumption 5.3.18 the following robust pricing–hedging duality holds

$$\widetilde{V}^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = V^{(p)}_{\vec{\mu},\mathfrak{P}}(G) = P_{\vec{\mu},\mathfrak{P}}(G) = \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G), \tag{5.3.16}$$

where p is the same as in Assumption 5.3.20.

5.3.4 Martingale optimal transport duality for unbounded claims

We want to extend Theorems 5.3.17 and 5.3.21 to unbounded exotic options, including a lookback option. However, the admissibility condition considered so far, and given by (5.2.1), is too restrictive and has to be relaxed. To see this consider $d = 1$, $K = 0$, $\mathcal{X}$ is given by (5.3.13) and $G(\mathbb{S}) = \sup_{0 \le t \le T_n} \mathbb{S}_t$. If G could be superreplicated by an admissible trading strategy $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$ then, following similar arguments as for (5.3.1), we see that

$$\mathbf{P}_{\mathcal{I}}(G - X) \le 0.$$

This is clearly impossible since X is bounded and there exists $\mathbb{P} \in \mathcal{M}$ such that $\mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] = \infty$. The argument is similar if instead of puts we took all call options. We conclude that we need to enlarge the set of dynamic trading strategies $\mathcal{A}$.

We fix $p > 1$ and, following Dolinsky and Soner [39], define the following admissibility condition: γ is admissible if $\gamma : \Omega \to \mathcal{D}[0, T_n]$ is progressively measurable and of bounded variation, satisfying

$$\int_0^t \gamma_u(S) \cdot \mathrm{d}S_u \ge -M\big(1 + \sup_{0 \le s \le t} |S_s|^p\big), \quad \forall\, S \in \mathcal{I},\ t \in [0, T_n],\ \text{for some } M > 0. \tag{5.3.17}$$

To avoid confusion, we denote by $\mathcal{A}^{(p)}$ the set of all such γ. We also say $(X, \gamma) \in \mathcal{A}^{(p)}_{\mathcal{X}}$ if $\gamma \in \mathcal{A}^{(p)}$ and $X = a_0 + \sum_{i=1}^m a_i X_i$, for some m and $X_i \in \mathcal{X}^{(p)}$ given by

$$\mathcal{X}^{(p)} := \{f(\mathbb{S}^{(i)}_{T_j}) : |f(x)| \le K(1+|x|^p) \text{ for some } K > 0,\ \text{for } j = 1, \ldots, n, i = 1, \ldots, d\}.$$

As previously with $\mathcal{X}$ in (5.3.13), the above set $\mathcal{X}^{(p)}$ is large enough to determine uniquely the marginal distributions of $\mathbb{S}^{(i)}_{T_j}$. That is $\mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ implies that there exist unique probability measures $\mu^{(i)}_j$ such that $\mathcal{L}_{\mathbb{P}}(\mathbb{S}^{(i)}_{T_j}) = \mu^{(i)}_j$, $i = 1, \dots, d$, $j = 1, \dots, n$ for any $\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathfrak{P}} = \mathcal{M}_{\vec{\mu},\mathfrak{P}}$. We write $V^{(p)}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ for the superreplication cost $V_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$ but with $(X, \gamma) \in \mathcal{A}^{(p)}_{\mathcal{X}}$ and $\widetilde{V}^{(p)}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ for the approximative value.

Theorem 5.3.22. Let $\vec{\mu}$ satisfy Assumption 5.3.20, $\mathfrak{P}$ be a measurable subset of $\mathcal{I}$ such that for any $\eta > 0$, $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta} \neq \varnothing$. Then, under Assumption 5.3.1, for any uniformly continuous G that satisfies

$$|G(\mathbb{S})| \leq L(1 + \sup_{0 \leq t \leq T_n} |\mathbb{S}_t|^p),$$

the following robust pricing–hedging duality holds

$$\widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G),$$

where p is the same as in Assumption 5.3.20.

5.4 First proofs

We present in this section proofs of all the results except Theorem 5.3.2 which is shown in Section 5.5.

5.4.1 Proof of Corollary 5.3.3 and Remark 5.3.4

Note that any (X, γ) that superreplicates $G - N\lambda_{\mathfrak{P}}$ also superreplicates $G - 1/N$ on $\mathfrak{P}^{\frac{1}{N^2}}$. It follows that

$$\begin{aligned}\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) =& \inf\Big\{\mathcal{P}(X) \,:\, \exists (X,\gamma) \in \mathcal{A}^{sp}_{\mathcal{X}} \text{ s.t. } (X,\gamma) \text{ superreplicates } G \text{ on } \mathfrak{P}^{\epsilon} \text{ for some } \epsilon > 0\Big\} \\ \leq& \frac{1}{N} + \inf\Big\{\mathcal{P}(X) \,:\, \exists (X,\gamma) \in \mathcal{A}^{sp}_{\mathcal{X}} \text{ s.t. } (X,\gamma) \text{ superreplicates } G - N\lambda_{\mathfrak{P}} \text{ on } \mathcal{I}\Big\}.\end{aligned}$$

Since it holds for any N, we have

$$\begin{aligned}\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) &\leq \inf_{N \geq 0} \inf\Big\{\mathcal{P}(X) \,:\, \exists (X,\gamma) \in \mathcal{A}^{sp}_{\mathcal{X}} \text{ s.t. and } (X,\gamma) \text{ superreplicates } G - N\lambda_{\mathfrak{P}} \text{ on } \mathcal{I}\Big\} \\ &= \inf_{N \geq 0} \widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G - N\lambda_{\mathfrak{P}}).\end{aligned}$$

Note that by the same argument above we have

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(H) \leq \inf_{N \geq 0} \widetilde{V}_{\mathcal{X},\mathcal{P},\mathcal{I}}(H - N\lambda_{\mathfrak{P}}) \quad \text{and} \quad \widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(H) \leq \inf_{N \geq 0} \widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathcal{I}}(H - N\lambda_{\mathfrak{P}}) \quad (5.4.1)$$

hold for every bounded measurable H.

Notice that

$$
\begin{aligned}
&\inf\Big\{\mathcal{P}(X)\ :\ \exists(X,\gamma)\in\mathcal{A}^{sp}_{\mathcal{X}}\text{ s.t. }(X,\gamma)\text{ superreplicates }G-N\lambda_{\mathfrak{P}}\text{ on }\mathcal{I}\Big\}\\
=&\inf_{X\in\mathrm{Lin}(\mathcal{X})}\Big\{\mathcal{P}(X)+\inf\Big\{x\ :\ \exists\gamma\in\mathcal{A}^{sp}\text{ s.t. }\gamma\text{ superreplicates }G-N\lambda_{\mathfrak{P}}-X-x\text{ on }\mathcal{I}\Big\}\Big\}\\
=&\inf_{X\in\mathrm{Lin}(\mathcal{X})}\big\{\mathcal{P}(X)+\mathbf{V}^{sp}_{\mathcal{I}}(G-N\lambda_{\mathfrak{P}}-X)\big\}\\
=&\inf_{X\in\mathrm{Lin}(\mathcal{X})}\Big\{\mathcal{P}(X)+\mathbf{P}_{\mathcal{I}}(G-X-N\lambda_{\mathfrak{P}})\Big\},
\end{aligned}
$$

where the last equality is justified by Theorem 5.3.2 as $\lambda_{\mathfrak{P}}$ and X are bounded and uniformly continuous. Hence, we have

$$
\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)\leq\inf_{N\geq 0,\,X\in\mathrm{Lin}(\mathcal{X})}\Big\{\mathcal{P}(X)+\mathbf{P}_{\mathcal{I}}(G-X-N\lambda_{\mathfrak{P}})\Big\}.
$$

On the other hand, given any $(X,\gamma)\in\mathcal{A}_{\mathcal{X}}$ and $\epsilon>0$ such that (X,γ) superreplicates G on $\mathfrak{P}^{\epsilon}$, by the admissibility of $(X,\gamma)\in\mathcal{A}_{\mathcal{X}}$ and boundedness of X and G, if $N>0$ is sufficiently large then

$$
X(S)+\int_0^{T_n}\gamma_u(S)\cdot\mathrm{d}S_u\geq G(S)-N,
$$

and hence (X,γ) superreplicates $G-N\lambda_{\mathfrak{P}}$. It follows that

$$
\begin{aligned}
\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)=&\inf\Big\{\mathcal{P}(X)\ :\ \exists(X,\gamma)\in\mathcal{A}_{\mathcal{X}}\text{ s.t. }(X,\gamma)\text{ superreplicates }G\text{ on }\mathfrak{P}^{\epsilon}\text{ for some }\epsilon>0\Big\}\\
\geq&\inf_{N\geq 0}\inf\Big\{\mathcal{P}(X)\ :\ \exists(X,\gamma)\in\mathcal{A}_{\mathcal{X}}\text{ s.t. }(X,\gamma)\text{ superreplicates }G-N\lambda_{\mathfrak{P}}\text{ on }\mathcal{I}\Big\}\\
=&\inf_{N\geq 0,\,X\in\mathrm{Lin}(\mathcal{X})}\big\{\mathcal{P}(X)+\mathbf{V}_{\mathcal{I}}(G-X-N\lambda_{\mathfrak{P}})\big\}\\
=&\inf_{N\geq 0,\,X\in\mathrm{Lin}(\mathcal{X})}\Big\{\mathcal{P}(X)+\mathbf{P}_{\mathcal{I}}(G-X-N\lambda_{\mathfrak{P}})\Big\},
\end{aligned}
$$

where the last equality is again justified by Theorem 5.3.2. As $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)\leq\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$, this establishes the equality in (5.3.4).

Note that by the same argument above we can argue that

$$
\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(H)\geq\inf_{N\geq 0}\widetilde{V}_{\mathcal{X},\mathcal{P},\mathcal{I}}(H-N\lambda_{\mathfrak{P}})\quad\text{and}\quad\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(H)\geq\inf_{N\geq 0}\widetilde{V}^{sp}_{\mathcal{X},\mathcal{P},\mathcal{I}}(H-N\lambda_{\mathfrak{P}})
$$

hold for every bounded measurable H. Therefore, combining this with (5.4.1), we show (5.3.5).

5.4.2 Proof of Theorem 5.3.6

To establish (5.3.6), we consider a $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$ that superreplicates G on $\mathfrak{P}^\epsilon$ for some $\epsilon > 0$, i.e.,

$$X(\mathbb{S}) + \int_0^{T_n} \gamma_u \mathrm{d}\mathbb{S}_u \geq G(\mathbb{S}) \text{ on } \mathfrak{P}^\epsilon.$$

Since X is bounded, by the definition of admissibility there exists a $M > 0$ such that

$$X(\mathbb{S}) + \int_0^{T_n} \gamma_u \mathrm{d}\mathbb{S}_u \geq G(\mathbb{S}) - M\lambda_{\mathfrak{P}}(\mathbb{S}). \tag{5.4.2}$$

Next, for each $N \geq 1$, we pick $\mathbb{P}^{(N)} \in \mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] \geq \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N}.$$

Since γ is progressively measurable, the integral $\int_0^{\cdot} \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u$, defined pathwise via integration by parts, agrees a.s. with the stochastic integral under any $\mathbb{P}^{(N)}$. Then, by (5.2.1), the stochastic integral is a $\mathbb{P}^{(N)}$–super-martingale and hence $\mathbb{E}_{\mathbb{P}^{(N)}}\left[\int_0^{T_n} \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u\right] \leq 0$. Therefore, from (5.4.2), we can derive that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S})] \geq \mathbb{E}_{\mathbb{P}^{(N)}}\big[G(\mathbb{S}) - M\lambda_{\mathfrak{P}}(\mathbb{S})\big] \geq \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N} - \frac{M}{N}. \tag{5.4.3}$$

Also note that X takes the form of $a_0 + \sum_{i=1}^m a_i X_i$. Then by definition of $\mathcal{M}^\eta_{\mathcal{X},\mathcal{P},\mathfrak{P}}$

$$\big|\mathcal{P}(X) - \mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S})]\big| \to 0 \text{ as } N \to \infty.$$

This, together with (5.4.3), yields

$$\mathcal{P}(X) \geq \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G).$$

As $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$ is arbitrary, we therefore establish (5.3.6).

To show (5.3.7), we first deduce from Theorem 5.3.2 and (5.3.4) that

$$\begin{aligned}
\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) &= \inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0} \Big\{\mathbf{P}_{\mathcal{I}}(G - X - N\lambda_{\mathfrak{P}}) + \mathcal{P}(X)\Big\} \\
&= \lim_{N\to\infty} \inf_{X\in\mathrm{Lin}_N(\mathcal{X})} \Big\{\sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\} \\
&= \lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G] \\
&= \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G),
\end{aligned} \tag{5.4.4}$$

where the crucial third equality follows from (5.4.5) in Lemma 5.4.1 below.

Last, we show that $\underline{\mathcal{M}}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ for any $\eta > 0$. By the above and equality between $\mathbf{P}_{\mathcal{I}} = \underline{\mathbf{P}}_{\mathcal{I}}$ in Theorem 5.3.2 we have

$$\begin{aligned}
&\inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\}\\
&= \inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\} = \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G).
\end{aligned}$$

Then, taking $G = 0$, as $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ for any $\eta > 0$,

$$\inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[-X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\} = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(0) = 0.$$

Therefore, it follows from the equivalence in Lemma 5.4.1, with $\mathcal{M}_s = \underline{\mathcal{M}}$, that $\underline{\mathcal{M}}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \neq \varnothing$ for any $\eta > 0$. In addition, by (5.4.6) in Lemma 5.4.1 below,

$$\begin{aligned}
\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) &= \inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\}\\
&= \lim_{N\to\infty}\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}}\mathbb{E}_{\mathbb{P}}[G] = \underline{\widetilde{P}}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G).
\end{aligned}$$

This completes the proof of Theorem 5.3.6. It remains to argue the following which is stated in a general form and also used in subsequent proofs.

Lemma 5.4.1. Let $\mathfrak{P}$ be a measurable subset of $\mathcal{I}$, $\mathcal{X}$ satisfy Assumption 5.3.5 and $\mathcal{M}_s$ be a non-empty convex subset of $\mathcal{M}_{\mathcal{I}}$. Then the following two are equivalent:

(i) for any $\eta > 0$, $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \bigcap \mathcal{M}_s \neq \varnothing$.

(ii) $\inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[-X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\} = 0$.

Further, under (i) or (ii), for any uniformly continuous and bounded $G : \Omega \to \mathbb{R}$ we have:

$$\inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)\Big\} = \lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G]. \tag{5.4.5}$$

Moreover, for any $\alpha, \beta \geq 0$ and $D \in \mathbb{N}$.

$$\begin{aligned}
&\inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \alpha\wedge(\beta\sqrt{m^{(D)}(\mathbb{S})}) - X(\mathbb{S}) - N\lambda_{\mathfrak{P}}(\mathbb{S})] + \mathcal{P}(X)\Big\}\\
&\leq \lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \alpha\wedge(\beta\sqrt{m^{(D-2)}(\mathbb{S})})],
\end{aligned} \tag{5.4.6}$$

where $m^{(D)}$ is defined in Definition 4.5.1.

Proof. Choose $\kappa > 2 \vee (\|G\|_\infty + \alpha)$. We first observe that

$$\begin{aligned}
&\inf_{X \in \text{Lin}(\mathcal{X}),\, N \geq 0} \Big\{ \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta \sqrt{m^{(D)}}) - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \Big\} \\
&= \lim_{N \to \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \Big\{ \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta \sqrt{m^{(D)}}) - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \Big\}. \qquad (5.4.7)
\end{aligned}$$

Define the function $\mathcal{G} : \text{Lin}_N(\mathcal{X}) \times \mathcal{M}_s \to \mathbb{R}$ by

$$\begin{aligned}
\mathcal{G}(X, \mathbb{P}) :=& \lim_{\epsilon \searrow 0} \inf_{\tilde{\mathbb{P}} \in \underline{\mathcal{M}}_{\mathcal{I}},\, d_p(\tilde{\mathbb{P}}, \mathbb{P}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[G(\mathbb{S}) - \alpha \wedge (\beta \sqrt{m^{(D-2)}}) - x(\mathbb{S}) - N\lambda_{\mathfrak{P}}(\mathbb{S})\Big] + \mathcal{P}(X) \\
=& \lim_{\epsilon \searrow 0} \inf_{\tilde{\mathbb{P}} \in \underline{\mathcal{M}}_{\mathcal{I}},\, d_p(\tilde{\mathbb{P}}, \mathbb{P}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D-2)}})\Big] + \mathbb{E}_{\mathbb{P}}[G - N\lambda_{\mathfrak{P}} - X] + \mathcal{P}(X).
\end{aligned}$$

Then by (4.5.1) in Remark 4.5.2, for any sequence $(\mathbb{P}^{(k)})_{k \geq 1}$ converging to $\mathbb{P}$ weakly,

$$\mathbb{E}_{\mathbb{P}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D)}(\mathbb{S})})\Big] \leq \liminf_{k \to \infty} \mathbb{E}_{\mathbb{P}^{(k)}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D-2)}(\mathbb{S})})\Big].$$

and hence

$$\begin{aligned}
&\lim_{N \to \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \Big\{ \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta \sqrt{m^{(D)}}) - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \Big\} \\
&\leq \lim_{N \to \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \sup_{\mathbb{P} \in \mathcal{M}_s} \mathcal{G}(X, \mathbb{P}),
\end{aligned}$$

with equality when $\alpha = \beta = 0$.

The next step is to interchange the order of the infimum and supremum. Notice that when we fix $\mathbb{P}$, $\mathcal{G}$ is affine in the first variable and continuous due to the bounded convergence theorem. In addition, by definition $\mathcal{G}$ is lower-semi continuous in the second variable. Furthermore, $\mathcal{G}$ is convex in the second variable. To justify this, we notice that $\mathbb{P} \mapsto \mathbb{E}_{\mathbb{P}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D-2)}(\mathbb{S})})\Big]$ is a linear functional and it follows that for each $\epsilon > 0$ and $\lambda \in [0, 1]$

$$\begin{aligned}
&\inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s,\, d_p(\tilde{\mathbb{P}}, \lambda\mathbb{P}^{(1)} + (1-\lambda)\mathbb{P}^{(2)}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D-2)}(\mathbb{S})})\Big] \\
\leq& \lambda \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s,\, d_p(\tilde{\mathbb{P}}, \mathbb{P}^{(1)}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D-2)}(\mathbb{S})})\Big] \\
&+ (1 - \lambda) \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s,\, d_p(\tilde{\mathbb{P}}, \mathbb{P}^{(2)}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[- \alpha \wedge (\beta \sqrt{m^{(D-2)}(\mathbb{S})})\Big].
\end{aligned}$$

Since $\text{Lin}_N(\mathcal{X})$ is convex and compact, it follows that we can now apply min–max Theorem (see Corollary 2 in Terkelsen [102]) to $\mathcal{G}$ and derive

$$\lim_{N \to \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \sup_{\mathbb{P} \in \mathcal{M}_s} \mathcal{G}(X, \mathbb{P}) = \lim_{N \to \infty} \sup_{\mathbb{P} \in \mathcal{M}_s} \inf_{X \in \text{Lin}_N(\mathcal{X})} \mathcal{G}(X, \mathbb{P}).$$

Therefore, we have

$$\begin{aligned}
&\lim_{N\to\infty}\inf_{X\in \mathrm{Lin}_N(\mathcal{X})}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D)}})-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\}\\
&\le \lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}_s}\inf_{X\in \mathrm{Lin}_N(\mathcal{X})}\Big\{\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-2)}})-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\},
\end{aligned}\tag{5.4.8}$$

with equality when $\alpha=\beta=0$.

Now first consider the case: $\alpha=\beta=0$. In this case, it follows from above that

$$\begin{aligned}
&\lim_{N\to\infty}\inf_{X\in \mathrm{Lin}_N(\mathcal{X})}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\}\\
&=\lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}_s}\Big\{\inf_{X\in \mathrm{Lin}_N(\mathcal{X})}\mathbb{E}_{\mathbb{P}}[G-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\}.
\end{aligned}\tag{5.4.9}$$

Suppose that for any $\eta>0$, $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\bigcap\mathcal{M}_s\neq\varnothing$. Then for any N, $\mathbb{P}\in\mathcal{M}^{1/N^2}_{\mathcal{X},\mathcal{P},\mathfrak{P}}$ and $X\in\mathrm{Lin}_N(\mathcal{X})$,

$$|\mathbb{E}_{\mathbb{P}}[X]-\mathcal{P}(X)|\le\sum_{i=1}^{m}|a_i|\big|\mathbb{E}_{\mathbb{P}}[X_i]-\mathcal{P}(X_i)\big|\le\frac{N}{N^2}=\frac{1}{N},$$

where X takes the form of $a_0+\sum_{i=1}^m a_iX_i$, for some $m\in\mathbb{N}$, $X_i\in\mathcal{X}$ and $a_i\in\mathbb{R}$ such that $\sum_{i=0}^m|a_i|\le N$. In addition, $\lambda_{\mathfrak{P}}\le\frac{1}{N^2}\mathbb{1}_{\{\mathbb{S}\in\mathfrak{P}^{1/N^2}\}}+\mathbb{1}_{\{\mathbb{S}\notin\mathfrak{P}^{1/N^2}\}}$ leads to

$$\mathbb{E}_{\mathbb{P}}[N\lambda_{\mathfrak{P}}]\le\frac{1}{N}\mathbb{P}(\mathbb{S}\in\mathfrak{P}^{1/N^2})+N\mathbb{P}(\mathbb{S}\notin\mathfrak{P}^{1/N^2})\le\frac{2}{N}.$$

Therefore, we can deduce that

$$\begin{aligned}
&\lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}^{1/N^2}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s}\Big\{\inf_{X\in \mathrm{Lin}_N(\mathcal{X})}\mathbb{E}_{\mathbb{P}}[G-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\}\\
&\ge\lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}^{1/N^2}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s}\Big\{\mathbb{E}_{\mathbb{P}}[G]-\frac{3}{N}\Big\}=\lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G].
\end{aligned}\tag{5.4.10}$$

Consequently, by taking $G=0$, using (5.4.7)–(5.4.10) and noting that considering a sup over a larger set increases its value, we have

$$\inf_{X\in\mathrm{Lin}(\mathcal{X}),\,N\ge0}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\}\ge0,$$

which leads to

$$\inf_{X\in\mathrm{Lin}(\mathcal{X}),\,N\ge0}\Big\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\Big\}=0.$$

On the other hand, if

$$\inf_{X\in \mathrm{Lin}(\mathcal{X}),\, N\geq 0}\left\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\right\}=0, \tag{5.4.11}$$

then we will argue in the following that in the sup term of (5.4.9) it suffices to consider probability measures $\mathbb{P}\in\mathcal{M}^{2\kappa/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s$. Suppose $\mathbb{P}\in(\mathcal{M}_s\setminus\mathcal{M}^{2\kappa/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}})$, then either there exist $X\in\mathcal{X}$ such that $\mathbb{E}_{\mathbb{P}}[X]-\mathcal{P}(X)>2\kappa/N$ or $\mathbb{P}(\mathbb{S}\notin\mathfrak{P}^{2\kappa/N})\geq 2\kappa/N$. In the former case, since $NX\in\mathrm{Lin}(\mathcal{X})$,

$$\mathbb{E}_{\mathbb{P}}[G-NX-N\lambda_{\mathfrak{P}}]+\mathcal{P}(NX)\leq\mathbb{E}_{\mathbb{P}}[G]-N(\mathbb{E}_{\mathbb{P}}[X]-\mathcal{P}(X))<-\kappa,$$

and in the latter case, $\mathbb{E}_{\mathbb{P}}[G-N\lambda_{\mathfrak{P}}]<\kappa-2\kappa=-\kappa$, while

$$\begin{aligned}&\lim_{N\to\infty}\sup_{\mathbb{P}\in\mathcal{M}_s}\left\{\inf_{X\in\mathrm{Lin}_N(\mathcal{X})}\mathbb{E}_{\mathbb{P}}[G-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\right\}\\ =&\inf_{X\in \mathrm{Lin}(\mathcal{X}),\, N\geq 0}\left\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\right\}\\ \geq&\inf_{X\in \mathrm{Lin}(\mathcal{X}),\, N\geq 0}\left\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[-\kappa-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\right\}=-\kappa,\end{aligned}$$

where the last equality follows from (5.4.9). This argument also implies that $\mathcal{M}^{2\kappa/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s\neq\varnothing$ for any $N\in\mathbb{N}$. Therefore we have the equivalence between

$$\forall\eta>0,\ \mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\bigcap\mathcal{M}_s\neq\varnothing\ \text{ and }\ \inf_{X\in \mathrm{Lin}(\mathcal{X}),\, N\geq 0}\left\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}(-X-N\lambda_{\mathfrak{P}})+\mathcal{P}(X)\right\}=0.$$

Now consider the general case: $\alpha,\beta\geq 0$. We begin to verify (5.4.5) and (5.4.6). Since $\mathcal{M}^{\eta}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\bigcap\mathcal{M}_s\neq\varnothing\ \forall\eta>0$,

$$\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[X-N\lambda_{\mathfrak{P}}]-\mathcal{P}(X)\geq 0,\quad \forall\, X\in\mathcal{X},\ N\in\mathbb{R}_+.$$

Hence for every N and $X\in\mathrm{Lin}_N(X)$

$$\begin{aligned}&\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D)}})-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\\ \geq&-\|G\|_\infty-\alpha+\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[X-N\lambda_{\mathfrak{P}}]-\mathcal{P}(X)\geq-\kappa,\end{aligned}$$

and therefore

$$\lim_{N\to\infty}\inf_{X\in\mathrm{Lin}_N(\mathcal{X})}\left\{\sup_{\mathbb{P}\in\mathcal{M}_s}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D)}})-X-N\lambda_{\mathfrak{P}}]+\mathcal{P}(X)\right\}\geq-\kappa.$$

Then, by using the same argument as above, we can argue that in the sup term of (5.4.8) it suffices to consider probability measures $\mathbb{P} \in \mathcal{M}^{2\kappa/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}} \cap \mathcal{M}_s$ and hence we have

$$\begin{aligned}
&\lim_{N\to\infty} \inf_{X\in\mathrm{Lin}_N(\mathcal{X})} \Big\{ \sup_{\mathbb{P}\in\mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D)}}) - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \Big\} \\
&\leq \lim_{N\to\infty} \inf_{X\in\mathrm{Lin}_N(\mathcal{X})} \Big\{ \sup_{\mathbb{P}\in\mathcal{M}^{2\kappa/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-2)}}) - X] + \mathcal{P}(X) \Big\} \\
&\leq \lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}^{2\kappa/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-2)}})],
\end{aligned}$$

where the second inequality follows from the fact that $-X \in \mathrm{Lin}_N(\mathcal{X})$ for every $X \in \mathrm{Lin}_N(\mathcal{X})$. This completes the verification of (5.4.6). In the case that $\alpha = \beta = 0$, combining the inequality above with (5.4.10), we then conclude that

$$\inf_{X\in\mathrm{Lin}(\mathcal{X}),\, N\geq 0} \Big\{ \sup_{\mathbb{P}\in\mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \Big\} = \lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G].$$

□

5.4.3 Proof of Theorem 5.3.10

From Theorem 5.3.6, as worked out in Example 5.3.7, we know that

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \widetilde{P}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G]. \tag{5.4.12}$$

Now for every positive integer N, we pick $\mathbb{P}^{(N)} \in \mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathcal{I}}$ such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G] + 1/N \geq \sup_{\mathbb{P}\in\mathcal{M}^{1/N}_{\mathcal{X},\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G].$$

We write

$$p^{(N)}_{k,i,j} := \mathbb{E}_{\mathbb{P}^{(N)}}[(K^{(i)}_{k,j} - S^{(i)}_{k,j})^+]$$

for any $i = 1, \ldots d$, $j = 1, \ldots, n$, $k = 1, \ldots, m(i,j)$, and define $\tilde{p}^{(N)}_{k,i,j}$'s by

$$\tilde{p}^{(N)}_{k,i,j} = \sqrt{N}\big(p_{k,i,j} - (1 - 1/\sqrt{N})p^{(N)}_{k,i,j}\big).$$

Note that

$$|\tilde{p}^{(N)}_{k,i,j} - p_{k,i,j}| = (\sqrt{N} - 1)|p_{k,i,j} - p^{(N)}_{k,i,j}| \leq \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \quad \forall i, j, k. \tag{5.4.13}$$

Then, it follows from Assumption 5.3.8 that when N is large enough there exists a $\tilde{\mathbb{P}}^{(N)} \in \mathcal{M}_{\mathcal{I}}$ such that

$$\tilde{p}_{k,i,j}^{(N)} := \mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] \quad \forall i,j,k.$$

Now we consider $\mathbb{Q} := (1-1/\sqrt{N})\mathbb{P}^{(N)} + \tilde{\mathbb{P}}^{(N)}/\sqrt{N}$. It follows that

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] =& (1-1/\sqrt{N})\mathbb{E}_{\mathbb{P}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] + \frac{1}{\sqrt{N}}\mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] \\ =& (1-1/\sqrt{N})p_{k,i,j}^{(N)} + \tilde{p}_{k,i,j}^{(N)}/\sqrt{N} = p_{k,i,j}\end{aligned}$$

and hence $\mathbb{Q} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}$. In addition,

$$\left|\mathbb{E}_{\mathbb{Q}}[G] - \mathbb{E}_{\mathbb{P}^{(N)}}[G]\right| \leq \frac{1}{\sqrt{N}}(\mathbb{E}_{\mathbb{P}^{(N)}}[|G|] + \mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[|G|]) \leq \frac{2\|G\|_\infty}{\sqrt{N}}.$$

Therefore, we have

$$\sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G] \leq \sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G] - \frac{2\|G\|_\infty}{\sqrt{N}} - \frac{1}{N},$$

which leads us to conclude

$$\widetilde{P}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G] \leq P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G).$$

Together with (5.4.12) and (5.3.1) this completes the proof.

5.4.4 Proof of Theorem 5.3.14

Let

$$\mathcal{Y} = \{f(\mathbb{S}_T^{(1)},\dots,\mathbb{S}_T^{(d)}) \,:\, f \in \mathcal{C}(\mathbb{R}_+^d,\mathbb{R}) \text{ s.t. } \sup_{\vec{x}\neq\vec{y}} \frac{|f(\vec{x}) - f(\vec{y})|}{|\vec{x}-\vec{y}|} \leq 1,\ \|f\|_\infty \leq 1\}.$$

Then, as $\mathrm{Lin}(\mathcal{X})$ is dense in $\mathrm{Lin}(\mathcal{Y})$,

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = V_{\mathcal{Y},\mathcal{P},\mathcal{I}}(G). \tag{5.4.14}$$

We now consider $\mathcal{Y}_M := \{f(\mathbb{S}_T^{(1)} \wedge M,\dots,\mathbb{S}_T^{(d)} \wedge M) \,:\, f \in \mathcal{Y}\}$. $\mathcal{Y}_M$ is a subset of $\mathcal{Y}$ for each $M \in \mathbb{N}$, and in consequence,

$$V_{\mathcal{Y},\mathcal{P},\mathcal{I}}(G) \leq V_{\mathcal{Y}_M,\mathcal{P},\mathcal{I}}(G), \quad \forall\, M \in \mathbb{N}. \tag{5.4.15}$$

Observe that, for any $X \in \mathcal{X}$,

$$\left|X(\mathbb{S}_T^{(1)} \wedge M, \dots, \mathbb{S}_T^{(d)} \wedge M) - X(\mathbb{S}_T^{(1)}, \dots, \mathbb{S}_T^{(d)})\right| \le \frac{2}{M-1} \sum_{i=1}^{d} \mathbb{S}_T^{(i)}$$

and hence

$$\Big| \int_{\mathbb{R}_+^d} X(s_1 \wedge M, \dots, s_d \wedge M) \pi(\mathrm{d}s_1, \dots, \mathrm{d}s_d) - \int_{\mathbb{R}_+^d} X(s_1, \dots, s_d) \pi(\mathrm{d}s_1, \dots, \mathrm{d}s_d) \Big| \le \frac{2d}{M}.$$

Note that by definition $\mathcal{Y}_M$ is closed and convex. Also, by Arzelá-Ascoli theorem, $\mathcal{Y}_M$ is compact. Hence $\mathrm{Lin}_1(\mathcal{Y}_M)$ satisfies Assumption 5.3.5. Therefore, applying Theorem 5.3.6 to $\mathcal{Y}_M$, we have

$$V_{\mathcal{Y}_M,\mathcal{P},\mathcal{I}}(G) = \underline{\widetilde{P}}_{\mathcal{Y}_M,\mathcal{P},\mathcal{I}}(G) \quad \forall M > 0. \tag{5.4.16}$$

Then, by putting (5.4.14), (5.4.15) and (5.4.16) together

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) \le \lim_{M\to\infty} \underline{\widetilde{P}}_{\mathcal{Y}_M,\mathcal{P},\mathcal{I}}(G) = \lim_{M\to\infty} \lim_{N\to\infty} \sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/N}_{\mathcal{Y}_M,\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G] \le \lim_{M\to\infty} \sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G].$$

For every $N \in \mathbb{N}$, take $\mathbb{P}^{(N)} \in \underline{\mathcal{M}}^{1/N}_{\mathcal{Y}_N,\mathcal{P},\mathcal{I}}$ such that

$$\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/N}_{\mathcal{Y}_N,\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G] \le \mathbb{E}_{\mathbb{P}^{(N)}}[G] + \frac{1}{N}.$$

Let $\pi^{(N)}$ be the law of $(\mathbb{S}_T^{(1)}, \dots, \mathbb{S}_T^{(d)})$ under $\mathbb{P}^{(N)}$. It is a probability measure on $\mathbb{R}_+^d$ with mean equal to 1. It follows that the family $\{\pi^{(N)}\}_{N\ge1}$ is tight. By Prokhorov's theorem, there exists $\{\pi^{(N_k)}\}_{k\ge1}$, a subsequence of $\{\pi^{(N)}\}_{N\ge1}$, converging to some $\tilde{\pi}$. In the following, we are going to argue that $\tilde{\pi}$ is in fact π. For any $X \in \mathcal{Y}$ and $N \in \mathbb{N}$,

$$\begin{aligned}
&\Big|\mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S}_T)] - \int_{\mathbb{R}_+^d} X(s_1, \dots, s_d) \pi(\mathrm{d}s_1, \dots, \mathrm{d}s_d)\Big| \\
\le &\Big|\mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S}_T^{(1)} \wedge N, \dots, \mathbb{S}_T^{(d)} \wedge N)] - \int_{\mathbb{R}_+^d} X(s_1 \wedge N, \dots, s_d \wedge N) \pi(\mathrm{d}s_1, \dots, \mathrm{d}s_d)\Big| \\
&+ 2\mathbb{E}_{\mathbb{P}^{(N)}}\left[\mathbb{1}_{\left\{\max_{1\le d\le n}\{\mathbb{S}_T^{(i)} > N\}\right\}}\right] + 2\int_{\mathbb{R}_+^d} \mathbb{1}_{\left\{\max_{1\le d\le n}\{s_i > N\}\right\}} \pi(\mathrm{d}s_1, \dots, \mathrm{d}s_d) \\
\le &\frac{1}{N} + \frac{2d}{N} + \frac{2d}{N} = \frac{4d+1}{N}.
\end{aligned}$$

By weak convergence of $\pi^{(N)}$, along a subsequence of $\{\pi^{(N)}\}_{N\ge1}$, for every $X \in \mathcal{Y}$

$$\int_{\mathbb{R}_+^d} X(s_1, \dots, s_d) \pi^{(N)}(\mathrm{d}s_1, \dots, \mathrm{d}s_d) \to \int_{\mathbb{R}_+^d} X(s_1, \dots, s_d) \tilde{\pi}(\mathrm{d}s_1, \dots, \mathrm{d}s_d) \quad \text{as } N \to \infty.$$

Therefore, for every $X \in \mathcal{Y}$

$$\int_{\mathbb{R}_+^d} X(s_1, \ldots, s_d)\pi(\mathrm{d}s_1, \ldots, \mathrm{d}s_d) = \int_{\mathbb{R}_+^d} X(s_1, \ldots, s_d)\tilde{\pi}(\mathrm{d}s_1, \ldots, \mathrm{d}s_d),$$

which implies that $\pi = \tilde{\pi}$ as $\mathcal{Y}$ is rich enough to guarantee uniqueness of π.

It follows that

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) \leq \lim_{N\to\infty} \sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/N}_{\mathcal{Y}_N,\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G] \leq \limsup_{k\to\infty} \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\pi^{(N_k)},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G] \leq \underline{P}_{\pi,\mathcal{I}}(G),$$

where the last inequality follows from the following lemma.

Lemma 5.4.2. Assume $\pi^{(N)}$ and π are probability measures on $\mathbb{R}_+^d$ such that $\pi^{(N)}$ and π satisfies (5.3.12) and $\pi^{(N)}$ converges to π weakly. Then, for any bounded and uniformly continuous G, $\alpha, \beta \geq 0$ and $D \in \mathbb{N}$.

$$\limsup_{N\to\infty} \underline{P}_{\pi^{(N)},\mathcal{I}}\big(G(\mathbb{S}) - \alpha \wedge (\beta\sqrt{m^{(D)}(\mathbb{S})})\big) \leq \underline{P}_{\pi,\mathcal{I}}\big(G(\mathbb{S}) - \alpha \wedge (\beta\sqrt{m^{(D-2)}(\mathbb{S})})\big),$$

where $m^{(D)}$ is defined in Definition 4.5.1.

Proof. Recall that $X_1^{(c)}, \ldots, X_K^{(c)}$ are the options which can traded continuously. Choose $f_e : \mathbb{R}_+ \to \mathbb{R}_+$ such that $|G(\omega) - G(\upsilon)| \leq f_e(\|\omega - \upsilon\|)$ for any $\omega, \upsilon \in \Omega$, $|X_i^{(c)}(\vec{x}) - X_i^{(c)}(\vec{y})| \leq f_e(|\vec{x} - \vec{y}|)$ for any $\vec{x}, \vec{y} \in \mathbb{R}_+^d$ and $\lim_{x\searrow 0} f_e(x) = 0$. Now fix N and $\mathbb{P}^{(N)} \in \underline{\mathcal{M}}_{\pi^{(N)},\mathcal{I}}$. By definition of $\underline{\mathcal{M}}$, there exists a complete probability space $(\Omega^W, \mathcal{F}_T^W, \mathbb{F}^W, P^W)$ together a finite dimensional Brownian motion $(W_t)_{t\geq 0}$ and the natural filtration $\mathcal{F}_t^W = \sigma\{W_s | s \leq t\}$, and a continuous martingale M defined on $(\Omega^W, \mathcal{F}_T^W, \mathbb{F}^W, P^W)$ such that $\mathbb{P}^{(N)} = P^W \circ M^{-1}$.

Write $\epsilon_N := d_p(\pi^{(N)}, \pi)$. That $\pi^{(N)}$ converges to π weakly is equivalent to saying that $\epsilon_N \to 0$ as $N \to \infty$. Fix N. If $\epsilon_N = 0$, then it is trivially true that $\underline{P}_{\pi^{(N)},\mathcal{I}}(G) = \underline{P}_{\pi,\mathcal{I}}(G)$. Therefore, we only consider the case that $\epsilon_N > 0$. It follows from Strassen's Theorem, Corollary of Theorem 11 on page 438 in Strassen [101] or Theorem 4 on page 358 in Shiryaev [96], and Theorem 1 in Skorokhod [98], by enlarging the Wiener space if necessary, we can find a $\mathcal{F}_T^W$ measurable random variable Λ such that $\Lambda^{(d+i)} = X_i^{(c)}(\Lambda^{(1)}, \ldots, \Lambda^{(1)})$ for every $i \leq K$,

$$(\Lambda^{(1)}, \ldots, \Lambda^{(d)}) \sim_{P^W} \pi \quad \text{and} \quad P^W(|\Lambda^{(i)} - M_T^{(i)}| > 2\epsilon_N) < 2\epsilon_N \quad \forall i \leq d. \tag{5.4.17}$$

We now construct a continuous martingale from Λ by taking conditional expectation, i.e.,

$$\Gamma_t = E^W[\Lambda | \mathcal{F}_t^W], \quad t \in [0, T],$$

where E^W is the expectation with respect to P^W. Note that by uniform continuity of $X_i^{(c)}$

$$|\Lambda^{(d+i)} - M_T^{(d+i)}| \leq f_e(2\epsilon_N) \quad \forall i \leq K, \quad \text{whenever } |\Lambda^{(j)} - M_T^{(j)}| \leq 2\epsilon_N \quad \forall j \leq d$$

Hence, for every $i \leq K$

$$P^W\big(|\Lambda^{(i+d)} - M_T^{(i+d)}| > f_e(2\epsilon_N)\big) \leq 2d\epsilon_N. \tag{5.4.18}$$

Observe that $E^W[\Lambda^{(i)}] = E^W[M_T^{(i)}] = 1$ and $\Lambda^{(i)} \geq 0$ P^W-a.s. $\forall\, i$. Then, using (5.4.17),

$$\begin{aligned}
E^W[|\Lambda^{(i)} - M_T^{(i)}|] &= 2E^W[(\Lambda^{(i)} - M_T^{(i)})^+] - E^W[\Lambda^{(i)} - M_T^{(i)}] \\
&= 2E^W[(\Lambda^{(i)} - M_T^{(i)})^+] \\
&\leq 4\epsilon_N + 2E^W[\Lambda^{(i)} \mathbb{1}_{\{|\Lambda^{(i)} - M_T^{(i)}| > 2\epsilon_N\}}] \\
&\leq 4\epsilon_N + 2E^W[\Lambda^{(i)} \mathbb{1}_{\{|\Lambda^{(i)} - M_T^{(i)}| > 2\epsilon_N\}} \mathbb{1}_{\{\Lambda > 1/\sqrt{\epsilon_N}\}}] + 4\sqrt{\epsilon_N} \\
&\leq 4\epsilon_N + 2\int_{\{x_i \geq \frac{1}{\sqrt{\epsilon_N}}\} \cap \mathbb{R}_+^d} x_i \pi(\mathrm{d}x_1, \dots, \mathrm{d}x_d) + 4\sqrt{\epsilon_N}, \quad \forall i = 1, \dots, d.
\end{aligned}$$

Similarly, for every $i \leq K$,

$$\begin{aligned}
E^W[|\Lambda^{(d+i)} - M_T^{(d+i)}|] =& 2E^W[(\Lambda^{(d+i)} - M_T^{(d+i)})^+] \\
\leq& 2f_e(2\epsilon_N) + 2E^W[\Lambda^{(i)} \mathbb{1}_{\{|\Lambda^{(d+i)} - M_T^{(d+i)}| > f_e(2\epsilon_N)\}}] \\
\leq& 2f_e(2\epsilon_N) + 4d \frac{\|X_i^{(c)}\|_\infty}{\mathcal{P}(X_i^{(c)})} \epsilon_N.
\end{aligned}$$

Now define η_N by

$$\eta_N = 2f_e(2\epsilon_N) + 4\epsilon_N + 4d \sum_{i=1}^{K} \frac{\|X_i^{(c)}\|_\infty}{\mathcal{P}(X_i^{(c)})} \epsilon_N + 2\sum_{i=1}^{d} \int_{\{x_i \geq \frac{1}{\sqrt{\epsilon_N}}\} \cap \mathbb{R}_+^d} x_i \pi(\mathrm{d}x_1, \dots, \mathrm{d}x_d) + 4\sqrt{\epsilon_N}$$

and note that $\eta_N \to 0$ as $N \to \infty$. Then by Doob's martingale inequality

$$P^W(\|\Gamma - M\| \geq \eta_N^{1/2}) \leq \eta_N^{-1/2} \sum_{i=1}^{d+K} E^W[|\Lambda^{(i)} - M_T^{(i)}|] \leq (d+K)\eta_N^{1/2}. \tag{5.4.19}$$

It follows that

$$\begin{aligned}
\big|E^W[G(\Gamma) - G(M)]\big| &\leq 2(d+K)\|G\|_\infty \eta_N^{1/2} + E^W\big[\big|G(\Gamma) - G(M)\big| \mathbb{1}_{\{\|\Gamma - M\| < \eta_N^{1/2}\}}\big] \\
&\leq 2(d+K)\|G\|_\infty \eta_N^{1/2} + f_e(\eta_N^{1/2}).
\end{aligned}$$

Note that by (4.5.1) in Remark 4.5.2 for N sufficiently large,

$$E^W[\alpha \wedge (\beta\sqrt{m^{(D)}(M)})] \geq E^W[\alpha \wedge (\beta\sqrt{m^{(D-2)}(\Gamma)})] - \alpha\eta_N^{1/2}.$$

As $\mathbb{P}^{(N)} \in \underline{\mathcal{M}}_{\pi^{(N)},\mathcal{I}}$ is arbitrary,

$$\begin{aligned}&\underline{P}_{\pi^{(N)},\mathcal{I}}\big(G - \alpha \wedge (\beta\sqrt{m^{(D)}})\big)\\ \leq\, &\underline{P}_{\pi,\mathcal{I}}\big(G - \alpha \wedge (\beta\sqrt{m^{(D-2)}})\big) - \Big((d+K)\|G\|_\infty\eta_N^{1/2} + f_e(\eta_N^{1/2}) + \alpha\eta_N^{1/2}\Big).\end{aligned}$$

Therefore, we can conclude that

$$\limsup_{N\to\infty} \underline{P}_{\pi^{(N)},\mathcal{I}}\big(G - \alpha \wedge (\beta\sqrt{m^{(D)}})\big) \leq \underline{P}_{\pi,\mathcal{I}}\big(G - \alpha \wedge (\beta\sqrt{m^{(D-2)}})\big), \qquad \text{as required.}$$

□

5.4.5 Proof of Theorem 5.3.17

From Theorem 5.3.6, we know that

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \geq \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G).$$

We also have the observation that $\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) \leq \widetilde{P}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G)$. Then to establish Theorem 5.3.17, it suffices to show that

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \leq \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G). \tag{5.4.20}$$

This follows as a special case ($\alpha = \beta = 0$) of the following crucial lemma which also be used to prove Theorem 5.3.21 below.

Lemma 5.4.3. Let $\mathfrak{P}$ be a measurable subset of $\mathcal{I}$, $\mathcal{X}$ be given by (5.3.13) and $\mathcal{P}$ be such that, for any $\eta > 0$, $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta} \neq \varnothing$, where $\vec{\mu}$ is defined via (3.3.3). Then for any uniformly continuous and bounded G and $\alpha, \beta \geq 0$

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\Big(G - \alpha \wedge (\beta\sqrt{m^{(D)}})\Big) \leq \widetilde{P}_{\vec{\mu},\mathfrak{P}}\Big(G - \alpha \wedge (\beta\sqrt{m^{(D-10)}})\Big).$$

where $m^{(D)}$ is defined in Definition 4.5.1.

Proof. Define

$$\begin{aligned}\mathfrak{G}_N([-M,M]) := \Big\{ f \in \mathcal{C}(\mathbb{R}_+^d, \mathbb{R}) \,:\, \sup_{\vec{x}\neq\vec{y}} \frac{|f(\vec{x}) - f(\vec{y})|}{|\vec{x}-\vec{y}|} \leq N,\ \|f\|_\infty \leq N,\\ \text{and } f(x_1,\ldots,x_d) = f(x_1 \wedge N^2, \ldots, x_d \wedge N^2)\ \forall (x_1,\ldots,x_d) \in \mathbb{R}_+^d \Big\}\end{aligned}$$

and let $\mathfrak{G}_N([-M,M]) = \cup_{N>0}\mathfrak{G}_N([-M,M])$.

Let $\mathcal{Z}_M = \{f(\mathbb{S}^{(i)}_{T_n}) : f \in \mathfrak{G}_M(\mathbb{R}),\, i = 1,\dots,d\}$ and $\mathcal{Y}_M = \{f(\mathbb{S}^{(i)}_{T_j}) : f \in \mathfrak{G}_M(\mathbb{R}),\, i = 1,\dots,d,\; j = 1,\dots,n-1\}$. We also write

$$\mathcal{Z} = \bigcup_{M\geq 0} \mathcal{Z}_M \;\text{ and }\; \mathcal{Y} = \bigcup_{M\geq 0} \mathcal{Y}_M.$$

Notice that given any $f \in C_b(\mathbb{R}_+,\mathbb{R})$, $\epsilon > 0$ and a measure μ on $\mathbb{R}_+$ which has finite first moment, there is some $u : \mathbb{R}_+ \to \mathbb{R}$ taking the form $a_0 + \sum_{i=1}^n a_i(s-K_i)^+$ such that $u \geq f$ and $\int (u-f)d\mu < \epsilon$. It follows that

$$\begin{aligned}
&\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\Big(G - \alpha \wedge (\beta\sqrt{m^{(D)}})\Big)\\
=&\widetilde{V}_{\mathcal{Z}\cup\mathcal{Y},\mathcal{P},\mathfrak{P}}\Big(G - \alpha \wedge (\beta\sqrt{m^{(D)}})\Big) && (5.4.21)\\
=&\inf_{X\in\mathrm{Lin}(\mathcal{Z}\cup\mathcal{Y}),\,N\geq 0}\Big\{\mathbf{V}_{\mathcal{I}}\Big(G - X - \alpha\wedge(\beta\sqrt{m^{(D)}}) - N\lambda_{\mathfrak{P}}\Big) + \mathcal{P}(X)\Big\} && (5.4.22)\\
\leq&\inf_{X\in\mathrm{Lin}(\mathcal{Z}\cup\mathcal{Y}),\,N\geq 0}\Big\{\underline{\mathbf{P}}_{\mathcal{I}}\Big(G - X - \alpha\wedge(\beta\sqrt{m^{(D-4)}}) - N\lambda_{\mathfrak{P}}\Big) + \mathcal{P}(X)\Big\} && (5.4.23)\\
=&\inf_{Y\in\mathrm{Lin}(\mathcal{Y}),\,N\geq 0}\;\inf_{M\geq 0}\;\inf_{Z\in\mathrm{Lin}(\mathcal{Z}_M)}\Big\{\underline{\mathbf{P}}_{\mathcal{I}}\Big(G - Y - Z - \alpha\wedge(\beta\sqrt{m^{(D-4)}}) - N\lambda_{\mathfrak{P}}\Big) + \mathcal{P}(Y+Z)\Big\}\\
& && (5.4.24)\\
\leq&\inf_{Y\in\mathrm{Lin}(\mathcal{Y})}\;\inf_{M\geq 0,\,N\geq 0}\;\lim_{L\to\infty}\;\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/L}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[G - \alpha\wedge(\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)]\\
& && (5.4.25)\\
\leq&\inf_{Y\in\mathrm{Lin}(\mathcal{Y})}\;\inf_{M\geq 0,\,N\geq 0}\;\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/M}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[G - \alpha\wedge(\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)].\\
& && (5.4.26)
\end{aligned}$$

where the equality between (5.4.21) and (5.4.22) follows from Remark 5.3.4, the inequality between (5.4.22) and (5.4.23) is justified by Theorem 5.5.1. Finally the inequality between (5.4.24) and (5.4.25) is given by Lemma 5.4.1. To justify this, we first observe that $\mathfrak{G}_1(\mathbb{R})$ is a convex and compact subset of $\mathcal{C}(\mathbb{R}_+,\mathbb{R})$. Then, since $\mathrm{Lin}_1(\mathcal{Z}_M) = \mathcal{Z}_M$ for any M, $\mathrm{Lin}_1(\mathcal{Z}_M)$ satisfies Assumption 5.3.5. Therefore, by keeping Y and N fixed and applying Lemma 5.4.1 to

$$\inf_{Z\in\mathrm{Lin}(\mathcal{Z}_M)}\Big\{\underline{\mathbf{P}}_{\mathcal{I}}(G - \alpha\wedge(\beta\sqrt{m^{(D-4)}}) - Y - Z - N\lambda_{\mathfrak{P}}) + \mathcal{P}(Y) + \mathcal{P}(Z)\Big\},$$

we establish the inequality.

For any $\mathbb{P} \in \underline{\mathcal{M}}^{1/M}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}$, let $\epsilon_{\mathbb{P}} = \max\{d_p(\mu^{(i)}_n, \mathcal{L}_{\mathbb{P}}(\mathbb{S}^{(i)}_{T_n})) : 1 \leq i \leq d\}$. Since d_p, the Lévy–Prokhorov's metric on probability measures on $\mathbb{R}^d_+$, is given by

$$d_p(\mu,\nu) := \sup_{f\in\mathfrak{G}^b_1(\mathbb{R}^d_+)}\Big|\int f\mathrm{d}\nu - \int f\mathrm{d}\mu\Big|,$$

where $\mathfrak{G}_1^b(\mathbb{R}_+^d) := \{f \in C(\mathbb{R}_+^d, \mathbb{R}) : \|f\| \leq 1 \text{ and } |f(\vec{x}) - f(\vec{y})| \leq |\vec{x} - \vec{y}| \ \forall \vec{x} \neq \vec{y}\}$ (for more details, see Bogachev [11], Chapter 8, Theorem 8.3.2), we can pick $g \in \mathfrak{G}_1^b(\mathbb{R}_+)$ such that

$$\Big| \int_{\mathbb{R}_+} g(x)\mu_n^{(i)}(\mathrm{d}x) - \mathbb{E}_{\mathbb{P}}[g(\mathbb{S}_{T_n}^{(i)})] \Big| > \epsilon_{\mathbb{P}}/2 \quad \text{for some } i = 1, \dots, d,$$

and define $\hat{g} \in \mathfrak{G}_M^b(\mathbb{R}_+)$ via $\hat{g}(x) = Mg(x \wedge M^2)$. Then,

$$\begin{aligned} &\Big| \int_{\mathbb{R}_+} \hat{g}(x)\mu_n^{(i)}(\mathrm{d}x) - \mathbb{E}_{\mathbb{P}}[\hat{g}(\mathbb{S}_{T_n}^{(i)})] \Big| \qquad (5.4.27) \\ \geq & M\Big| \int g\mathrm{d}\mu_n^{(i)} - \mathbb{E}_{\mathbb{P}}[g(\mathbb{S}_{T_n}^{(i)})] \Big| - (M+1)\mu_n^{(i)}(\{|x| > M^2\}) - (M+1)\mathbb{P}(|\mathbb{S}_{T_n}^{(i)}| > M^2) \\ \geq & M\epsilon_{\mathbb{P}}/2 - \frac{2(M+1)}{M^2}. \end{aligned}$$

By definition of $\underline{\mathcal{M}}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}^{1/M}$,

$$\Big| \int_{\mathbb{R}_+} \hat{g}(x)\mu_n^{(i)}(\mathrm{d}x) - \mathbb{E}_{\mathbb{P}}[\hat{g}(\mathbb{S}_{T_n}^{(i)})] \Big| \leq 1/M.$$

Hence, $\epsilon_{\mathbb{P}} \leq 1/M^2 + 2(M+1)/M^3 \leq 2/M$ when M is sufficiently large. It follows that $\mathbb{P} \in \underline{\mathcal{M}}_{\mu_n,\mathcal{I},2/M}$ and hence $\underline{\mathcal{M}}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}^{1/M} \subseteq \underline{\mathcal{M}}_{\mu_n,\mathcal{I},2/M}$ when M is sufficiently large. In consequence

$$\begin{aligned} &\inf_{M\geq 0} \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}^{1/M}} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)] \\ \leq & \inf_{M\geq 0} \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mu_n,\mathcal{I},2/M}} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)]. \end{aligned}$$

For every $M \in \mathbb{N}_+$, take $\mathbb{P}^{(M)} \in \underline{\mathcal{M}}_{\mu_n,\mathcal{I},2/M}$ such that

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}^{(M)}}\Big[G - \alpha \wedge (\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)\Big] \\ \geq & \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mu_n,\mathcal{I},2/M}} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)] - \frac{1}{M}. \end{aligned}$$

Let $\pi_n^{(M)}$ be the law of $(\mathbb{S}_{T_n}^{(1)}, \dots, \mathbb{S}_{T_n}^{(d)})$ under $\mathbb{P}^{(M)}$. It is a probability measure on $\mathbb{R}_+^d$ with mean 1. It follows that the family $\{\pi_n^{(M)}\}_{M\geq 1}$ is tight. By Prokhorov theorem, there exists a subsequence $\{\pi_n^{(M_k)}\}$ converging to some π_n. Note that the marginal distributions of π_n are $\mu_n^{(i)}$'s. By Lemma 5.4.2, it follows that

$$\begin{aligned} &\lim_{M\to\infty} \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mu_n,\mathcal{I},2/M}} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-6)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)] \\ \leq & \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mu_n,\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-8)}}) - Y - N\lambda_{\mathfrak{P}} + \mathcal{P}(Y)]. \end{aligned}$$

With the result above, we continue with (5.4.26):

$$
\begin{aligned}
\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) &\le \inf_{Y\in\mathrm{Lin}(\mathcal{Y})}\ \inf_{M\ge 0,\, N\ge 0}\ \sup_{\mathbb{P}\in\underline{\mathcal{M}}^{1/M}_{\mathcal{Z}_M,\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-6)}})-Y-N\lambda_{\mathfrak{P}}+\mathcal{P}(Y)] \\
&\le \inf_{Y\in\mathrm{Lin}(\mathcal{Y}),\, N\ge 0}\ \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mu_n,\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-8)}})-Y-N\lambda_{\mathfrak{P}}+\mathcal{P}(Y)] \\
&\le \inf_{M\ge 0}\ \inf_{Y\in\mathrm{Lin}(\mathcal{Y}_M),N\ge 0}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-8)}})-Y-N\lambda_{\mathfrak{P}}+\mathcal{P}(Y)] \\
&\le \inf_{M\ge 0}\ \lim_{N\to\infty}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}\cap\mathcal{M}^{1/N}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})] \\
&\le \inf_{M\ge 0}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}\cap\mathcal{M}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})].
\end{aligned}
$$

To justify the second last inequality, we first notice that $\mathcal{M}^{\eta}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}\neq\varnothing$ for any $M\in\mathbb{N}$ and $\eta>0$ since $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}\neq\varnothing$ for any $\eta>0$, and hence it follows from Lemma 5.4.1, with $\mathcal{M}_s=\mathcal{M}_{\mu_n,\mathcal{I}}$, that

$$
\begin{aligned}
&\inf_{M\ge 0}\ \inf_{Y\in\mathrm{Lin}(\mathcal{Y}_M),N\ge 0}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-8)}})-Y-N\lambda_{\mathfrak{P}}+\mathcal{P}(Y)] \\
&\le \inf_{M\ge 0}\ \lim_{N\to\infty}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}\cap\mathcal{M}^{1/N}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})] \\
&\le \inf_{M\ge 0}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}\cap\mathcal{M}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})].
\end{aligned}
$$

Next we are going to argue that $\mathcal{M}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_{\mu_n,\mathcal{I}}\subseteq\mathcal{M}_{\vec{\mu},\mathfrak{P},2/M}$ when M is large enough. Fix $\mathbb{P}\in\mathcal{M}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_{\mu_n,\mathcal{I}}$ and let $\tilde{\epsilon}_{\mathbb{P}}=\max\{d_p(\mu^{(i)}_j,\mathcal{L}_{\mathbb{P}}(\mathbb{S}^{(i)}_{T_j})):\ i\le d,\ j\le n\}$. We can pick $g\in\mathfrak{G}^b_1(\mathbb{R}_+)$ such that

$$
\Big|\int_{\mathbb{R}_+}\hat{g}(x)\mu^{(i)}_j(\mathrm{d}x)-\mathbb{E}_{\mathbb{P}}[\hat{g}(\mathbb{S}^{(i)}_{T_j})]\Big|>\tilde{\epsilon}_{\mathbb{P}}/2 \quad \text{for some } i\le d,\ j\le n-1.
$$

By following the same argument as above, we can show that $\tilde{\epsilon}_{\mathbb{P}}\le 1/M^2+2(M+1)/M^3$. Hence, $\mathcal{M}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}\cap\mathcal{M}_{\mu_n,\mathcal{I}}\subseteq\mathcal{M}_{\vec{\mu},\mathfrak{P},2/M}$ when M is sufficiently large. Therefore, we have

$$
\begin{aligned}
\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) &\le \inf_{M\ge 0}\ \sup_{\mathbb{P}\in\mathcal{M}_{\mu_n,\mathcal{I}}\cap\mathcal{M}^{1/M}_{\mathcal{Y}_M,\mathcal{P},\mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})] \\
&\le \lim_{N\to\infty}\ \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},2/N}} \mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})]=\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G-\alpha\wedge(\beta\sqrt{m^{(D-10)}})).
\end{aligned}
$$

□

5.4.6 Proof of Theorem 5.3.21

We first make two simple observations.

Remark 5.4.4. If $\mathfrak{P}$ is a non-empty closed subset of Ω with respect to sup norm, then

$$\mathfrak{P} = \bigcap_{\epsilon>0} \mathfrak{P}^\epsilon = \bigcap_{\epsilon>0} \overline{\mathfrak{P}^\epsilon},$$

where $\overline{\mathfrak{P}^\epsilon}$ is the closure of $\mathfrak{P}^\epsilon$.

Lemma 5.4.5. If $\mathfrak{P}$ is time invariant, then for every $\epsilon > 0$ $\mathfrak{P}^\epsilon$ is also time invariant.

Proof. Fix $\epsilon > 0$, $S \in \mathfrak{P}^\epsilon$ and a nondecreasing continuous function $f : [0, T_n] \to [0, T_n]$ such that $f(0) = 0$ and $f(T_i) = T_i$ for any $i = 1, \dots, n$. By definition, there exist $S^{(N)} \in \mathfrak{P}$ such that

$$\|S^{(N)} - S\| \leq \epsilon + \frac{1}{N}$$

Now write $\tilde{S}_t = S_{f(t)}$ and $\tilde{S}^{(N)}_t = S^{(N)}_{f(t)}$. Note that $\tilde{S}^{(N)} \in \mathfrak{P}$ as $\mathfrak{P}$ is time invariant. Then it is clear that

$$\|\tilde{S}^{(N)} - \tilde{S}\| = \|S^{(N)} - S\| \leq \epsilon + \frac{1}{N},$$

which implies that $\tilde{S} \in \mathfrak{P}^\epsilon$. Since $S \in \mathfrak{P}^\epsilon$ and f are arbitrary, we can therefore conclude that $\mathfrak{P}^\epsilon$ is time invariant. □

We now proceed with the proof the Theorem 5.3.21. Note that the inequalities $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \geq V_{\vec{\mu},\mathfrak{P}}(G) \geq P_{\vec{\mu},\mathfrak{P}}(G)$ hold in general. In addition, according to Theorem 5.3.17, $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G)$. Therefore, we only need to show $\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) = P_{\vec{\mu},\mathfrak{P}}(G)$. Our proof of this equality is divided into six steps. First, using Lemma 5.4.3, we argue that it suffices to consider measures with "good control" on the expectation of $m^{(D)}(\mathbb{S})$. Next, we perform three time changes within each trading period $[T_i, T_{i+1}]$. The resulting time change of $\mathbb{S}$, denoted $\ddot{\mathbb{S}}$, allows for a "good control" over its quadratic variation process. At the same time, we keep $G(\mathbb{S})$ and $G(\ddot{\mathbb{S}})$ "close" and given a measure $\mathbb{P} \in \mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}$ with good control on $\mathbb{E}_\mathbb{P}[m^{(D)}(\mathbb{S})]$, since $\mathfrak{P}^\eta$ is time invariant, the law of the time-changed price process $\ddot{\mathbb{S}}$ remains an element of $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}$. Then, in Step 5, given a sequence of models with improved calibration precisions, we show tightness of quadratic variation process of the time-changed price process $\ddot{\mathbb{S}}$

under these measures. This then leads to tightness of images measures via $\ddot{\mathbb{S}}$. In Step 6, we deduce the duality $\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) = P_{\vec{\mu},\mathfrak{P}}(G)$ from tightness and conclude.

Step 1: Reducing to measures $\mathbb{P}$ with good control on $\mathbb{E}_{\mathbb{P}}[m^{(D)}(\mathbb{S})]$.

Let G be bounded and satisfy Assumption 5.3.18. Choose $\kappa \in \mathbb{R}_+$ such that $\|G\| \leq \kappa$ and let $f_e : \mathbb{R}_+^{d+K} \to \mathbb{R}_+$ be the modulus of continuity of G, i.e.

$$|G(\omega) - G(\upsilon)| \leq f_e(|\omega - \upsilon|) \quad \text{for any } \omega, \upsilon \in \Omega$$

with $\lim_{x \to 0} f_e(x) = 0$. Fix $D \in \mathbb{N}$. Consider a random variable

$$\begin{aligned} X_D(\mathbb{S}) =& \sqrt{\sum_{j=1}^{m^{(D)}(\mathbb{S}) \wedge (2^{6D}\kappa^2)} \sum_{i=1}^{d+K} |\mathbb{S}^{(i)}_{\tau_j^{(D)}(\mathbb{S})} - \mathbb{S}^{(i)}_{\tau_{j-1}^{(D)}(\mathbb{S})}|^2} \\ \geq& 2^{-D}\sqrt{m^{(D)}(\mathbb{S}) \wedge (2^{6D}\kappa^2) - 1} \\ \geq& \Big(2^{-D}(\sqrt{m^{(D)}(\mathbb{S}) \wedge 2^{6D}\kappa^2} - 1)\Big) = \kappa 2^{2D} \wedge \frac{\sqrt{m^{(D)}(\mathbb{S})}}{2^D} - 2^{-D} \end{aligned}$$

where $\tau_i^{(D)}$'s and $m^{(D)}$ are defined in Definition 4.5.1.

It follows from the proof of Lemma 5.4 in Dolinsky and Soner [41] that there exists a $\gamma \in \mathcal{A}$ such that

$$\int_0^{\tau^{(D)}_{m^{(D)}(\mathbb{S}) \wedge (2^{6D}\kappa^2)}} \gamma_u \mathrm{d}\mathbb{S}_u + 3(d+K) \max_{0 \leq j \leq \left(m^{(D)}(\mathbb{S}) \wedge (2^{6D}\kappa^2)\right)} |\mathbb{S}_{\tau_j^{(D)}}| \geq X_D(\mathbb{S})$$

Hence, $V_{\vec{\mu},\mathfrak{P}}(X_D(\mathbb{S})) \leq 3(d+K) V_{\vec{\mu},\mathfrak{P}}\Big(\|\mathbb{S}\| \wedge (\kappa^2 2^{5D} + 1)\Big)$, and therefore

$$0 \leq V_{\vec{\mu},\mathfrak{P}}(X_D(\mathbb{S})) \leq 3(d+K) P_{\mu_n,\mathcal{I}}(\|\mathbb{S}\|) < \infty.$$

since $V_{\vec{\mu},\mathfrak{P}}\Big(\|\mathbb{S}\| \wedge (\kappa^2 2^{5D} + 1)\Big) \leq V_{\mu_n,\mathcal{I}}\Big(\|\mathbb{S}\| \wedge (\kappa^2 2^{5D} + 1)\Big) = P_{\mu_n,\mathcal{I}}\Big(\|\mathbb{S}\| \wedge (\kappa^2 2^{5D} + 1)\Big)$, where the equality is due to Theorem 5.3.17.

It follows that from the sub-linearity of $\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(\cdot)$ and the estimate above

$$\begin{aligned} \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G(\mathbb{S})) \leq& \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\Big(G(\mathbb{S}) - \kappa 2^D \wedge \frac{X_D(\mathbb{S})}{2^D}\Big) + \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(X_D(\mathbb{S})/2^D) \\ \leq& \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\Big(G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D)}(\mathbb{S})}}{2^{2D}}\Big) + c_2/2^D \\ \leq& \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}\Big(G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D)}(\mathbb{S})}}{2^{2D}}\Big) + c_2/2^D \\ \leq& \widetilde{P}_{\vec{\mu},\mathfrak{P}}\Big(G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big) + c_2/2^D \\ =& \lim_{N \to \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}\Big[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big] + c_2/2^D. \end{aligned}$$

where c_2 is a constant and the last inequality follows from Lemma 5.4.3.

Next we denote $\widetilde{\mathcal{M}}_{\mathcal{I}}$ the set of $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$ such that

$$\mathbb{E}_{\mathbb{P}}\Big[\kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big] \leq 2\kappa + 2. \tag{5.4.28}$$

We notice that if $\mathbb{P} \in \mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}$ such that $\mathbb{P} \notin \widetilde{\mathcal{M}}_{\mathcal{I}}$, then

$$\mathbb{E}_{\mathbb{P}}\Big[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big] < \kappa - 2\kappa - 2 = -\kappa - 2.$$

While for N sufficiently large,

$$\begin{aligned} &\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}\Big[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big] \\ &\geq \widetilde{P}_{\vec{\mu},\mathfrak{P}}\Big(G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big) \\ &\geq V^{(p)}_{\vec{\mu},\mathfrak{P}}(G(\mathbb{S})) - c_2/2^D \geq -\kappa - 1 \quad \text{ for a large } D\ . \end{aligned}$$

It follows that

$$\begin{aligned} &\lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}\Big[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big] \\ = &\lim_{N\to\infty} \sup_{\mathbb{P}\in\widetilde{\mathcal{M}}_{\mathcal{I}}\cap\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}\Big[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\Big]. \end{aligned}$$

In particular, $\widetilde{\mathcal{M}}_{\mathcal{I}} \cap \mathcal{M}_{\vec{\mu},\mathfrak{P},1/N} \neq \varnothing$ for N large enough.

Step 2: First time change: "squeezing paths and adding constant paths".

Now for every $N \in \mathbb{N}$, take $\mathbb{P}^{(N)} \in \widetilde{\mathcal{M}}_{\mathcal{I}} \cap \mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}$ such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] \geq \sup_{\mathbb{P}\in\tilde{\mathcal{M}}_{\mathcal{I}}\cap\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - 1/N.$$

Define an increasing function $f : [0, T_n] \mapsto [0, T_n]$ by

$$f(t) = \sum_{i=1}^{n} \Big(T_i \wedge \Big(T_{i-1} + \frac{(T_i - T_{i-1})(t - T_{i-1})}{T_i - T_{i-1} - 1/D}\Big)\Big) \mathbb{1}_{\{T_{i-1}<t\leq T_i\}}$$

and then a process $(\tilde{\mathbb{S}}_t)_{t\in[0,T_n]}$ by a time change of $\mathbb{S}$ via f, i.e., $\tilde{\mathbb{S}}_t = \mathbb{S}_{f(t)}$. It follows from (5.3.15) that

$$\begin{aligned} &|G(\mathbb{S}) - G(\tilde{\mathbb{S}})| \\ &\leq |G(\mathbb{S}) - G(F^{(D)}(\mathbb{S}))| + |G(\tilde{\mathbb{S}}) - G(F^{(D)}(\tilde{\mathbb{S}}))| + |G(F^{(D)}(\mathbb{S})) - G(F^{(D)}(\tilde{\mathbb{S}}))| \\ &\leq 2f_e(2^{-D+11}) + \frac{2nL\|\mathbb{S}\|}{D}. \end{aligned} \tag{5.4.29}$$

In addition, $\mathbb{S}_{T_i} = \tilde{\mathbb{S}}_{T_i}$ $\forall i \le n$. In particular, $\mathcal{L}_{\mathbb{P}^{(N)}}(\mathbb{S}_{T_i}) = \mathcal{L}_{\mathbb{P}^{(N)}}(\tilde{\mathbb{S}}_{T_i})$ $\forall i \le n$ and further $\mathbb{P}^{(N)} \circ (\tilde{\mathbb{S}}_t)^{-1} \in \mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}$ as $\mathfrak{P}^{1/N}$ is time invariant, by Lemma 5.4.5.

Step 3: Second time change: introducing lower bound on time step.

For ease of notation it is helpful to rename the elements of the set

$$\{\tau_j^{(N)} \,:\, j \le m_j^{(N)}\} \cup \{T_i \,:\, i = 1, \dots, n\}$$

as follows. We define a sequence of stopping times $\tau_{i,j}^{(N)} : \Omega \to [T_{i-1}, T_i]$ and $m_i^{(N)} : \Omega \to \mathbb{N}_+$ in a recursive manner: set $m_0^{(N)}(\mathbb{S}) = 0$ and $\tau_{0,-1}^{(N)}(\mathbb{S}) = 0$, and $\forall i = 1, \dots, n$, set $\tau_{i,0}^{(N)}(\mathbb{S}) = T_{i-1}$ and let

$$\tau_{i,1}^{(N)}(\mathbb{S}) = \inf\Big\{t \ge T_{i-1} : |\mathbb{S}_t - \mathbb{S}_{\tau^{(N)}_{i-1,m_{i-1}^{(N)}(\mathbb{S})-1}(\mathbb{S})}| = \tfrac{1}{2^N}\} \wedge T_i,$$

$$\tau_{i,k}^{(N)}(\mathbb{S}) = \inf\Big\{t \ge \tau_{i,k-1}(\mathbb{S}) : |\mathbb{S}_t - \mathbb{S}_{\tau^{(N)}_{i,k-1}(\mathbb{S})}| = \tfrac{1}{2^N}\} \wedge T_i,$$

$$m_i^{(N)}(\mathbb{S}) = m_{i-1}^{(N)}(\mathbb{S}) + \min\{k \in \mathbb{N} : \tau_{i,k}^{(N)}(\mathbb{S}) = T_i\}.$$

It follows that for any $S \in \mathcal{I}$

$$m^{(D-10)}(S) \le m_n^{(D-10)}(S) \le m^{(D-10)}(S) + n - 1. \tag{5.4.30}$$

Set $\Theta = 2\lceil \kappa^2 2^{6D}\rceil + n$ and $\delta = 1/(4D\Theta^2)$. We now define a sequence of stopping times $\sigma_{i,j} : \Omega \to [0, T_n]$. Fix any $S \in \Omega$ as follows. Firstly, set $\sigma_{i,0}(S) = T_{i-1}$ and $\sigma_{i,\Theta+1}(S) = T_i$. Then, for $j \le \Theta$, $\sigma_{i,j}(S) = \Big(\tau_{i,j}^{(D-10)}(S) + \delta j\Big) \wedge \big(T_i - 1/(2D)\big)$ if $j < m_i^{(D-10)}(S)$, and $\sigma_{i,j}(S) = T_i - 1/(2D)$ otherwise.

Then it follows from the definition that $T_{i-1} = \sigma_{i,0}(\mathbb{S}) \le \sigma_{i,1}(\mathbb{S}) \le \ldots \le \sigma_{i,\Theta}(\mathbb{S}) < \sigma_{i,\Theta+1}(\mathbb{S}) = T_i$. We also note that since $\tilde{\mathbb{S}}$ is always constant on $[T_i - 1/D, T_i]$, $\tau_{i,j}^{(D-10)}(\tilde{\mathbb{S}}) \le T_i - 1/D$ and hence for $j \le \Theta \wedge (m_i^{(D-10)}(\tilde{\mathbb{S}}) - 1)$

$$\sigma_{i,j}(\tilde{\mathbb{S}}) \le \tau^{(D-10)}_{i,m_i^{(D-10)}-1}(\tilde{\mathbb{S}}) + \delta(\Theta - 1) \le T_i - \frac{1}{D} + \frac{1}{4D\Theta} < T_i - \frac{1}{2D}.$$

Therefore, $\forall j = 1, \dots, \Big(\Theta \wedge (m_i^{(D-10)}(\tilde{\mathbb{S}}) - 1)\Big)$

$$\sigma_{i,j}(\tilde{\mathbb{S}}) - \sigma_{i,j-1}(\tilde{\mathbb{S}}) = \delta + \big(\tau_{i,j}^{(D-10)}(\tilde{\mathbb{S}}) - \tau_{i,j-1}^{(D-10)}(\tilde{\mathbb{S}})\big) \ge \delta. \tag{5.4.31}$$

Define a process $\check{\mathbb{S}}$ by

$$\check{\mathbb{S}}_t = \sum_{i=0}^{n-1} \sum_{j=0}^{\Theta-1} \Big\{\tilde{\mathbb{S}}_{\tau_{i,j}^{(D-10)}(\tilde{\mathbb{S}}) + (t - \sigma_{i,j}(\tilde{\mathbb{S}}) - \delta)^+} \mathbb{1}_{[\sigma_{i,j}(\tilde{\mathbb{S}}), \sigma_{i,j+1}(\tilde{\mathbb{S}}))}(t)$$
$$+ \tilde{\mathbb{S}}_{\big(\tau_{i-1,\Theta}^{(D-10)}(\tilde{\mathbb{S}}) + \frac{1}{T_i - t} - \frac{1}{T_i - \sigma_{i,\Theta}(\tilde{\mathbb{S}})}\big) \wedge T_i} \mathbb{1}_{[\sigma_{i,\Theta}(\tilde{\mathbb{S}}), T_i]}(t)\Big\}.$$

Equivalently, $\check{\mathbb{S}}$ can be obtained by time changing $\tilde{\mathbb{S}}$ via an increasing and continuous process $g : [0, T_n] \times \mathcal{I} \to [0, T_n]$, defined by

$$\begin{aligned} g_t(S) = \sum_{i=0}^{n-1} \sum_{j=0}^{\Theta-1} \Big\{ \Big(& \tau_{i,j}^{(D-10)}(S) + (t - \sigma_{i,j}(S) - \delta)^+ \Big) \mathbb{1}_{[\sigma_{i,j}(S), \sigma_{i,j+1}(S))}(t) \\ & + T_i \wedge \Big(\tau_{i,\Theta-1}^{(D-10)}(S) + (\sigma_{i,\Theta}(S) - \sigma_{i,\Theta-1}(S) - \delta)^+ \\ & \qquad + \frac{1}{T_i - t} - \frac{1}{T_i - \sigma_{i,\Theta}(\tilde{\mathbb{S}})} \Big) \mathbb{1}_{[\sigma_{i,\Theta}(S), T_i]}(t) \Big\}. \end{aligned}$$

In particular, it follows from (5.4.31) that

$$\begin{aligned} g_t(\tilde{\mathbb{S}}) = \sum_{i=0}^{n-1} \sum_{j=0}^{\Theta-1} \Big\{ \Big(& \tau_{i,j}^{(D-10)}(\tilde{\mathbb{S}}) + (t - \sigma_{i,j}(\tilde{\mathbb{S}}) - \delta)^+ \Big) \mathbb{1}_{[\sigma_{i,j}(\tilde{\mathbb{S}}), \sigma_{i,j+1}(\tilde{\mathbb{S}}))}(t) \\ & + T_i \wedge \Big(\sigma_{i,\Theta}(\tilde{\mathbb{S}}) + \frac{1}{T_i - t} - \frac{1}{T_i - \sigma_{i,\Theta}(\tilde{\mathbb{S}})} \Big) \mathbb{1}_{[\sigma_{i,\Theta}(S), T_i]}(t) \Big\}. \end{aligned}$$

Furthermore, g is adapted to $\mathbb{F}$ and hence predictable with respect to $\mathbb{F}^{\mathbb{P}^{(N)}}$ – the usual augmentation of $\mathbb{F}$ (since g is continuous). Therefore, it is clear that $\check{\mathbb{S}}$ is a local martingale with respect to $\mathbb{F}^{\mathbb{P}^{(N)}}$ under $\mathbb{P}^{(N)}$. Moreover, $\tilde{\mathbb{S}}_{T_i} = \check{\mathbb{S}}_{T_i} = \mathbb{S}_{T_i}$ for any $i \leq n$. This implies that $\check{\mathbb{S}}$ is a martingale with respect to $\mathbb{F}^{\mathbb{P}^{(N)}}$ under $\mathbb{P}^{(N)}$ and further $\mathbb{P}^{(N)} \circ (\check{\mathbb{S}}_t)^{-1} \in \mathcal{M}_{\vec{\mu}, \mathfrak{P}, 1/N}$.

Observe that for any $S \in \Omega$ such that $m_n^{(D-10)}(S) \leq \Theta$ it follows from (5.3.15) – the time continuity property of G

$$\begin{aligned} & |G(\tilde{\mathbb{S}}(S)) - G(\check{\mathbb{S}}(S))| \\ \leq & |G(\tilde{\mathbb{S}}(S)) - G(F^{(D-10)}(\tilde{\mathbb{S}}(S)))| + |G(\check{\mathbb{S}}(S)) - G(F^{(D-10)}(\check{\mathbb{S}}(S)))| \\ & + |G(F^{(D-10)}(\mathbb{S})(S)) - G(F^{(D-10)}(\check{\mathbb{S}}(S)))| \\ \leq & 2 f_e(2^{-D+11}) + 2nL \|\tilde{\mathbb{S}}(S)\| \Theta \delta \leq 2 f_e(2^{-D+11}) + 2nL \|\mathbb{S}(S)\| / D, \end{aligned} \tag{5.4.32}$$

when D is sufficiently large, where $F^{(D-10)}$ is defined in (4.5.2). From (5.4.28), Markov inequality gives

$$\mathbb{P}^{(N)}(\{S \in \mathcal{I} : m^{(D-10)}(S) \geq \Theta - n + 2\}) \leq \frac{2\kappa + 2}{\kappa D}. \tag{5.4.33}$$

and hence by (5.4.30)

$$\mathbb{P}^{(N)}(\{S \in \mathcal{I} : m_n^{(D-10)}(S) \geq \Theta + 1\}) \leq \frac{2\kappa + 2}{\kappa D}. \tag{5.4.34}$$

Furthermore, by (5.4.32) and (5.4.34)

$$\begin{aligned}&\left|\mathbb{E}_{\mathbb{P}^{(N)}}[G(\tilde{\mathbb{S}})] - \mathbb{E}_{\mathbb{P}^{(N)}}[G(\check{\mathbb{S}})]\right| \\ &\leq 2\kappa\mathbb{P}^{(N)}(m^{(D-10)}(\tilde{\mathbb{S}}) > \Theta) + 2f_e(2^{-D+11}) + 2nL\mathbb{E}_{\mathbb{P}^{(N)}}[\|\mathbb{S}\|]/D \\ &\leq \frac{4\kappa+4}{2^D} + 2f_e(2^{-D+11}) + 2nLV^{(p)}_{\mu_n,\mathcal{I}}(\|\mathbb{S}\|)/D.\end{aligned} \tag{5.4.35}$$

Step 4: Third time change: controlling increments of quadratic variation.

We say $\omega \in \mathcal{C}([0,T_n],\mathbb{R})$ admits quadratic variation if

$$\sum_{k=0}^{m^{(N)}(\omega)-1}\left(\omega_{\tau_k^{(N)}(\omega)} - \omega_{\tau_{k+1}^{(N)}(\omega)}\right)^2 \text{ converges to a limit as } N \to \infty \text{ for any } i \leq d+K.$$

We let $\langle\omega\rangle$ be that limit if ω admits quadratic variation and zero otherwise. In addition, for $S \in \Omega$, we say S admits quadratic variation if $S^{(i)}$ admits quadratic variation for any $i \leq d+K$.

It follows from Theorem 4.30.1 in Rogers and Williams [91] and its proof that for any $\mathbb{P} \in \mathcal{M}$, $\langle\mathbb{S}\rangle := \big(\langle\mathbb{S}^{(1)}\rangle, \ldots, \langle\mathbb{S}^{(d+K)}\rangle\big)$ agrees $\mathbb{P}$-a.s. with the classical definition of quadratic variation of $\mathbb{S}$ under $\mathbb{P}$.

Now Doob's inequality gives $\forall i \leq d$

$$\mathbb{E}_{\mathbb{P}^{(N)}}[\|\check{\mathbb{S}}^{(i)}\|^p] \leq \left(\frac{p}{p-1}\right)^p \int_{[0,\infty)} x^p \mu_n^{(i)}(\mathrm{d}x). \tag{5.4.36}$$

And, by BDG-inequalities, we know there exist constants $c_p, C_p \in (0,\infty)$ such that

$$c_p\mathbb{E}_{\mathbb{P}^{(N)}}\big[\langle\check{\mathbb{S}}^{(i)}\rangle_{T_n}^{p/2}\big] \leq \mathbb{E}_{\mathbb{P}^{(N)}}[\|\check{\mathbb{S}}^{(i)}\|^p] \leq C_p\mathbb{E}_{\mathbb{P}^{(N)}}\big[\langle\check{\mathbb{S}}^{(i)}\rangle_{T_n}^{p/2}\big]. \tag{5.4.37}$$

It follows that

$$\mathbb{E}_{\mathbb{P}^{(N)}}\Big[\sum_{i=1}^{d+K}\langle\check{\mathbb{S}}^{(i)}\rangle_{T_n}^{p/2}\Big] \leq K_1, \tag{5.4.38}$$

where $K_1 := \left(\frac{1}{c_p}\right)\left(\left(\frac{p}{p-1}\right)^p \sum_{i=1}^d \int_{[0,\infty)} x^p\mu_n^{(i)}(\mathrm{d}x) + K\kappa^p\right)$.

In the following we want to modify $\check{\mathbb{S}}$ on

$$\tilde{\mathcal{I}} := \{S \in \mathcal{I} \,:\, \check{\mathbb{S}}(S) \text{ admits quadratic variation }\} = \{S \in \mathcal{I} \,:\, S \text{ admits quadratic variation }\}$$

using time change technique to obtain another process $\ddot{\mathbb{S}}$, the law of which is in $\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}$. In fact, $\ddot{\mathbb{S}}$ is a time change of $\check{\mathbb{S}}$ on each interval $[\sigma_{i,j}(\tilde{\mathbb{S}}), \sigma_{i,j+1}(\tilde{\mathbb{S}}))$. Then by continuity of G, it follows that

$$|G(\check{\mathbb{S}}(S)) - G(\ddot{\mathbb{S}}(S))| \leq f_e(2^{-D+11}) \qquad \forall\, S \in \tilde{\mathcal{I}} \cap \{\tilde{S} \in \mathcal{I} \,:\, m_n^{(D-10)}(\tilde{\mathbb{S}}(\tilde{S})) \leq \Theta\}.$$

This, together with (5.4.34) and the fact that $\mathbb{P}(\tilde{\mathcal{I}}) = 1$ for any $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$, yields

$$\begin{aligned}\left|\mathbb{E}_{\mathbb{P}^{(N)}}[G(\check{\mathbb{S}}) - G(\ddot{\mathbb{S}})]\right| &\le f_e(2^{-D+11}) + 2\kappa\mathbb{P}^{(N)}(\{S \in \mathcal{I} : m_n^{(D-10)}(\tilde{\mathbb{S}}(S)) \ge \Theta + 1\}) \\ &\le f_e(2^{-D+11}) + \frac{4\kappa + 4}{D}.\end{aligned}$$

Hence, by (5.4.29) and (5.4.35),

$$\left|\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S}) - G(\ddot{\mathbb{S}})]\right| \le 5f_e(2^{-D+11}) + \frac{2nL\|\mathbb{S}\|}{D} + \frac{8\kappa + 8}{2^D} + \frac{2nLV^{(p)}_{\mu_n,\mathcal{I}}(\|\mathbb{S}\|)}{D}. \quad (5.4.39)$$

First, for every i, j, k, define $\rho^{(i,j,k)} : \Omega \to [0, T_n]$ by $\rho^{(i,j,k)}(S) = \sigma_{i,j}(\check{\mathbb{S}}(S)) + \delta(1 - 2^{-k+1})$ $\forall\, S \in \Omega$. Then, $\forall\, i = 1, \dots, n$, $j = 0, 1, \dots$, and $k = 1, 2, \dots$, consider change of time $\theta^{(i,j,k)} : \mathcal{I} \times [0, T_n] \to [0, T_n]$ defined as follows: if $S \in \tilde{\mathcal{I}}$, $\theta_t^{(i,j,k)}(S) = t$ $\forall\, t \le \rho^{(i,j,k)}(S)$ and for $t > \rho^{(i,j,k)}(S)$

$$\begin{aligned}\theta_t^{(i,j,k)}(S) = \inf\{u \ge \rho^{(i,j,k)}(S) : &\sum_{l=1}^{d+K} \left(\langle \check{\mathbb{S}}^{(l)}(S)\rangle_u - \langle \check{\mathbb{S}}^{(l)}(S)\rangle_{\rho^{(i,j,k)}}\right) \\ &> 2^k(t - \rho^{(i,j,k)})/\delta\} \wedge \rho^{(i,j,k+1)} \wedge \sigma_{i,j+1}(\check{\mathbb{S}}(S)),\end{aligned}$$

otherwise $\theta_t^{(i,j,k)}(S) = t$ on $[0, T_n]$.

Now, by considering $\ddot{\mathbb{S}}$ – a time change of $\check{\mathbb{S}}$ via $\theta^{(i,j,k)}$'s, defined by $\ddot{\mathbb{S}}_t := \check{\mathbb{S}}_{(\theta_t^{(i,j,k)}(\mathbb{S}))^{-1}}$ on $[\rho^{(i,j,k)}(\mathbb{S}), \rho^{(i,j,k+1)}(\mathbb{S}))$ $\forall i, j, k$, we see that for any i, j, k and any $S \in \tilde{\mathcal{I}}$ the quadratic variation of $\ddot{\mathbb{S}}(S)_t$ grows linearly at rate 2^k on $[\rho^{(i,j,k)}(S), \rho^{(i,j,k+1)}(S))$ with $\rho^{(i,j,k+1)}(S) - \rho^{(i,j,k)}(S) = 2^{-k}$ if $\sigma_{i,j+1}(\tilde{\mathbb{S}}(S)) - \sigma_{i,j}(\tilde{\mathbb{S}}(S)) > 0$ and 0 otherwise. It follows that $\ddot{\mathbb{S}}$ is a continuous process on $\tilde{\mathcal{I}}$ and furthermore

$$\sum_{l=1}^{d+K} \left(\langle \ddot{\mathbb{S}}^{(l)}(S)\rangle_t - \langle \ddot{\mathbb{S}}^{(l)}(S)\rangle_s\right) \le 2^k|t-s|/\delta \qquad \forall s, t \text{ s.t. } \sigma_{i,j}(\tilde{\mathbb{S}}(S)) \le s \le t \le \sigma_{i,j+1}(\tilde{\mathbb{S}}(S)),$$

whenever $S \in \tilde{\mathcal{I}}$ is such that $\sum_{i=1}^{d+K} \langle \check{\mathbb{S}}^{(i)}(S)\rangle_{T_n} \le k$. Therefore, on $\{S \in \tilde{\mathcal{I}}, : \sum_{i=1}^{d+K} \langle \check{\mathbb{S}}^{(i)}(S)\rangle_{T_n} \le k\}$,

$$\sum_{l=1}^{d+K} \left(\langle \ddot{\mathbb{S}}^{(l)}\rangle_t - \langle \ddot{\mathbb{S}}^{(l)}\rangle_s\right) \le 2^{k+1}|t - s|/\delta \quad \forall s, t \in [0, T_n] \text{ with } |t - s| \le \delta. \qquad (5.4.40)$$

Hence by Markov's inequality

$$\begin{aligned}\mathbb{P}^{(N)}\Big(\sum_{i=1}^{d+K} \langle \ddot{\mathbb{S}}^{(i)}\rangle_{T_n} > k\Big) &= \mathbb{P}^{(N)}\Big(\sum_{i=1}^{d+K} \langle \check{\mathbb{S}}^{(i)}\rangle_{T_n} > k\Big) \\ &\le \sum_{i=1}^{d+K} \mathbb{P}^{(N)}\Big(\langle \check{\mathbb{S}}^{(i)}\rangle_{T_n} > k/(d+K)\Big) \\ &\le \frac{\mathbb{E}_{\mathbb{P}^{(N)}}\big[\sum_{i=1}^{d+K} \langle \check{\mathbb{S}}^{(i)}\rangle_{T_n}^{p/2}\big](d+K)^{p/2}}{k^{p/2}} \le (d+K)^{p/2} K_1 k^{-p/2}.\end{aligned}$$

Step 5: Tightness of measures through tightness of quadratic variation processes.

Together with (5.4.40), by Arzelá-Ascoli theorem, this implies that $\{\mathbb{P}^{(N)} \circ (\langle \ddot{\mathbb{S}} \rangle_t)^{-1}\}_{N\in\mathbb{N}}$ is tight (in $\mathcal{C}([0,T_n],\mathbb{R}^d)$. Then by Lemma 6.4.13 in Jacod and Shiryaev [64], $\{\mathbb{P}^{(N)} \circ (\ddot{\mathbb{S}}_t)^{-1}\}_{N\in\mathbb{N}}$ is tight (in $\mathbb{D}([0,T_n],\mathbb{R}^d)$), which by Theorem 6.3.21 in Jacod and Shiryaev [64] implies that $\forall \epsilon > 0, \eta > 0$, there are $N_0 \in \mathbb{N}$ and $\theta > 0$ with

$$N \geq N_0 \;\Rightarrow\; \mathbb{P}^{(N)}(w'_{T_n}(\mathbb{S},\theta) \geq \eta) \leq \epsilon,$$

where w'_{T_n} is defined by

$$w'_{T_n}(S,\theta) = \inf\left\{ \max_{i\leq r} \sup_{t_{i-1}\leq s\leq t<t_i} |S_t - S_s| : \, 0 = t_0 < \ldots < t_r = T_n, \, \inf_{i<r}(t_i - t_{i-1}) \geq \theta \right\}.$$

Note that for S such that $w'_{T_n}(S,\theta) > 0$, there exist $t_0, \ldots, t_r$ with $0 = t_0 < \ldots < t_r = T_n$ and $\inf_{i<r}(t_i - t_{i-1}) \geq \theta$ such that

$$\max_{i\leq r} \sup_{t_{i-1}\leq s\leq t<t_i} |S_t - S_s| \leq 2w'_{T_n}(S,\theta).$$

which by continuity of S implies that

$$w_{T_n}(S,\theta) := \sup\{|S_t - S_s| : \, 0 \leq s < t \leq T_n, \, t - s \geq \theta\} \leq 4w'_{T_n}(S,\theta).$$

Then we have

$$N \geq N_0 \;\Rightarrow\; \mathbb{P}^{(N)}(w_{T_n}(\ddot{\mathbb{S}},\theta) \geq 4\eta) \leq \epsilon,$$

which then by Theorem 6.1.5 in Jacod and Shiryaev [64] implies that $\{\mathbb{P}^{(N)} \circ (\ddot{\mathbb{S}}_t)^{-1}\}$ is tight (in $\mathcal{C}([0,T_n],\mathbb{R}^d)$).

Step 6: Tightness gives exact duality.

Then there exists a converging subsequence $\{\mathbb{P}^{(N_k)} \circ (\ddot{\mathbb{S}}_t)^{-1}\}$ such that $\mathbb{P}^{(N_k)} \circ (\ddot{\mathbb{S}}_t)^{-1} \to \mathbb{P}$ weakly for some probability measure $\mathbb{P}$ on Ω. Consequently,

$$\lim_{k\to\infty} \mathbb{E}_{\mathbb{P}^{(N_k)}}[G(\ddot{\mathbb{S}})] = \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})].$$

In addition, if $\mathbb{P}$ is an element of $\mathcal{M}_{\vec{\mu},\mathfrak{P}}$, then

$$\begin{aligned} V_{\vec{\mu},\mathfrak{P}}(G) \leq \widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) \leq & \lim_{N\to\infty} \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathcal{I},1/N}} \mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-10)}(\mathbb{S})}}{2^{2D}}\right] + c_2/2^D \\ \leq & \liminf_{N\to\infty} \mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] + c_2/2^D \\ \leq & \liminf_{N\to\infty} \mathbb{E}_{\mathbb{P}^{(N)}}[G(\ddot{\mathbb{S}})] + e(D) \\ \leq & \lim_{k\to\infty} \mathbb{E}_{\mathbb{P}^{(N_k)}}[G(\ddot{\mathbb{S}})] + e(D) \\ \leq & \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] + e(D) \leq P_{\vec{\mu},\mathfrak{P}}(G) + e(D). \end{aligned} \tag{5.4.41}$$

where $e(x) := 5f_e(2^{-x+11}) + \frac{2nL\|\mathbb{S}\|}{x} + \frac{c_2+8\kappa+8}{2^x} + \frac{2nLV^{(p)}_{\mu_n,\mathcal{I}}(\|\mathbb{S}\|)}{x}$ and the third inequality follows from (5.4.39).

It remains to argue that $\mathbb{P}$ is an element of $\mathcal{M}_{\vec{\mu},\mathfrak{P}}$. First, it is straightforward to see that $\mathbb{S}$ is a $\mathbb{P}$–martingale and $\mathcal{L}_{\mathbb{P}}(S_{T_i}) = \mu_i$ for any $i \leq n$. To show that $\mathbb{P}(\{\mathbb{S} \in \mathfrak{P}\}) = 1$, notice that by the portemanteau theorem, for every $\epsilon > 0$

$$\mathbb{P}(\{\mathbb{S} \in \overline{\mathfrak{P}^\epsilon}\}) \geq \limsup_{k\to\infty} \mathbb{P}^{(N_k)}(\{\mathbb{S} \in \overline{\mathfrak{P}^\epsilon}\}) \geq \limsup_{k\to\infty} \mathbb{P}^{(N_k)}(\{\mathbb{S} \in \mathfrak{P}^{1/N_k}\}) = 1.$$

Therefore, it follows from Remark 5.4.4 and the monotone convergence theorem that

$$\mathbb{P}(\{\mathbb{S} \in \mathfrak{P}\}) = \lim_{\epsilon>0} \mathbb{P}(\{\mathbb{S} \in \overline{\mathfrak{P}^\epsilon}\}) = 1,$$

and hence $\mathbb{P} \in \mathcal{M}_{\vec{\mu},\mathfrak{P}}$.

To conclude, as D is arbitrary, (5.4.41) yields that

$$V_{\vec{\mu},\mathfrak{P}}(G) \leq P_{\vec{\mu},\mathfrak{P}}(G),$$

which then implies that

$$\widetilde{V}_{\mathcal{X},\mathcal{P},\mathfrak{P}}(G) = V_{\vec{\mu},\mathfrak{P}}(G) = P_{\vec{\mu},\mathfrak{P}}(G) = \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G).$$

5.4.7 Proof of Theorem 5.3.22

We start with a key lemma, analogous to the one obtained in Dolinsky and Soner [39].

Lemma 5.4.6. Consider

$$\alpha_D(S) := \big(\max_{1\leq i\leq d} \|S^{(i)}\|^p + 1\big)\mathbb{1}_{\{\max_{1\leq i\leq d}\|S^{(i)}\|+1\geq D\}} + \frac{\max_{1\leq i\leq d}\|S^{(i)}\|^p}{D}. \tag{5.4.42}$$

Then, given that $(\mu_j^{(i)})$ satisfies Assumption 5.3.20, for any $\mathbb{P} \in \mathcal{M}$ such that $\mathcal{L}_{\mathbb{P}}(\mathbb{S}^{(i)}_{T_j}) = \mu_j^{(i)}$ $\forall i \leq d,\ j \leq n$

$$\mathbb{E}_{\mathbb{P}}[\alpha_D(\mathbb{S})] \leq e_2(\vec{\mu}_n, D), \tag{5.4.43}$$

where $e_2(\vec{\mu}_n, D) := \left(\frac{p}{p-1}\right)^p \sum_{i=1}^d \left(2\int_{|x|\geq(\frac{p-1}{p})(D-1)} |x|^p\mu_n^{(i)}(\mathrm{d}x) + \frac{1}{K}\int |x|^p\mu_n^{(i)}(\mathrm{d}x)\right) \to 0$ as $D \to \infty$.

Proof. First define $h_D : \mathbb{R} \to \mathbb{R}$ by

$$h_D(x) = pD^{p-1}\Big(|x| - \Big(\frac{p-1}{p}\Big)D\Big)\mathbb{1}_{\{(\frac{p-1}{p})D\leq|x|<D\}} + |x|^p\mathbb{1}_{\{|x|\geq D\}}.$$

Notice that h_K is convex and satisfies

$$|x|^p \mathbb{1}_{\{|x|\geq D\}} \leq h_D(x) \leq |x|^p \mathbb{1}_{\left\{|x|\geq (\frac{p-1}{p})D\right\}}, \qquad \text{for any } D \geq 1.$$

For any $\mathbb{P} \in \mathcal{M}$ such that $\mathcal{L}_{\mathbb{P}}(\mathbb{S}^{(i)}_{T_j}) = \mu^{(i)}_j \ \forall i \leq d,\ j \leq n$, $\{h_D(\mathbb{S}_t) := (h^{(1)}_D(\mathbb{S}_t), \ldots, h^{(d)}_D(\mathbb{S}_t))\}_{t\geq 0}$ is a sub-martingale under $\mathbb{P}$ since h_D is convex. Therefore by Doob's inequality

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[\alpha_D(\mathbb{S})] &\leq \mathbb{E}_{\mathbb{P}}\big[2\|h_{D-1}(\mathbb{S})\|\big] + \frac{1}{D}\mathbb{E}_{\mathbb{P}}[\|\mathbb{S}\|^p] \\
&\leq \sum_{i=1}^{d} \Big(\frac{p}{p-1}\Big)^p \Big(\mathbb{E}_{\mathbb{P}}[2h^{(i)}_{D-1}(\mathbb{S}_{T_n})] + \frac{1}{K}\mathbb{E}_{\mathbb{P}}[|\mathbb{S}^{(i)}_{T_n}|^p]\Big) \\
&\leq \sum_{i=1}^{d} \Big(\frac{p}{p-1}\Big)^p \Big(2\int_{|x|\geq(\frac{p-1}{p})(D-1)} |x|^p \mu^{(i)}_n(\mathrm{d}x) + \frac{1}{K}\int |x|^p \mu^{(i)}_n(\mathrm{d}x)\Big) \\
&= e_2(\vec{\mu}_n, D).
\end{aligned}$$

□

We now proceed with the proof the Theorem 5.3.22. We first show that $\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) \leq \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G)$.

Given $(\mathcal{X}, \gamma) \in \mathcal{A}^{(p)}_{\mathcal{X}}$ such that $(\mathcal{X}, \gamma)$ superreplicates G on $\mathfrak{P}^\epsilon$ for some $\epsilon > 0$, since X is bounded, it follows from the definition of $\mathcal{A}^{(p)}$ that there exists $M_1 > 0$ such that

$$X(\mathbb{S}^{(1)}, \ldots, \mathbb{S}^{(d)}) + \int_0^{T_n} \gamma_u \mathrm{d}\mathbb{S}_u \geq G(\mathbb{S}) - M_1(1 + \sup_{0\leq t\leq T_n} |\mathbb{S}_t|^p)\mathbb{1}_{\{\mathbb{S}\notin\mathfrak{P}^\epsilon\}} \ \text{ on } \mathcal{I}. \quad (5.4.44)$$

Next, for each $N \geq 1$, we pick $\mathbb{P}^{(N)} \in \mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}$ such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] \geq \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N}.$$

We first notice that $X(\mathbb{S}^{(1)}, \ldots, \mathbb{S}^{(d)})$ is of the form $\sum_{i=1}^d \sum_{j=1}^n f_{i,j}(\mathbb{S}^{(i)}_j)$ for some $f_{i,j}$ such that $\forall i \leq d,\ j \leq n$ $f_{i,j}$ is continuous and bounded by $M_2(1+|\mathbb{S}^{(i)}_{T_j}|)$ for some M_2. Since by Jensen's inequality, for any $\mathbb{P} \in \mathcal{M}$ such that $\mathcal{L}_{\mathbb{P}}(\mathbb{S}^{(i)}_{T_n}) = \mu^{(i)}_n \ \forall i \leq d$

$$\mathbb{E}_{\mathbb{P}}[|\mathbb{S}^{(i)}_{T_j}|^p] \leq \mathbb{E}_{\mathbb{P}}[|\mathbb{S}^{(i)}_{T_n}|^p] \leq \int_{[0,\infty)} x^p \mu^{(i)}_n(\mathrm{d}x) < \infty \quad \forall i \leq d,\ j \leq n,$$

it follows from weak convergence of measures, the definition of $\mathcal{M}_{\vec{\mu},\mathfrak{P},\epsilon}$ and Lemma 5.4.6 that

$$\big|\mathcal{P}(X) - \mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S}^{(1)}, \ldots, \mathbb{S}^{(d)})]\big| \to 0 \quad \text{as } N \to \infty. \qquad (5.4.45)$$

Since γ is progressively measurable, the integral $\int_0^\cdot \gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u$, defined pathwise via integration by parts, agrees a.s. with the stochastic integral under any $\mathbb{P}^{(N)}$. Then, by (5.2.1), the stochastic integral is a $\mathbb{P}^{(N)}$ super-martingale and hence $\mathbb{E}_{\mathbb{P}^{(N)}}\Big[\int_0^{T_n}\gamma_u(\mathbb{S})\mathrm{d}\mathbb{S}_u\Big]\leq 0$. Therefore, by Lemma 5.4.6

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S})] \geq & \mathbb{E}_{\mathbb{P}^{(N)}}\big[G(\mathbb{S}) - M_1(1+|\mathbb{S}|^p)\mathbb{1}_{\{\mathbb{S}\notin\mathfrak{P}^\epsilon\}}\big] \\
\geq & \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N} - M_1\mathbb{E}_{\mathbb{P}^{(N)}}\big[(1+|\mathbb{S}|^p)\mathbb{1}_{\{\mathbb{S}\notin\mathfrak{P}^\epsilon\}}\mathbb{1}_{\{\|\mathbb{S}\|\leq N^{1/2p}\}}\big] \\
& - M_1\mathbb{E}_{\mathbb{P}^{(N)}}\big[(1+|\mathbb{S}|^p)\mathbb{1}_{\{\mathbb{S}\notin\mathfrak{P}^\epsilon\}}\mathbb{1}_{\{\|\mathbb{S}\|> N^{1/2p}\}}\big] \\
\geq & \sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N} - \frac{M_1(1+\sqrt{N})}{N} - \frac{M_1}{N} - e_2(\vec{\mu}_n, N^{1/2p}).
\end{aligned}
$$

This, together with (5.4.45), yields

$$\mathcal{P}(X) \geq \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G).$$

As $(X,\gamma)\in\mathcal{A}_{\mathcal{X}}$ is arbitrary, we therefore establish $\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G)\leq \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G)$.

Let $D>1$ and define G_D by $G_D = G\wedge D\vee(-D)$. Then it is clear that G_D is bounded and uniformly continuous. Therefore, by Theorem 5.3.17

$$
\begin{aligned}
\widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G_{L+D}) = & \lim_{\eta\searrow 0}\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}} \mathbb{E}_{\mathbb{P}}[G_{L+D}(\mathbb{S})] \\
\leq & \lim_{\eta\searrow 0}\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] + 2L\lim_{\eta\searrow 0}\sup_{\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}} \mathbb{E}_{\mathbb{P}}\Big[(1+\|\mathbb{S}\|^p)\mathbb{1}_{\{\|\mathbb{S}\|\geq(\frac{D}{L})^{1/p}\}}\Big],
\end{aligned}
$$

where the second inequality follows from $G_{L+D}(\mathbb{S})\leq G(\mathbb{S})+2L\big(1+\|\mathbb{S}\|^p\mathbb{1}_{\{\|\mathbb{S}\|\geq(\frac{D}{L})^{1/p}\}}\big)$. We know from Assumption 5.3.1 that any $S\in\mathcal{I}$ satisfies $\|S^{(i)}\|\leq\kappa\ \forall i>d$, where κ is the smallest number such that $X_i^{(c)}/\mathcal{P}(X_i^{(c)})$'s are bounded by κ. It follows from Lemma 5.4.6 that for any $D\geq L\kappa^p$ and $\mathbb{P}\in\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta}$

$$
\begin{aligned}
& \mathbb{E}_{\mathbb{P}}\Big[(1+\|\mathbb{S}\|^p)\mathbb{1}_{\{\|S\|\geq(\frac{D}{L})^{1/p}\}}\Big] \\
= & \mathbb{E}_{\mathbb{P}}\Big[\max_{1\leq i\leq d}\|\mathbb{S}^{(i)}\|^p\mathbb{1}_{\{\max_{1\leq i\leq d}\|\mathbb{S}^{(i)}\|^p\geq(\frac{D}{L})^{1/p}\}}\Big]\leq e_2(\vec{\mu}_n, D/L)\to 0, \quad \text{as } D\to\infty,
\end{aligned}
$$

and therefore,

$$\limsup_{D\to\infty}\widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G_{D+L})\leq\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G).$$

On the other hand, by the linearity of the market,

$$\widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G) \ \leq\ \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G_{D+L}) + \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(2L(1+\|\mathbb{S}\|^p)\mathbb{1}_{\{\|\mathbb{S}\|\geq(\frac{D}{L})^{1/p}\}}).$$

Since $\mathcal{M}_{\vec{\mu},\mathfrak{P},\eta} \neq \varnothing$ for any $\eta > 0$, $\widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(\alpha_D) \geq \widetilde{P}_{\vec{\mu},\mathfrak{P}}(\alpha_D) \geq 0$. Then it follows from Lemma 4.1 in Dolinsky and Soner [39] and the obvious fact that $\widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\Omega}(\alpha_D) \geq \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(\alpha_D)$ that

$$\limsup_{D\to\infty} \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(\alpha_D) = 0.$$

Hence we conclude that

$$\widetilde{P}_{\vec{\mu},\mathfrak{P}}(G) \leq \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G) \leq \limsup_{D\to\infty} \widetilde{V}^{(p)}_{\mathcal{X}^{(p)},\mathcal{P},\mathfrak{P}}(G_{D+L}) \leq \widetilde{P}_{\vec{\mu},\mathfrak{P}}(G)$$

and therefore we have equalities throughout.

5.5 Proof of Theorem 5.3.2

In this section, we prove an improved version of Theorem 5.3.2.

Theorem 5.5.1. For any $\alpha, \beta \geq 0$, $D \in \mathbb{N}$

$$\mathbf{V}_{\mathcal{I}}(G - \alpha \wedge (\beta\sqrt{m^{(D)}})) \leq \mathbf{P}_{\mathcal{I}}(G - \alpha \wedge (\beta\sqrt{m^{(D-4)}})),$$

Proof. Let $f_e : \mathbb{R}_+ \to \mathbb{R}_+$ be the modulus of continuity of G and choose $\kappa \geq 1 \vee \alpha$ be such that $X_i^{(c)}/\mathcal{P}(X_i^{(c)})$ is bounded by $\kappa - 1$ for any i and G is bounded by κ.

We first notice that

$$\begin{aligned}
\mathbf{V}_{\mathcal{I}}(G - \alpha \wedge (\beta\sqrt{m^{(D)}})) &\leq \inf_{N\geq 0} \mathbf{V}_{\Omega}(G - \alpha \wedge (\beta\sqrt{m^{(D)}}) - N\lambda_{\mathcal{I}}) \\
&= \inf_{N\geq 0} \underline{\mathbf{P}}_{\Omega}(G - \alpha \wedge (\beta\sqrt{m^{(D)}}) - N\lambda_{\mathcal{I}}) \\
&= \inf_{N\geq 0} \sup_{\mathbb{P}\in\underline{\mathcal{M}}} \mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D)}}) - N\lambda_{\mathcal{I}}]. \qquad (5.5.1)
\end{aligned}$$

Define the function $\mathcal{G} : \mathbb{R}_+ \times \underline{\mathcal{M}} \to \mathbb{R}$ by

$$\begin{aligned}
\mathcal{G}(N, \mathbb{P}) :=& \lim_{\epsilon\searrow 0} \inf_{\tilde{\mathbb{P}}\in\underline{\mathcal{M}},\, d_p(\tilde{\mathbb{P}},\mathbb{P})<\epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[G(\mathbb{S}) - \alpha \wedge (\beta\sqrt{m^{(D-2)}}) - N\lambda_{\mathcal{I}}(\mathbb{S})\Big] \\
=& \lim_{\epsilon\searrow 0} \inf_{\tilde{\mathbb{P}}\in\underline{\mathcal{M}},\, d_p(\tilde{\mathbb{P}},\mathbb{P})<\epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}\Big[- \alpha \wedge (\beta\sqrt{m^{(D-2)}})\Big] + \mathbb{E}_{\mathbb{P}}[G - N\lambda_{\mathcal{I}}].
\end{aligned}$$

Then by (4.5.1) in Remark 4.5.2, for any sequence $(\mathbb{P}^{(k)})_{k\geq 1}$ converging to $\mathbb{P}$ weakly,

$$\mathbb{E}_{\mathbb{P}}\Big[- \alpha \wedge (\beta\sqrt{m^{(D)}(\mathbb{S})})\Big] \leq \liminf_{k\to\infty} \mathbb{E}_{\mathbb{P}^{(k)}}\Big[- \alpha \wedge (\beta\sqrt{m^{(D-2)}(\mathbb{S})})\Big].$$

and hence

$$\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D)}})-N\lambda_{\mathcal{I}}]\leq\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\mathcal{G}(N,\mathbb{P}).$$

The next step is to interchange the order of the infimum and supremum. Notice that when we fix $\mathbb{P}$, $\mathcal{G}$ is affine in the first variable and continuous due to bounded convergence theorem. In addition, by definition $\mathcal{G}$ is lower-semi continuous in the second variable. Furthermore, $\mathcal{G}$ is convex in the second variable. To justify this, we notice that $\mathbb{P}\mapsto\mathbb{E}_{\mathbb{P}}\Big[-\alpha\wedge(\beta\sqrt{m^{(D-2)}(\mathbb{S})})\Big]$ is a linear functional and it follows that for each $\epsilon>0$ and $\lambda\in[0,1]$

$$\begin{aligned}&\inf_{\tilde{\mathbb{P}}\in\mathcal{M}_s,\,d_p(\tilde{\mathbb{P}},\lambda\mathbb{P}^{(1)}+(1-\lambda)\mathbb{P}^{(2)})<\epsilon}\mathbb{E}_{\tilde{\mathbb{P}}}\Big[-\alpha\wedge(\beta\sqrt{m^{(D-2)}(\mathbb{S})})\Big]\\ &\leq\lambda\inf_{\tilde{\mathbb{P}}\in\mathcal{M}_s,\,d_p(\tilde{\mathbb{P}},\mathbb{P}^{(1)})<\epsilon}\mathbb{E}_{\tilde{\mathbb{P}}}\Big[-\alpha\wedge(\beta\sqrt{m^{(D-2)}(\mathbb{S})})\Big]\\ &\quad+(1-\lambda)\inf_{\tilde{\mathbb{P}}\in\mathcal{M}_s,\,d_p(\tilde{\mathbb{P}},\mathbb{P}^{(2)})<\epsilon}\mathbb{E}_{\tilde{\mathbb{P}}}\Big[-\alpha\wedge(\beta\sqrt{m^{(D-2)}(\mathbb{S})})\Big].\end{aligned}$$

It follows that we can now apply min–max Theorem (see Corollary 2 in Terkelsen [102]) to $\mathcal{G}$ restricted to $[0,N]\times\underline{\mathcal{M}}$ and deduce that

$$\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\mathcal{G}(X,\mathbb{P})=\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\inf_{0\leq L\leq N}\mathcal{G}(L,\mathbb{P}).$$

Therefore, we have

$$\begin{aligned}\mathbf{V}_{\mathcal{I}}(G-\alpha\wedge(\beta\sqrt{m^{(D)}}))&=\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D)}})-N\lambda_{\mathcal{I}}]\\ &\leq\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\mathcal{G}(N,\mathbb{P})\\ &\leq\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\inf_{0\leq L\leq N}\mathcal{G}(L,\mathbb{P})\\ &=\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\inf_{0\leq L\leq N}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-2)}})-L\lambda_{\mathcal{I}}].\end{aligned}\tag{5.5.2}$$

Then we will argue that in the following that in the supremum term of (5.5.2), it suffices to consider $\mathbb{P}\in\underline{\mathcal{M}}$ such that

$$\mathbb{E}_{\mathbb{P}}[\lambda_{\mathcal{I}}]\leq 3\kappa/N.\tag{5.5.3}$$

Denote by $\underline{\mathcal{M}}^c_{\mathcal{I}}$ the set of $\mathbb{P}\in\underline{\mathcal{M}}$ such that $\mathbb{E}_{\mathbb{P}}[\lambda_{\mathcal{I}}]\geq c$. Suppose $\mathbb{P}\notin\underline{\mathcal{M}}^{3\kappa/N}$. Then

$$\inf_{0\leq L\leq N}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-2)}})-L\lambda_{\mathcal{I}}]\leq\kappa-N\mathbb{E}_{\mathbb{P}}[\lambda_{\mathcal{I}}]<-2\kappa,$$

while

$$\mathbf{V}_{\mathcal{I}}(G-\alpha\wedge(\beta\sqrt{m^{(D)}}))\geq\mathbf{P}_{\mathcal{I}}(G-\alpha\wedge(\beta\sqrt{m^{(D)}}))\geq-2\kappa.$$

Therefore, it suffices to consider $\mathbb{P}\in\underline{\mathcal{M}}^{3\kappa/N}$ and we can then deduce that

$$\mathbf{V}_{\mathcal{I}}(G-\alpha\wedge(\beta\sqrt{m^{(D)}})\leq\inf_{N\geq 0}\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{2\kappa/N}}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-2)}}].$$

Now fix N and take $\mathbb{P}^{(N)}\in\underline{\mathcal{M}}_{\mathcal{I}}^{3\kappa/N}$ such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G-\alpha\wedge(\beta\sqrt{m^{(D-2)}}]\geq\sup_{\mathbb{P}\in\underline{\mathcal{M}}^{3\kappa/N}}\mathbb{E}_{\mathbb{P}}[G-\alpha\wedge(\beta\sqrt{m^{(D-2)}}]-1/N.$$

By definition of $\underline{\mathcal{M}}$, we know there exists a complete probability space $(\Omega^W,\mathcal{F}_T^W,\mathbb{F}^W,P^W)$ together with a finite dimensional Brownian motion $(W_t)_{t\geq 0}$ and the natural filtration $\mathcal{F}_t^W=\sigma\{W_s|s\leq t\}$, and a continuous martingale M defined on $(\Omega^W,\mathcal{F}_T^W,\mathbb{F}^W,P^W)$ such that $\mathbb{P}^{(N)}=P^W\circ M^{-1}$.

Let $\tau=\{t\geq 0:\max_{d<i\leq d+K}M_t^{(i)}\geq\kappa+1\}\wedge T$ be the first time that $M^{(i)}$ hits $\kappa+1$ for some $i>d$ and define a stopped process $\Gamma:=M^{\tau}$. Note that since $X_i^{(c)}/\mathcal{P}(X_i^{(c)})$ is bounded by $\kappa-1$ for any i, it follows from (5.5.3) that

$$P^W(\tau<T)\leq 3\kappa/N.$$

Hence

$$|E^W[G(\Gamma)]-E^W[G(M)]|\leq 6\kappa^2/N.$$

We now construct a new process $\tilde{\Gamma}$ from Γ such that the law of $\tilde{\Gamma}$ under P^W is an element of $\underline{\mathcal{M}}_{\mathcal{I}}$.

We write $\eta_N=3\kappa/N$ and

$$p_i^{(N)}:=\mathbb{E}_{\mathbb{P}^{(N)}}[X_i^{(c)}(\mathbb{S}_T^{(1)},\dots,\mathbb{S}_T^{(d)})]$$

for any $i=1,\dots K$, and define $\tilde{p}_i^{(N)}$'s by

$$\tilde{p}_i^{(N)}=\frac{\mathcal{P}(X_i^{(c)})-(1-\sqrt{\eta_N})p_i^{(N)}}{\sqrt{\eta_N}}.$$

Note that $\mathbb{P}^{(N)}\in\underline{\mathcal{M}}_{\mathcal{I}}^{\eta_N}$ implies that

$$P^W(|X_i^{(c)}(\Gamma_T^{(1)},\dots,\Gamma_T^{(d)})/\mathcal{P}(X_i^{(c)})-\Gamma_T^{(d+i)}|>\eta_N)\leq\eta_N. \tag{5.5.4}$$

Hence we can deduce that

$$
\begin{aligned}
&\big|\mathcal{P}(X_i^{(c)}) - p_i^{(N)}\big| \\
\leq& E^W[|X_i^{(c)}(\Gamma_T^{(1)},\dots,\Gamma_T^{(d)}) - \mathcal{P}(X_i^{(c)})\Gamma_T^{(d+i)}|] \\
\leq& \mathcal{P}(X_i^{(c)})\eta_N + E^W\Big[|X_i^{(c)}(\Gamma_T^{(1)},\dots,\Gamma_T^{(d)}) - \mathcal{P}(X_i^{(c)})\Gamma_T^{(d+i)}|\mathbb{1}_{\left\{|X_i^{(c)}(\Gamma_T^{(1)},\dots,\Gamma_T^{(d)})/\mathcal{P}(X_i^{(c)})-\Gamma_T^{(d+i)}|>\eta_N\right\}}\Big] \\
\leq& \mathcal{P}(X_i^{(c)})\eta_N + 2(\kappa+1)\mathcal{P}(X_i^{(c)})\eta_N, \quad \forall i=1,\dots,K.
\end{aligned}
$$

It follows immediately that

$$
\begin{aligned}
\big|\tilde{p}_i^{(N)} - \mathcal{P}(X_i^{(c)})\big| =& \Big(\frac{1}{\sqrt{\eta_N}} - 1\Big)\Big|\mathcal{P}(X_i^{(c)}) - p_i^{(N)}\Big| \\
\leq& \frac{2(\kappa+1)\mathcal{P}(X_i^{(c)})\eta_N}{\sqrt{\eta_N}} = 2(\kappa+1)\mathcal{P}(X_i^{(c)})\sqrt{\eta_N} \quad \forall i \leq K. \qquad (5.5.5)
\end{aligned}
$$

Then, it follows from Assumption 5.3.1 that when N is large enough there exists a $\tilde{\mathbb{P}}^{(N)} \in \underline{\mathcal{M}}_{\tilde{\mathcal{I}}}$ such that

$$
\tilde{p}_i^{(N)} := \mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[X_i(\mathbb{S}_T^{(1)},\dots,\mathbb{S}_T^{(d)})] \quad \forall i \leq K.
$$

Enlarging Wiener space $(\Omega^W, \mathcal{F}^W, P^W)$ if necessary, then there are continuous martingales Γ and $\tilde{M}$ which have laws equal to $\mathbb{P}^{(N)}$ and $\tilde{\mathbb{P}}^{(N)}$ respectively, and an $\mathcal{F}_T^W$-measurable random variable $\xi \in \{0,1\}$ that is independent of Γ and $\tilde{M}$, with

$$
P^W(\xi = 1) = 1 - \sqrt{\eta_N} \text{ and } P^W(\xi = 0) = \sqrt{\eta_N}.
$$

Define $\mathcal{F}_T^W$-measurable random variables $\tilde{\Lambda}^{(i)}$ by

$$
\begin{aligned}
\tilde{\Lambda}^{(i)} &= \Gamma_T^{(i)}\mathbb{1}_{\{\xi=1\}} + \tilde{M}_T^{(i)}\mathbb{1}_{\{\xi=0\}} \quad \forall i=1,\dots,d, \\
\tilde{\Lambda}^{(i)} &= X_{i-d}(\tilde{\Lambda}^{(1)},\dots,\tilde{\Lambda}^{(d)})/\mathcal{P}(X_{i-1}^{(c)}) \quad \forall i > d.
\end{aligned}
$$

We now construct a continuous martingale from $\tilde{\Lambda}$ by taking conditional expectations:

$$
\tilde{\Gamma}_t = E^W[\tilde{\Lambda}|\mathcal{F}_t^W], \quad t \in [0,T].
$$

It follows from the fact that ξ is independent of M and $\tilde{M}$

$$
\begin{aligned}
\tilde{\Gamma}_0^{(i)} =& E^W[\tilde{\Gamma}_T^{(i)}|\mathcal{F}_0^W] \\
=& (1-\sqrt{\eta_N})E^W[X_i(\Gamma_T^{(1)},\dots,\Gamma_T^{(d)})/\mathcal{P}(X_i^{(c)})] + \sqrt{\eta_N}E^W[X_i(\tilde{M}_T^{(1)},\dots,\tilde{M}_T^{(d)})/\mathcal{P}(X_i^{(c)})] \\
=& \frac{(1-\sqrt{\eta_N})p_i^N + \sqrt{\eta_N}\tilde{p}_i^N}{\mathcal{P}(X_i^{(c)})} = 1 \quad \forall i > d
\end{aligned}
$$

and

$$\tilde{\Gamma}_0^{(i)} = E^W[\tilde{\Gamma}_T^{(i)}|\mathcal{F}_0^W] = E^W[\tilde{\Lambda}_T^{(i)}|\mathcal{F}_0^W] = (1-\eta_N)E^W[\Gamma_T^{(i)}] + \eta_N E^W[\tilde{M}_T^{(i)}] = 1 \quad \forall i \leq d.$$

Hence $\tilde{\mathbb{P}} := P^W \circ (\tilde{\Gamma}_t)^{-1} \in \underline{\mathcal{M}}_{\mathcal{I}}$. Also by independence between ξ and $(M, \tilde{M})$, we have

$$E^W[|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}|] = \sqrt{\eta_N} E^W[|\tilde{M}_T^{(i)} - \Gamma_T^{(i)}|] \leq 2\sqrt{\eta_N} \quad \forall i \leq d$$

and by (5.5.4)

$$P^W(|\Gamma_T - \tilde{\Lambda}^{(i)}| > \eta_N) \leq \eta_N + \sqrt{\eta} \leq 2\sqrt{\eta_N} \quad \forall i > d,$$

which implies that

$$\begin{aligned} E^W[|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}|] &= 2E^W[(\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)})^+] - E^W[\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}] \\ &= 2E^W[(\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)})^+] \\ &\leq 2\eta_N + 2E^W\left[\Lambda^{(i)} \mathbb{1}_{\left\{|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}| > \eta_N\right\}}\right] \\ &\leq 2\eta_N + 4(\kappa+2)\sqrt{\eta_N} \leq 14\kappa\sqrt{\eta_N}, \quad \forall i = d+1, \ldots, K. \end{aligned}$$

Then by Doob's martingale inequality

$$P^W(\|\tilde{\Gamma} - \Gamma\| \geq \kappa\eta_N^{1/4}) \leq \frac{1}{\kappa\eta_N^{1/4}} \sum_{i=1}^{d+K} E^W[|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}|] \leq 14(d+K)\eta_N^{1/4}$$

and hence

$$\begin{aligned} \left|\mathbb{E}_{\tilde{\mathbb{P}}}[G(\mathbb{S})] - \mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})]\right| \leq& \left|E^W[G(\tilde{\Gamma}) - G(\Gamma)]\right| + \left|E^W[G(\Gamma) - G(M)]\right| \\ \leq& f_e(\kappa\eta_N^{1/4}) + E^W\left[|G(\Gamma) - G(\Gamma)| \mathbb{1}_{\left\{\|\tilde{\Gamma} - \Gamma\| \geq \kappa\eta_N^{1/4}\right\}}\right] + 6\kappa^2/N \\ \leq& f_e(\kappa\eta_N^{1/4}) + 28\kappa(d+K)\eta_N^{1/4} + 6\kappa^2/N. \end{aligned}$$

In addition, by Remark 4.5.2 the inequality $m^{(D-2)}(\Gamma) \geq m^{(D-4)}(\tilde{\Gamma})$ holds on $\big\{\omega \in \Omega^W : \|\Gamma(\omega) - \tilde{\Gamma}(\omega)\| < \kappa\eta_N^{1/4}\big\}$.

It follows that

$$\begin{aligned} \mathbb{E}_{\tilde{\mathbb{P}}}[\alpha \wedge (\beta\sqrt{m^{(D-2)}(\mathbb{S})})] =& E^W[\alpha \wedge (\beta\sqrt{m^{(D-2)}(\Gamma)})] \\ \geq& E^W[\alpha \wedge (\beta\sqrt{\hat{m}^{(D-4)}(\tilde{\Gamma})})] - 14\alpha(d+K)\eta_N^{1/4}. \end{aligned}$$

Therefore,

$$
\begin{aligned}
&\mathbb{E}_{\mathbb{P}^{(N)}}[G - \alpha \wedge (\beta\sqrt{m^{(D-2)}(\mathbb{S})})] - \mathbb{E}_{\tilde{\mathbb{P}}}[G - \alpha \wedge (\beta\sqrt{m^{(D-4)}(\mathbb{S})})] \\
\le\, & f_e(\kappa\eta_N^{1/4}) + 28\kappa(d+K)\eta_N^{1/4} + 14\alpha(d+K)\eta_N^{1/4} + 6\kappa^2/N.
\end{aligned}
$$

which leads to

$$
\begin{aligned}
&\sup_{\mathbb{P}\in\underline{\mathcal{M}}}\inf_{0\le L\le N}\mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-2)}}) - L\lambda_{\mathcal{I}}] \\
=\, & \sup_{\mathbb{P}\in\underline{\mathcal{M}}^{2\kappa/N}}\mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-2)}}] \\
\le\, & \mathbb{E}_{\mathbb{P}^{(N)}}[G - \alpha \wedge (\beta\sqrt{m^{(D-2)}(\mathbb{S})})] + \frac{1}{N} \\
\le\, & \sup_{\mathbb{P}\in\underline{\mathcal{M}}_{\mathcal{I}}}\mathbb{E}_{\mathbb{P}}[G - \alpha \wedge (\beta\sqrt{m^{(D-4)}(\mathbb{S})})] \\
&+ \frac{1}{N} + f_e(\kappa\eta_N^{1/4}) + 28\kappa(d+K)\eta_N^{1/4} + 14\alpha(d+K)\eta_N^{1/4} + 6\kappa^2/N.
\end{aligned}
$$

As N is arbitrary, we can then conclude that

$$
\mathbf{V}_{\mathcal{I}}(G - \alpha \wedge (\beta\sqrt{m^{(D)}})) \le \underline{\mathbf{P}}_{\mathcal{I}}(G - \alpha \wedge (\beta\sqrt{m^{(D-4)}}).
$$

□

Chapter 6

Research outlook

In this final chapter we present some interesting problems which remain to be studied and further questions which have arisen from this and previous works, and also discuss the possibilities of extending our results to answer some of the questions that we left open.

6.1 Enlargement of filtration

The prime focus of our work is on the case when trading strategies are adapted to the natural filtration $\mathbb{F}$ of the price process $\mathbb{S}$, or a technical enlargement of it. In general, the information stream can contain not only the price information of the traded risky assets so far, but also some additional information which can come initially, at a later time or progressively. For an example of a progressive enlargement of filtration, there are economic indicators and broad based indices which can be relevant information for traders to make their trading decisions. On the other hand, a company announcement which may affect share price can be modelled as an initial enlargement of filtration.

The consideration of an enlarged filtration is not only natural but also meaningful. In particular, the enlargement of filtration can be used to study an informed agent problem. In continuous time, a general framework built on our robust framework in Chapter 4 was introduced in Aksamit et al. [2], and the pricing–hedging duality for a regular agent, who only observes the stock prices, can be therein extended to agents with additional information. Moreover, the framework therein allows us to quantify the value of additional information. Although there are already a few interesting results, the research on robust pricing–hedging with respect to an enlarged filtration is still in its infancy. Only initial enlargement of filtration and the case that the filtration is enlarged at a later specific time are considered. Therefore, one interesting question here is whether we can extend the analysis there further to deal with a progressive enlargement of filtration.

Also, it would be interesting to extend our discrete-time results in Chapter 2 to a general filtration.

6.2 Trading restriction

In Chapter 3, we study discrete-time robust pricing and hedging under a short selling ban. In discrete time, robust pricing and hedging with frictions has undergone rapid development in recent years. Several authors have made important contributions to robust pricing and hedging with trading restrictions, see Fahim and Huang [46] and Bayraktar and Zhou [5]. Robust pricing and hedging with transaction costs were also considered. See Dolinsky and Soner [38], Burzoni [16] and the references therein.

In contrast, few studies of continuous-time robust pricing and hedging with frictions have been done so far. Recently, Dolinsky and Soner [40] have extended their discrete-time results in Dolinsky and Soner [38] to continuous time via an asymptotic approximation. However, as far as we know, there has been no study of continuous-time robust pricing and hedging with trading restrictions. It would be interesting to see how we could fill the gap in the literature.

It is one of our on-going projects to establish a version of robust pricing–hedging duality with tradings restrictions. By examining the proof of Theorem 4.3.1 in Chapter 4, we see that it involves the discretisation of both the primal and the dual problems, and applying classical probabilistic results of Föllmer and Kramkov [48] to connect the discrete counterparts. The classical results of Föllmer and Kramkov [48] concern trading restrictions. However, there are some difficulties of adapting the proof of Theorem 4.3.1 to address the pricing and hedging problem with trading restrictions. The first one is the discretisation of the primal and the dual problems with trading restrictions. In order to define the discretised problems, a good discretisation of the trading restrictions is needed. The second one, which seems also the more difficult one, is to connect the discretised problems with the original problems. That is to show asymptotically the value of discretised primal (dual) problem is equal to that of the original primal (dual) problem. Although so far only a small class of general convex portfolio constraints can be adapted into our new construction, a preliminary result, nevertheless, has been obtained and it already includes many interesting cases such as a short selling ban and a total trading ban on a selected number of stocks.

6.3 Dynamic hedging versus static hedging

In Chapter 2 and 4, we included both static and dynamic hedging into the robust framework and establish pricing–hedging duality in both discrete time and continuous time. One thing that has not been discussed is a comparison of static hedging and dynamic hedging. While it is intuitively true that being able to continuously trade derivatives can enrich the trading strategies and hence benefits the hedging, in practice, due to potentially higher costs incurred by active derivative trading, there is a need for a proper understanding of the true advantages of dynamic hedging over static hedging on the theoretical level.

Given that the same finite set of put options that can be either statically hedged or continuously traded, in the absence of a prediction set $\mathfrak{P}$, by comparing Theorem 5.3.2 with Theorem 5.3.10, we see that the primal value for both cases is the same and hence it is clear that the superhedging price of a derivative with stocks as underlying is unchanged. In discrete time, the same conclusion can be drawn from a study of Theorem 2.4.2 for the two cases. In fact, it is even true that if beliefs only concerns stock prices but not derivative prices then the superhedging price remains unchanged. However, when there are beliefs on how derivative prices will move in a response to a move in the prices of underlying, the ability of being able to re-adjust derivative portfolios will be an advantage in terms of lower minimal initial capital required. This gives one reason for preferring dynamic hedging to static hedging.

Another exciting thing about dynamic hedging is the possibility of achieving a better result for another problem among all the superhedging strategies requiring the same initial capital. We here give a (motivating) example, which comes from the idea of Bruno Dupire. Let us consider the following strategy: We start with an initial superhedging strategy at time 0; At any later time, if we find out that in the market there is a cheaper superhedging strategy in terms of capital required, then we can close our previous positions and build new positions according to a cheaper superhedging strategy. Every time we spot such an opportunity, we only need to invest part of the money into the new superhedging strategy, and the rest of the money is freed up and can be reallocated to other business. The improvement of cash flow from this dynamic hedging strategy can make it much more appealing than the initial superhedging strategy. So far, there has been every little research along this direction, and we think it could be a very interesting topic to study.

Bibliography

[1] Acciaio, B., Beiglböck, M., Penkner, F., and Schachermayer, W. (2016). A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. *Mathematical Finance*, 26(2):233–251.

[2] Aksamit, A., Hou, Z., and Obłój, J. (2016). Robust framework for quantifying the value of information in pricing and hedging. *arXiv:1605.02539*.

[3] Avellaneda, M., Levy, A., and Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Math Finance*, 2(2):73–88.

[4] Battalio, R. and Schultz, P. (2011). Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets. *The Journal of Finance*, 66(6):2013–2053.

[5] Bayraktar, E. and Zhou, Z. (2015). On arbitrage and duality under model uncertainty and portfolio constraints. *Mathematical Finance, to appear.* DOI: 10.1111/mafi.12104.

[6] Beiglböck, M., Henry-Labordère, P., and Penkner, F. (2013). Model-independent bounds for option prices: a mass transport approach. *Finance and Stochastics*, 17(3):477–501.

[7] Beiglböck, M., Nutz, M., and Touzi, N. (2016). Complete duality for martingale optimal transport on the line. *The Annals of Probability, to appear.*

[8] Biagini, S., Bouchard, B., Kardaras, C., and Nutz, M. (2015). Robust fundamental theorem for continuous processes. *Mathematical Finance, to appear.* DOI: 10.1111/mafi.12110.

[9] Bichteler, K. (1981). Stochastic integration and $\mathbf{L}^p$-theory of semimartingales. *The Annals of Probability*, 9(1):49–89.

[10] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654.

[11] Bogachev, V. I. (2007). *Measure theory. Vol. I, II.* Springer-Verlag, Berlin.

[12] Bouchard, B. and Nutz, M. (2015). Arbitrage and duality in nondominated discrete-time models. *The Annals of Applied Probability*, 25(2):823–859.

[13] Breeden, D. T. and Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices'. *The Journal of Business*, 51(4):621–651.

[14] Brown, H., Hobson, D., and Rogers, L. C. G. (2001). Robust hedging of barrier options. *Math. Finance*, 11(3):285–314.

[15] Bruckner, A. M., Thomson, B. S., and Bruckner, J. B. (2008). *Real Analysis*. Pearson.

[16] Burzoni, M. (2016). Arbitrage and hedging in model-independent markets with frictions. *arXiv:1512.01488*.

[17] Burzoni, M., Frittelli, M., Hou, Z., Maggis, M., and Obłój, J. (2016a). Pathwise arbitrage theory in discrete time. *arXiv:1612.07618*.

[18] Burzoni, M., Frittelli, M., and Maggis, M. (2015). Model-free superhedging duality. *Mathematical Finance, to appear.*

[19] Burzoni, M., Frittelli, M., and Maggis, M. (2016b). Robust arbitrage theory of capital asset pricing. Private communication.

[20] Burzoni, M., Frittelli, M., and Maggis, M. (2016c). Universal arbitrage aggregator in discrete-time markets under uncertainty. *Finance and Stochastics*, 20(1):1–50.

[21] Carr, P., Fisher, T., and Ruf, J. (2014). On the hedging of options on exploding exchange rates. *Finance and Stochastics*, 18(1):115–144.

[22] Cassese, G. (2008). Asset pricing with no exogenous probability measure. *Math. Finance*, 18(1):23–54.

[23] Cohn, D. L. (1980). *Measure theory*, volume 165. Springer.

[24] Cox, A. M., Hou, Z., and Obłój, J. (2016). Robust pricing and hedging under trading restrictions and the emergence of local martingale models. *Finance and Stochastics*, 20(3):669–704.

[25] Cox, A. M. and Obłój, J. (2011). Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 15(3):573–605.

[26] Cox, A. M. G. and Hobson, D. G. (2005). Local martingales, bubbles and option prices. *Finance and Stochastics*, 9(4):477–492.

[27] Cox, A. M. G. and Wang, J. (2013). Root's barrier: Construction, optimality and applications to variance options. *The Annals of Applied Probability*, 23(3):859–894.

[28] Cvitanić, J. and Karatzas, I. (1993). Hedging contingent claims with constrained portfolios. *The Annals of Applied Probability*, 3(3):652–681.

[29] Cvitanić, J., Pham, H., and Touzi, N. (1999). Super-replication in stochastic volatility models under portfolio constraints. *Journal of Applied Probability*, 36(2):523–545.

[30] Dalang, R. C., Morton, A., and Willinger, W. (1990). Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 29(2):185–201.

[31] Davis, M. H. and Hobson, D. G. (2007). The range of traded option prices. *Mathematical Finance*, 17(1):1–14.

[32] Davis, M. H. A., Obłój, J., and Raval, V. (2014). Arbitrage bounds for prices of weighted variance swaps. *Mathematical Finance*, 24:821–854.

[33] De Finetti, B. (1970). *Theory of probability: a critical introduction treatment*. John Wiley.

[34] Delbaen, F. and Schachermayer, W. (1994a). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520.

[35] Delbaen, F. and Schachermayer, W. (1994b). Arbitrage and free lunch with bounded risk for unbounded continuous processes. *Mathematical Finance*, 4(4):343–348.

[36] Dellacherie, C. and Meyer, P. (1978). *Probabilities and Potential. Chapters I–IV*. North-Holland Mathematics Studies.

[37] Denis, L. and Martini, C. (2006). A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *The Annals of Applied Probability*, 16(2):827–852.

[38] Dolinsky, Y. and Soner, H. (2014). Robust hedging with proportional transaction costs. *Finance and Stochastics*, 18(2):327–347.

[39] Dolinsky, Y. and Soner, H. M. (2013). Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields*, 160(1-2):391–427.

[40] Dolinsky, Y. and Soner, H. M. (2015a). Convex duality with transaction costs. *arXiv:1502.01735*.

[41] Dolinsky, Y. and Soner, H. M. (2015b). Martingale optimal transport in the Skorokhod space. *Stochastic Processes and their Applications*, 125(10):3893–3931.

[42] Dupire, B. (1993). Model art. *Risk*, 6(9):118–124.

[43] Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1):18–20.

[44] Ekström, E. and Janson, S. (2016). The inverse first-passage problem and optimal stopping. *The Annals of Applied Probability*, 26(5):3154–3177.

[45] El Karoui, N. and Tan, X. (2013). Capacities, measurable selection and dynamic programming part I: abstract framework. *arXiv:1310.3363*.

[46] Fahim, A. and Huang, Y. (2016). Model-independent superhedging under portfolio constraints. *Finance and Stochastics*, 20(1):51–81.

[47] Föllmer, H. (1981). Calcul d'Itô sans probabilités. *Séminaire de probabilités de Strasbourg*, 15:143–150.

[48] Föllmer, H. and Kramkov, D. (1997). Optional decompositions under constraints. *Probability Theory and Related Fields*, 109(1):1–25.

[49] Galichon, A., Henry-Labordère, P., and Touzi, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312–336.

[50] Guasoni, P. and Rásonyi, M. (2015). Fragility of arbitrage and bubbles in local martingale diffusion models. *Finance and Stochastics*, 19(2):215–231.

[51] Guo, G., Tan, X., and Touzi, N. (2016). Tightness and duality of martingale transport on the Skorokhod space. *Stochastic Processes and their Applications, to appear.*

[52] Harrison, J. M. and Kreps, D. M. (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics*, 92(2):323–336.

[53] Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic theory*, 20(3):381–408.

[54] Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic processes and their applications*, 11(3):215–260.

[55] Hendershott, T., Namvar, E., and Phillips, B. (2013). The intended and collateral effects of short-sale bans as a regulatory tool. *The Journal of Investment Management*, 11:5–13.

[56] Henry-Labordère, P., Obłój, J., Spoida, P., and Touzi, N. (2016). The maximum maximum of a martingale with given n marginals. *The Annals of Applied Probability*, 26(1):1–44.

[57] Herdegen, M. and Schweizer, M. (2016). Strong bubbles and strict local martingales. *International Journal of Theoretical and Applied Finance, to appear.* Available at SSRN: http://ssrn.com/abstract=2566815.

[58] Heston, S. L., Loewenstein, M., and Willard, G. A. (2007). Options and bubbles. *Review of Financial Studies*, 20(2):359–390.

[59] Hobson, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*. Springer.

[60] Hobson, D. G. (1998). Robust hedging of the lookback option. *Finance and Stochastics*, 2(4):329–347.

[61] Hou, Z. and Obłój, J. (2015). On robust pricing–hedging duality in continuous time. *arXiv:1503.02822*.

[62] Huberman, G. (1989). A simple approach to arbitrage pricing theory. *Theory of Valuation. Frontiers of Modern Financial Theory*, 1:289–297.

[63] Hugonnier, J. (2012). Rational asset pricing bubbles and portfolio constraints. *Journal of Economic Theory*, 147(6):2260–2302.

[64] Jacod, J. and Shiryaev, A. N. (2002). *Limit theorems for stochastic processes*, volume 288. Springer Berlin.

[65] Jarrow, R. and Protter, P. (2012). Discrete versus continuous time models: Local martingales and singular processes in asset pricing theory. *Finance Research Letters*, 9(2):58–62.

[66] Jarrow, R., Protter, P., and Shimbo, K. (2007). Asset price bubbles in complete markets. In Fu, M. C., Jarrow, R. A., Yen, J.-Y. J., and Elliott, R. J., editors, *Advances in Mathematical Finance*. Birkhäuser Boston.

[67] Jarrow, R. A., Protter, P., and Shimbo, K. (2010). Asset price bubbles in incomplete markets. *Mathematical Finance*, 20(2):145–185.

[68] Jouini, E. and Kallal, H. (1995). Arbitrage in securities markets with short-sales constraints. *Mathematical Finance*, 5(3):197–232.

[69] Karandikar, R. L. et al. (1995). On pathwise stochastic integration. *Stochastic Processes and their applications*, 57(1):11–18.

[70] Kellerer, H. G. (1984). Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 67(4):399–432.

[71] Knight, F. (1921). *Risk, Uncertainty and Profit.* Boston: Houghton Mifflin.

[72] Kreps, D. M. (1981). Arbitrage and equilibrium in economies with infinitely many commodities. *Journal of Mathematical Economics*, 8(1):15–35.

[73] Lewis, A. L. (2000). *Option valuation under stochastic volatility: with Mathematica code.* Finance Press.

[74] Loewenstein, M. and Willard, G. A. (2000). Local martingales, arbitrage, and viability free snacks and cheap thrills. *Economic Theory*, 16(1):135–161.

[75] Lyons, T. J. (1995). Uncertain volatility and the risk-free synthesis of derivatives. *Applied Math Finance*, 2(2):117–133.

[76] Malkiel, B. G. (2007). *A random walk down Wall Street: the time-tested strategy for successful investing.* WW Norton & Company.

[77] Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics*, 4(1):141–183.

[78] Mykland, P. A. (2003). Financial options and statistical prediction intervals. *Ann. Statist.*, 31(5):1413–1438.

[79] Nadtochiy, S. and Obłój, J. (2015). Robust pricing and hedging of barrier options with beliefs on implied volatility. Private communication.

[80] Neufeld, A. and Nutz, M. (2013). Superreplication under volatility uncertainty for measurable claims. *Electron. J. Probab.*, 18(48):1–14.

[81] Nutz, M. (2012). Pathwise construction of stochastic integrals. *Electron. Commun. Probab*, 17(24):1–7.

[82] Nutz, M. and van Handel, R. (2013). Constructing sublinear expectations on path space. *Stochastic Processes and their Applications*, 123(8):3100–3121.

[83] Obłój, J. (2010). Skorokhod Embedding. In Cont, R., editor, *Encyclopedia of Quantitative Finance*. Wiley.

[84] Obłój, J. and Ulmer, F. (2012). Performance of robust hedges for digital double barrier options. *International J.Theoretical and Applied Finance*, 15(1):1–34.

[85] Peng, S. (2013). Nonlinear expectation theory and stochastic calculus under Knightian uncertainty. *Real Options, Ambiguity, Risk and Insurance*, 5:144–184.

[86] Pham, H. and Touzi, N. (1999). The fundamental theorem of asset pricing with cone constraints. *Journal of Mathematical Economics*, 31(2):265–279.

[87] Possamaï, D., Royer, G., and Touzi, N. (2013). On the robust superhedging of measurable claims. *Electron. Commun. Probab.*, 18:1–13.

[88] Protter, P. (2013). A mathematical theory of financial bubbles. In Henderson, V. and Sircar, R., editors, *Paris-Princeton Lectures on Mathematical Finance 2013*, number 2081 in Lecture Notes in Mathematics. Springer.

[89] Pulido, S. (2014). The fundamental theorem of asset pricing, the hedging problem and maximal claims in financial markets with short sales prohibitions. *The Annals of Applied Probability*, 24(1):54–75.

[90] Riedel, F. (2015). Financial economics without probabilistic prior assumptions. *Decisions in Economics and Finance*, 38(1):75–91.

[91] Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes and martingales: Volume 2, Itô calculus*, volume 2. Cambridge university press.

[92] Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory*, 13(3):341–360.

[93] Ross, S. A. (1977). Return, risk and arbitrage. In Friend, I. and Bicksler, J., editors, *Risk and Return in Finance*. Ballinger Cambridge, Massachusetts.

[94] Samuelson, P. (1965). Rational theory of warrant pricing. *Industrial Management Rev.*, 6:13–31.

[95] Scheinkman, J. A. and Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111(6):1183–1220.

[96] Shiryaev, A. N. (1984). *Probability.* Springer-Verlag, New York.

[97] Sin, C. A. (1998). Complications with stochastic volatility models. *Advances in Applied Probability*, 30(1):256–268.

[98] Skorokhod, A. V. (1977). On a representation of random variables. *Theory of Probability and Its Applications*, 21(3):628–632.

[99] Soner, M., Touzi, N., and Zhang, J. (2011). Quasi-sure stochastic analysis through aggregation. *Electron. J. Probab.*, 16(2):1844–1879.

[100] Spoida, P. (2014). *Robust pricing and hedging beyond one marginal.* PhD thesis, University of Oxford. Available at http://ora.ox.ac.uk/objects/uuid:0315824b-52f7-4e44-9ac6-0a688c49762c.

[101] Strassen, V. (1965). The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439.

[102] Terkelsen, F. (1972). Some minimax theorems. *Mathematica Scandinavica*, 31:405–413.

[103] Villani, C. (2003). *Topics in optimal transportation.* Number 58. American Mathematical Soc.

[104] Villani, C. (2008). *Optimal transport: old and new*, volume 338. Springer Science & Business Media.

[105] Vovk, V. (2012). Continuous-time trading and the emergence of probability. *Finance and Stochastics*, 16(4):561–609.

[106] Whalley, A. E. and Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7(3):307–324.

[107] Xiong, W. and Yu, J. (2011). The chinese warrants bubble. *American Economic Review*, 101:2723–2753.