Infinite-Dimensional Linear Programming and Model-Independent Hedging of Contingent Claims

a thesis presented for the degree of

Doctor of Philosophy of Imperial College London

and the

Diploma of Imperial College

by

Sergey Badikov

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2BZ

July 2017

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: ______________________________

Copyright

Thesis advisors: Professor Mark H. A. Davis
Dr Antoine Jacquier

Sergey Badikov

Infinite-Dimensional Linear Programming and Model-Independent Hedging of Contingent Claims

Abstract

We consider model-independent pathwise hedging of contingent claims in discrete-time markets, in the framework of infinite-dimensional linear programmes (LP). The dual problem can be formulated as optimization over the set of martingale measures subject to market constraints. Absence of model-independent arbitrage plays a crucial role in ensuring that both the primal and the dual problems are well posed and there is no duality gap. In fact we show that different notions of model-independent arbitrage are required to prove duality results in various settings. We then specialize this duality theory to the situation where European Call options are traded on the market. In particular we consider hedging portfolios that consist of static positions in traded options and a dynamic trading strategy. The dual variables are then constrained to martingale measures consistent with prices of traded options. When only finitely many Call options are traded, the notion of weak arbitrage introduced in Davis and Hobson (2007) is sufficient to ensure absence of duality gap between the primal and the dual problems. In this case the set of feasible dual variables is not closed, and extrapolation of Call option prices (equivalently of the implied volatility smile) is required. We finally provide numerical examples to support our theoretical claims. By discretizing the infinite-dimensional LPs, we compute arbitrage-free price bounds for Forward-Start options. We further perform a sensitivity analysis of the aforementioned extrapolation and find that in the case of Forward-Start options it does not significantly influence arbitrage bounds obtained by numerically solving discretized problems.

To my mother and Jackie.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisors Prof. Mark Davis and Dr Antoine Jacquier for their continual support and encouragement. Our many discussions never failed to inspire me and their guidance helped me in all the time of research through to the final writing of my thesis.

I would like to thank my mother, for selflessly giving me more than I ever could have asked for. Her strength and wisdom is my constant inspiration in life. Thank you for encouraging me to always pursue new challenges.

I would also like to thank Jackie for encouraging me to be the best version of myself. Thank you for putting up with me!

I am grateful to have met such wonderful friends during my PhD. The experiences we shared along with countless laughs and strong cups of coffee will always be one of my fondest memories!

Finally, I would like to recognize Rogge Global Partners Ltd for providing financial support for my PhD. I would especially like to thank Dr Igor Pikovsky for giving me the opportunity to pursue my PhD and his invaluable advice along the way.

List of figures

List of tables

Contents

Notations

- The set of extended real numbers is $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$.
- The set of non-negative real numbers is $\mathbb{R}_+ := [0, +\infty)$.
- The set of extended non-negative real numbers is $\overline{\mathbb{R}}_+ := [0, +\infty]$.
- The set of positive real numbers is $\mathbb{R}_{++} := (0, +\infty)$.
- For a vector $\mathrm{x} \in \mathbb{R}^n$ with $n < \infty$ the l_1-norm is $\|\mathrm{x}\|_1 := \sum_{i=1}^n |x_i|$.
- Dimension of a set C is denoted by $\dim C$.
- The space of continuous functions on a space X is $C(X)$.
- The space of functions on a space X differentiable once is $C^1(X)$.
- The weighted Banach space of continuous functions dominated by a positive continuous function $g : X \to \mathbb{R}_{++}$ is defined as

$$C_g(X) := \left\{ f \in C(X) : \|f\|_g := \sup_{x \in X} \frac{|f(x)|}{g(x)} < \infty \right\}.$$

- The space of upper semi-continuous functions bounded above by a positive continuous function (up to a constant multiple) to $g : X \to \mathbb{R}_{++}$ is $U_g(X)$.
- The space of lower semi-continuous functions bounded below by a positive continuous function (up to a constant multiple) $g : X \to \mathbb{R}_{++}$ is $L_g(X)$.
- The space of signed Borel measures of bounded variation on a space X is $\mathcal{M}(X)$.

- The space of Borel probability measures on a space X is $\mathcal{P}(X)$.
- The space of Borel probability measures on a space X with finite integrals of a positive continuous function $g : X \to \mathbb{R}_{++}$ is defined as

$$\mathcal{P}_g(X) := \left\{ \pi \in \mathcal{P}(X) : \int_X g(x)\pi(\mathrm{d}x) < \infty \right\}.$$

- The space of signed Borel measures of bounded variation on a space X with that integrate a positive continuous function $g : X \to \mathbb{R}_{++}$ to a finite constant is defined as

$$\mathcal{M}_g(X) := \left\{ \mu \in \mathcal{M}(X) : \int_X g(x)\mu(\mathrm{d}x) < \infty \right\}.$$

- For a sequence of sets $(A_n)_{n\in\mathbb{N}}$ the limit is defined as

$$\lim_{n\to\infty} A_n := \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_n.$$

- The indicator of a set C is defined to be

$$\chi_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

- Algebraic difference of two sets B and C is defined as

$$B - C := \{b - c : b \in B,\, c \in C\}.$$

- Interior of a set C is

$$\mathrm{int}\, C := \{x \in C : \text{ there is a neighbourhood } U \text{ of } x \text{ with } U \subset C\}.$$

- Affine hull of a set C is

$$\mathrm{aff}\, C := \left\{ \sum_{i=1}^{n} a_i x_i : \text{ for any } n > 0 \text{ such that } x_i \in C,\ \sum_{i=1}^{n} a_i = 1 \right\}.$$

- Relative interior of a set C is

  $$\operatorname{ri} C := \{x \in \operatorname{aff} C : \text{ there is a neighbourhood } U \text{ of } x \text{ with } U \cap \operatorname{aff} C \subset C\}.$$

- The linear space (or linear span) generated by a set C is

  $$\operatorname{Span} C := \{\lambda(x-y) : \lambda \in \mathbb{R}_+, \text{ for all } x, y \in C\}$$

- Effective domain of a function $g : X \to Y$ is

  $$\operatorname{dom} g := \{x \in X : g(x) < \infty\} \subset Y.$$

- Lower semi-continuous hull of a function $g : X \to Y$ at $x \in X$ is

  $$\operatorname{lsc} g(x) := \min \left\{ \liminf_{x' \to x} g(x'), g(x) \right\}.$$

- For a finite index set $I = \{i(1), \ldots, i(n)\}$ a real vector

  $$\left(a_1^{i(1)}, \ldots, a_{i(1)}^{i(1)}, a_1^{i(2)}, \ldots, a_{i(2)}^{i(2)}, \ldots, a_1^{i(n)}, \ldots, a_{i(n)}^{i(n)} \right)$$

  in $\mathbb{R}^d$ with $d := \sum_{l=1}^n i(l)$ is denoted by $(a_1^i, \ldots, a_i^i)_{i \in I}$.

0

Introduction

The main objective of this thesis is to investigate model-independent hedging of contingent claims in the framework of infinite-dimensional linear programming (LP). The classical approach to pricing and hedging contingent claims with a finite horizon T is to postulate existence of a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathcal{F}, \mathbb{P})$ such that the underlying price process $S := (S_t)_{t\in[0,T]}$ defined on this space, becomes a martingale when appropriately discounted under the measure $\mathbb{P}$. In complete markets a consequence of the Fundamental Theorem of Asset Pricing (Harrison and Pliska, 1981; Kreps, 1981; Delbaen and Schachermayer, 1994) is that the price of a contingent claim can be expressed as the expectation of the discounted pay-off of the claim under $\mathbb{P}$. Moreover the value of the expectation is the unique fair price for the claim if the market is frictionless and the model for the evolution of the price process is chosen appropriately. The celebrated Black-Scholes formula is based precisely on this argument. Moreover given the price of a single European Call or Put option the Black-Scholes model can be calibrated and used to price other contingent claims.

In practice traded options exhibit a range of volatilities that cannot be explained by the Black-Scholes model. This inefficiency resulted in the development of stochastic volatility models (see Gatheral (2006), Fouque et al. (2011) for a detailed account). However, introduction of the instantaneous volatility process in stochastic volatility models makes the market incomplete (unless more assets are traded, see Romano and Touzi (1997) and Davis and Obłój (2008)) and thus not all derivatives can be perfectly replicated. However the range of possible prices is constrained by arbitrage conditions. More precisely, given a contingent claim, the super-replication problem consists in finding an initial amount of capital $x_0 \in \mathbb{R}$ and a trading strategy $\phi \in \mathcal{A}$ such that its pay-off $x_0 + \int_0^T \phi_t \mathrm{d}S_t$ dominates the claim's pay-off H a.s.

$$\pi(H) := \inf \left\{ x_0 \in \mathbb{R} : \text{ there exists } \phi \in \mathcal{A} x_0 + \int_0^T \phi_t \mathrm{d}S_t \geq H \text{ a.s. } \right\},$$

where $\mathcal{A}$ is the set of admissible trading strategies ϕ such that the stochastic integral $\int \phi_t \mathrm{d}S_t$ is bounded from below to preclude arbitrage (Delbaen and Schachermayer (2006) and Pham (2009)). Because the market is incomplete but arbitrage-free there exists a family of equivalent martingale measures $\mathbb{M}$ such that the discounted price process is a $\mathbb{Q}$-martingale for each $\mathbb{Q} \in \mathbb{M}$. The dual problem is then formulated as finding an element in $\mathbb{M}$ that produces the largest arbitrage-free price of the claim. Delbaen (1992) showed that when the process S is a continuous and bounded semi-martingale the duality

$$\pi(H) = \sup_{\mathbb{Q} \in \mathbb{M}} \mathbb{E}^{\mathbb{Q}} \{H\}$$

holds. Kramkov (1996) extended the result to the case of locally bounded processes and showed that for every model there exists a super-replicating strategy. Föllmer and Kabanov (1997) subsequently dispensed with the local boundedness assumption.

However these approaches to super-replication assume that the model for the underlying asset price is known (although parameters of the model might not be specified exactly, for instance the uncertain volatility model developed by Avellaneda et al. (1995), Lyons (1995) and Leblanc and Martini (2000)). An idea to finding arbitrage-free bounds for prices of options goes back to Merton (1973), who derived bounds for the prices of European options without any assumptions on the model of the underlying price. Lo (1987) derived a semi-parametric bound on the price of European Call options depending only on the first two moments of the distribution of the asset price at maturity. Boyle and Lin (1997) extended Lo's result to the case of several assets. Curran (1994) and Rogers and Shi (1995) proposed a method of conditioning to determine the lower bound of the Asian Call price in the Black-Scholes model. Deelstra, Liinev, and Vanmaele (2004) obtained lower and upper bounds on the price of arithmetic Basket options using similar conditioning techniques and provided partially analytical bounds in the Black-Scholes case. Albrecher et al. (2008) used the idea of conditioning to provide model-independent lower bounds for the price of a discretely monitored arithmetic Asian Call option. Chen, Deelstra, Dhaene, and Vanmaele (2008) showed how to construct robust super-replicating strategies for European Call options on a weighted sum of asset prices via stochastic ordering and co-monotonic risks.

New developments in model-free approach to pricing and super-replication were spurred by Hobson, who in his highly influential paper (Hobson, 1998), presented the idea of super-replication for a wide class of arbitrage-free models using only European Call and Put options on the same underlying asset as hedging instruments. The author used the Hardy-Littlewood transform of the density of the asset price at maturity to obtain the bounds for the price of a Lookback option. The solution was obtained under the assumption that European Call or Put options on the underlying asset (assets) are traded for all required maturities and a continuum of strikes. This assumption is equivalent to knowing the marginal distributions of the asset price for each maturity exactly as was shown by Breeden and Litzenberger (1978). Brown, Hobson, and Rogers (2001) computed arbitrage-free bounds on the price of Barrier options under the same assumption using maximal martingale inequalities. Separately Hodges and Neuberger (2000) found arbitrage-free bounds on prices of Barrier options by restricting the set of martingale measures to two-stage jump processes and assuming full marginals of the underlying asset price were known. In this case the optimal martingale measure turns out to be solution to an LP and the optimal super-hedging portfolio consists of static positions in Call options maturing at the final maturity of the Barrier option. In the case when only finitely many options are available the optimal bound can be obtained by a static super-replicating portfolio in available options.

Another strand of literature has focused on providing model-independent bounds on prices of exotic derivatives as solutions to the Skorokhod embedding problem first introduced and solved by Skorokhod (1965), see comprehensive surveys by Obłój (2004) and Hobson (2010). Cox and Obłój (2011a,b) used such an approach to find upper and lower bounds for prices of the double touch and double no-touch Barrier options respectively. The super- and sub-hedging strategies attaining those bounds were also provided. Semi-analytical expressions for the arbitrage-free bounds for Forward-Start options were derived in Hobson and Neuberger (2012). Cox and Wang (2013) obtained arbitrage-free bounds on Variance options. Hobson and Klimmek (2012) studied arbitrage-free bounds on discretely monitored Variance Swaps with jumps in the price of the underlying. Recently Cox and Kinsley (2017) provided arbitrage bounds on prices of options on leveraged exchange-traded funds.

Recently, the problem of obtaining arbitrage-free bounds for exotic options has been studied through the theory of optimal transportation, introduced first by Monge and later developed by Kantorovich (1942). Beiglböck, Henry-Labordère, and Penkner (2013) show that when full marginals are known the problem of finding arbitrage-free bounds on prices of exotic derivatives can be formulated as a martingale optimal transportation (MOT) problem, an optimal transportation problem with added linear constraints that any feasible measure is a martingale (see Zaev (2015) for discussion of optimal transportation problems with additional linear constraints). Restriction of spaces of dual variables is used in Beiglböck et al. (2013, Section 1.5) to obtain duality results. Stebegg (2014) provides robust bounds for prices of Asian options in the MOT framework. De Marco and Henry-Labordère (2015) obtain arbitrage-free bounds on VIX options and super- and sub-hedging semi-static portfolios in discrete two-time period ($0 < t_1 < t_2$) setting. The problem is formulated as the MOT problem with the additional constraints that any feasible martingale measure consistent with marginals at times t_1 and t_2 also correctly prices the value of VIX at time t_1 and the price is equal to its current level. On the other hand, Guyon, Menegaux, and Nutz (2017) derive bounds for the prices of VIX futures under the assumption of full marginal distributions for the price process and only partial knowledge of the distribution of the VIX process. The problem is similar to the MOT set-up in nature, however as the authors do not assume the knowledge of the full VIX distribution, the MOT framework is not directly applicable. Therefore the problem is formulated as an infinite-dimensional LP that consists of finding the cheapest super-replicating or the most expensive sub-replicating portfolio in options and a dynamic hedging strategy. The dual problem is then to find an optimal martingale measure for the joint underlying price and VIX processes. Recently several authors including Dolinsky and Soner (2014); Galichon, Henry-Labordère, and Touzi (2014); Beiglböck, Nutz, and Touzi (2016) extended the MOT theory to cases of quasi-sure rather than pathwise hedging (for a family of probability measures $\mathcal{P}$, a property holds quasi-surely if it holds everywhere outside a polar set A, i.e. a set such that $\mathbb{P}(A) = 0$ for all $\mathbb{P} \in \mathcal{P}$, see Denis and Martini (2006)). A full survey of the MOT theory and its applications in finance (including numerical examples) can be found in Henry-Labordère (2017).

When prices of only finitely many European options are known the MOT framework does not appear to be helpful. Bertsimas and Popescu (2002) obtain bounds on the price of a European Call option given at least the first two moments of the distribution of the asset price at maturity as well as bounds on the price of exotic European type derivatives given finitely many Call option prices as a solution to a semi-infinite LP. The authors also propose an efficient algorithm to solve the resulting semi-infinite programmes in polynomial time. Laurence and Wang (2005) show that robust static hedging of Basket options can be done via semi-infinite LP using finitely many prices of European options on individual stocks in the basket. Independently d'Aspremont and Ghaoui (2006) derive similar results by adopting the Lagrangian approach to prove duality results. Hobson, Laurence, and Wang (2005) also obtain arbitrage-free bounds on the price of a Basket option via Lagrangian approach when only finitely many options are available. However the result is derived as a corollary to the case when full marginals are known and linear interpolation between given Call option prices was used to synthesize prices at strikes needed for the static portfolios. Davis, Obłój, and Raval (2014) provide upper and lower bounds on the price of continuously monitored weighted variance swaps as solutions to semi-infinite LPs.

A different approach is taken in Cox, Hou, and Obłój (2016); Hou and Obłój (2015); Aksamit, Hou, and Obłój (2016) where restriction of the state space via addition of modelling assumptions (or beliefs) is incorporated in the duality framework. In another development, Nadtochiy and Obłój (2017) similarly restrict the state space by bounding the set of possible future skews of implied volatility surface (the skew in this case is understood as the ratio of out-of-the-money implied volatilities).

Obtaining arbitrage-free robust bounds on prices of exotic options requires knowledge of market prices of European Call or Put options as inputs to the problem and therefore it is important to understand whether the supplied prices contain arbitrage opportunities. Cousot (2007) uses finitely many observable prices of European Call options to construct discrete martingale measures consistent with those prices. Davis and Hobson (2007) show that given a finite set of traded European Call options one can provide necessary and sufficient conditions to preclude arbitrage. Acciaio, Beiglböck, Penkner, and Schachermayer (2016) prove a model-free version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem under the assumption that, in addition to traded European-type options, an option with a super-linear convex pay-off is available on the market at some finite (although possibly large) price. In a more general setting, Riedel (2015) proved equivalence between absence of arbitrage and existence of a full support martingale measure. The super-replication result was also proved in the framework of infinite-dimensional LPs. Other authors including Cassese (2008); Alvarez et al. (2013); Bouchard and Nutz (2015); Ibraimi et al. (2015); Burzoni et al. (2016) studied equivalence between absence of arbitrage and existence of a martingale measure in the model-independent setting. It must be noted that there are different definitions of arbitrage opportunities in the model-free setting and each definition introduces subtle variations in statements and proofs of the Fundamental Theorem of Asset Pricing.

This thesis is organised as follows.

In Chapter 1 we study duality results in general discrete-time markets. We assume that the pay-offs of traded securities are continuous functions on a locally compact Polish space and are available to buy and sell at market prices without frictions. As a result the set of traded securities is a linear space and we define a pricing functional on that space mapping pay-offs to market prices. In Section 1.2 we show that absence of strong model-independent arbitrage introduced in Definition 1.5 is equivalent to the pricing functional being strictly positive, linear and uniquely defined. Section 1.3 is devoted to proving Fundamental Theorems of Asset Pricing, formulated as existence of a strictly positive linear extension to the pricing functional. This approach is similar to Kreps (1981), where the existence of such extension was called market viability see also Clark (1993, 2000) for study of arbitrage conditions in general markets. We assume that there exists a strictly positive continuous function h with a bounded reciprocal such that all traded securities pay-offs grow at most as fast as that function. That allows us to embed the linear space of traded securities into the weighted Banach space of continuous functions of order $\mathcal{O}(h)$. Assuming that the function h is itself a traded asset, we prove that absence of strong model-independent arbitrage is equivalent to the existence of a strictly positive linear extension of the pricing functional. Under additional technical assumption this extension can be identified with a Borel probability measure μ such that $\int h \mathrm{d}\mu < \infty$. This approach is similar to the one in Acciaio et al. (2016), although we do not require h to be super-linear. We also show that h needs not be a traded asset to state the existence of a strictly positive linear extension. However absence of strong model-independent arbitrage does not suffice and we introduce the notion of weak free lunch in Definition 1.10, which can be seen as a topological equivalent of strong model-independent arbitrage. Weak free lunch can also be seen as an application of "free lunch" introduced in Kreps (1981) to the model-independent setting. Moreover we assume that in this case all traded securities are $o(h)$ in order to represent the linear extension as a Borel probability measure. In Section 1.4 we present super- and sub-replication results for a wide class of options with upper and lower semi-continuous functions respectively. We formulate hedging problems as infinite-dimensional LPs following Clark (2003), and their duals are formulated as optimization problems over the set of Borel probability measures subject to market constraints. It was shown in (Clark, 2003, Lemma 1) that existence of a feasible solution to the dual problem is equivalent to existence of a positive linear extension of the pricing functional, and we employ this to prove duality results.

Chapter 2 presents application of the duality theory in general markets to the situation where European Call options are traded. The space of traded securities in this case consists of semi-static portfolios with static positions in Call options and dynamic trading strategies in the underlying asset. We distinguish between two cases. First in Section 2.2.1 we consider the case where infinitely many Call options are traded. In this case absence of weak free lunch is equivalent to existence of a strictly positive linear extension of the pricing functional which can be represented as a martingale measure consistent with observable prices. We contrast the notion of weak free lunch with a notion of "weak free lunch with vanishing risk" introduced in (Cox and Obłój, 2011b, Definition 2.1), which can be viewed as a model-independent counterpart to the classical notion of free lunch with vanishing risk introduced in Delbaen and Schachermayer (1994). We show that in our set-up absence of weak free lunch with vanishing risk implies absence of weak free lunch. On the other hand we consider the case when only finitely many Call options in Section 2.2.2 are traded we present another notion of the weak arbitrage first introduced in (Davis and Hobson, 2007, Definition 2.3) and extended in (Cox and Obłój, 2011b, Definition 2.3). Absence of weak arbitrage is equivalent to existence of a positive linear extension of the pricing functional as shown in (Davis and Hobson, 2007, Theorem 4.2). We show that the duality results still hold in this case but the modification of the cone of positive claims is required, similarly to Clark (2006). However the set of market models is not closed under no weak arbitrage as shown in Lemma 2.25. We propose arbitrage-free extrapolation of Call options prices (or equivalently of implied volatility smiles) to remedy this shortcoming as discussed in Section 2.2.4. The duality results hold in this case but require addition of untraded Call options to the set of static positions with prices consistent with the aforementioned extrapolation.

In Chapter 3 we approximate infinite-dimensional LPs defined in previous chapters by two-step discretization procedure. First we discretize the delta-hedging strategies by considering an appropriate basis in Section 3.1. We also discretize the set of static positions by considering only a finite subset of traded Call options where appropriate. As a result we obtain semi-infinite LPs and show that in the limit their values converge to the values of the original infinite-dimensional problems. The theory of semi-infinite linear programming is interesting in its own right as arbitrage conditions in that case simplify considerably (see Shapiro (2009) for a survey of semi-infinite optimization). In particular the dual problem has only finitely many constraints in this case (even though the optimization variables are probability measures which are infinite-dimensional objects) and arbitrage conditions are equivalent to the vector of input market prices being in relative interior of the range of the linear constraints map (see Davis, Obłój, and Raval (2014) and Bartl, Kupper, Prömel, and Tangpi (2017) for applications). We then introduce another level of discretization by considering finite subsets of the state space. Such discretization is performed so that the obtained finite-dimensional primal and dual problems are consistent with input market prices and original infinite-dimensional constraints (namely the martingale condition). In the case where Call options are traded for each strike $K \in \mathbb{R}_+$ and each maturity, we provide in Section 3.2 a way to consistently discretize the marginal distributions of the underlying price process so that the discretized marginal distributions satisfy no-arbitrage conditions. In Section 3.3 we provide an application to computing arbitrage bounds on the price of Forward-Start options (of Type I and of Type II), which are among the simplest products amenable to model-independent hedging techniques. The upper bound price for the at-the-money Type-II Forward-Start Straddle (with pay-off $|S_{t+\tau} - S_t|$ for some $t, \tau > 0$) was computed in Hobson and Neuberger (2012), where the support of the optimal martingale measure is a binomial tree. Unfortunately the optimal measure and the associated super-hedging portfolio are not available analytically. The martingale optimal transference plan for the lower bound price of the at-the-money Type-II Forward-Start Straddle has been characterised semi-analytically in Hobson and Klimmek (2014), and the transference plan (supported along a trinomial tree) is found by solving a set of coupled ODEs. Recently, Campi, Laachir, and Martini (2017) studied the change of numeraire in these two-dimensional optimal transport problems and showed that, under some technical conditions, the lower bound for the Type-I at-the-money Forward-Start Straddle is also attained by the Hobson-Klimmek transference plan. In the lower bound at-the-money case we numerically solve, in Section 3.4, the coupled ordinary differential equations associated with the Hobson-Klimmek transference plan (Hobson and Klimmek, 2014), and show that it is in striking agreement with the LP solution of the sub-hedging primal problem.

In both examples the range of forward smiles consistent with the marginal laws is large. Using European options to 'lock-in' (replicate) forward volatility or hedge forward volatility dependent claims seems illusory. Forward-Start options should be seen as fundamental building blocks for exotic pricing and not decomposable (or approximately decomposable) into European options. Models used for forward volatility dependent exotics should have the capability of calibration to Forward-Start option prices and at a minimum should produce realistic forward smiles that are consistent with trader expectations and observable prices. We conclude the chapter by discussing convergence of the finite-dimensional LPs as the discretization mesh of the state space goes to zero in Section 3.4.4.

Chapter 4 is devoted to a sensitivity analysis of the semi-infinite programmes defined in Chapter 3. We embed the primal and the dual problems into a family of perturbed problems where the perturbation is performed on the input market prices. In particular we wish to investigate how Assumptions 10 and 11 on arbitrage-free extrapolation of the total implied variance influences the optimal values of the dual programme. The value function of the parametrized dual problem turns out be convex and proper as a consequence of absence of arbitrage. Moreover it is directionally differentiable and we employ the theory of directional differentiability of convex functions (reviewed in Section 4.1) to obtain first-order expansion of the value function around the unperturbed input market prices similar to Goberna and López (2014). When the perturbation is itself parametrized we refine the result to obtain a first-order expansion of the value function in terms of those parameters.

Sensitivity analysis of the arbitrage-free price bounds of at-the-money Type-II Forward-Start Straddle is discussed in Section 4.3. First in Section 4.3.1, we consider the case where only prices of at-the-money Call options are observable on the market. The Black-Scholes model is fitted to the observed prices, which results in extrapolation of the total implied variance being constant in log-moneyness for each maturity. We enlarge the set of options available for static hedging by including non-traded options with prices consistent with the flat extrapolation and treat such prices as unperturbed input data. We then assume that the model was misspecified and in fact extrapolation of the total implied variance is linear in log-moneyness for each maturity and each wing. The prices of the options that result from linear extrapolation are treated as perturbations of the prices consistent with the Black-Scholes model. The primal problem is solved with perturbed prices as inputs and the obtained values are compared with the values obtained by first-order expansion of the value function. As expected the expansion becomes less accurate with the size of the perturbation.

Section 4.3.2 is devoted to discussing the case where a range of strikes is traded for each maturity and we assume that the observable prices are consistent with the Heston model. Additional untraded Call options are included in the set of static positions with prices given by the Heston model. Perturbation of input prices is obtained similarly to the Black-Scholes case by extrapolating the total implied variance linearly in log-moneyness for each wing and each maturity (extrapolations are of course constrained by arbitrage conditions derived in Chapter 2). The numerical analysis is then performed as before. The accuracy of the obtained expansion deteriorates with the size of the perturbation similarly to the Black-Scholes case.

It is of interest to note that although the accuracy of the expansion decreases as the size of the perturbation increases, the sensitivity of the dual problem to perturbations is rather small, thus providing further evidence that Forward-

Start options cannot be replicated by positions in the spot implied volatility and should rather be treated as building blocks of any model that is used to price exotics on forward volatility.

1

Duality Theory in General Markets

This chapter is devoted to establishing super-hedging duality in general markets as an application of infinite-dimensional linear programming. We present main ideas of the theory of infinite-dimensional LPs in Section 1.1 as well as specialising the general definition of the primal (1.2) and the dual (1.3) programmes following Clark (2003). In Section 1.2 we set-up a general market that consists of securities with continuous pay-offs φ_i indexed by some general set $\mathcal{I}$ and traded at prices c_i. As we assume that the market is frictionless the set of traded securities becomes a subspace of the space of continuous functions and we define a pricing functional on the subspace so that it maps pay-offs of traded securities to their market prices. We also introduce a notion of strong model-independent arbitrage and show in Proposition 1.6 that absence of such arbitrage is equivalent to the pricing functional having some desirable properties. Section 1.3 is devoted to proving Fundamental Theorems of Asset Pricing, which we formulate as a question of existence of a strictly positive linear extension to the pricing functional in spirit of Kreps (1981). Additional assumptions are made in order to represent this extension as a Borel probability measure on the state space Ω. We conclude with Section 1.4 where we prove duality results for super- and sub-hedging problems for a large class of upper and lower semi-continuous functions.

1.1 Linear Programmes in infinite-dimensional spaces

Let X be a topological vector space endowed with a partial order induced by a positive convex cone X_+. For two elements x_1 and $x_2 \in X$ we say that $x_1 \geq x_2$ if and only if $x_1 - x_2 \in X_+$. We shall call $X_{++} := X_+ \setminus \{0\}$ the strictly positive cone. For any subspace $V \subset X$ we define a positive cone $V_+ := V \cap X_+$ and a strictly positive cone $V_{++} = V \cap X_{++}$.

Let Y be another topological vector space that is placed in the separating duality with X, i.e. there is a bilinear form $(x, y) \mapsto \langle x, y \rangle$ from $X \times Y$ to $\mathbb{R}$ such that it is linear in x and y and $\langle x, y \rangle = 0$ for all $y \in Y$ implies $x = 0$ and similarly it implies $y = 0$ if it is zero for all $x \in X$ (Aliprantis and Border, 2007, Definition 5.90). If the space Y consists of all continuous functionals on X defined by $y(x) := \langle x, y \rangle_{XY}$ for $y \in Y$ then Y is called the topological dual space of X and the pair (X, Y) is called a dual pair.

The space X is endowed with a weak topology generated by elements of Y denoted by $\sigma(X, Y)$. It is a locally convex Hausdorff topology where a net of elements $\{x_\alpha\}_{\alpha \in \mathcal{A}} \subset X$ converges weakly to an element $x \in X$ if and only if $\langle x_\alpha, y \rangle$ is convergent to $\langle x, y \rangle$ in $\mathbb{R}$ for all $y \in Y$. Similarly one defines a weak* topology on Y induced by elements of X, denoted by $\sigma(Y, X)$. A net $\{y_\alpha\}_{\alpha \in \mathcal{A}} \subset Y$ is weak* convergent to $y \in Y$ if and only if $\langle x, y_\alpha \rangle \to \langle x, y \rangle$ in $\mathbb{R}$ for each $x \in X$.

A positive convex cone X_+ has a dual counterpart in Y, called a negative polar in notation of Anderson and Nash (1987). It is defined as follows

$$Y_+ := \{y \in Y : \langle x, y \rangle_{XY} \geq 0 \text{ for all } x \in X_+\}, \tag{1.1}$$

and it induces a partial order on the space Y in the same way that the positive cone X_+ induces an ordering on X. The negative polar is a non-empty as it contains zero. It is also convex and $\sigma(Y, X)$-closed (Aliprantis and Border, 2007, Lemma 5.102). Note that the statement of the lemma involves a different definition of the negative polar. Nonetheless the result of the lemma stays true. Indeed, for any two $y_1, y_2 \in Y_+$ and a constant $\lambda \in (0,1)$ we have that $\langle x, \lambda y_1 + (1-\lambda) y_2 \rangle = \lambda \langle x, y_1 \rangle + (1-\lambda)\langle x, y_2 \rangle \geq 0$ for any $x \in X_+$ and hence Y_+ is convex. To see that it is closed let $\{y_\alpha\}_{\alpha \in \mathcal{A}}$ be net in Y_+ that converges to $y \in Y$ in $\sigma(Y, X)$-topology. Then for any $x \in X_+$ we have $\langle x, y \rangle = \lim_\alpha \langle x, y_\alpha \rangle \geq 0$.

We call an element $y \in Y_+$ a positive linear functional on X. If $y(x) > 0$ for all $x \in X_+ \setminus \{0\}$ the functional is strictly positive on X. As a negative polar is itself a convex cone in Y one can take its negative polar, a bipolar, as a subset of X

$$X_+^{**} := \{x \in X : \langle x, y \rangle_{XY} \geq 0 \text{ for all } y \in Y_+\}.$$

The next result is significant in the development of theory of linear programming and is referred to in the literature as Bipolar Theorem (Aliprantis and Border, 2007, Theorem 5.103). We present a special case of the theorem when the subset of the space X is a positive cone.

Proposition 1.1. *The bipolar X_+^{**} of a positive cone X_+ is a $\sigma(X, Y)$-closed convex hull of X_+. In particular if X_+ is $\sigma(X, Y)$-closed then $X_+ = X_+^{**}$.*

As the definition of the negative polar is different in the statement of Bipolar Theorem we outline the proof below.

Outline of proof. As X_+^{**} is a negative polar of the cone X_+^* it is convex and $\sigma(X, Y)$-closed. Clearly $X_+ \subset X_+^{**}$. Moreover defining the set C as the intersection of all convex and $\sigma(X, Y)$-closed sets that contain X_+, we see that $C \subset X_+^{**}$. On the other hand if there is $x \in X_+^{**}$ such that $x \notin C$ then by Hahn-Banach Theorem there exists $y \in Y$ such that $\langle x, y \rangle < 0$ and $\langle c, y \rangle \geq 0$ for all $c \in C$. Hence $y \in X_+^*$ but also that $x \notin X_+^{**}$, a contradiction. □

We can also endow X with a stronger topology without simultaneously enlarging its topological dual Y.

Definition 1.2. (Aliprantis and Border, 2007, Definition 5.96) A locally convex topology τ on X is compatible with the dual pair (X, Y) if the topological dual of (X, τ) is precisely Y. Similar definition holds for Y.

It is must be noted that convex subsets of X have exactly the same closures in all topologies compatible with the dual pair (X, Y).

Theorem 1.3. *(Aliprantis and Border, 2007, Theorem 5.98) All locally convex topologies compatible with the given dual pair have the same collection of closed convex sets.*

This result will allow us in the sequel to vary topologies of the underlying spaces involved in the definition of an LP without altering the programme's structure.

1.1.1 Linear Programme formulation

Let us now introduce another dual pair of topological vector spaces (Z, W). As before a continuous linear functional on Z is defined as $w(z) := \langle z, w \rangle_{ZW}$ for $w \in W$, a weak topology on Z is denoted by $\sigma(Z, W)$ and the weak* topology on W is denoted by $\sigma(W, Z)$.

Fix elements $c \in Y$, $b \in Z$ and let $A : X \to Z$ be a linear map between two spaces. The primal problem is defined as

$$\mathrm{val}(P) := \inf \left\{ \langle x, c \rangle_{XY} : x \in X,\ Ax - b \in Z_+ \right\}. \tag{1.2}$$

This formulation corresponds to inequality constrained primal problem in (Anderson and Nash, 1987, Section 3.3). The dual problem is then formulated as an equality constrained problem

$$\mathrm{val}(D) := \sup \left\{ \langle b, w \rangle_{ZW} : w \in Z_+^*,\ A^* w = c \right\}, \tag{1.3}$$

where the linear map $A^* : W \to Y$ is an adjoint operator to A defined implicitly be relation $\langle Ax, w \rangle_{ZW} = \langle x, A^* w \rangle_{XY}$. We can also form a double dual problem as follows

$$\text{val}(DD) := \inf \left\{ \langle x, c \rangle_{XY} : x \in X,\ Ax - b \in Z_+^{**} \right\}, \tag{1.4}$$

As $Z_+ \subseteq Z_+^{**}$ the feasible set of the primal problem (1.2) is a subset of the feasible set of the double dual problem (1.4) and thus it follows that $\text{val}(P) \geq \text{val}(DD)$ where the equality is attained if Z_+ is $\sigma(Z, W)$-closed by Proposition 1.1. In this case the double dual programme is equal to the primal problem and the primal (1.2) and the dual (1.3) are symmetric, i.e. one is exactly the dual problem of the other.

We say that a problem is consistent if its feasible set is non-empty and any element of the feasible set is called a feasible solution. Note that the value of the primal problem (1.2) is always greater or equal to the value of the dual problem (1.3), i.e. weak duality always holds. Indeed, for any feasible pair of solutions (x_0, w_0) the following holds

$$\langle x_0, c \rangle_{XY} = \langle x_0, A^* w_0 \rangle_{ZW} = \langle Ax_0 - b + b, w_0 \rangle_{ZW} \geq \langle b_0, w \rangle_{ZW}.$$

If the inequality is strict we say there exists a duality gap between the primal and the dual programmes. On the other hand if the values of the two programmes are equal to each other then we say that there is no duality gap and moreover we will say that the strong duality holds if the optimal values for both programmes exist. One of the most basic ways to verify that the strong duality holds between the two programmes is complementary slackness, i.e. if for any feasible pair of solutions (x_0, w_0) the following equality holds

$$\langle Ax_0 - b, w_0 \rangle_{ZW} = 0,$$

then x_0 is optimal for the primal problem (1.2) and w_0 is optimal for the dual problem (1.3), (Anderson and Nash, 1987, Theorem 3.2).

Following Clark (2003) we define the range $M := \{Ax : x \in X\}$ of the linear map $A : X \to Z$ and the valuation functional $\rho : M \to \mathbb{R}$ defined by $\rho(m) = \{\langle x, c \rangle_{XY} : Ax = m\}$. For now we assume that ρ is a well-defined, linear functional and $\rho(m)$ is a singleton for each $m \in M$. In the sequel we will show that certain geometric conditions ensure that it is indeed the case. We call an element $w \in W$ a positive linear extension of ρ if $\langle m, w\rangle_{ZW} = \rho(m)$ for all $m \in M$ and $w \in Z_+^*$. It turns out that constructing a positive linear extension w of the functional ρ to the whole space Z is equivalent to finding a feasible solution to the dual problem (1.3). If such extension exists then ρ is itself linear as shown in Lemma 1.4 below.

Lemma 1.4. *(Clark, 2003, Lemma 1) An element $w \in W$ is a feasible solution to the dual problem* (1.3) *if and only if it is a positive linear extension of the functional ρ.*

Outline of proof. For any $m \in M$ and $w \in W$ one has

$$\langle m, w\rangle_{ZW} = \langle Ax, w\rangle_{ZW} = \langle x, A^*w\rangle_{XY}.$$

If w is a feasible solution to the dual problem (1.3) then $\langle x, A^*w\rangle_{XY} = \langle x, c\rangle_{XY} = \rho(m)$ by definition. On the other hand if $w \in Z_+^*$ is a positive linear extension of ρ then $\langle x, A^*w\rangle = \langle Ax, w\rangle = \rho(Ax) = \langle x, c\rangle$ and hence $A^*w = c$, i.e. w is feasible. $\square$

The topological duality proposed by Clark (Clark, 2003, Section 4) seeks to re-define the primal (1.2) and the dual (1.3) problems by closing the range M of the operator A in $\sigma(Z, W)$ topology. Existence of a positive linear extension $\overline{\rho}$ of ρ to the $\sigma(Z, W)$-closure $\bar{M}$ is discussed in the sequel. The primal problem is reformulated as follows

$$\mathrm{val}(P) = \inf\left\{\overline{\rho}(m) : m \in \bar{M} \text{ and } m - b \in Z_+\right\} \tag{1.5}$$

and the dual problem now reads

$$\mathrm{val}(D) = \sup\left\{\langle b, w\rangle_{ZW} : w \in Z_+^* \text{ and } w(m) = \overline{\rho}(m) \text{ for all } m \in \bar{M}\right\} \tag{1.6}$$

It is clear that the value of the programme (1.5) is smaller than the value of the programme (1.2) as the the feasible set is bigger and equality is achieved if M is closed in the weak topology $\sigma(Z, W)$. On the other hand $\{w \in Z_+^* : w(m) = \overline{\rho}(m) \text{ for all } m \in \bar{M}\} = \{w \in Z_+^* : A^*w = c\}$ by Lemma 1.4 and continuity of w. Thus the values of the dual programmes (1.3) and (1.6) coincide.

1.2 Market set-up

Let us now introduce the market and its structure. We shall define the market in abstract terms similarly to Kreps (1981). From now on assume that the underlying state space $\Omega := \mathbb{R}^n$ where $n < \infty$. Let us assume we are given an index set $\mathcal{I}$ (not necessarily finite) and a collection of functions $\varphi_i \in C(\Omega)$ for each $i \in \mathcal{I}$ representing pay-offs of securities available on the market at finite prices $c_i \in \mathbb{R}$. We assume that the market is frictionless, i.e. there are no transaction costs associated with buying and selling securities, there are no liquidity constraints and market participants are allowed to buy and sell any position in a security or a portfolio of securities. Denote by $\mathfrak{M}$ the space of traded claims, i.e. the set of portfolios of securities that can be bought and sold freely on the market, as

$$\mathfrak{M} := \left\{ \sum_{n=1}^{N} \alpha_n \varphi_{i_n} : N \in \mathbb{N} \text{ and } i_1, \ldots, i_N \in \mathcal{I} \right\}.$$

Note that $\mathfrak{M}$ is a linear subspace of $C(\Omega)$ as the trading is frictionless. We also define a pricing functional $\rho : \mathfrak{M} \to \mathbb{R}$ that maps the pay-offs to their market prices

$$\rho(m) := \left\{ \sum_{n=1}^{N} \alpha_n c_{i_n} : m = \sum_{n=1}^{N} \alpha_n \varphi_{i_n} \text{ for some } N \in \mathbb{N},\, i_1, \ldots, i_N \in \mathcal{I} \right\}. \tag{1.7}$$

We note that the pricing functional is a set function. However we show below that absence of arbitrage is equivalent to certain properties of the pricing functional, including ρ being single-valued. Before we proceed we make a regularity assumption on the market so that we can establish separating duality in the sequel.

Assumption 1. There exists a continuous function $h : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ such that its level sets $\{\omega \in \Omega : h(\omega) \leq c\}$ are compact for any $c < +\infty$ and $1/h$

is bounded on Ω. We assume that $\varphi_i = \mathcal{O}(h)$ for all $i \in \mathcal{I}$.

Let us now define a weighted space

$$C_h(\Omega) := \left\{ f \in C(\Omega) : \|f\|_h := \sup_{\omega\in\Omega} \frac{|f(\omega)|}{h(\omega)} < \infty \right\},$$

with a positive cone $(C_h)_+(\Omega)$. Assumption 1 implies that $\mathfrak{M} \subset C_h(\Omega)$. Note that $C_h(\Omega)$ endowed with $\|\cdot\|_h$ is a Banach lattice and the order unit in $C_h(\Omega)$ is h, see Appendix A. The negative polar of the positive cone $(C_h)_+(\Omega)$ can be identified with the space of Borel probability measures that integrate h to a finite constant

$$\mathcal{P}_h(\Omega) := \left\{ \mu \in \mathcal{P}(\Omega) : \int_\Omega h(\omega)\mu(\mathrm{d}\omega) < \infty \right\},$$

see Appendix B. The bilinear form $\langle \cdot, \cdot \rangle : C_h(\Omega) \times \mathcal{P}_h(\Omega) \to \mathbb{R}$ is defined

$$\langle f, \pi \rangle := \int_\Omega f(\omega)\pi(\mathrm{d}\omega),$$

for all $f \in C_h(\Omega)$ and $\pi \in \mathcal{P}_h(\Omega)$. We shall also sometime use notation $\pi(f) = \langle f, \pi \rangle$. We now define a notion of arbitrage in this abstract market.

Definition 1.5. We say that ρ does not admit strong model-independent arbitrage on $\mathfrak{M}$ if $\inf \rho(m) \geq 0$ for all $m \in \mathfrak{M}_+$, and $\inf \rho(m) > 0$ for all $m \in \mathfrak{M}_{++}$.

Let us define a set of traded claims available at non-positive prices as

$$\mathfrak{F} := \{m \in \mathfrak{M} : \sup \rho(m) \leq 0\}. \tag{1.8}$$

We can equivalently state absence of strong model-independent arbitrage as $\mathfrak{F} \cap (C_h)_+(\Omega) = \{0\}$. In order to avoid a degenerate situation where $\rho(m) = 0$ for all $m \in \mathfrak{M}_+$ we make the following assumption.

Assumption 2. There exists a traded claim $m_0 \in \mathfrak{M}$ with $m_0(\omega) > 0$ for all $\omega \in \Omega$ such that $\rho(m_0) > 0$.

In particular Assumption 2 is satisfied when there is a riskless bond available on the market. In general ρ is a set-valued function as mentioned above. However following (Clark, 1993, Theorem 3) we show that absence of strong model-independent arbitrage is equivalent to the pricing functional being strictly positive, linear and uniquely defined for each $m \in \mathfrak{M}$.

Proposition 1.6. *(Clark, 1993, Theorem 3) Suppose Assumption 2 holds. Then absence of strong model-independent arbitrage holds if and only if the functional $\rho : \mathfrak{M} \to \mathbb{R}$ defined in (1.7) is strictly positive, linear and uniquely defined for each $m \in \mathfrak{M}$.*

Outline of the proof. When ρ is a strictly positive linear functional uniquely defined on $\mathfrak{M}$ the absence of strong model-independent arbitrage is immediate.

On the other hand assume that there is no strong model-independent arbitrage. Assume that ρ is not uniquely determined for each $m \in \mathfrak{M}$. Then there exists $\hat{m} \in \mathfrak{M}$ such that $p, p^* \in \rho(\hat{m})$ with $p < p^*$. Set up a portfolio such that one unit of $\hat{m}$ is sold for p^*, buy one unit of $\hat{m}$ for p and buy $(p^* - p - \varepsilon)/\rho(m_0)$ units of m_0, where $0 < \varepsilon < p^* - p$. The immediate profit is ε and at maturity the portfolio yields $(p^* - p - \varepsilon)/\rho(m_0)m_0$ which is positive for all $\omega \in \Omega$. Thus we constructed a strong model-independent arbitrage opportunity, a contradiction. Linearity of ρ follows by similar arguments. The functional ρ is clearly positive, for if there were an element $m \in \mathfrak{M}_+$ such that $\rho(m) < 0$ then it would constitute a strong model-independent arbitrage. Now assume that ρ is not strictly positive. Then there exist $m_1, m_2 \in \mathfrak{M}_{++}$ with $m_1 - m_2 > 0$ but $\rho(m_1) = c = \rho(m_2)$ for some non-negative real constant c. Then $\rho(m_1 - m_2) = 0$ and hence $\{0, m_1 - m_2\} \subset \mathfrak{F} \cap \mathfrak{M}_+$, a contradiction. $\square$

Proposition 1.6 implies that $\rho(0) = 0$ and as it is strictly positive and unique on $\mathfrak{M}$ we can show the following representation of $\mathfrak{M}$.

Lemma 1.7. *Under Assumption 2 the space of traded claims $\mathfrak{M}$ can be represented as a linear span of m_0 and the set $\mathfrak{F}$, i.e. $\mathfrak{M} = \text{Span}\,\{m_0, \mathfrak{F}\}$.*

Proof. It is immediate to see that $\text{Span}\,\{m_0, \mathfrak{F}\} \subseteq \mathfrak{M}$. On the other hand for any $m \in \mathfrak{M}$ available at price $\rho(m)$ define $f := m - [\rho(m)/\rho(m_0)]m_0$ with

$\rho(f) = 0$ and thus $f \in \mathfrak{F}$. Then m can be trivially represented as a linear combination $f + [\rho(m)/\rho(m_0)]m_0$ and the reverse inclusion follows. $\square$

1.3 Fundamental Theorems of Asset Pricing

In this section we prove Fundamental Theorem of Asset Pricing in two separate case: when the function h defined in Assumption 1 is a traded asset and when it is not. The second case requires a stronger notion of arbitrage that we shall present in the sequel.

1.3.1 The function h is a traded asset

We begin with the following assumptions.

Assumption 3. The function h defined in Assumption 1 belongs to the subspace of traded claims $\mathfrak{M}$ and $\rho(h) > 0$.

Assumption 4. If there exists a continuous linear extension $\pi : C_h(\Omega) \to \mathbb{R}$ of ρ, then for all $(f_n)_{n\in\mathbb{N}} \in C_h(\Omega)$ decreasing pointwise to 0 the following limit exists:

$$\lim_{n\to\infty} \pi(f_n) = 0$$

Theorem 1.8. *Suppose Assumptions 1, 2 and 3 hold. Then absence of strong model-independent arbitrage holds if and only if there exists a strictly positive linear extension $\pi : C_h(\Omega) \to \mathbb{R}$ of the functional ρ such that $\pi(m) = \rho(m)$ for all $m \in \mathfrak{M}$. Moreover if Assumption 4 holds, then π can be represented as an integral with respect to a unique Borel probability measure $\mu \in \mathcal{P}_h(\Omega)$.*

Proof. If there exists a strictly positive linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ satisfying conditions in the statement of the theorem, then absence of strong model-independent arbitrage is ensured by Proposition 1.6.

On the other hand assume there is no strong model-independent arbitrage. Then ρ is linear and strictly positive by Proposition 1.6. Let us define an isometry $T : C_h(\Omega) \to C_b(\Omega)$ with $T(f) = f/h$. The space $C_h(\Omega)$ can be identified with $C_b(\Omega)$. Define $\mathfrak{M}_h := \{m/h : m \in \mathfrak{M}\} \subset C_b(\Omega)$, the set of traded claims as a subset of $C_b(\Omega)$. Define a strictly positive linear functional $\rho_h : \mathfrak{M}_h \to \mathbb{R}$ as $\rho_h(f) := \rho(m)/\rho(h)$ for all $f \in \mathfrak{M}_h$ (note that by Assumption 3 the function $h \in \mathfrak{M}$). The functional ρ_h possesses the same properties as ρ and $\rho_h(h) = 1$. Let us now define a convex functional

$$p(f) := \inf \{\rho_h(m) : m \in \mathfrak{M}_h,\, m - f \in (C_b)_+(\Omega)\}.$$

As $\mathbb{1}_\Omega = h/h \in \mathfrak{M}_h$ which is an order unit of the vector lattice $C_b(\Omega)$ it follows that p is a proper convex function and $p(m) = \rho_h(m)$ for all $m \in \mathfrak{M}_h$. Thus by Hahn-Banach extension Theorem there exists a linear functional $\tilde{\pi} : C_b(\Omega) \to \mathbb{R}$ that extends ρ_h and is dominated by p. For any $f \in (C_b)_+(\Omega)$ we have that $\tilde{\pi}(-f) \leq p(-f) \leq \rho(0) = 0$ and hence $\tilde{\pi}$ is a positive linear functional. As $C_b(\Omega)$ is a Banach lattice it follows that $\tilde{\pi}$ is continuous by Theorem A.6. By Assumption 4 and Corollary B.8 it follows that $\tilde{\pi}$ can be represented as

$$\tilde{\pi}(f) := \int_\Omega f(\omega)\tilde{\mu}(\mathrm{d}\omega),$$

where $\tilde{\mu} \in \mathcal{P}(\Omega)$, a Borel probability measure. Define a measure $\mu \in \mathcal{P}_h(\Omega)$ as

$$\frac{\mathrm{d}\mu}{\mathrm{d}\tilde{\mu}} := \frac{\mathcal{C}}{h},$$

where $\mathcal{C} := \left(\int_\Omega 1/h(\omega)\tilde{\mu}(\mathrm{d}\omega)\right)^{-1}$, and a positive linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ as

$$\pi(f) := \int_\Omega f(\omega)\mu(\mathrm{d}\omega) = \int_\Omega \frac{f}{h}\mathcal{C}\tilde{\mu}(\mathrm{d}\omega) = \tilde{\pi}(f)\mathcal{C}.$$

As $\mathbb{1}_\Omega \in \mathfrak{M}$ and hence belongs to $\mathfrak{M}_h$ we have

$$1 = \pi(1) = \tilde{\pi}(1)\mathcal{C} = \frac{\rho(1)}{\rho(h)}\mathcal{C},$$

and assuming without loss of generality that $\rho(1) = 1$ we have that

$$\rho(h) = \mathcal{C}.$$

It is then easy to see that π agrees with ρ on $\mathfrak{M}$.

To see that the obtained extension π is strictly positive, it suffices to note that absence of strong model-independent arbitrage is equivalent to $(\mathfrak{F} - (C_h)_+(\Omega)) \cap (C_h)_+(\Omega) = \{0\}$ and as $0 \in \mathfrak{F}$ it follows that $\pi(f) > 0$ for all $f \in (C_h)_{++}(\Omega)$. $\square$

Remark 1.9. The weighted Banach spaces $C_h(\Omega)$ are used in Acciaio et al. (2016). The authors assume that there exists a super-linear convex function g (the function h in our notation) that belongs to the set of traded claims. They then prove Fundamental Theorem of Asset Pricing by first considering linear functionals on finite subsets of the set of traded claims and then prove that intersection of sets of such linear functionals is non-empty. Theorem 1.8 is similar in spirit when the function g in (Acciaio et al., 2016, Theorem 1.3) is replaced by the function h. Assumption 1 is weaker as it allows for functions with only linear growth, however Assumption 4 is needed to show that the extension can be represented as an integral with respect to a unique measure $\mu \in \mathcal{P}_h(\Omega)$.

1.3.2 The function h is not a traded asset

When the function h is not in the subspace $\mathfrak{M}$ of traded claims the situation is more subtle. We now introduce another notion of arbitrage.

Definition 1.10. We say there is a weak free lunch if there exists a sequence $(g_n)_{n\in\mathbb{N}} \subset C_h(\Omega)$ converging weakly to $g \in (C_h)_{++}(\Omega)$ and there is a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathfrak{F}$ with $f_n \geq g_n$ for all $n \in \mathbb{N}$.

Before we proceed let us first show an auxiliary result.

Lemma 1.11. *Let $\mathfrak{F} - (C_h)_+(\Omega) := \{f - g : f \in \mathfrak{F}$ and $g \in (C_h)_+(\Omega)\}$ be the algebraic difference of sets $\mathfrak{F}$ and $(C_h)_+(\Omega)$. Then the following equivalent representation holds*

$$\mathfrak{F} - (C_h)_+(\Omega) = \{g \in C_h(\Omega) : \text{ there exists } f \in \mathfrak{F} \text{ such that } f \geq g\}.$$

Proof. Let $G := \{g \in C_h(\Omega) :$ there exists $f \in \mathfrak{F}$ such that $f \geq g\}$. For any $g \in G$ there exists $f \in \mathfrak{F}$ such that $f - g \in (C_h)_+(\Omega)$ or equivalently $g - f \leq 0$. As $0 \in \mathfrak{F}$ we have that $0 - (f - g) \in \mathfrak{F} - (C_h)_+(\Omega)$ hence $G \subseteq \mathfrak{F} - (C_h)_+(\Omega)$.

On the other hand let $f \in \mathfrak{F}$ and $z \in (C_h)_+(\Omega)$. Let $g := f - z$ and note that $f \geq g$. Hence $g \in G$ and therefore $\mathfrak{F} - (C_h)_+(\Omega) \subseteq G$. $\square$

Note that the result of Lemma 1.11 still applies if the positive cone $(C_h)_+(\Omega)$ is restricted to $\mathfrak{M}_+$. It follows that Definition 1.10 can equivalently be stated as $\overline{(\mathfrak{F} - (C_h)_+(\Omega))} \cap (C_h)_+(\Omega) = \{0\}$, where the closure is taken with respect to the weak topology on $C_h(\Omega)$. We now show that absence of weak free lunch is equivalent to the existence of a strictly positive continuous linear functional on $C_h(\Omega)$ that can be identified with an element of $\mathcal{P}_h(\Omega)$ via integral representation. In order to do that we replace Assumption 4 with a stronger version of Assumption 1.

Assumption 5. Let the function $h : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ be as in Assumption 1. We assume that $h \notin \mathfrak{M}$ and $\varphi_i = o(h)$ (as $\|\omega\| \to \infty$) for all $i \in \mathcal{I}$.

Note that the direct consequence of the above assumption is that $m_0 = o(h)$. We are now ready to state and prove the main theorem.

Theorem 1.12. *Suppose Assumptions 1 and 2 hold. Then absence of weak free lunch holds if and only if there exists a strictly positive linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ that extends ρ. Moreover if Assumption 5 holds then π can be represented as an integral with respect to a unique Borel probability measure $\mu \in \mathcal{P}_h(\Omega)$.*

Proof. Suppose there exists a strictly positive linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ that extends ρ. As $C_h(\Omega)$ is a Banach lattice it implies that π is continuous by Theorem A.6. It is also evident that it implies absence of weak free lunch.

Conversely, assume that there is no weak free lunch. It then follows that $m_0 \notin \overline{\mathfrak{F} - (C_h)_+(\Omega)}$ by Assumption 2 (recall that $m_0 \in \mathfrak{M}_{++}$ such that $\rho(m_0) > 0$). As $\{m_0\}$ is compact by virtue of being finite-dimensional and $\overline{\mathfrak{F} - (C_h)_+(\Omega)}$ is closed in the weak topology, it implies by the Strong Separating Hyperplane Theorem (Aliprantis and Border, 2007, Theorem 5.79) that there exists a non-zero continuous linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ such that $\pi(m_0) > 0$ and $\pi(f - g) \leq 0$ for all $f \in \mathfrak{F}$ and $g \in (C_h)_+(\Omega)$. As $0 \in \mathfrak{F}$ it follows that $\pi(-g) \leq 0$ for all $g \in (C_h)_+(\Omega)$ and hence π is positive. Moreover

$\pi(g) > 0$ for all $g \in (C_h)_{++}(\Omega)$. Otherwise there exists $g \in (C_h)_{++}(\Omega)$ such that $\pi(g) = 0$, i.e. $g \in \mathfrak{F}$ and hence $g \in \overline{(\mathfrak{F} - (C_h)_+(\Omega))} \cap (C_h)_+(\Omega)$ which contradicts absence of weak free lunch. Similarly as $0 \in (C_h)_+(\Omega)$ one has that $\pi(f) \leq 0$ for all $f \in \mathfrak{F}$. Therefore there exists a constant $\xi \in \mathbb{R}$ such that $\xi\pi(m) = \rho(m)$ for all $m \in \mathfrak{M}$. As $\xi\pi(m_0) = \rho(m_0) > 0$ implies that $\xi > 0$ and without loss of generality one can take $\xi = 1$. Thus we have shown existence of a strictly positive continuous and linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ that extends ρ.

Now suppose Assumption 5 holds. The map $T : C_h(\Omega) \to C_b(\Omega)$ defined by $T(f) := f/h$ is an isometry between two spaces. Define a functional $\tilde{\pi} : C_b(\Omega) \to \mathbb{R}$ by $\tilde{\pi}(f) := \mathcal{C}\pi(T^{-1}(f))$ for all $f \in C_b(\Omega)$, where $\mathcal{C}$ is a strictly positive constant. Note that $\tilde{\pi}$ is continuous, linear and strictly positive by definition. The space $C_b(\Omega)$ can be identified with $C(\breve{\Omega})$, the space of continuous functions on $\breve{\Omega}$ which is the Stone-Čech compactification of Ω, see Appendix C. As the dual of $C(\breve{\Omega})$ can be identified with the space of regular signed Borel measures of bounded variation (Aliprantis and Border, 2007, Theorem 14.12), the following representation holds

$$\tilde{\pi} \circ T(f) = \int_{\breve{\Omega}} \breve{T}(f)(\omega)\nu(\mathrm{d}\omega),$$

where $\breve{T}(f)$ is the unique extension of $T(f) \in C_b(\Omega)$.

Observe that ν is positive as $0 < \mathcal{C}\pi(f) = \tilde{\pi} \circ T(f) = \int_{\breve{\Omega}} \breve{T}(f)(\omega)\nu(\mathrm{d}\omega)$ for all $f \in (C_h)_{++}(\Omega)$. Let $\nu = \nu^r + \nu^s$ where ν^r is a measure with support in Ω and ν^s is a measure with support in $\breve{\Omega} \setminus \Omega$. For each $i \in \mathcal{I}$ the extension $\breve{T}(\varphi_i)$ is continuous (see Appendix C) and hence by Assumption 5 we have that $\breve{T}(\varphi_i)(\omega) = 0$ for all $\breve{\Omega} \setminus \Omega$. Therefore we have

$$\begin{aligned}
\tilde{\pi} \circ T(\varphi_i) &= \int_{\breve{\Omega}} \breve{T}(f)(\omega)\nu(\mathrm{d}\omega) \\
&= \int_{\Omega} \breve{T}(f)\nu^r(\mathrm{d}\omega) + \int_{\breve{\Omega}\setminus\Omega} \breve{T}(f)\nu^s(\mathrm{d}\omega) \\
&= \int_{\Omega} T(f)\nu^r(\mathrm{d}\omega),
\end{aligned}$$

for all $i \in \mathcal{I}$. The last equality follows from the fact that the extension $\breve{T}(f)$ coincides with $T(f)$ on Ω for all $f \in C_h(\Omega)$. Note also that $\nu^r \neq 0$ otherwise one would have

$$0 < \mathcal{C}\pi(m_0) = \tilde{\pi} \circ T(m_0) = \int_\Omega \breve{T}(m_0)(\omega)\nu^r(\mathrm{d}\omega) + \int_{\breve{\Omega}\setminus\Omega} \breve{T}(m_0)(\omega)\nu^s(\mathrm{d}\omega) = 0,$$

where the last equality follows from the fact that $m_0 \in o(h)$ and we arrive at a contradiction. We can then define a probability measure on Ω as $\eta := \nu^r/\|\nu^r\|$ and

$$\tilde{\pi} \circ T(f) = \int_\Omega f(\omega)\eta(\mathrm{d}\omega),$$

for all $f \in C_h(\Omega)$. Moreover defining a probability measure μ such that

$$\frac{\mathrm{d}\mu}{\mathrm{d}\eta} := \frac{1}{h}\left(\int_\Omega \frac{1}{h(\omega)}\eta(\mathrm{d}\omega)\right)^{-1},$$

and setting

$$\mathcal{C} := \int_\Omega \frac{1}{h(\omega)}\eta(\mathrm{d}\omega),$$

we see that $\mu \in \mathcal{P}_h(\Omega)$ and

$$\pi(f) = \int_\Omega f(\omega)\mu(\mathrm{d}\omega),$$

for all $f \in C_h(\Omega)$. □

Remark 1.13. It must be noted that when Assumption 3 holds then absence of strong model-independent arbitrage implies absence of weak free-lunch, as by Theorem 1.8 there exists a strictly positive linear functional that extends ρ, separating $\mathfrak{F}$ and $(C_h)_{++}(\Omega)$ and hence $\mathfrak{F} - (C_h)_+(\Omega) \cap (C_h)_{++}(\Omega) = \emptyset$. Then by continuity the functional also separates the closure $\overline{\mathfrak{F} - (C_h)_+(\Omega)}$ and $(C_h)_{++}(\Omega)$.

Let us now define an extension of ρ to the weak closure of the set of traded claims $\bar{\mathfrak{M}}$. For a sequence $(m_n)_{n\in\mathbb{N}} \subset \mathfrak{M}$ converging weakly to $m \in \bar{\mathfrak{M}}$ define $\overline{\rho}(m) := \lim_{n\to\infty} \rho(m_n)$. In comparison with the definition of $\overline{\rho}$ introduced in Clark (1993), where nets were used, it is sufficient to consider sequences only as the space $C_h(\Omega)$ is metrizable.

By definition $\overline{\rho}$ is a continuous functional and we now show that absence of weak free lunch implies that $\overline{\rho}$ is strictly positive and linear. It can be seen as a corollary to Theorem 1.12 and similar in flavour to (Clark, 1993, Theorem 2).

Corollary 1.14. *Let Assumptions 1 and 2 hold. Absence of weak free lunch implies that the functional $\overline{\rho} : \overline{\mathfrak{M}} \to \mathbb{R}$ is strictly positive and linear.*

Proof. By Theorem 1.12 absence of weak free lunch implies existence of a strictly positive, continuous linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ that extends ρ. It follows that ρ is continuous and it implies that $\overline{\rho} : \bar{\mathfrak{M}} \to \mathbb{R}$ is a continuous positive linear functional. Moreover it is unique as $C_h(\Omega)$ is Hausdorff in the weak topology.

To show that $\overline{\rho}$ is strictly positive we follow closely the argument presented in (Clark, 1993, Proof of Theorem 2). Assume that $\overline{\rho}$ is not strictly positive. Then there exists $m \in \bar{\mathfrak{M}}_{++}$ with $\overline{\rho}(m) = 0$ and a sequence $(m_n)_{n\in\mathbb{N}} \subset \mathfrak{M}$ converging weakly to m. Then by continuity of $\overline{\rho}$ it follows that $\lim_n \overline{\rho}(m_n) = 0$. Define $\alpha_n := \overline{\rho}(m_n)/\overline{\rho}(m_0)$, where $\overline{\rho}(m_0) = \rho(m_0) > 0$ and let $\tilde{m}_n := m_n - \alpha_n m_0$ for all $n \in \mathbb{N}$ with $\overline{\rho}(\tilde{m}_n) = 0$. We then have a sequence $(\tilde{m}_n)_{n\in\mathbb{N}} \subset \bar{\mathfrak{F}}$ that converges to m as $\lim_{n\to\infty} \alpha_n = 0$, which constitutes a weak free lunch, a contradiction. $\square$

Remark 1.15. It must be noted that absence of weak free lunch in particular implies that $\overline{(\mathfrak{F} - \mathfrak{M}_+)} \cap \mathfrak{M}_+ = \{0\}$ as $\mathfrak{M}_+ \subset (C_h)_+(\Omega)$. As $\bar{\mathfrak{M}}$ is a closed linear subspace in $C_h(X)$ one can extend ρ to $\overline{\rho}$ using arguments similar to ones used in the proof of Theorem 1.12 to show the result in Corollary 1.14. Moreover if the function h belongs to the closure of the set of marketed claims $\mathfrak{M}$ but not the set itself, then one can extend $\overline{\rho}$ from $\bar{\mathfrak{M}}$ to the whole space by Theorem 1.8.

1.4 Super-Replication Theorem

The main result of this section is based on (Clark, 2003, Theorem 3) and (Zălinescu, 2008, Proposition 2.3), however unlike the authors we employ absence of weak free lunch to show separation of subsets by a positive continuous linear functional. We prove the result in case when the order unit is not present in the set of traded claims and under assumption of absence of weak free lunch. When the order unit is present by Remark 1.13, it is sufficient to consider only absence of strong model-independent arbitrage.

We formulate the super- and sub-hedging problems as infinite-dimensional LPs. The dual problem consists of finding a Borel probability measure subject to market constraints that maximizes the price of a derivative to be hedged. We formulate the super-hedging problem for an option with pay-off represented by an upper semi-continuous function $\Phi \in U_h(\Omega)$ as

$$\overline{\vartheta}_p(\Phi) := \inf\left\{\overline{\rho}(m) : m \in \bar{\mathfrak{M}}, m(\omega) \geq \Phi(\omega), \text{ for all } \omega \in \Omega\right\}, \tag{1.9}$$

and the dual problem is formulated as follows

$$\overline{\vartheta}_d(\Phi) := \sup\left\{\langle \Phi, \mu\rangle : \mu \in \mathcal{P}_h(\Omega), \langle m, \mu\rangle = \overline{\rho}(m), m \in \bar{\mathfrak{M}}\right\}. \tag{1.10}$$

On the other hand the sub-hedging problem for an option with a lower semi-continuous pay-off $\Phi \in L_h(\Omega)$ can be stated as

$$\underline{\vartheta}_p(\Phi) := \sup\left\{\overline{\rho}(m) : m \in \bar{\mathfrak{M}}, m(\omega) \leq \Phi(\omega), \text{ for all } \omega \in \Omega\right\}, \tag{1.11}$$

and its dual problem is written as follows

$$\underline{\vartheta}_d(\Phi) := \inf\left\{\langle \Phi, \mu\rangle : \mu \in \mathcal{P}_h(\Omega), \langle m, \mu\rangle = \overline{\rho}(m), m \in \bar{\mathfrak{M}}\right\}. \tag{1.12}$$

It is easily seen that the weak duality holds, i.e. $\underline{\vartheta}_p(\Phi) \leq \underline{\vartheta}_d(\Phi)$ for $\Phi \in L_h(\Omega)$ and $\overline{\vartheta}_d(\Phi) \leq \overline{\vartheta}_p(\Phi)$ for $\Phi \in U_h(\Omega)$. Moreover, the inequality $\underline{\vartheta}_d(\Phi) \leq \overline{\vartheta}_d(\Phi)$ holds when $\Phi \in C_h(\Omega)$. As the function h might not necessarily be a traded claim, we make the following assumption on the pay-off Φ to avoid degeneracy of the primal problem (1.9).

Assumption 6. For a fixed pay-off $\Phi \in U_h(\Omega)$, there exists $m \in \bar{\mathfrak{M}}$ such that $m(\omega) \geq \Phi(\omega)$ for all $\omega \in \Omega$, i.e. $\underline{\vartheta}_p(\Phi) < \infty$.

By equality $\underline{\vartheta}_p(-\Phi) = -\overline{\vartheta}_p(\Phi)$ we see that the sub-hedging problem (1.11) is feasible for a pay-off Φ if $-\Phi$ satisfy Assumption 6.

Theorem 1.16. *Suppose Assumptions 2 and 5 hold. Then under Assumption 6 absence of weak free lunch implies no duality gap between the primal* (1.9) *and the dual* (1.10) *problems.*

Proof. Let us first specialise to the case where $\Phi \in C_h(\Omega)$. Absence of weak free lunch and Assumption 5 imply existence of a strictly positive linear extension $\pi_0 : C_h(\Omega) \to \mathbb{R}$ of $\overline{\rho}$ by Theorem 1.12. Moreover it can be represented as an integral with respect to a unique Borel probability measure $\mu_0 \in \mathcal{P}_h(\Omega)$. It is clear that $\underline{\vartheta}_p(\Phi) \leq \overline{\vartheta}_p(\Phi)$. If $\Phi \in \bar{\mathfrak{M}}$ then $\overline{\vartheta}_p(\Phi) = \underline{\vartheta}_p(\Phi)$ and hence there is no duality gap between the primal (1.9) and the dual (1.10) programmes.

Assume $\Phi \notin \bar{\mathfrak{M}}$ and fix some $\alpha \in (\pi_0(\Phi), \overline{\vartheta}_p(\Phi))$. Let $L := \text{Span}\ \{\bar{\mathfrak{M}}, \Phi\} \subset C_h(\Omega)$ and note that as $\bar{\mathfrak{M}}$ is a subspace then any $l \in L$ can be represented as $l = m + \lambda\Phi$ for some $m \in \bar{\mathfrak{M}}$ and $\lambda \in \mathbb{R}$. Let us now define a functional $\eta : L \to \mathbb{R}$ as $\eta(l) = \eta(m + \lambda\Phi) := \overline{\rho}(m) + \lambda\alpha$. Clearly η is linear and we now show that η is strictly positive on $L_{++} := L \cap (C_h)_{++}(\Omega)$. Let $z = m + \lambda\Phi \in L_{++}$ where $m \in \bar{\mathfrak{M}}$ and $\lambda \in \mathbb{R}$ and consider three cases. If $\lambda = 0$ then $\eta(z) = \overline{\rho}(m) > 0$. On the other hand if $\lambda < 0$ then $m > -\lambda\Phi$ and $\overline{\rho}(m/(-\lambda)) \geq \overline{\vartheta}_p(\Phi) > \alpha$ by assumption. Then $\eta(z) = \overline{\rho}(m) + \lambda\alpha = -\lambda((-\lambda)^{-1}\overline{\rho}(m) - \alpha) > 0$. Finally if $\lambda > 0$ then by above we have $(-\lambda)^{-1}\overline{\rho}(m) < \alpha$ and $\eta(z) > 0$.

We now define a set $G := \{l \in L : \eta(l) \leq 0\}$. Observe that $G \cap (C_h)_+(\Omega) = \{0\}$ as η is strictly positive. We now show that $m_0 \notin \overline{G - (C_h)_+(\Omega)}$. On the contrary assume that $m_0 \in \overline{G - (C_h)_+(\Omega)}$. Then there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset C_h(\Omega)$ such that $f_n \to m_0$ as $n \to \infty$ such that $(g_n)_{n\in\mathbb{N}} \subset G$ with $g_n = m_n + \lambda_n\Phi$ for $(m_n)_{n\in\mathbb{N}} \subset \bar{\mathfrak{M}}$, $(\lambda_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ and $g_n \geq f_n$ for all $n \in \mathbb{N}$. Clearly the following holds $m_n + \lambda_n\Phi - m_0 \geq f_n - m_0 \to 0$ as $n \to \infty$ and hence $\liminf_n \eta(m_n + \lambda_n\Phi - m_0) \geq 0$ or equivalently $\limsup_n -\eta(g_n) + \overline{\rho}(m_0) \leq 0$. Thus we obtain that $0 \geq \liminf_n \eta(g_n) \geq \overline{\rho}(m_0) > 0$, a contradiction.

Therefore there exists a non-zero continuous linear functional $\pi : C_h(\Omega) \to \mathbb{R}$ such that $\pi(m_0) > 0 \geq \pi(g - f)$ for all $g \in G$ and $f \in (C_h)_+(\Omega)$ and by argument used in the proof of Theorem 1.12 we can see that π extends the functional η, i.e. $\pi(l) = \eta(l) = \overline{\rho}(m) + \lambda\alpha$ for all $l \in L$. In particular π extends $\overline{\rho}$ and hence is a feasible solution to the dual programme (1.10) and $\pi(\Phi) = \alpha$. Moreover as π is a feasible solution it follows that $\alpha \leq \overline{\vartheta}_d(\Phi)$. As $\alpha \in (\pi_0(\Phi), \overline{\vartheta}_p(\Phi))$ was chosen arbitrarily it implies that $\overline{\vartheta}_d(\Phi) = \overline{\vartheta}_p(\Phi)$.

Let Φ now be an upper semi-continuous function with $\Phi = \mathcal{O}(h)$. Then Φ can be expressed as an infimum over continuous functions f_n for $n \in \mathbb{N}$ that dominate Φ and by Assumption 6 we can take $(f_n)_{n \in \mathbb{N}}$ such that $\overline{\vartheta}_p(f_n) < \infty$ for all $n \in \mathbb{N}$. As it was shown above the no-duality gap holds for all $f \in C_h(\Omega)$ with $\overline{\vartheta}_p(f) < \infty$ and hence duality result carries over to the case of upper semi-continuous functions. $\square$

If $\Phi : \Omega \to \mathbb{R}$ is a lower semi-continuous function such that $\Phi = \mathcal{O}(h)$, then one can also show that there is no duality gap between problems (1.11) and (1.12).

Corollary 1.17. *Suppose Assumptions 2 and 5 hold. Assume also that Assumption 6 holds for $-\Phi$, i.e. $\underline{\vartheta}_p(\Phi) > -\infty$. Then absence of weak free lunch implies no duality gap between the primal* (1.11) *and the dual* (1.12) *problems.*

Proof. If Φ is lower semi-continuous then $-\Phi$ is upper semi-continuous and $\underline{\vartheta}_p(\Phi) = -\overline{\vartheta}_p(-\Phi)$ then the result follows by the Super-Replication Theorem 1.16. $\square$

If the function h is actually a traded asset then we have the following corollary.

Corollary 1.18. *Suppose Assumptions 1, 2, 3 and 4 hold. Then absence of strong model-independent arbitrage implies there is no duality gap between the primal* (1.9) *and the dual* (1.10) *problems.*

Proof. As $h \in \mathfrak{M}$ by Assumption 3 the primal problem (1.9) is feasible for all $\Phi \in U_h(\Omega)$. Moreover Assumptions 1, 2 and 3 together with absence of strong model-independent arbitrage imply existence of a strictly positive linear functional $\pi_0 : C_h(\Omega) \to \mathbb{R}$ that extends ρ and Assumption 4 implies that it can be represented as an integral with respect to a unique Borel probability measure $\mu_0 \in \mathcal{P}_h(\Omega)$ by Theorem 1.8. It also implies absence of weak free lunch by reverse implication of Theorem 1.12 and hence the result follows by the Super-Replication Theorem 1.16. $\square$

Similar result holds for the sub-hedging primal (1.11) and the dual (1.12) problems. We conclude this chapter with a remark.

Remark 1.19. It must be noted that if h is not a traded asset then Assumptions 5 and 6 imply that $\Phi = o(h)$.

2

Duality Theory in Markets with Call Options

In this chapter we apply general theory presented in Chapter 1 to markets with traded European Call options. We start discussion by reviewing the Black-Scholes formula and the notion of implied variance in Section 2.1.1. Bounds on the total implied variance introduced in Lee (2004) are presented in Section 2.1.2 along with their refinement by Benaim and Friz (2008, 2009). One of the consequences of the bounds is that the total implied variance is at most linear in log-moneyness and any extrapolation of the total implied variance that satisfies this condition is free of static arbitrage. In Section 2.2 we discuss how notions of arbitrage presented in Chapter 1 apply in case where European Call options are traded. We relax the notion of strong model-independent arbitrage in Section 2.2.2 and introduce weak arbitrage similar to Cox and Obłój (2011b, Definition 2.3). The pricing functional defined in (1.7) is required to only be positive in this case. We prove the duality result by modifying the positive cone of claims similar to Clark (2006). In the case when only finitely many options are traded, the set of market models is not closed under no weak arbitrage and we discuss how extrapolation of the total implied variance discussed in Section 2.1.3 is used to remedy this shortcoming in Section 2.2.4.

2.1 Black-Scholes formula and implied volatility

2.1.1 Black-Scholes formula

For a European Call or Put option with strike $\mathfrak{K}$ and maturity T the Black-Scholes formula allows us to compute its price explicitly. Let us define $D(0,t)$ to be the price at time 0 of a bank account that pays one unit of currency at time t. The value of the bank account at time t paying at time $T > t$ is $D(t,T) := D(0,T)/D(0,t)$. Define $\mathrm{Fwd}(0,t)$ to be the fair price of a forward contract struck at time 0 that delivers one unit of asset $\tilde{S}$ at time t. The price of a forward contract for immediate delivery equals the spot price of the asset, i.e. $\mathrm{Fwd}(0,0) = S_0$ and the price of a forward contract at time t delivered at a later time T is $\mathrm{Fwd}(t,T) := S_t \mathrm{Fwd}(0,T)/\mathrm{Fwd}(0,t)$. We also denote by $S_t := \tilde{S}_t/\mathrm{Fwd}(0,t)$ the normalised asset price at time t. Under the risk-neutral measure $\mathbb{Q}$ the asset price S is the exponential martingale $S_t = \exp(-\sigma^2/2t + \sigma W_t)$, where $(W_t)_{t\geq 0}$ is a $\mathbb{Q}$-Brownian motion, σ denotes the volatility parameter. The Black-Scholes formula for the normalised price of a European Call option (we normalise option prices by dividing by $D(t,T)\mathrm{Fwd}(t,T)$) at time $t < T$ is then as follows

$$c_{BS}(k,a) = \Phi(d(k,a)) - e^k \Phi(d(k,a) - a), \tag{2.1}$$

where $k := \log(\mathfrak{K}/\mathrm{Fwd}(t,T))$ denotes the forward log-forward moneyness of the option (forward moneyness is then $K := \exp(k)$), $a := \sigma\sqrt{T-t}$ denotes the total volatility and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. The function $d : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is defined as follows

$$d(k,a) = \frac{a}{2} - \frac{k}{a}. \tag{2.2}$$

In practice traded options exhibit a range of volatilities that cannot be explained by the Black-Scholes model. This inefficiency resulted in development of stochastic volatility models (see Gatheral (2006), Fouque et al. (2011) for detailed accounts). The correlation between the asset price and its instantaneous volatility in stochastic volatility models makes the market incomplete and thus not all derivatives can be perfectly replicated. Nonetheless, the validity of the formula (2.1) is not nullified in practice. Instead it can be viewed as a function that maps the price of a Call option, denoted $c(k,T)$, given market data k, T and S to the total volatility parameter a, called the total implied volatility $I(k,T-t)$ which in general depends on the log-moneyness of the option and time to maturity (we write $I(k) := I(k,T-t)/\sqrt{T-t}$ for implied volatility). In particular the total implied volatility $I(k,T)$ is the unique solution the following equation

$$c(k,T) = c_{BS}(k, I(k,T)). \tag{2.3}$$

The two-dimensional map $w : \mathbb{R} \times \mathbb{R}_+ \to (0,\infty)$ is called the total implied variance surface and defined as $w(k,t) := I^2(k,t)$. For a fixed maturity t we shall call $w(\cdot,t) : \mathbb{R} \to (0,\infty)$ a slice of the total implied variance surface. We shall assume it satisfies the following conditions.

Assumption 7. For a fixed $k \in \mathbb{R}$ the function $w(k,\cdot) \in C^1(\mathbb{R}_+)$. For a fixed maturity $t > 0$ the function $w(\cdot,t) \in C(\mathbb{R})$ is strictly positive, differentiable except possibly at finitely many points and its first derivative $\partial_k w(k,t)$ is an essentially bounded measurable function.

2.1.2 Lee's moment formula

Having defined the total implied volatility in Equation (2.3), it is possible to study the behaviour of $I(k,T)$ as $k \to \pm\infty$. In his seminal work Lee (2004) showed that Call and Put option prices are bounded above by a function of k

Theorem 2.1. *(Lee, 2004, Theorem 2.1) Fix $T \geq 0$. For each $p > 0$ and for all $k > -\infty$ the Call price bound reads*

$$c(k,T) \leq \frac{1}{(1+p)} \mathbb{E}^{\mathbb{Q}} \left\{ S_T^{1+p} \right\} \left(\frac{p}{1+p} \right)^p e^{-pk}. \tag{2.4}$$

Similarly, for each $q > 0$ and for all $k > -\infty$ the Put price bound reads

$$p(k,T) \leq \frac{1}{(1+q)}\mathbb{E}^{\mathbb{Q}}\left\{S_T^{-q}\right\}\left(\frac{q}{1+q}\right)^q e^{(1+q)k}. \tag{2.5}$$

Bounds (2.4) and (2.5) are non-trivial only if S_T possesses some moments higher than 1 as and some negative moments smaller than 0 respectively.

We now provide heuristic argument for the asymptotic behaviour of total implied volatility. Let us assume that the total implied volatility $I(\cdot,T)$ behaves like $\sqrt{\beta|k|}$ as $k \to \pm\infty$ for some $\beta \geq 0$. Then plugging it in the Black-Scholes formula (2.1) yields the following

$$c_{BS}(k,\sqrt{\beta k}) = \Phi(d(k,\sqrt{\beta k})) - e^k\Phi(d(k,\sqrt{\beta k}) - \sqrt{\beta k}),$$

with $d(k,\sqrt{\beta k}) = \sqrt{\beta k}/2 - \sqrt{k/\beta} = -\sqrt{k}(1/\sqrt{\beta} - \sqrt{\beta}/2) =: -f_1(\beta)\sqrt{k}$ and $d(k,\sqrt{\beta k}) - \sqrt{\beta k} = -\sqrt{k}(1/\sqrt{\beta} + \sqrt{\beta}/2) =: -f_2(\beta)\sqrt{k}$. Using asymptotic behaviour of the standard normal cumulative distribution function

$$\Phi(z) \sim \frac{e^{-z^2/2}}{\sqrt{2\pi}z}, \quad z \to \infty,$$

the expression for the Call price can be written as follows

$$c_{BS}(k,\sqrt{\beta k}) \sim \frac{1}{\sqrt{2\pi}}\left(\frac{e^{-f_1^2(\beta)k/2}}{f_1(\beta)\sqrt{k}} - \frac{e^k e^{-f_2^2(\beta)k/2}}{f_2(\beta)\sqrt{k}}\right), \tag{2.6}$$

and using the following equalities

$$f_1^2(\beta) = 1/\beta + \beta/4 - 1, \quad f_2^2(\beta) - 2 = f_1^2(\beta),$$

Equation (2.6) can be re-written as

$$c_{BS}(k,\sqrt{\beta k}) \sim \sqrt{\frac{\beta}{2\pi k}}\frac{e^{-f_1^2(\beta)k/2}}{f_1(\beta)f_2(\beta)}. \tag{2.7}$$

On the other hand the bound (2.4) implies that $c_{BS}(k,\sqrt{\beta k}) = \mathcal{O}(e^{-pk})$ and matching this to the asymptotics provided in Equation (2.7) one obtains

$$p = f_1^2(\beta)/2,$$

or

$$\beta = 2 - 4(\sqrt{p^2 + p} - p).$$

Define the strictly decreasing extended real-valued function $\psi : \overline{\mathbb{R}}_+ \to [0, 2]$ as

$$\psi(x) := 2 - 4\left(\sqrt{x^2 + x} - x\right). \tag{2.8}$$

Then the moment formula for the right wing of the total implied variance is stated in the theorem below.

Theorem 2.2. *(Lee, 2004, Theorem 3.2) For any $T \geq 0$ let*

$$\tilde{p} := \sup\left\{p \in \mathbb{R}_+ : \mathbb{E}^{\mathbb{Q}}\{S_T^{1+p}\} < \infty\right\}. \tag{2.9}$$

Then

$$\limsup_{k\to\infty} \frac{w(k, T)}{k} = \psi(\tilde{p}). \tag{2.10}$$

Similar arguments can applied to Put prices to arrive at the moment formula for the left wing.

Theorem 2.3. *(Lee, 2004, Theorem 3.4) For any $T \geq 0$ let*

$$\tilde{q} := \sup\left\{q \in \mathbb{R}_+ : \mathbb{E}^{\mathbb{Q}}\{S_T^{-q}\} < \infty\right\}. \tag{2.11}$$

Then

$$\limsup_{k\to\infty} \frac{w(-k, T)}{k} = \psi(\tilde{q}). \tag{2.12}$$

These bounds imply that the total implied variance $w(\cdot, T)$ is at most linear in log-moneyness as $k \to \pm\infty$. Also the total implied variance cannot grow sub-linearly unless S_T has moments of all orders (Benaim et al., 2008, Section 2.2). One example is the Black-Scholes model where S_T has moments of all orders and $w(\cdot, T)$ is constant for each $T > 0$.

In fact these upper bounds can be shown to be the limiting behaviour of the total variance at extreme strikes. Let denote by F the distribution of log-returns $\log S_t$ under the risk-neutral measure $\mathbb{Q}$ (we write $\bar{\mathrm{F}} = 1 - \mathrm{F}$). Benaim and Friz (2009) showed that limits superior in (2.10) and (2.12) are actually limits given asymptotic behaviour of the distribution F. The authors extended their discussion to models with marginal distributions at a fixed maturity possessing known moment generating function (MGF) defined as

$$M(u) := \int_{-\infty}^{\infty} e^{ux} \mathrm{dF}(x),$$

where u is a real exponent (Benaim and Friz, 2008). We define two exponents p^* and q^* such that

$$p^* := \sup\{s \in \mathbb{R}_+ : M(s+1) < \infty\}, \quad q^* := \sup\{s \in \mathbb{R}_+ : M(-s) < \infty\}.$$

Such exponents provide bounds on the effective domain of the MGF. The authors also provide certain conditions on the behaviour of the MGF near p^* and q^* which allows them to derive smile asymptotics behaviour in terms of these exponents (Benaim and Friz, 2008).

Before we proceed let us define the notation $f \in R_0$ which stems from regular variation theory. For our present purposes it suffices to provide only basic definitions.

Definition 2.4. (Benaim and Friz, 2009, Definition 2.1) A positive real-valued measurable function f is regularly varying with index α if for all $\lambda > 0$

$$\lim_{x \to \infty} \frac{f(\lambda x)}{f(x)} = \lambda^\alpha.$$

We denote it symbolically as $f \in R_\alpha$. We say the function f is slowly varying if $f \in R_0$.

Lemma 2.5. *(Benaim and Friz, 2008, Criterion 2)*

(i) If $\log M(-q^* + s) \sim s^{-\rho} l_1(1/s)$ for some $\rho > 0$ and $l_1 \in R_0$ as $s \downarrow 0$ then

$$\log \mathrm{F}(-x) \sim -q^* x, \quad x \uparrow +\infty;$$

(ii) *If* $\log M(p^* + 1 - s) \sim s^{-\rho} l_1(1/s)$ *for some* $\rho > 0$ *and* $l_1 \in R_0$ *as* $s \downarrow 0$ *then*

$$\log \bar{\mathrm{F}}(x) \sim -(1 + p^*)x, \quad x \uparrow +\infty;$$

We now state smile asymptotics expressed as conditions on the implied variance.

Theorem 2.6. *(Benaim and Friz, 2008, Theorem 2) If* $q^* \in \mathbb{R}_+$ *and the moment generating function* M *satisfies the assumption in Lemma 2.5(i) then*

$$w(-k, T)/k \sim \psi(-\frac{\log \mathrm{F}(-k)}{k}) \sim \psi(q^*).$$

Similarly if $p^* \in \mathbb{R}_+$ *and* M *satisfies Lemma 2.5(ii) then*

$$w(k, T)/k \sim \psi(-1 - \frac{\log \bar{\mathrm{F}}(k)}{k}) \sim \psi(p^*).$$

2.1.3 Arbitrage-free extrapolation of variance

In practice only finitely many option prices are quoted on the market and hence the total implied variance function cannot be uniquely specified based on the market quotes alone. In the sequel we will concentrate our attention on extrapolation of the total implied variance $w(\cdot, t)$ for a fixed maturity t. Theorem 2.2 states that a slice of the total variance $w(k, t)$ can be at most linear in k as k tends to $+\infty$ with slope equal to $\psi(\tilde{p})$ defined in Equation (2.8) with $\tilde{p}$ defined in (2.9). Similarly Theorem 2.3 states that the slice $w(k, t)$ is at most linear in k with slope $\psi(\tilde{q})$, where $\tilde{q}$ is defined in (2.11), as k tends to $-\infty$. Theorem 2.6 provides a further refinement of this result under additional conditions on the moment generating function M of the log-returns distribution. However any extrapolation of $w(\cdot, t)$ has to be performed so that absence of arbitrage is preserved. We have introduced notion of strong model-independent arbitrage in Definition 1.5, which can be specialised to the case when European Call or Put options are traded on the market. Strong model-independent static arbitrage in presence of options is equivalent to absence of calendar and butterfly spread arbitrages, that are understood as absence of arbitrage opportunities across option maturities for a fixed strike and absence of arbitrage opportunities across different strikes for a fixed maturity respectively.

Absence of static arbitrage can equivalently be stated as conditions on the shape of the total implied variance $w(\cdot,\cdot)$ as shown in Gatheral and Jacquier (2014) and Guo, Jacquier, Martini, and Neufcourt (2016). In particular under proportional dividends absence of calendar spread arbitrage is equivalent to $\partial_t w(k,t) \geq 0$ for all $k \in \mathbb{R}$ and $t > 0$ (Gatheral and Jacquier, 2014, Lemma 2.1). This is equivalent to the function $c_{BS}(k, \sqrt{w(k,\cdot)})$ being non-decreasing for a fixed $k \in \mathbb{R}$. For a fixed maturity t butterfly spread arbitrage is precluded if and only if the function $g : \mathbb{R} \to \mathbb{R}$ defined by

$$g(k) := \left(1 - \frac{k\partial_k w(k,t)}{2w(k,t)}\right)^2 - \frac{\partial_k^2 w(k,t)}{4}\left(\frac{1}{w(k,t)} + \frac{1}{4}\right) + \frac{\partial_{kk} w(k,t)}{2}, \quad (2.13)$$

is a positive distribution for all $k \in \mathbb{R}$, where the second derivative $\partial_{kk} w(k,t)$ is defined in the distributional sense, which in turn is equivalent to the function $c_{BS}(\cdot, \sqrt{w(\cdot,t)})$ being convex (Guo et al., 2016, Proposition 4.8). Assumption 7 ensures that $\partial_t w(k,t)$ is well-defined for all $t > 0$ and $\partial_k w(k,t)$ can be taken to be right or left derivatives at k if w is not differentiable at k. Any valid extrapolation of the total implied variance for a fixed maturity must satisfy conditions of Theorem 2.2 and Theorem 2.3 as well as be arbitrage-free. We show below that in the case when the total implied variance grows linearly in log-moneyness absence of butterfly spread arbitrage is equivalent to the slope of the slice of the total implied variance lying between 0 and 2.

Proposition 2.7. *Fix a maturity $t > 0$, constants $a_0, a_1 \in \mathbb{R}_+$ and consider the function $w(k,t) \mapsto a_1 k + a_0$. Then the function g defined in (2.13) is non-negative on $[k^*(a_0,a_1), \infty)$ if and only if $a_1 \in [0,2]$, where $k^*(a_0,a_1)$ is a positive constant that depends on a_0 and a_1 defined in (2.14) when $a_1 = 2$ and (2.15) otherwise.*

Proof. Let us fix $a_0 \in \mathbb{R}_+$ and $a_1 \in [0,2]$. If $a_1 = 0$ then $w(k,t) = a_0$ for all $k \in \mathbb{R}$ and g is constant equal to 1. So in the sequel we assume that $a_1 \in (0, 2]$. As $w(k, t)$ is linear in k by definition, the function g can be written as

$$g(k) = \left(1 - \frac{a_1 k}{2w(k,t)}\right)^2 - \frac{a_1^2}{4}\left(\frac{1}{w(k,t)} + \frac{1}{4}\right)$$

$$= \left(\frac{w(k,t) + a_0}{2w(k,t)}\right)^2 - \frac{a_1^2}{4}\left(\frac{4 + w(k,t)}{4w(k,t)}\right).$$

Let us denote $x := w(k, t)$. Then the above expression becomes

$$g\left(\frac{x - a_0}{a_1}\right) = \frac{1}{16x^2}\left(4x^2 + 8a_0 x + 4a_0^2 - 4a_1^2 x - a_1^2 x^2\right)$$

$$= \frac{(4 - a_1^2)x^2 + 4(2a_0 - a_1^2)x + 4a_0^2}{16x^2}.$$

If $a_1 = 2$ then the numerator is linear in x. Solving for x yields the root

$$x = \frac{-4a_0^2}{8(a_0 - 2)}.$$

It can be seen that g is non-negative if $a_0 < 2$ and substituting k back produces the expression

$$k^*(a_0, a_1) = \frac{a_0(8 - 6a_0)}{8(a_0 - 2)}, \tag{2.14}$$

which is positive if $a_0 \in (4/3, 2)$.

Consider now the case when $a_1 \in (0, 2)$. The numerator in the expression for g above is quadratic in x and solving for x yields two roots

$$x_\pm = \frac{-2(2a_0 - a_1^2) \pm 2a_1\sqrt{a_0^2 - 4a_0 + a_1^2}}{4 - a_1^2}.$$

As $x = a_1 k + a_0$ the corresponding values of k are

$$k_\pm = \frac{a_1(a_0 + 2) - \frac{8a_0}{a_1} \pm 2\sqrt{a_0^2 - 4a_0 + a_1^2}}{4 - a_1^2}, \tag{2.15}$$

and both roots are real if and only if $a_0 \in \mathbb{R} \setminus (2 - \sqrt{4 - a_1^2}, 2 + \sqrt{4 - a_1^2})$ for $a_1 \in (0, 2]$. If $a_0 \geq 2 - \sqrt{4 - a_1^2}$ then substituting the lower bound for a_0 into the expression for g above we get

$$
\begin{aligned}
&g\left(\frac{x - a_0}{a_1}\right) \\
&\geq \frac{(4 - a_1^2)x^2 + 4(4 - 2\sqrt{4 - a_1^2} - a_1^2)x + 4(4 - 4\sqrt{4 - a_1^2} + 4 - a_1^2)}{16x^2} \\
&= \frac{(4 - a_1^2)(x + 2)^2 - 8x\sqrt{4 - a_1^2} + 16}{16x^2} \\
&= \frac{(x\sqrt{4 - a_1^2} - 4)^2}{16x^2} \geq 0,
\end{aligned}
$$

for all $k > 0$. On the other hand if $a_0 < 2 - \sqrt{4 - a_1^2}$ then g is strictly positive for all $k > k_+$ and setting $k^*(a_0, a_1) = k_+$ we obtain the result.

Suppose now that $g(k) \geq 0$ for all $k \in [k^*(a_0, a_1), \infty)$. The second derivative of the Black-Scholes formula (2.1) with respect to e^k gives for any $k \in [k^*(a_0, a_1), \infty)$ the Call price $c(k, t)$ expressed as

$$
c(k, t) = \frac{g(k)}{\sqrt{2\pi w(k, t)}} \exp\left(-\frac{\left(d(k, \sqrt{w(k, t)}) - \sqrt{w(k, t)}\right)^2}{2}\right),
$$

which is non-negative by assumption on g. As $k \to \infty$ by assumption we have that $w(k, t) \sim a_1 k$ and note that $d(k, \sqrt{a_1 k}) - \sqrt{a_1 k} = -(1/\sqrt{a_1} + \sqrt{a_1}/2)\sqrt{k}$. Recalling calculations in Section 2.1.2 together with the bound (2.4) that holds for all $p \geq 0$ (with $p = 0$ being the trivial bound), we obtain that $a_1 = \psi(p)$, i.e. $a_1 \in [0, 2]$. $\square$

As Theorem 2.3 implies that the total variance increases as $k \to -\infty$ we have the following simple corollary to Proposition 2.7.

Corollary 2.8. *Fix maturity $t > 0$, constants $a_0, a_1 \in \mathbb{R}_+$ and consider the function $w(k, t) \mapsto a_1|k| + a_0$. Then the function g defined in (2.13) is non-negative on $(-\infty, k^*(a_0, a_1)]$ if and only if $a_1 \in [0, 2]$, where $k^*(a_0, a_1)$ is a negative constant that depends on a_0 and a_1.*

As the total implied variance grows at most linearly with slope $\psi(\tilde{p})$ defined in Theorem 2.2, following the heuristic arguments in Benaim and Friz (2009, Appendix A), one can show that

$$\log c(k,t) = -\frac{d(k, I(k,t))^2}{2} + \varepsilon(k),$$

where $\varepsilon(k) = \mathcal{O}(\log k)$ and $d(k, I(k,t))$ is defined in (2.2). Moreover by definition of d

$$\frac{\log c(k,t)}{k} = \frac{1}{2} - \frac{k}{2w(k,t)} - \frac{w(k,t)}{8k} + \frac{\varepsilon(k)}{k},$$

and for large k we have

$$\frac{\log c(k,t)}{k} \sim -\frac{1}{2}\left(\frac{1}{\psi(\tilde{p})} + \frac{\psi(\tilde{p})}{4} - 1\right),$$

where function ψ is defined in (2.8). Simple calculation yields

$$\frac{1}{2}\left(\frac{1}{\psi(\tilde{p})} + \frac{\psi(\tilde{p})}{4} - 1\right) = \tilde{p},$$

and hence $c(k,t) = \mathcal{O}(e^{-\tilde{p}k})$. Therefore once the slope of the total variance is specified for large values of k the bound (2.4) on Call option prices can be computed. Similar argument under conditions of Theorem 2.3 can be applied to bound Put option prices at small strikes.

In general an additional condition that the limit $\lim_{k\to\infty} d(k, \sqrt{w(k,t)}) = -\infty$ for each maturity $t > 0$ of the function d defined in (2.2) is needed to ensure that for each $t > 0$ the limit $\lim_{k\to\infty} c_{BS}(k, \sqrt{w(k,t)})$ is equal to 0 (Guo et al., 2016, Theorem 2.2). We show below that this condition is satisfied when $w(k,t)$ is linear in k for large values of k with the value of the slope in $[0,2)$.

Lemma 2.9. *Fix maturity $t > 0$. Suppose assumptions of Proposition 2.7 are satisfied. Then the limit $\lim_{k\to\infty} d(k, \sqrt{w(k,t)}) = -\infty$ holds if the slope of the total implied variance $a_1 \in [0,2)$.*

Proof. If $w(k,t)$ is constant for all $k \in \mathbb{R}$, then the statement of the lemma holds trivially. So assume that $w(k,t)$ is linear in k for large values of k with slope $a_1 \in (0,2)$. By definition of d we have

$$d(k, \sqrt{w(k,t)}) = \frac{\sqrt{w(k,t)}}{2} - \frac{k}{\sqrt{w(k,t)}} = \sqrt{\frac{k}{a_1}}\left(\frac{a_1}{2} - 1\right) + \mathcal{O}(\frac{1}{\sqrt{k}}).$$

As $a_1 \in (0,2)$ we have that $a_1/2 - 1 \in (-1,0)$. Hence the statement follows. $\square$

Lemma 2.10. *Fix maturity $t > 0$. Suppose assumptions of Corollary 2.8 are satisfied. Then the limit $\lim_{k\to-\infty}(d(k, \sqrt{w(k,t)}) - \sqrt{w(k,t)}) = +\infty$ holds if the slope of the total implied variance $a_1 \in [0,2)$.*

Proof. As before, if $a_1 = 0$ then the statement of the lemma follows immediately. Note that $d(k,a) - a = -a/2 - k/a$ and hence direct calculation yields

$$\begin{aligned} d(k, \sqrt{w(k,t)}) - \sqrt{w(k,t)} &= -\frac{\sqrt{w(k,t)}}{2} + \frac{-k}{\sqrt{w(k,t)}} \\ &= -\sqrt{\frac{-k}{a_1}}\left(\frac{a_1}{2} - 1\right) + \mathcal{O}(\frac{1}{\sqrt{k}}). \end{aligned}$$

As $a_1 \in (0,2)$ implies that $a_1/2 - 1 \in (-1,0)$ the statement follows. $\square$

It must be noted that the value of the slope of $w(k,t)$ for a fixed $t > 0$ is equal to 2 for large positive (small negative) values of k if and only if there are no finite higher (or negative) moments of the underlying stock price process at t by Definition (2.8) of function ψ and Theorems 2.2, 2.3.

2.2 Application to option prices

In this section we specialize the theory presented in Chapter 1 to the case where European Call options are traded on the market and discuss how notions of arbitrage introduced in Definitions 1.5 and 1.10 translate to this setting. We also discuss how definition of strong model-independent arbitrage can be relaxed in this setting and derive super-replication result under weaker conditions. In the sequel we work in a discrete time setting with a finite time horizon $T < \infty$ and intermediate times $0 = t_0 < t_1 < \ldots < t_n = T$. The collection of times is defined to be $\mathcal{T}_0 := \{t_0, t_1, \ldots, t_n\}$, and we shall denote by $\mathcal{T} := \mathcal{T}_0 \setminus \{t_0\}$. The state space $\Omega := \prod_{t\in\mathcal{T}} \Omega_t$, where $\Omega_t := \mathbb{R}_+ = [0,\infty)$, is locally compact, and the coordinate process $\mathbb{S} : \Omega \times \mathcal{T}_0 \to \mathbb{R}_+$ is defined to be $\mathbb{S}_t(\omega) = \omega_t$ for all $\omega \in \Omega$ and $\omega_t \in \Omega_t$, and $\mathbb{S}_0(\omega) = s_0 \in \mathbb{R}_+$. The unique forward price $\mathrm{Fwd}(0,t)$ is defined in Subsection 2.1.1 with $\mathrm{Fwd}(0,0) = s_0$. Let us also define a transformation of the coordinate process as $S : \Omega \times \mathcal{T}_0 \to \mathbb{R}_+$ with $S_t(\omega) := \mathbb{S}_t(\omega)/\mathrm{Fwd}(0,t)$ for all $t \in \mathcal{T}_0$. In particular $S_0(\omega) = 1$ for all $\omega \in \Omega$.

We assume that for each maturity $t \in \mathcal{T}$ there are European Call options traded on the market with forward moneyness in the set $\mathfrak{K}_t$, which can be finite or infinite. For each $t \in \mathcal{T}$ we assume that a Call option with forward moneyness $K \in \mathfrak{K}_t$ is available on the market for a price $c(K,t)$ (which has been normalised as in Section 2.1.1). We also refer to forward log-moneyness $k = \log(K)$ in the sequel and it must be noted that as Call option prices are market data we can interchangeably use $c(k,t)$ and $c(K,t)$ to denote the time 0 price of a Call option with moneyness $K \in \mathfrak{K}_t$. We shall also refer to moneyness as strike. Because we operate with normalised prices it should not create any confusion.

Remark 2.11. Put options can be used instead and by Put-Call parity the subsequent analysis can be performed on either.

Let us define K_*^t for each $t \in \mathcal{T}$ to be moneyness of a Call option available on the market at zero cost

$$K_*^t := \inf\{K \in \mathfrak{K}_t : c(K,t) = 0\},$$

and by standard notation $K_*^t = \infty$ (or $k_*^t := \log(K_*^t)$ expressed as log-moneyness) if the set is empty.

Definition 2.12. A static position $\mathfrak{f} := (\varphi_t)_{t\in\mathcal{T}_0}$ is a finite collection of functions with $\varphi_{t_0} \in \mathbb{R}$ and for each $t \in \mathcal{T}$ and real vectors $(\alpha_i)_{i=1,\ldots,\kappa(t)}$, where

$\kappa(t) < \infty$, we have

$$\varphi_t := \sum_{i=1}^{\kappa(t)} \alpha_i (S_t - K_i^t)_+, \quad \text{and} \quad c_t := \sum_{i=1}^{\kappa(t)} \alpha_i c(K_i^t, t),$$

where $K_1^t, \ldots, K_{\kappa(t)}^t \in \mathfrak{K}_t$. The function φ_t represents the pay-off of a static position in $\kappa(t)$ Call options maturing at $t \in \mathcal{T}$ with time 0 price c_t and φ_{t_0} represents a static position in a riskless bond with pay-off 1 for all $\omega \in \Omega$ by normalisation of prices (see Section 2.1). The set of all static positions is denoted by $\mathcal{S}$.

Definition 2.13. A trading strategy is a vector $\Delta := (\Delta_t)_{t=t_0,\ldots,t_{n-1}} \in \mathcal{H}$, where $\mathcal{H} := \mathbb{R} \times \prod_{j=1}^{n-1} C_b(\mathbb{R}_+^j)$ denotes the set of trading strategies. The first component denotes the initial position in the stock and the other components are continuous and bounded functions. The stochastic integral is defined as

$$(\Delta \bullet S(\omega))_T := \sum_{i=0}^{n-1} \Delta_{t_i}(\omega) \left(S_{t_{i+1}}(\omega) - S_{t_i}(\omega) \right),$$

and represents the gains or losses obtained by trading according to Δ. We use notation $\Delta_{t_i}(\omega) := \Delta_{t_i}(\operatorname{Pr}\omega)$, where $\operatorname{Pr}\omega$ is the projection of $\omega \in \Omega$ onto $\mathbb{R}_+^i$ for each $i = 1, \ldots, n-1$.

Remark 2.14. The above definition includes the trivial strategy

$$\tilde{\Delta} = (1, 1, \ldots, 1, 1)$$

of entering a forward contract at time zero that matures at time T (or equivalently entering a forward contract with maturity t_1 and rolling it to the final maturity T), with payoff $(\tilde{\Delta} \bullet S)_T = S_T(\omega) - 1$ for all $\omega \in \Omega$. Also note that the pay-off of any trading strategy $\Delta \in \mathcal{H}$ is at most linear in $\omega \in \Omega$.

For a collection of static positions $\mathfrak{f} = (\varphi_t)_{t \in \mathcal{T}_0} \in \mathcal{S}$ and a trading strategy $\Delta \in \mathcal{H}$, we introduce the following notations for the initial cost of a semi-static portfolio $(\mathfrak{f}, \Delta)$ and its final pay-off:

$$\Pi_{t_0}(\mathfrak{f}, \Delta) := \varphi_{t_0} + \sum_{t \in \mathcal{T}} c_t \text{ and } \Pi_T(\mathfrak{f}, \Delta; \omega) := \varphi_{t_0} + \sum_{t \in \mathcal{T}} \varphi_t(S_t(\omega)) + (\Delta \bullet S(\omega))_T, \tag{2.16}$$

for all $\omega \in \Omega$. Note that it is possible to have a semi-static portfolio with the final maturity $t < T$. However as we work with normalised prices, one can represent the final pay-off of a portfolio maturing at time $t < T$ as a position in the riskless bond maturing at time T with the value of the position equal to the said pay-off. The set of traded claims $\mathfrak{M}$ (see Definition 1.2) is then defined as a collection of all semi-static portfolio pay-offs $\Pi_T(\mathfrak{f}, \Delta; \cdot)$ for a static position $\mathfrak{f} \in \mathcal{S}$ and a trading strategy $\Delta \in \mathcal{H}$

$$\mathfrak{M} = \{\Pi_T(\mathfrak{f}, \Delta; \cdot) : \mathfrak{f} \in \mathcal{S} \text{ and } \Delta \in \mathcal{H}\}.$$

We shall denote elements of $\mathfrak{M}$ as m if there is no confusion in doing so. As we assume that only European Call options are traded for each maturity $t \in \mathcal{T}$ and the pay-off of a trading strategy $\Delta \in \mathcal{H}$ is a continuous function that grows at most linearly in $\omega \in \Omega$, the set of traded claims $\mathfrak{M}$ consists of functions $m \in C(\Omega)$ such that $m = \mathcal{O}(1 + \|\omega\|_1)$ as $\|\omega\|_1$ tends to infinity. It is in fact a subspace of a set of continuous functions with at most linear growth $C_l(\Omega)$ where the function $l : \Omega \to \mathbb{R}_+$ is defined as follows

$$l(\omega) := 1 + \sum_{i=0}^{n} S_{t_i}(\omega). \tag{2.17}$$

Note that $l \in \mathfrak{M}$, as the semi-static portfolio $(\mathfrak{f}_*, \Delta_*)$ with $\mathfrak{f}_* := (2 + n, 0, \ldots, 0)$ and $\Delta_* := (n, n-1, \ldots, 1)$ has a final pay-off $\Pi_T(\mathfrak{f}_*, \Delta_*; \omega) = l(\omega)$ for all $\omega \in \Omega$. The space of Borel probability measures with finite first moments is denoted by $\mathcal{P}_l(\Omega) := \{\pi \in \mathcal{P}(\Omega) : \int_\Omega l(\omega)\pi(\mathrm{d}\omega) < \infty\}$. Similar to Section 1.2 we also define the pricing functional $\rho : \mathfrak{M} \to \mathbb{R}$ as

$$\rho(\Pi_T(\mathfrak{f}, \Delta; \cdot)) := \Pi_{t_0}(\mathfrak{f}, \Delta), \tag{2.18}$$

for any $\Pi_T(\mathfrak{f}, \Delta; \cdot) \in \mathfrak{M}$. By Proposition 1.6 absence of strong model-independent arbitrage is equivalent to ρ being linear, uniquely defined and strictly positive. In particular it implies that Call Spreads, Butterfly Spreads and Calendar Spreads can only be traded at strictly positive prices. We also define a market model similarly to Cox and Obłój (2011b, Definition 1.1).

Definition 2.15. We define a model as a probability measure $\pi \in \mathcal{P}_l(\Omega)$ and a market model such that it is a positive linear extension of the pricing operator ρ defined in (2.18) from $\mathfrak{M}$ to the whole space $C_l(\Omega)$.

The set of all martingale measures will be denoted by $\mathbb{M}$. A sufficient condition that ensures that S is a martingale under a measure $\mathbb{P} \in \mathbb{M}$ is

$$\int_\Omega (\Delta \bullet S(\omega))_T \, \mathbb{P}(\mathrm{d}\omega) = 0, \tag{2.19}$$

for all $\Delta \in \mathcal{H}$. By definition the process S is a martingale in its own filtration $\mathbb{F}$ under a measure $\mu \in \mathcal{P}_l(\Omega)$ if

$$\int_\Omega \sum_{i=0}^{n-1} \mathbb{1}_{B_{t_i}}(\omega)(S_{t_{i+1}}(\omega) - S_{t_i}(\omega))\mu(\mathrm{d}\omega) = 0,$$

for all Borel sets $B_{t_i} \subset \Omega_{t_i}$ for all $i = 1, \ldots, n-1$. To see sufficiency of Condition (2.19) note that the Borel σ-algebra is generated by open sets and the indicator function of an open set is a lower semi-continuous function as it can be approximated pointwise by a non-decreasing sequence of continuous and bounded functions from below. By Lebesgue Monotone Convergence Theorem the definition of a martingale follows. It also implies from Definition 2.15 and Equation (2.18) that a market model π is also a martingale measure.

2.2.1 Case of infinitely many options

When number of traded Call options is infinite, i.e. $|\mathfrak{K}_t| = \infty$ for all $t \in \mathcal{T}$, it might be possible that there exists a sequence of portfolios in $\mathfrak{M}$ that converge (in some sense - see below) to a position with non-negative pay-off and negative price. Such situation is discussed in Cox and Obłój (2011b, Proof of Proposition 2.2), where the authors assume the existence, for any

$t \in \mathcal{T}$, of a sequence $(K_n)_{n\in\mathbb{N}} \subset \mathfrak{K}_t$ with $K_n \to \infty$ such that the limit $a := \lim_{n\to\infty} C(K_n, t)$ is strictly positive, and the existence of short positions in Call options with strikes $(K_n)_{n\in\mathbb{N}}$ converging pointwise to zero with limiting price equal to $-a$. The authors introduced the notion of "weak free lunch with vanishing risk" (Cox and Obłój, 2011b, Definition 2.1) and show equivalence between absence of weak free lunch with vanishing risk and existence of a market model (Cox and Obłój, 2011b, Proposition 2.2)

Recall that absence of strong model-independent arbitrage can be expressed as $\mathfrak{F}\cap(C_l)_+(\Omega) = \{0\}$ or equivalently $(\mathfrak{F}-(C_l)_+(\Omega))\cap(C_l)_+(\Omega) = \{0\}$, where the set $\mathfrak{F}$ is defined in (1.8). It is clear that a strong model-independent arbitrage opportunity is also a weak free lunch with vanishing risk. On the other hand, absence of weak free lunch introduced in Definition 1.10 states that the weak closure of the set $\mathfrak{F} - (C_l)_+(\Omega)$ has an empty intersection with the strictly positive cone $(C_l)_{++}(\Omega)$. We show below that in presence of Call options absence of weak free lunch with vanishing risk implies absence of weak free lunch.

Lemma 2.16. *When $|\mathfrak{K}_t| = \infty$ for each maturity $t \in \mathcal{T}$ absence of weak free lunch with vanishing risk implies absence of weak free lunch.*

Proof. Assume for contradiction that there exists a weak free lunch, i.e. there exist sequences $(f_n)_{n\in\mathbb{N}} \subset \mathfrak{F}$ and $(g_n)_{n\in\mathbb{N}} \subset C_l(\Omega)$ such that $f_n \geq g_n$ for all $n \in \mathbb{N}$ and $g_n \to g$ with $g \in (C_l)_{++}(\Omega)$ weakly. In particular $g_n = \mathcal{O}(l)$ and as $\delta_\omega \in \mathcal{P}_l(\Omega)$ for all $\omega \in \Omega$ we have that $g_n(\omega) \to g(\omega)$ for all $\omega \in \Omega$. Hence $(g_n)_{n\in\mathbb{N}}$ is also a weak free lunch with vanishing risk. $\square$

Suppose there exist sequences $(K_n)_{n\in\mathbb{N}} \subset \mathfrak{K}_t$ for each $t \in \mathcal{T}$ such that $K_n \to \infty$ as $n \to \infty$. In this case it is possible to relax the assumption that Φ is bounded by linear growth (above or below). Before we proceed we present an auxiliary lemma that expands the remark in (Acciaio et al., 2016, Section 4).

Lemma 2.17. *Fix a probability measure μ on $\mathbb{R}_+$ and let $g : \mathbb{R}_+ \to \mathbb{R}$ be a convex function such that $\lim_{x\to\infty} \dfrac{g(x)}{x} = \infty$ and $\int_{\mathbb{R}_+} g(x)\mu(\mathrm{d}x) < \infty$. Then* *there exist a sequence $(K_n)_{n\in\mathbb{N}}$ with $K_n \uparrow \infty$ and a sequence of non-negative numbers $(\alpha_n)_{n\in\mathbb{N}}$ together with a constant $\beta \in \mathbb{R}_+$ such that*

$$g(x) \leq \beta + \sum_{n=1}^{\infty} \alpha_n (x - K_n)_+, \quad \sum_{n=1}^{\infty} \alpha_n = \infty, \quad \sum_{n=1}^{\infty} \alpha_n c_n < \infty,$$

where $c_n := \int_{\mathbb{R}_+} (x - K_n)_+ \mu(\mathrm{d}x)$.

Proof. As $g : \mathbb{R}_+ \to \mathbb{R}$ is the proper convex function, it is continuous on the interior of its domain. It is sufficient to take an increasing sequence $(K_n)_{n\in\mathbb{N}}$ such that $K_1 > 0$ and $g(K_1) < \infty$. Set $\beta := g(K_1)$ and take

$$\alpha_n := \frac{g(K_{n+1}) - g(K_n)}{K_{n+1} - K_n} - \frac{g(K_n) - g(K_{n-1})}{K_n - K_{n-1}}, \quad \text{and} \quad \alpha_1 := \frac{g(K_2) - g(K_1)}{K_2 - K_1}.$$

It is clear that $\sum_{n=1}^{\infty} \alpha_n = \infty$. Then $g(x) \leq \beta + \sum_{n=1}^{\infty} \alpha_n (x - K_n)_+$ for all $x \in [K_i, K_{i+1}]$ for all $i \in \mathbb{N}$ and equality is attained at the end points of each interval. Integrating the sum on the right-hand side yields

$$\int_{\mathbb{R}_+} \sum_{n=1}^{\infty} \alpha_n (x - K_n)_+ \mu(\mathrm{d}x) = \sum_{n=1}^{\infty} \alpha_n c_n < \infty,$$

and the result follows. □

If the traded Call options with strike in $\mathfrak{K}_t$ for each $t \in \mathcal{T}$ satisfy no weak free lunch with vanishing risk then any market model $\mathbb{P}$ is consistent with prices of those traded options. Moreover Lemma 2.17 implies that there exists a convex super-linear function $g_t : \mathbb{R}_+ \to \mathbb{R}$ integrable with respect to μ_t. We can then define the function $h : \Omega \to \mathbb{R}$ as

$$h(\omega) := \sum_{t\in\mathcal{T}} g_t(S_t(\omega)). \tag{2.20}$$

Naturally the function h is integrable with respect to $\mathbb{P}$. In this case the duality results in Section 1.4 hold for any $\Phi \in U_h(\Omega)$ in case of super-hedging and any $\Phi \in L_h(\Omega)$ in case of sub-hedging. It must be noted that in this case we can also enlarge the space of trading strategies $\mathcal{H}$ to include strategies

$\Delta \in \mathbb{R} \times \prod_{j=1}^{n-1} C(\mathbb{R}_+^j)$ such that $(\Delta \bullet S)_T = \mathcal{O}(h)$ in the similar manner to (Acciaio et al., 2016, Definition 2.2).

2.2.2 Case of finitely many options

Assume now that only finitely many Call options are traded on the market. Let $\mathfrak{K}_t = \{K_1^t, \dots, K_{\kappa(t)}^t\}$ with $\kappa(t) < \infty$ and $K_1^t < \cdots < K_{\kappa(t)}^t$ for each maturity $t \in \mathcal{T}$. Denote by $\mathfrak{C} := \{c(K,t) : K \in \mathfrak{K}_t \text{ for all } t \in \mathcal{T}\}$ the collection of prices of traded Call options. It is clear that the span of traded options forms a subspace of $\mathfrak{M}$. Let us define the set of market models when there are only finitely many Call options are traded on the market as

$$\mathbb{M}_{\mathfrak{C}} := \Big\{\mathbb{P} \in \mathcal{P}_l(\Omega) : \int_\Omega \Pi_T(\mathfrak{f}, \Delta; \omega)\mathbb{P}(\mathrm{d}\omega) = \Pi_{t_0}(\mathfrak{f}, \Delta) \text{ for } (\mathfrak{f}, \Delta) \in \mathcal{S} \times \mathcal{H}\Big\}. \tag{2.21}$$

It is clear that the set of static positions $\mathcal{S}$ includes all traded Call options and therefore for any $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ one has that $c(K,t) = \int_\Omega (S_t(\omega) - K)_+ \mathbb{P}(\mathrm{d}\omega)$ for any $t \in \mathcal{T}$ and each $K \in \mathfrak{K}_t$.

As shown in Proposition 1.6 absence of strong model-independent arbitrage implies that ρ is a strictly positive linear functional. This is a rather strict assumption as it is possible to have Butterfly Spreads traded on the market at zero price and find a corresponding market model as shown in Davis and Hobson (2007, Theorems 3.1, 4.2). In the sequel we relax the definition of strong model-independent arbitrage similarly to definitions presented in Davis and Hobson (2007, Definition 2.1) and Cox and Obłój (2011b, Definition 1.2) that require the pricing functional ρ to be only positive on $\mathfrak{M}$.

Definition 2.18. We say that the functional ρ defined in (2.18) admits no model-independent arbitrage on $\mathfrak{M}$ if $\rho(m) \geq 0$ for all $m \in \mathfrak{M}_+$.

Arguments of Proposition 1.6 can be amended in a straightforward manner to show that absence of model-independent arbitrage is equivalent to ρ being uniquely defined positive linear functional on $\mathfrak{M}$. It is clear that absence of model-independent arbitrage is weaker than absence of strong model-independent arbitrage as Definition 2.18 implies that an element $m \in \mathfrak{M}_+$ with $\rho(m) = 0$ is not a model-independent arbitrage opportunity, yet constitutes strong model-independent arbitrage. We thus introduce a notion of weak arbitrage similar to Cox and Obłój (2011b, Definition 2.3).

Definition 2.19. We say that the pricing functional ρ defined in (2.18) admits weak arbitrage on $\mathfrak{M}$ if for any model $\mathbb{P} \in \mathbb{M}$, there exists $m \in \mathfrak{M}$ such that $\rho(m) \leq 0$, but $\mathbb{P}(\{\omega \in \Omega : m(\omega) \geq 0\}) = 1$ and $\mathbb{P}(\{\omega \in \Omega : m(\omega) > 0\}) > 0$.

It is easily seen that both strong model-independent and model-independent arbitrage opportunities are also weak arbitrage opportunities. This definition of weak arbitrage allows the use of the result by Davis and Hobson (2007, Theorem 4.2) stating that when only finitely many options are traded on the market, absence of weak arbitrage is equivalent to existence of a market model. In the sequel, we concentrate our attention on Butterfly Spreads traded on the market that can create weak arbitrage opportunities as shown in (Davis and Hobson, 2007, Theorem 3.1, 4.2).

It is easily seen that absence of weak arbitrage implies that if there exists a claim $m \in \mathfrak{M}_+$ with market price $\rho(m) = 0$ then for any market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ it follows that $\mathbb{P}(\{\omega \in \Omega : m(\omega) > 0\}) = 0$ as by definition of a market model we have $\rho(m) = \langle m, \mathbb{P} \rangle$ for all $m \in \mathfrak{M}$. Let us define $\mathfrak{F}_0 := \{m \in \mathfrak{M} : \rho(m) = 0\}$ the set of all traded claims that are available on the market at price 0.

Assumption 8. Define the set

$$W := \mathfrak{F}_0 \cap (C_l)_+(\Omega), \tag{2.22}$$

and assume that it is a convex cone generated by finitely many Butterfly Spreads traded on the market at price 0 for each $t \in \mathcal{T}$.

We assume that W is spanned by finitely many Butterfly Spreads in order to simplify further analysis. Indeed, for a fixed maturity $t \in \mathcal{T}$ and three strikes

$K^t_{i-1} < K^t_i < K^t_{i+1}$ (with $1 < i < \kappa(t)$) the pay-off of a butterfly spread is

$$\alpha(S_t - K^t_{i-1})_+ - (\alpha+\beta)(S_t - K^t_i)_+ + \beta(S_t - K^t_{i+1})_+,$$

where $\alpha := 1/(K^t_i - K^t_{i-1})$ and $\beta := 1/(K^t_{i+1} - K^t_i)$. If such butterfly spread is traded on the market at price 0 then absence of weak arbitrage implies that any market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ places no mass on the open interval (K^t_{i-1}, K^t_{i+1}). If another butterfly spread comprised of Call options with strikes $K^t_i < K^t_{i+1} < K^t_{i+2}$ is available on the market at price 0 then similarly any market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ places no mass on the open interval (K^t_i, K^t_{i+2}). Moreover we have that $\mathbb{P}((K^t_{i-1}, K^t_{i+1}) \cup (K^t_i, K^t_{i+2})) = 0$ and $\mathbb{P}((K^t_{i-1}, K^t_{i+1}) \cap (K^t_i, K^t_{i+2})) = 0$ for any market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$, thus the collection of such sets is closed under taking finite intersections and unions, which is sufficient in our situation by Assumption 8.

If there are Butterfly Spreads traded on the market at price 0 the duality results presented in Section 1.4 can be amended to hold under condition of absence of weak arbitrage. Let us introduce a new ordering on the space $C_l(\Omega)$ by defining a "trans-positive" convex cone (first introduced in Clark (2006))

$$J := \overline{(C_l)_+(\Omega) - W}, \tag{2.23}$$

where the closure is taken with respect to the norm topology on $C_l(\Omega)$. It is clear that J is a closed convex cone and as $0 \in J$ we can introduce a new ordering "$\succeq$" on $C_l(\Omega)$ such that the relation $f_1 \succeq f_2$ holds if and only if $f_1 - f_2 \in J$. The negative polar $J^* \subset \mathcal{P}_l(\Omega)$ is defined in the usual manner (see Equation (1.1)). As any $j \in J$ can be written as $j = f - w$ for $f \in (C_l)_+(\Omega)$ and $w \in W$, it must be noted that for any $\pi \in \mathbb{M}_{\mathfrak{C}}$ we have

$$0 \leq \langle j, \pi \rangle = \langle f - w, \pi \rangle = \langle f, \pi \rangle - \langle w, \pi \rangle = \langle f, \pi \rangle,$$

where the last equality follows from the fact that for any $w \in W$ we have $\langle \pi, w \rangle = \rho(w) = 0$. This observation is summarised in the lemma below.

Lemma 2.20. *The set of market models $\mathbb{M}_{\mathfrak{C}}$ is a subset of the negative polar J^* of the trans-positive cone J defined in* (2.23).

Proof. If $\mathbb{M}_{\mathfrak{C}}$ is an empty set, the inclusion is trivial. So assume that $\mathbb{M}_{\mathfrak{C}} \neq \emptyset$. For a market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ we have that $\langle f, \mathbb{P}\rangle \geq 0$ for all $f \in (C_l)_+(\Omega)$. On the other hand for any $w \in W$ the equality holds $\langle w, \mathbb{P}\rangle = 0$ by definition of a market model. So for any $f \in (C_l)_+(\Omega)$ and $w \in W$ we have the following

$$0 \leq \langle f, \mathbb{P}\rangle = \langle f, \mathbb{P}\rangle - \langle w, \mathbb{P}\rangle = \langle j, \mathbb{P}\rangle\,,$$

where $j = f - w$ and hence is an element in J. The statement of the lemma follows by definition of the negative polar J^*. $\square$

The above analysis also remains the same for any j on the boundary of J. In particular let $j := \lim_{n\to\infty} j_n = \lim_{n\to\infty}(f_n - w_n)$ and by linearity of the inner product for any $\pi \in \mathbb{M}_{\mathfrak{C}}$ we have

$$\begin{gathered}\langle j, \pi\rangle = \left\langle \lim_{n\to\infty}(f_n - w_n), \pi\right\rangle = \\ \lim_{n\to\infty}\langle f_n - w_n, \pi\rangle = \lim_{n\to\infty}\left(\langle f_n, \pi\rangle - \langle w_n, \pi\rangle\right) = \\ \left\langle \lim_{n\to\infty} f_n, \pi\right\rangle = \langle f, \pi\rangle\,,\end{gathered}$$

where $f \in (C_l)_+(\Omega)$ as the positive cone is closed in the weak topology.

Let us now define the super-hedging problem for an option with pay-off $\Phi \in C_l(\Omega)$ as

$$^*\vartheta_p(\Phi) := \inf\left\{\overline{\rho}(m) : m \in \bar{\mathfrak{M}},\, m - \Phi \in J\right\}, \tag{2.24}$$

and its dual problem reads

$$^*\vartheta_d(\Phi) := \sup\left\{\langle \Phi, \mu\rangle : \mu \in \mathbb{M}_{\mathfrak{C}}\right\}. \tag{2.25}$$

The sub-hedging primal problem can be defined as $_*\vartheta_p(\Phi) = -^*\vartheta_p(-\Phi)$ and similarly the sub-hedging dual problem is defined via $_*\vartheta_d(\Phi) = -_*\vartheta_d(-\Phi)$. We then have the following duality result.

Theorem 2.21. *Suppose Assumption 4 holds. Absence of weak arbitrage implies no duality gap between the primal* (2.24) *and the dual* (2.25) *problems, i.e.* $^*\vartheta_p(\Phi) = {}^*\vartheta_d(\Phi)$ *for* $\Phi \in C_l(\Omega)$.

Proof. The proof below follows closely the arguments in proof of Theorem 1.16. We assume that $\Phi \notin \bar{\mathfrak{M}}$, otherwise the statement of the theorem trivially follows. Absence of weak arbitrage implies there exists a market model $\mathbb{P}_0 \in \mathbb{M}_{\mathfrak{C}}$ such that $\mathbb{E}^{\mathbb{P}_0}\{\Phi\} := \langle \Phi, \mathbb{P}_0 \rangle \leq {}^*\vartheta_p(\Phi)$ and fix $\lambda \in (\mathbb{E}^{\mathbb{P}_0}\{\Phi\}, {}^*\vartheta_p(\Phi))$. Let $G :=$ Span $\{\bar{\mathfrak{M}}, \Phi\}$ and define $\eta : G \to \mathbb{R}$ as $\eta(g) := \eta(m + t\Phi) = \overline{\rho}(m) + t\lambda$. We now show that η is positive on $J_G := J \cap G$. Let $g = m + t\Phi \in J_G$ and consider three cases. If $t = 0$ then $\eta(g) = \overline{\rho}(m) \geq 0$. If $t < 0$ then $(-t)^{-1}m \succeq \Phi$ and $(-t)^{-1}\overline{\rho}(m) \geq {}^*\vartheta_p(\Phi) > \lambda$. Similarly if $t > 0$ then $\Phi \succeq (-t)^{-1}m$ and hence $\overline{\rho}(m) > -t\lambda$. It also follows that if $t \neq 0$ then $\eta(g) > 0$.

As η is linear and dominated by a convex function ${}^*\vartheta_p$ (as the function l defined in (2.17) is an element of $\mathfrak{M}$, the function $-\infty < {}^*\vartheta_p(f) < \infty$ for all $f \in C_l(\Omega)$) hence by Hahn-Banach Extension Theorem there exists a linear extension of π to the whole space $C_l(\Omega)$ such that π is dominated by ${}^*\vartheta_p$. For a $j \in J$ we have $0 \succeq -j$ and $\pi(-j) \leq {}^*\vartheta_p(-j) \leq \overline{\rho}(0) = 0$ thus $\pi(j) \geq 0$ by linearity of π. As $0 \in W$ it implies that π is a positive linear functional and as $C_l(\Omega)$ is a Banach lattice it follows by Theorem A.6 that π is continuous and Assumption 4 implies that π can be represented as an integral with respect to a unique Borel probability measure $\mu \in \mathcal{P}_l(\Omega)$, see Appendix B. Moreover π also extends ρ and hence is a market model.

By construction we have $\pi(\Phi) = \eta(\Phi) = \lambda$. As π is a market model, it is a feasible solution to the dual problem (2.25) and hence $\lambda = \pi(\Phi) \leq {}^*\vartheta_d(\Phi)$. As $\lambda \in (\mathbb{E}^{\mathbb{P}_0}\{\Phi\}, {}^*\vartheta_p(\Phi))$ was chosen arbitrarily it follows that ${}^*\vartheta_d(\Phi) = {}^*\vartheta_p(\Phi)$. □

Corollary 2.22. *Suppose Assumption 4 holds. Absence of weak arbitrage implies no duality gap between the sub-hedging problem and its dual, i.e.* ${}_*\vartheta_p(\Phi) = {}_*\vartheta_d(\Phi)$ *for* $\Phi \in C_l(\Omega)$.

Proof. As ${}_*\vartheta_p(\Phi) = -{}^*\vartheta_p(-\Phi)$ and ${}_*\vartheta_d(\Phi) = -{}^*\vartheta_d(-\Phi)$ by above, the statement follows as a direct consequence of Theorem 2.21. □

The primal (2.24) and the dual (2.25) problems can be extended to the case when $\Phi \in U_l(\Omega)$ by defining the extension to the primal problem $\overline{\vartheta}_p : U_l(\Omega) \to$

$\overline{\mathbb{R}}$ as follows

$$\overline{\vartheta}_p(\Phi) := \inf\left\{{}^*\vartheta_p(f) : f \in C_l(\Omega),\ f(\omega) - \Phi(\omega) \geq 0 \text{ for all } \omega \in \Omega\right\}. \quad (2.26)$$

The corresponding extension to the dual problem $\overline{\vartheta}_d : U_l(\Omega) \to \overline{\mathbb{R}}$ is defined as

$$\overline{\vartheta}_d(\Phi) := \sup\left\{\langle \Phi, \mu\rangle : \mu \in \mathbb{M}_{\mathfrak{C}}\right\}. \quad (2.27)$$

Naturally the sub-hedging primal problem can be extended to $\Phi \in L_l(\Omega)$ so that $\underline{\vartheta}_p : L_l(\Omega) \to \overline{\mathbb{R}}$ is the value function of the sub-hedging primal problem defined as

$$\underline{\vartheta}_p(\Phi) := \sup\left\{{}_*\vartheta_p(f) : f \in C_l(\Omega),\ f(\omega) - \Phi(\omega) \leq 0 \text{ for all } \omega \in \Omega\right\}, \quad (2.28)$$

and the equality $\underline{\vartheta}_p(\Phi) = -\overline{\vartheta}_p(-\Phi)$ holds. The sub-hedging dual problem value function $\underline{\vartheta}_d : L_l(\Omega) \to \overline{\mathbb{R}}$ is similarly defined as

$$\underline{\vartheta}_d(\Phi) := \inf\left\{\langle \Phi, \mu\rangle : \mu \in \mathbb{M}_{\mathfrak{C}}\right\}, \quad (2.29)$$

and $\underline{\vartheta}_d(\Phi) = -\overline{\vartheta}_d(-\Phi)$. Duality results follow from arguments in proof of Theorem 1.16.

Remark 2.23. It must be noted that if the convex cone W defined in (2.22) is trivial, i.e. $W = \{0\}$, then the trans-positive cone is reduced to the positive cone $(C_l)_+(\Omega)$, i.e. $J = \overline{(C_l)_+(\Omega) - W} = \overline{(C_l)_+(\Omega)} = (C_l)_+(\Omega)$ as the positive cone $(C_l)_+(\Omega)$ is closed in the norm topology, see Appendix A. Then the definitions of the primal (2.26) and the dual (2.27) coincide with the definitions of the primal (1.9) and the dual (1.10) programmes. In particular the super-hedging primal problem for any $\Phi \in C_l(\Omega)$ is written as

$${}^*\vartheta_p(\Phi) := \inf\left\{\overline{\rho}(m) : m \in \bar{\mathfrak{M}},\ m - \Phi \in (C_l)_+(\Omega)\right\},$$

and hence coincides with the value $\overline{\vartheta}_p(\Phi)$. The sub-hedging problems are likewise reduced to programmes (1.11) and (1.12).

2.2.3 Consequences of no arbitrage

Absence of weak free lunch with vanishing risk in case when Call options are traded for all $K \in \mathbb{R}_+$ for each maturity $t \in \mathcal{T}$ (Cox and Obłój, 2011b, Proposition 2.2) or absence of weak arbitrage (Davis and Hobson, 2007, Theorem 4.2) in case when only finitely many Call options are traded is equivalent to existence of a convex function $C : \mathbb{R}_+ \times \mathcal{T} \to \mathbb{R}$ satisfying the following conditions

(i) $C(K^t, t) = c(K^t, t)$, for all $t \in \mathcal{T}$ and each $K^t \in \mathfrak{K}_t$,

(ii) $\partial_+ C(K, \cdot)|_{K=0} \geq -1$,

(iii) $\lim_{K \to 0} C(K, \cdot) = 1$,

(iv) $\lim_{K \to \infty} C(K, \cdot) = 0$,

(v) $C(K_1, \cdot) > C(K_2, \cdot)$ for all $K_1 < K_2 \in \mathbb{R}_+$,

(vi) $C(\cdot, t) \leq C(\cdot, r)$ for all $t \leq r \in \mathcal{T}$.

Remark 2.24. Equivalence between absence of weak free lunch with vanishing risk and existence of a market model is proved in Cox and Obłój (2011b, Proposition 2.2) for a single maturity. Therefore for Condition (vi) to be satisfied the risk-neutral measures of normalised asset returns μ_t for each $t \in \mathcal{T}$ must be placed in convex order as in this case Strassen's theorem (Strassen, 1965) implies that there exists at least one martingale measure $\mathbb{P} \in \mathbb{M}$ on Ω such that its marginals coincide with μ_t for all $t \in \mathcal{T}$.

A sufficient condition for any two Borel probability measures μ and ν on $\mathbb{R}_+$ to be placed in convex order (Baker, 2012) is that they have equal means and the following inequality holds for all $K \in \mathbb{R}_+$

$$\int_0^\infty (x - K)_+ \mu(\mathrm{d}x) \leq \int_0^\infty (x - K)_+ \nu(\mathrm{d}x). \tag{2.30}$$

Breeden and Litzenberger (1978) showed that the risk-neutral measure of normalised asset returns μ_t at maturity $t \in \mathcal{T}$ is equal to the right derivative of option prices c with respect to moneyness K plus one:

$$\mu_t([0, K]) = 1 + \partial_+ c(K, t). \tag{2.31}$$

It can then be shown that Inequality (2.30) imposes absence of calendar spread arbitrages for any two maturities $t_1, t_2 \in \mathcal{T}$ with $t_1 < t_2$ which is equivalent to Condition (vi).

The function C satisfying Conditions (i)-(vi) will be referred to as the Call price surface (or function) in the sequel. When Call options are traded for each moneyness $K \in \mathbb{R}_+$ for all maturities $t \in \mathcal{T}$, Conditions (i)-(vi) are equivalent to existence of a martingale measure $\mathbb{P} \in \mathbb{M}$ with marginals at each maturity $t \in \mathcal{T}$ uniquely determined by the prices of traded options. However when there are only finitely many Call options traded per maturity as it is often the case in practice (i.e. $|\mathfrak{K}_t| < \infty$ for each $t \in \mathcal{T}$) then absence of weak arbitrage produces a collection of feasible functions C satisfying Conditions (i)-(vi) that interpolate between given Call prices and lie in in the region bounded by dash-dotted lines shown in Figure 2.1. The first price (square) lies at the intersection of the following two lines: linear extrapolation to the left of the first two observed (triangles) points, and the segment $K \mapsto (1-K)_+$. The last prices (circle) is, similarly, at the intersection of the linear extrapolation of the last two observed prices and the horizontal axis.

The set of market models $\mathbb{M}_{\mathfrak{C}}$ is not closed under condition of no weak arbitrage. In particular there exists a sequence of market models that convergences to a probability measure that is not a market model as the lemma below shows.

Lemma 2.25. *When only finitely many options are traded, the class of market models* (2.21) *is not closed under conditions of weak arbitrage if* $K_*^t = \infty$ *for some maturity* $t \in \mathcal{T}$.

In particular the statement of the lemma implies that there exists a sequence of market models $(\mathbb{P}_n)_{n\in\mathbb{N}} \subset \mathbb{M}_{\mathfrak{C}}$ that do not admit weak arbitrage but the limit $\bar{\mathbb{P}} := \lim_{n\to\infty} \mathbb{P}_n$ is not a market model and there is a weak arbitrage opportunity.

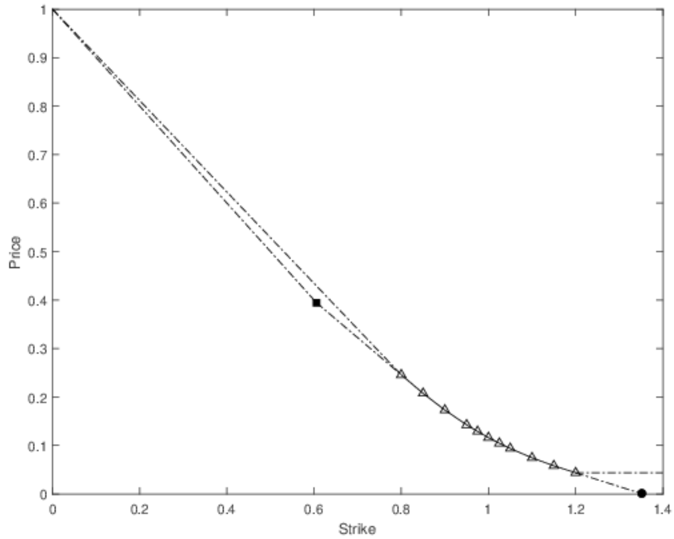


Figure 2.1: European Call options on the DAX index with maturity 2 years as of 27/05/2016 (triangles) and feasible regions for arbitrage-free extrapolations of the Call price function (dash-dotted lines). The first possible Call option with price 0 (circle), The last possible Call option with price equal to intrinsic value (square).

Proof. Let us define a sequence of market models $(\mathbb{P}_n)_{n\in\mathbb{N}} \subset \mathbb{M}_{\mathfrak{C}}$ such that for each $n \in \mathbb{N}$ the model $\mathbb{P}_n$ has a marginal distribution μ_t^n of S at some fixed maturity $t \in \mathcal{T}$ that satisfies the following conditions

(i) $$\int_\Omega (S_t(\omega) - K)_+ \mathbb{P}_n(\mathrm{d}\omega) = \frac{c(K_i^t, t) - c(K_{i-1}^t, t)}{K_i^t - K_{i-1}^t}(K - K_{i-1}^t) + c(K_{i-1}^t, t),$$

for all $K \in [K_{i-1}^t, K_i^t)$, $i = 1, \ldots, \kappa(t)$ and $K_0^t := 0$;

(ii) $$\int_\Omega (S_t(\omega) - K)_+ \mathbb{P}_n(\mathrm{d}\omega) = \max\left\{0, c(K_{\kappa(t)}^t, t) - \frac{1}{n}(K - K_{\kappa(t)}^t)\right\},$$

for all $K \geq K_{\kappa(t)}^t$.

It can be easily seen that each model $\mathbb{P}_n$ corresponds to a function C that satisfies Conditions (i)-(vi). As $\lim_{n\to\infty} \max\left\{0, c(K_{\kappa(t)}^t, t) - \frac{1}{n}(K - K_{\kappa(t)}^t)\right\} = c(K_{\kappa(t)}^t, t) > 0$ and by assumption of the lemma $K_*^t = \infty$ it follows that

$\bar{\mathbb{P}} := \lim_{n\to\infty} \mathbb{P}_n$ is such that $\mathbb{E}^{\bar{\mathbb{P}}}\{(S_t - K)_+\} = c(K^t_{\kappa(t)}, t)$ for all $K \geq K^t_{\kappa(t)}$, which is clearly not possible hence $\bar{\mathbb{P}}$ is not a market model.

On the other hand if for any $K \geq K^t_{\kappa(t)}$ the price of a Call option struck at K is equal to $c(K^t_{\kappa(t)}, t)$ there exists a weak arbitrage opportunity as proved in Davis and Hobson (2007, Theorem 3.2). □

Remark 2.26. The assumption that $K^t_* = \infty$ for some $t \in \mathcal{T}$ in Lemma 2.25 is needed to preclude the trivial case when $c(K^t_{\kappa(t)}, t) = 0$ for all $t \in \mathcal{T}$ and hence any feasible function C is identically zero for all $K \geq K^t_{\kappa(t)}$. Indeed, if $c(K^t_{\kappa(t)}, t) = 0$ for all $t \in \mathcal{T}$ then as shown in Davis, Obłój, and Raval (2014, Lemma 2.2) it follows (as a consequence of no weak arbitrage) that any market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ places no mass on the intervals $(K^t_{\kappa(t)}, \infty)$ for all $t \in \mathcal{T}$. Thus the state space Ω can be restricted to a compact subset $\prod_{t\in\mathcal{T}}[0, K^t_{\kappa(t)}]$ in which case the duality result as well as arbitrage conditions simplify significantly.

Remark 2.27. It must be noted that by simple modification of arguments of Lemma 2.25 it follows that the set of market model $\mathbb{M}_{\mathfrak{C}}$ is not closed under strong model-independent arbitrage.

2.2.4 Extrapolation of variance

We propose to restrict the set of market models $\mathbb{M}_{\mathfrak{C}}$ when only finitely many options are traded for each maturity by imposing conditions on arbitrage-free extrapolation of the total implied variance. It was discussed in Section 2.1.2 that for a fixed maturity behaviour of the total implied variance at extreme strikes is linked to existence of higher (or negative) moments of the price process S under a market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$.

Assumption 9. There exist constants $p^* > 0$ and $q^* > 0$ such that there is at least one market model $\mathbb{P} \in \mathbb{M}_{\mathfrak{C}}$ under which the process S possesses finite higher moments of order at most $1 + p^*$ and finite negative moments of order at most q^* up to the final maturity T.

The set of martingale measures that satisfies Assumption 9 is defined as

$$\mathbb{M}^{p^*,q^*} := \mathbb{M} \cap \Big\{\mathbb{P} \in \mathcal{P}(\Omega) : \\ \int_\Omega \omega^{1+p^*}\mathbb{P}(\mathrm{d}\omega) < \infty, \int_\Omega \omega^{-q^*}\mathbb{P}(\mathrm{d}\omega) < \infty\Big\}, \tag{2.32}$$

and the set of market models that satisfies Assumption 9 is then defined as

$$\mathbb{M}_{\mathfrak{C}}^{p^*,q^*} := \mathbb{M}_{\mathfrak{C}} \cap \mathbb{M}^{p^*,q^*}.$$

Let us define a function $f : \mathbb{R}_+ \to \mathbb{R}$ as $f(x) := x^{1+p^*} + x^{-q^*}$ and define a function $h : \Omega \to \mathbb{R}$ as

$$h(\omega) := \sum_{t\in\mathcal{T}} f(S_t(\omega)). \tag{2.33}$$

We shall use the same notation as in Equation (2.20) as there should be no confusion arising from it. From now on we work on the weighted Banach space $C_h(\Omega)$ and its topological dual $\mathcal{P}_h(\Omega)$. As the function l introduced in (2.17) is such that $l = \mathcal{O}(h)$, hence $M \subset C_l(\Omega) \subset C_h(\Omega)$ and it follows that the set of restricted models can be equivalently written as

$$\mathbb{M}_{\mathfrak{C}}^{p^*,q^*} = \left\{\mathbb{P} \in \mathcal{P}_h(\Omega) : \int_\Omega \Pi_T(\mathfrak{f}, \Delta; \omega)\mathbb{P}(\mathrm{d}\omega) = \Pi_{t_0}(\mathfrak{f}, \Delta) \text{ for } (\mathfrak{f}, \Delta) \in \mathcal{S} \times \mathcal{H}\right\}.$$

Recall that the total implied variance is denoted by $w : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}_+$. We now state assumption regarding extrapolation of the total implied variance for each maturity $t \in \mathcal{T}$. Assumptions are formulated in log-moneyness k to be consistent with notation of Section 2.1.2.

Assumption 10. Given $K_1^t > 0$ for some $t \in \mathcal{T}$, the extrapolation of the total implied variance for all $t \in \mathcal{T}$ and $k < k_1^t$, where $k_1^t := \log(K_1^t)$, reads

$$w(k, t) := f_L(k - k_1^t, t) + w(k_1^t, t),$$

where the function $f_L : \mathbb{R} \times \mathcal{T} \to \mathbb{R}_+$ is defined such that

(A) $f_L(0, \cdot) = 0$;

(B) $f_L(k,\cdot) = \mathcal{O}(\psi(q)k)$ for small enough k and $0 < q < q^*$ such that function g defined in (2.13) is non-negative for all $k < k_1^t$ with $t \in \mathcal{T}$;

(C) $\partial_t f_L(\cdot, t) \geq 0$ for all $t \in \mathcal{T}$.

Assumption 11. Given $K_*^t = \infty$ for some $t \in \mathcal{T}$, the extrapolation of the total implied variance for all $t \in \mathcal{T}$ and $k > k_{\kappa(t)}^t$, where $k_{\kappa(t)}^t := \log(K_{\kappa(t)}^t)$, reads

$$w(k, s) := f_R(k - k_{\kappa(t)}^t, t) + w(k_{\kappa(t)}^t, t),$$

where the function $f_R : \mathbb{R} \times \mathcal{T} \to \mathbb{R}_+$ is defined such that

(A) $f_R(0, \cdot) = 0$;

(B) $f_R(k, \cdot) = \mathcal{O}(\psi(p)k)$ for large enough k and $0 < p < p^*$ such that function g defined in (2.13) is non-negative for all $k > k_{\kappa(s)}^s$ and all $s \leq t$;

(C) $\partial_t f_R(\cdot, t) \geq 0$ for all $t \in \mathcal{T}$.

Assumptions 10 and 11 imply that extrapolation can be done linearly as long as the resulting total implied variance surface is consistent with the observed market prices (Assumptions 10(A) and 11(B)) and free of arbitrage, i.e. Assumptions 10(B)-(C) and 11(B)-(C) are satisfied. In particular Assumptions 10(B) and 11(B) ensure the extrapolation is free of butterfly spread arbitrage and can be checked using results in Proposition 2.7 and Corollary 2.8. Assumptions 10(C) and 11(C) ensure the extrapolation is free of calendar spread arbitrage.

Remark 2.28. As the underlying itself can be treated as an option with moneyness equal to zero, one can interpolate linearly between the traded option with the smallest available moneyness and the option with the zero moneyness. Therefore Assumption 10 appears to be superfluous. However interpolating linearly is only a crude approximation of the marginal distribution's behaviour near zero, whereas specifying extrapolation of the left wing of the total implied variance allows for a finer approximation (albeit parametric).

Lemma 2.29. *Under Assumptions 10 and 11 the Call price surface C resulting from the total implied variance extrapolation is free of weak free lunch.*

Proof. Assumption 11 ensures that $\lim_{k\to+\infty} c_{BS}(k, \sqrt{w(k,t)}) = 0$ for each $t \in \mathcal{T}$ as by Lemma 2.9. Similarly the limit $\lim_{k\to-\infty} c_{BS}(k, \sqrt{w(k,t)})$ is equal to 1 for each $t \in \mathcal{T}$ as a consequence of Assumption 10 and Lemma 2.10. It in turn implies absence of weak free lunch with vanishing risk (see Remark 2.24) and hence absence of weak free lunch by Lemma 2.16. $\square$

Remark 2.30. It must be noted that the extrapolation of the total implied variance as discussed above results in restriction of the feasible set of the dual problems (2.27) and (2.29). In order to avoid emergence of a duality gap between the primal and the dual problems, the feasible sets of the primal problems (2.26) and (2.28) must be enlarged. In particular untraded Call options that are priced from the extrapolation must be added to the set of static positions $\mathcal{S}$ and hence the set becomes infinite-dimensional. Additional care must be taken as the cone W introduced in Assumption 8 could potentially be enlarged as well as a result of the extrapolation of the total implied variance.

3

Semi-Infinite Approximation of Model-Independent Hedging Problems

The linear programming problems considered so far were infinite-dimensional, in that decision variables were infinite-dimensional and the number of constraints was infinite. In this chapter we reduce the super- and sub-hedging problems to semi-infinite programmes by considering finite-dimensional subspaces of the space of traded securities $\mathfrak{M}$. In Section 3.1 we take a finite subset of traded Call options if $|\mathfrak{K}_t| = \infty$ for each maturity $t \in \mathcal{T}$, otherwise we take all available Call options. We also discretize trading strategies by taking a basis of functions in $\mathcal{H}$, Definition 2.13. This results in only finitely many decision variables (weights in a hedging portfolio) and finitely many constraints in the dual problem. Section 3.2 focuses on further reduction of the dimension of the problem, by discretizing the state space, which respects the consistency of the primal and the dual problems with observed option prices, and yields robust numerical results. Our discretization method differs from that presented in Henry-Labordère (2013), and requires far fewer points, thus reducing the complexity of the finite-dimensional LPs to be solved, thereby improving the algorithmic speed. We also discuss how to discretize given marginal distributions so that the obtained discrete distributions are placed in convex order in Section 3.2.2. The finite-dimensional approximations to the semi-infinite primal and the dual problems are presented in Section 3.2.3. An application to bounds on forward implied volatility is given in Section 3.3. In particular, we specialise our computation of the upper and lower bounds to the cases where the marginal distributions are generated from the Black-Scholes model (lognormal marginals) and from the Heston stochastic volatility model. In the lower bound at-the-money case we numerically solve, in Section 3.4, the coupled ordinary differential equations associated with the Hobson-Klimmek transference plan (Hobson and Klimmek, 2014), and show that it is in striking agreement with the LP solution of the primal problem. This chapter is based on Badikov, Jacquier, Liu, and Roome (2017) published in Quantitative Finance.

3.1 Reduction to the semi-infinite case

In this section we discuss reduction of infinite-dimensional super- and subhedging problems defined in (2.26) and (2.28) respectively to the semi-infinite case. We start by selecting a finite subset of traded options to approximate elements of the set of static positions $\mathcal{S}$ presented in Definition 2.12. In case when only finitely many Call options are traded on the market, we perform extrapolation of the total implied variance according to Assumptions 10 and 11 and include Call options with prices corresponding to such extrapolation. Note that those options may not be traded on the market. We define a vector of Call option pay-offs as

$$\mathrm{C} := \big((S_t - K_1^t)_+, \dots, (S_t - K_{\kappa(t)}^t)_+\big)_{t\in\mathcal{T}},$$

and a vector of corresponding market prices as

$$\mathrm{c} := (c(K_1^t, t), \dots, c(K_{\kappa(t)}^t, t))_{t\in\mathcal{T}} \in \mathbb{R}^d,$$

where $d := \sum_{t\in\mathcal{T}} \kappa(t) < \infty$. We shall also write

$$\mathrm{C}(\omega) := \big((S_t(\omega) - K_1^t)_+, \dots, (S_t(\omega) - K_{\kappa(t)}^t)_+\big)_{t\in\mathcal{T}} \in \mathbb{R}^d$$

to denote the evaluation of the Call options pay-offs at $\omega \in \Omega$.

Assumption 12. The Call option prices c satisfy absence of weak arbitrage and the cone W defined in Assumption 8 is trivial, i.e. $W = \{0\}$.

Remark 3.1. Note that under Assumption 12 the super- and sub-hedging problems defined in (2.26) and (2.28) are equivalent to problems defined in (1.9) and (1.11) respectively.

The set of approximate static positions is then defined as $\tilde{\mathcal{S}} := \mathbb{R} \times \mathrm{Span}\,\{\mathrm{C}\}$. As a consequence of the Breeden-Litzenberger formula (2.31) the risk-neutral marginal distributions μ_t of the asset process S can be uniquely recovered for every maturity $t \in \mathcal{T}$ from the traded Call option prices if options are traded for all $K \in \mathbb{R}_+$. If we have only finitely many options for each maturity $t \in \mathcal{T}$ then the right derivative in (2.31) can be approximated as

$$\partial_+ C(K_i^t, t) \simeq \frac{c(K_{i+1}^t, t) - c(K_i^t, t)}{K_{i+1}^t - K_i^t},$$

for all $i = 1, \dots, \kappa(t)$ and C is the Call price surface defined in Section 2.2.3. Naturally if one takes a sequence of rational numbers $(K_n)_{n\in\mathbb{N}}$ that becomes denser as $n \to \infty$ then one can estimate $\mu([0, K])$ for all $K \in \mathbb{R}_+$. Thus it is sufficient to take only a countable subset of all traded Call options for each maturity $t \in \mathcal{T}$ to recover μ_t.

We shall also discretize elements of the set of trading strategies $\mathcal{H}$ introduced in Definition 2.13. For a rational number $\alpha \in \mathbb{Q}$ let $K_\alpha^j := [0, \alpha]^j$ where $j = 1, \dots, n-1$ and define a set of functions $B := \{\theta_i^{t_j} \in C_b(\mathbb{R}_+^j),\ j = 1, \dots, n-1,\ i \in \mathbb{N}\}$ such that for each j and α the set $\{\mathbb{1}_{K_\alpha^j}\theta_i^{t_j},\ i \in \mathbb{N}\}$ is dense in $C(K_\alpha^j)$. Let us also define a finite subset $B_j := \{\theta_1^{t_j}, \dots, \theta_{l(t_j)}^{t_j}\}$ with $l(t_j) < \infty$ of elements in B for each $j = 1, \dots, n-1$ (for instance, one can take a set of monomials defined on K_α^j for each j and α and extend each element in the set to $\mathbb{R}_+^j$ such that the extension is equal to the maximum of the element on K^j_α on the complement of K^j_α and is equal to the element itself otherwise). Then a discretized trading strategy $\Theta \in \tilde{\mathcal{H}} := \mathbb{R}\times\prod_{j=1}^{n-1} \operatorname{Span} B_j$ is defined as follows

$$\Theta(\omega) := \left(a_0, \langle \mathrm{a}^{t_1}, \theta^{t_1}(\omega)\rangle, \ldots, \langle \mathrm{a}^{t_{n-1}}, \theta^{t_{n-1}}(\omega)\rangle\right),$$

for each $\omega \in \Omega$, where $a_0 \in \mathbb{R}$, $\mathrm{a}^{t_j} := (a_i^{t_j})_{i=1,\ldots,l(t_j)}$ are real vectors and $\theta^{t_j}(\omega) := (\theta_1^{t_j}(\omega), \ldots, \theta_{l(t_j)}^{t_j}) \in \mathbb{R}_+^{l(t_j)}$ are the evaluation vectors of functions for each time period t_j with $j = 1, \ldots, n-1$. Note that $\theta^{t_j}(\omega) := \theta^{t_j}(\operatorname{Pr}\omega)$, where $\operatorname{Pr}\omega$ is the projection of $\omega \in \Omega$ onto $\mathbb{R}_+^j$ for each $j = 1, \ldots, n-1$. The pay-off of a discretized trading strategy $\Theta \in \tilde{\mathcal{H}}$ then reads

$$(\Theta \bullet S)_T = a_0(S_{t_1} - s_0) + \sum_{j=1}^{n-1}\sum_{i=1}^{l(t_j)} a_i^{t_j}\theta_i^{t_j}\left(S_{t_{j+1}} - S_{t_j}\right). \tag{3.1}$$

Recall definitions of the initial cost $\Pi_{t_0}(\mathfrak{f}, \Delta)$ of a semi-static portfolio $(\mathfrak{f}, \Delta) \in \mathcal{S}\times\mathcal{H}$ and its final pay-off $\Pi_T(\mathfrak{f}, \Delta; \cdot)$ introduced in (2.16). In the present set-up we write the initial cost of a discretized hedging portfolio $(\tilde{\mathfrak{f}}, \Theta) \in \tilde{\mathcal{S}} \times \tilde{\mathcal{H}}$ as

$$\Pi_{t_0}(\tilde{\mathfrak{f}}, \Theta) = \langle \mathrm{c}, \mathrm{w}\rangle + \lambda,$$

where $\lambda \in \mathbb{R}$, the vector $\mathrm{w} = (w_1^t, \ldots, w_{\kappa(t)}^t)_{t\in\mathcal{T}} \in \mathbb{R}^d$ with entries denoting portfolio weights in available options and $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^d$. We also write the pay-off of the hedging portfolio $(\tilde{\mathfrak{f}}, \Theta)$ at the final maturity T $\Pi_T(\tilde{\mathfrak{f}}, \Theta) = A(\mathrm{w}, \lambda, \Theta)$ where the linear map A is defined as

$$A(\mathrm{w}, \lambda, \Theta; \cdot) := \lambda + \sum_{t\in\mathcal{T}}\sum_{i=1}^{\kappa(t)} w_i^t(S_t - K_i^t)_+ + (\Theta\bullet S)_T = \lambda + \mathrm{C}\mathrm{w} + (\Theta\bullet S)_T. \tag{3.2}$$

We can then write a problem of super-hedging an option with the upper semi-continuous pay-off $\Phi \in U_l(\Omega)$ as follows

$$\overline{\vartheta}_p(\Phi) := \inf\left\{\langle \mathrm{c}, \mathrm{w}\rangle + \lambda : (\mathrm{w}, \lambda, \Theta) \in \overline{\mathcal{F}}_p\right\}, \tag{3.3}$$

where the feasible set $\overline{\mathcal{F}}_p$ is defined as

$$\overline{\mathcal{F}}_p := \left\{(\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}} : A(\mathrm{w}, \lambda, \Theta; \omega) - \Phi(\omega) \geq 0 \text{ for all } \omega \in \Omega\right\}.$$

The dual problem is then formulated as

$$\overline{\vartheta}_d(\Phi) := \sup \left\{ \int_\Omega \Phi(\omega)\mu(\mathrm{d}\omega) : \mu \in \widetilde{\mathbb{M}}_{\mathfrak{C}}^{p^*,q^*} \right\}. \tag{3.4}$$

The set of Borel probability measures that re-price the discretized portfolios in $\tilde{\mathcal{S}} \times \tilde{\mathcal{H}}$ is defined as

$$\widetilde{\mathbb{M}}_{\mathfrak{C}}^{p^*,q^*} := \left\{ \mu \in \mathcal{P}_h(\Omega) : \int_\Omega \Pi_T(\tilde{\mathfrak{f}}, \Theta; \omega)\mu(\mathrm{d}\omega) = \Pi_0(\tilde{\mathfrak{f}}, \Theta),\ (\tilde{\mathfrak{f}}, \Theta) \in \tilde{\mathcal{S}} \times \tilde{\mathcal{H}} \right\},$$

where the function h is defined in (2.33) and real constants $p^* > 0$ and $q^* > 0$ are defined in Assumption 9. We define the sub-hedging primal problem for an option with the lower semi-continuous pay-off $\Phi \in L_l(\Omega)$ as follows

$$\underline{\vartheta}_p(\Phi) := \sup \left\{ \langle \mathrm{c}, \mathrm{w} \rangle + \lambda : (\mathrm{w}, \lambda, \Theta) \in \underline{\mathcal{F}}_p \right\}, \tag{3.5}$$

where the feasible set $\underline{\mathcal{F}}_p$ is defined as

$$\underline{\mathcal{F}}_p := \left\{(\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}} : A(\mathrm{w}, \lambda, \Theta; \omega) - \Phi(\omega) \leq 0 \text{ for all } \omega \in \Omega\right\}.$$

The dual problem is then written as follows

$$\underline{\vartheta}_d(\Phi) := \inf \left\{ \int_\Omega \Phi(\omega)\mu(\mathrm{d}\omega) : \mu \in \widetilde{\mathbb{M}}_{\mathfrak{C}}^{p^*,q^*} \right\}. \tag{3.6}$$

We now show that the primal (3.3) and the dual (3.4) programmes admit no duality gap.

Proposition 3.2. *Under Assumptions 9 and 12 there is no duality gap between the primal* (3.3) *and the dual* (3.4) *problems.*

Proof. By Lemma 2.29 Assumptions 9 and 12 imply absence of weak free lunch. As all portfolio pay-offs $A(\mathrm{w}, \lambda, \Theta; \cdot)$ and the option pay-off Φ grow linearly, they satisfy Assumption 5 where h is defined in (2.33). Moreover as the riskless bond satisfies Assumption 2 the no duality result follows by Theorem 1.16. □

As the sub-hedging primal problem (3.5) can be represented as $\underline{\vartheta}_p(\Phi) = -\overline{\vartheta}_p(-\Phi)$ and the sub-hedging dual problem (3.6) is represented in terms of the super-hedging dual problem (3.4) as $\underline{\vartheta}_d(\Phi) = -\overline{\vartheta}_d(-\Phi)$, we have the following straightforward corollary which is presented without proof.

Corollary 3.3. *Under Assumptions 9 and 12 there is no duality gap between the primal* (3.5) *and the dual* (3.6) *problems.*

We now show that if the number of trading strategies in $\tilde{\mathcal{H}}$ increases to infinity the semi-infinite primal (3.3) and the dual (3.4) problems converge to the values of the infinite-dimensional problems defined in (2.26) and (2.27) respectively (similarly the sub-hedging primal (3.5) and dual (3.6) converge to (2.28) and (2.29) respectively).

Lemma 3.4. *Let $r := \min_{t\in\mathcal{T}}\{l(t)\}$. As $r \to \infty$ the set $\widetilde{\mathbb{M}}_{\mathfrak{c}}^{p^*,q^*}$ converges to the set of martingale measures $\mathbb{M}_{\mathfrak{c}}^{p^*,q^*}$ consistent with observed Call option prices* c *traded on the market for each $t \in \mathcal{T}$.*

Proof. It is sufficient to show that the limit of sets $\widetilde{\mathbb{M}}_r^{p^*,q^*}$ defined as

$$\widetilde{\mathbb{M}}_r^{p^*,q^*} := \left\{\mu \in \mathcal{P}_h(\Omega) : \int_\Omega (\Theta \bullet S(\omega))_T \,\mu(\mathrm{d}\omega) = 0, \text{ for all } \Theta \in \tilde{\mathcal{H}}\right\},$$

the set of probability measures in $\mathcal{P}_h(\Omega)$ that integrate $(\Theta \bullet S)_T$ to zero for all $\Theta \in \tilde{\mathcal{H}}$ (where $\tilde{\mathcal{H}}$ is dependent on r via the choice of $l(t_j)$ for all $j = 1, \dots, n-1$) converges to the set of all martingale measures $\mathbb{M}^{p^*,q^*}$ defined in (2.32).

Let us define for each $j = 1, \dots, n-1$ the sets $B_j^\infty := \lim_{l(t_j)\to\infty} B_j$ and let $\tilde{\mathcal{H}}_\infty := \mathbb{R} \times \prod_{j=1}^{n-1} B_j^\infty$. It is clear that $B = \cup_{j=1}^{n-1} B_j^\infty$. Also define the limit $\widetilde{\mathbb{M}}_\infty^{p^*,q^*} := \lim_{r\to\infty} \widetilde{\mathbb{M}}_r^{p^*,q^*}$. It is clear that $\mathbb{M}^{p^*,q^*} \subseteq \widetilde{\mathbb{M}}_\infty^{p^*,q^*}$. To show the reverse inclusion define the gain of the trading strategy $\Theta \in \tilde{\mathcal{H}}$ at time $t \le T$ as

$$(\Theta \bullet S)_t := a_0(S_{t_1} - s_0) + \sum_{j=1}^{t} \sum_{i=1}^{l(t_j)} a_i^{t_j} \theta_i^{t_j} \left(S_{t_{j+1}} - S_{t_j}\right),$$

where the upper limit of the first sum t is understood as $\max\{t_j : j = 1, \dots, n-1, \text{ and } t_j \leq t\}$. Also define a stopping time $\tau_\alpha := \min\{t_j : j = 1, \dots, n,\ S_{t_j} > \alpha\}$. The set $\widetilde{\mathbb{M}}_\infty^{p^*,q^*}$ consists of all probability measures $\mu \in \mathcal{P}_h(\Omega)$ such that $\langle (\Theta \bullet S)_{T\wedge\tau_\alpha}, \pi \rangle = 0$ for each $\alpha \in \mathbb{Q}$. Moreover by the definition of the set B, for each α and j any function $f \in C_b(K_\alpha^j)$ can be approximated by the elements in B. In particular the indicator function $\mathbb{1}_{K_\alpha^j}$ can be approximated by elements in B and hence $\langle (\Theta_0 \bullet S)_{T\wedge\tau_\alpha}, \pi \rangle = 0$, where $\Theta_0 := (a_0, \mathbb{1}_{K_\alpha^1}, \dots, \mathbb{1}_{K_\alpha^{n-1}})$ and as for all $\alpha \in \mathbb{Q}$ and each j the sets K_α^j generate Borel sigma algebra on $\mathbb{R}_+^j$ it follows that $S_{T\wedge\tau_\alpha}$ is a martingale under π. Hence by definition of a local martingale it follows that $(S_{t_j})_{j=1,\dots,n}$ is a local martingale. Moreover as S_T is integrable with respect to any measure $\mu \in \mathcal{P}_h(\Omega)$ it follows from Jacod and Shirayev (1998, Theorem 2(b)) that it is a martingale under any $\pi \in \widetilde{\mathbb{M}}_\infty^{p^*,q^*}$ and hence $\widetilde{\mathbb{M}}_\infty^{p^*,q^*} \subseteq \mathbb{M}_\infty^{p^*,q^*}$. $\square$

We then show that convergence of the semi-infinite approximations to their infinite-dimensional counterparts.

Proposition 3.5. *Assume that there is no duality gap between the primal* (3.3) *and the dual* (3.4) *programmes for each $l(t)$ for all $t \in \mathcal{T}$. Also assume that there is no duality gap between the values of the infinite-dimensional primal* (2.26) *and the dual* (2.27) *problems. Then as $r = \min_{t\in\mathcal{T}}\{l(t)\} \to \infty$ the values of both semi-infinite programmes converge to the values of their infinite-dimensional counterparts.*

Proof. As in the proof of Lemma 3.4, we let $\tilde{\mathcal{H}}_\infty$ be a countable subset of $\mathcal{H}$. The sequence of nested sets $(\tilde{\mathcal{H}}_r)_{r\in\mathbb{N}}$ with $\tilde{\mathcal{H}}_r \subset \tilde{\mathcal{H}}_{r+1}$ represents the set of discretized trading strategies as the bases B_j increase for each $j = 1, \dots, n-1$ simultaneously. It is clear that $\tilde{\mathcal{H}}_\infty = \lim_{r\to\infty} \tilde{\mathcal{H}}_r$ as the limit of the sequence. Let us also denote the primal problem (3.3) over the set of primal variables $\mathbb{R}^{d+1} \times \tilde{\mathcal{H}}_r$ as $\overline{\vartheta}_p^r(\Phi)$ for all $r \in \mathbb{N}$. Let us also denote the dual problem (3.4) over the set of probability measures in $\mathcal{P}_h(\Omega)$ that re-price given Call options C and satisfy the martingale condition (2.19) for all $\Theta \in \tilde{\mathcal{H}}_r$ as $\overline{\vartheta}_d^r(\Phi)$ for all $r \in \mathbb{N}$. By assumption of the proposition there is no duality gap between the primal and the dual problems, i.e. $\overline{\vartheta}_p^r(\Phi) = \overline{\vartheta}_d^r(\Phi)$ for all $r \in \mathbb{N}$ and as both sequences $(\overline{\vartheta}_p^r(\Phi))_{r\in\mathbb{N}}$ and $(\overline{\vartheta}_d^r(\Phi))_{r\in\mathbb{N}}$ are non-increasing we have that limits of both sequences exist and $\lim_{r\to\infty}\overline{\vartheta}_p^r(\Phi) = \lim_{r\to\infty}\overline{\vartheta}_d^r(\Phi)$. We also define the limit $\overline{\vartheta}_d^\infty(\Phi) := \lim_{r\to\infty}\overline{\vartheta}_d^r(\Phi) > -\infty$ with

$$\overline{\vartheta}_d^\infty(\Phi) := \left\{ \int_\Omega \Phi(\omega)\mu(\mathrm{d}\omega) : \mu \in \widetilde{\mathbb{M}}_\infty^{p^*,q^*},\ \int_\Omega \mathrm{C}(\omega)\mu(\mathrm{d}\omega) = \mathrm{c} \right\},$$

where the set $\widetilde{\mathbb{M}}_\infty^{p^*,q^*}$ is defined in the proof of Lemma 3.4. We conclude that the value of the semi-infinite dual problem (3.4) converges to the value of the infinite-dimensional dual problem (2.27) as a consequence of Lemma 3.4. It follows that the value of the semi-infinite primal problem (3.3) also converges to the value of the infinite-dimensional primal problem (2.26) as $\lim_{r\to\infty}\overline{\vartheta}_p^r(\Phi) = \lim_{r\to\infty}\overline{\vartheta}_d^r(\Phi)$ and there is no duality gap between the infinite-dimensional primal (2.26) and the dual (2.27) problems by assumption of the proposition. □

The simple corollary below shows that the same convergence result holds for the sub-hedging problems.

Corollary 3.6. *Assume that there is no duality gap between the primal* (3.5) *and the dual* (3.6) *programmes for each $l(t)$ for all $t \in \mathcal{T}$. Also assume that there is no duality gap between the values of the infinite-dimensional primal* (2.28) *and the dual* (2.29) *problems. Then as $r = \min_{t\in\mathcal{T}}\{l(t)\} \to \infty$ the values of both semi-infinite programmes converge to the values of their infinite-dimensional counterparts.*

3.2 Arbitrage-consistent discretization of the primal and the dual programmes

In this section we show how to reduce semi-infinite linear problems (3.3), (3.4), (3.5) and (3.5) to finite-dimensional LPs. In order to do so we discretize the state space Ω. If the marginal distributions of the process S are uniquely specified at each maturity $t \in \mathcal{T}$ (i.e. Call options are traded for each strike $K \in \mathbb{R}_+$ at each maturity) we also provide a way to discretize the marginal distributions that preserves convex order. Following Strassen (1965), we shall say that the two measures μ and ν on $\mathbb{R}_+$ are in convex, or balayage, order (which we denote $\mu \leq_{\mathrm{cv}} \nu$) if they have equal means and satisfy $\int_{\mathbb{R}_+}(y-x)_+\mu(\mathrm{d}y) \leq \int_{\mathbb{R}_+}(y-x)_+\nu(\mathrm{d}y)$ for all $x \in \mathbb{R}_+$. This assumption ensures that there exists at least one martingale measure that connects the two measures μ and ν. However if one discretizes the two measures then in general there is no guarantee that the obtained discrete measures will still be in convex order, see for more details Section 3.2.2 and Baker (2012).

3.2.1 No-arbitrage discretization

We say that for each $t \in \mathcal{T}$ the price process S has a marginal law denoted by μ_t, which is the distribution of the random variable S_t. We note that the law μ_t of S_t is not necessarily uniquely determined. It can satisfy only finitely many constraints that ensure that μ_t is consistent with market prices of traded securities.

Fix $m > 1$ and suppose that we are given a set $\mathrm{x} = (x_1, \ldots, x_m) \in \mathbb{R}_+^m$ of points $0 < x_1 < x_2 < \cdots < x_m$ in the support of μ_t, and a discrete distribution q with atoms q_i placed at the point x_i (for example if μ_t admits density f then one can take $q_i = f(x_i)/\sum_{i=1}^m f(x_i)$). We wish to find a discrete distribution p, close to q, satisfying the 'martingale' condition $\langle \mathrm{p}, \mathrm{x}\rangle = 1$ and consistent with a given set of European option prices. Borwein et al. (2003) suggested to recover discrete probability distributions from observed market prices of European Call options by minimising the Kullback-Leibler divergence to the uniform distribution (they also comment that any prior distribution can be chosen). In addition if the first $l \leq m$ moments of μ_t are known we wish to find the discrete distribution p to match given moments. The moments can be specified by extrapolating the total implied variance as discussed in Section 2.1.3. Extrapolation is performed such that Assumptions 10 and 11. In particular let $\mathrm{T} : \mathbb{R}_+ \to \mathbb{R}_+^l$ be given by $\mathrm{T}(x) := (x, x^2, ..., x^l)$ and define the moment vector $\mathfrak{T} := \int_{\mathbb{R}_+} \mathrm{T}(x)\mu_t(\mathrm{d}x) \in \mathbb{R}_+^l$. Given the law μ_t of S_t and a set of European Call options maturing at t with given market prices $\mathfrak{P} \in \mathbb{R}_+^M$ satisfying absence of weak arbitrage (otherwise the problem will be infeasible), and strikes $K_1, \dots, K_M$ we can solve the minimisation problem:

$$\min_{\mathrm{p}\in[0,1]^m:\|\mathrm{p}\|=1} \sum_{i=1}^m p_i \log\left(\frac{p_i}{q_i}\right), \text{ subject to}$$

$$\left((\langle \mathrm{C}_j(\mathrm{x}), \mathrm{p}\rangle)_{j=1,\dots,M}, (\langle \mathrm{T}_j(\mathrm{x}), \mathrm{p}\rangle)_{j=1,\dots,l} \right) = (\mathfrak{P}, \mathfrak{T}).$$

for some prior discrete distribution q, where

$$\mathrm{C}_j(\mathrm{x}) := ((x_1 - K_j)_+, \dots, (x_m - K_j)_+), \text{ for } j = 1, \dots, M,$$

denotes the payoff vector of the options and $\mathrm{T}_j(\mathrm{x}) := (x_1^j, \dots, x_m^j)$ is the j-th moment vector ($j = 1, \dots, l$).

Definition 3.7. The discrete distribution p is consistent with $\mathfrak{P}$ whenever the following hold:

(i) $\sum_{i=1}^m p_i = 1$;

(ii) $\langle \mathrm{C}_j(\mathrm{x}), \mathrm{p}\rangle = \sum_{i=1}^m p_i \mathrm{C}_j(x_i) = \mathfrak{P}_j$ for all $j = 1, \dots, M$;

(iii) $\langle \mathrm{T}_1(\mathrm{x}), \mathrm{p}\rangle = \langle \mathrm{x}, \mathrm{p}\rangle = \mathfrak{T}_1 = 1$;

The last item in the definition above is just the martingale condition. It must be noted that if the full marginal distribution μ_t of S_t is known, any finite subset of European options can be chosen above and the price vector can be defined as $\mathfrak{P} := \int_{\mathbb{R}_+} \mathrm{C}(x)\mu(\mathrm{d}x)$, where $\mathrm{C}(x) := ((x - K_1)_+, \dots, (x - K_M)_+) \in \mathbb{R}_+^M$ for any $x \in \mathbb{R}_+$. In particular the solution to this problem can be obtained as a modification of the solution in Tanaka and Toda (2013), which itself is based on arguments presented in Borwein and Lewis (1991, Corollary 2.6):

$$p_i = \frac{q_i \exp\left(\langle \lambda^*, (\mathrm{C}(x_i), \mathrm{T}(x_i))\rangle\right)}{\sum_{i=1}^m q_i \exp\left(\langle \lambda^*, (\mathrm{C}(x_i), \mathrm{T}(x_i))\rangle\right)},$$

$$\lambda^* := \underset{\lambda\in\mathbb{R}^{M+l}}{\operatorname{argmin}} \left\{ -\langle \lambda, (\mathfrak{P}, \mathfrak{T})\rangle + \log\left(\sum_{i=1}^m q_i \exp\left(\langle \lambda, (\mathrm{C}(x_i), \mathrm{T}(x_i))\rangle\right)\right)\right\}. \tag{3.7}$$

We can now consider the following algorithm:

Algorithm 3.8.

(i) Choose the m points $0 < x_1 < ... < x_m$; various choices are possible, for instance:

  (a) *Binomial*: Let Σ denote the at-the-money lognormal volatility (for European options maturing at t).
  Set $\delta := t/(m-1)$, $u := 1 + \left(\mathrm{e}^{\delta\Sigma^2} - 1\right)^{1/2}$, $d := 1 - \left(\mathrm{e}^{\delta\Sigma^2} - 1\right)^{1/2}$, and $x_i := u^{i-1}d^{m-i}$;

  (b) *Gauss-Legendre*: $x_i := \mathrm{e}^{\tilde{x}_i}$, where $\tilde{x}_1, ..., \tilde{x}_m$ are the nodes of an N-point Gauss-Legendre quadrature.

(ii) For the discrete distribution q, we can follow several routes:

  (a) if μ admits a density f_μ, then, for $i = 1, \dots, m$, set $q_i := \dfrac{f_\mu(x_i)}{\sum_{j=1}^m f_\mu(x_j)}$;

  (b) alternatively, for $i = 1, ..., m$, let $q_i := \mu([x_{i-1}, x_i))$ (with $x_0 = 0$);

(iii) Compute the discretised measure p by solving (3.7).

Remark 3.9. As pointed out in Tanaka and Toda (2013) the choice of discretization points $x_1, \dots, x_m$ in Algorithm 3.8 dictates the discretization of the integral $\int_{\mathbb{R}_+} \mathrm{T}(x) f_\mu(x)\mathrm{d}x \approx \sum_{i=1}^m w(x_i)\mathrm{T}(x_i)f_\mu(x_i)$. The weights $w(\cdot)$ are chosen in accordance with a given quadrature rule; in the case of Algorithm 3.8 the weights can be chosen constant $w(x_i) = 1/\sum_{i=1}^m f_\mu(x_i)$ for all $i = 1, \dots, m$ or as weights in the Gauss-Legendre quadrature. The number of points in the N-point Gauss-Legendre quadrature is chosen so that integrals of the given moments are computed exactly.

In the particular case where the discretization nodes and the given strikes satisfy some specific ordering, it is possible to choose a more explicit weighting scheme p consistent with $\mathfrak{P}$ with minimal assumptions on the choice of discretization:

Lemma 3.10. *Let $m = M + 2$. Consistency with $\mathfrak{P}$ is ensured if both sets of conditions hold:*

(i) $x_1 < K_1$, $x_{M+2} \geq \left(\mathfrak{P}_{M-1}K_M - \mathfrak{P}_M K_{M-1}\right) / \left(\mathfrak{P}_{M-1} - \mathfrak{P}_M\right)$ *and* $x_{i+1} = K_i$ *for* $i = 1, \ldots, M$*;*

(ii) $p_{M+2} = \mathfrak{P}_M / \left(x_{M+2} - K_M\right)$, $p_1 = 1 - \sum_{j=2}^{M+2} p_j$, *and, for any* $i = M, \ldots, 1$,

$$p_{i+1} = \frac{1}{x_{i+1} - K_{i-1}} \left(\mathfrak{P}_{i-1} - \sum_{j=i+2}^{M+2} p_j (x_j - K_{i-1}) \right),$$

with the convention $K_0 := x_1$ *and* $\mathfrak{P}_0 := 1 - x_1$.

Proof. The consistency of the weighting scheme with $\mathfrak{P}$ can be written as $\mathrm{A}\mathrm{p}^{\mathrm{x}} = \overline{\mathfrak{P}}$, with

$$\mathrm{A} := \begin{pmatrix} 1 & 1 & \ldots & \ldots & 1 \\ x_1 & x_2 & \ldots & \ldots & x_{M+2} \\ 0 & (x_2 - K_1) & \ddots & \ddots & (x_{M+2} - K_1) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & (x_M - K_{M-1}) & (x_{M+2} - K_{M-1}) \\ 0 & \ldots & 0 & 0 & (x_{M+2} - K_M) \end{pmatrix},$$

$$\mathrm{p}^{\mathrm{x}} := \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \\ p_{M+2} \end{pmatrix}, \qquad \overline{\mathfrak{P}} := \begin{pmatrix} 1 \\ 1 \\ \mathfrak{P}_1 \\ \vdots \\ \mathfrak{P}_M \end{pmatrix}.$$

As the strikes are ordered and $\overline{\mathfrak{P}}$ satisfies no-arbitrage conditions by Assumption 12, the condition (i) in the statement of the lemma

$$x_{M+2} \geq \left(\mathfrak{P}_{M-1}K_M - \mathfrak{P}_M K_{M-1}\right) / \left(\mathfrak{P}_{M-1} - \mathfrak{P}_M\right)$$

immediately implies that $x_{M+2} > K_M$. Here A is a real upper triangular matrix, so that the system has a unique solution given in the lemma. It remains to check that the feasible solutions of the system satisfy the additional constraints $p_i \geq 0$ for $i = 1, \dots, M+2$.

Since $x_{M+2} \geq \frac{\mathfrak{P}_{M-1}K_M - \mathfrak{P}_M K_{M-1}}{\mathfrak{P}_{M-1} - \mathfrak{P}_M}$ it follows that

$$p_{M+2} = \frac{\mathfrak{P}_M}{x_{M+2} - K_M} \leq \frac{\mathfrak{P}_{M-1} - \mathfrak{P}_M}{K_M - K_{M-1}} \leq \frac{-K_{M-1} + K_M}{K_M - K_{M-1}},$$

where the last inequality follows from the Put-Call Parity and absence of arbitrage in $\overline{\mathfrak{P}}$. By definition we have

$$\begin{aligned} p_{M+1} &= \frac{1}{K_M - K_{M-1}} \left(\mathfrak{P}_{M-1} - p_{M+2}(x_{M+2} - K_{M-1})\right) \\ &= \frac{1}{K_M - K_{M-1}} \left(\mathfrak{P}_{M-1} - \mathfrak{P}_M - \mathfrak{P}_M \frac{K_M - K_{M-1}}{x_{M+2} - K_M}\right), \end{aligned}$$

and by assumption on x_{M+2} we have that $p_{M+1} \geq 0$. Similarly for p_M we have

$$\begin{aligned} p_M &= \frac{1}{K_{M-1} - K_{M-2}} \left[\mathfrak{P}_{M-2} - p_{M+1}(K_M - K_{M-2}) - p_M(x_{M+2} - K_{M-2})\right] \\ &= \frac{\mathfrak{P}_{M-2} - \mathfrak{P}_{M-1}}{K_{M-1} - K_{M-2}} - \frac{\mathfrak{P}_{M-1} - p_{M+2}(x_{M+2} - K_M)}{K_M - K_{M-1}} \\ &= \frac{\mathfrak{P}_{M-2} - \mathfrak{P}_{M-1}}{K_{M-1} - K_{M-2}} - \frac{\mathfrak{P}_{M-1} - \mathfrak{P}_M}{K_M - K_{M-1}}. \end{aligned}$$

Proceeding recursively we observe that for $i = 2, \dots, M-1$ we have

$$p_i = \frac{\mathfrak{P}_{i-2} - \mathfrak{P}_{i-1}}{K_{i-1} - K_{i-2}} - \frac{\mathfrak{P}_{i-1} - \mathfrak{P}_i}{K_i - K_{i-1}}.$$

Since the prices $\overline{\mathfrak{P}}$ satisfy no-arbitrage conditions, then $p_i \geq 0$ for $i = 1, \dots, M+2$. Note that, when $x_{M+2} = \frac{\mathfrak{P}_{M-1}K_M - \mathfrak{P}_M K_{M-1}}{\mathfrak{P}_{M-1} - \mathfrak{P}_M}$,

$$p_{M+2} = \frac{\mathfrak{P}_{M-1} - \mathfrak{P}_M}{K_M - K_{M-1}}, \qquad p_{M+1} = 0, \qquad p_i = \frac{\mathfrak{P}_{i-2} - \mathfrak{P}_{i-1}}{K_{i-1} - K_{i-2}} - \frac{\mathfrak{P}_{i-1} - \mathfrak{P}_i}{K_i - K_{i-1}},$$

for $i = 2, \dots, M$ and the discretization reduces to $M+1$ points, where $x_{M+1} = K_M$ is discarded. $\square$

3.2.2 Balayage-consistent discretization

Assume now that there are M European Call options $\mathfrak{P}^{t_1}$ available with maturity t_1 and N options $\mathfrak{P}^{t_2}$ available with maturity t_2 (with $t_1 < t_2$), and that the measures μ_{t_1} and μ_{t_2} are such that $\int_{\mathbb{R}_+}(x - K_i^{t_1})_+\mu_{t_1}(\mathrm{d}x) = \mathfrak{P}_i^{t_1}$ for all $i = 1, \dots, M$ and $\int_{\mathbb{R}_+}(y - K_j^{t_2})_+\mu_{t_2}(\mathrm{d}y) = \mathfrak{P}_j^{t_2}$ for all $j = 1, \dots, N$. We further assume that $\mathfrak{P}^{t_1}$ and $\mathfrak{P}^{t_2}$ satisfy absence of weak arbitrage and μ_{t_1} and μ_{t_2} satisfy the following assumption:

Assumption 13. The support of $\eta := (\mu_{t_1} - \mu_{t_2})_+$ is given by an interval $[a, b] \subset \mathbb{R}_+$, and the support of $\gamma := (\mu_{t_2} - \mu_{t_1})_+$ is given by $\mathbb{R}_+ \setminus [a, b]$.

For two discretization meshes $\mathrm{x} = (x_1, \dots, x_m) \in \mathbb{R}_+^m$ and $\mathrm{y} = (y_1, \dots, y_n) \in \mathbb{R}_+^n$, Algorithm 3.8, for instance, produces discrete distributions p^{x} and p^{y} supported on x and y. It is not however guaranteed a priori that resulting distributions p^{x} and p^{y} are in convex order. The following lemma provides general sufficient conditions ensuring this.

Lemma 3.11. *Let* p^{x} *and* p^{y} *be two discrete measures supported on* x *and* y. *The following conditions altogether ensure that* $\mathrm{p}^{\mathrm{x}} \leq_{\mathrm{cv}} \mathrm{p}^{\mathrm{y}}$*:*

(i) p^{x} *and* p^{y} *have the same mean;*

(ii) $y_1 \leq x_1$ *and* $y_n \geq x_m$*;*

(iii) Assumption 13 holds with $\mu_{t_1} = \mathrm{p}^{\mathrm{x}}$ *and* $\mu_{t_2} = \mathrm{p}^{\mathrm{y}}$.

Remark 3.12. In our framework, the discrete measures p^{x} and p^{y} arise as discretizations of the original measures μ_{t_1} and μ_{t_2}. If p^{x} and p^{y} are to be consistent with $\mathfrak{P}^{\mathrm{x}}$ and $\mathfrak{P}^{\mathrm{y}}$, then the discretization nodes x and y must be finer than the sets of input strikes, which we can write as

- $x_1 < K_1^{t_1}$, $y_1 < K_1^{t_2}$, $x_m > K_M^{t_1}$ and $y_n > K_N^{t_2}$;
- for all $i \in \{1, \dots, M-1\}$, there exists $k_i \in \{2, \dots, m-1\}$ such that $K_i^{\mathrm{x}} \leq x_{k_i} \leq K_{i+1}^{t_1}$, and for all $j \in \{1, \dots, N-1\}$, there exists $k_j \in \{2, \dots, n-1\}$ such that $K_j^{t_2} \leq y_{k_j} \leq K_{j+1}^{t_2}$.

In particular when there are $M+2$ discretization points for the measure μ_{t_1} and $N+2$ discretization points for the measure μ_{t_2}, Lemma 3.10 provides discretization that preserve convex order.

Proof of Lemma 3.11. To fix notation define $\mathrm{p}^{\mathrm{x}}(A) := \sum_{\{i:x_i\in A\}} p_i^{\mathrm{x}}$ for all $A \subset \mathbb{R}_+$ and $\mathrm{p}^{\mathrm{y}}(B) := \sum_{\{j:y_j\in B\}} p_j^{\mathrm{y}}$ for all $B \subset \mathbb{R}_+$. In particular for some $z \in [0, y_n]$ we have

$$\mathrm{p}^{\mathrm{y}}([0,z]) = \sum_{\{j\leq n:y_j\leq z\}} p_j^{\mathrm{y}}, \quad \text{and} \quad \mathrm{p}^{\mathrm{x}}([0,z]) = \sum_{\{i\leq m:x_i\leq z\}} p_i^{\mathrm{x}}.$$

Also define a function $\delta F : \mathbb{R}_+ \to \mathbb{R}$ by $\delta F(z) := \mathrm{p}^{\mathrm{y}}([0,z]) - \mathrm{p}^{\mathrm{x}}([0,z])$. Then Assumption 13 with $\mu_{t_1} = \mathrm{p}^{\mathrm{x}}$ and $\mu_{t_2} = \mathrm{p}^{\mathrm{y}}$ can be written as

$$\begin{cases} \eta(A) = (\mathrm{p}^{\mathrm{x}}(A) - \mathrm{p}^{\mathrm{y}}(A))_+ > 0, & \text{for any } A \subseteq [a,b], \\ \gamma(A) = (\mathrm{p}^{\mathrm{y}}(A) - \mathrm{p}^{\mathrm{x}}(A))_+ > 0, & \text{for any } A \subseteq [0,y_n] \setminus [a,b]. \end{cases}$$

Define now the functions $f_{\mathrm{x}}, g_{\mathrm{y}} : \mathbb{R}_+ \to \mathbb{R}_+$ by $f_{\mathrm{x}}(K) := \langle \mathrm{p}^{\mathrm{x}}, (\mathrm{x} - K)_+ \rangle$ and $g_{\mathrm{y}}(K) := \langle \mathrm{p}^{\mathrm{y}}, (\mathrm{y} - K)_+ \rangle$. By Baker (2012, Chapter 2, Definition 2.1.6), the balayage order $\mathrm{p} \leq_{\mathrm{cv}} \mathrm{p}^{\mathrm{y}}$ is satisfied as soon as f_{x}, g_{y} are convex and $f_{\mathrm{x}}(K) \leq g_{\mathrm{y}}(K)$ for all $K \in \mathbb{R}_+$. As p^{x} and p^{y} have non-negative components, the functions f_{x} and g_{y} are clearly convex, so we are left to show that $f_{\mathrm{x}}(K) \leq g_{\mathrm{y}}(K)$ for all $K \in \mathbb{R}_+$.

We first show that the conditions $y_1 \leq x_1$ and $y_n \geq x_m$ are necessary. If $x_1 < y_1$ then $f_{\mathrm{x}}(K) = g_{\mathrm{y}}(K) = 1 - K$ for all $K \in [0, x_1]$. Also, for each $K \in (x_1, x_2 \wedge y_1)$, $g_{\mathrm{y}}(K) = 1 - K$ and $f_{\mathrm{x}}(K) = \sum_{i=2}^m p_i^{\mathrm{x}}(x_i - K) = (1 - K) + p_1^{\mathrm{x}}(K - x_1) > (1 - K)$, which violates the balayage order. Similarly if $x_m > y_n$ let $i^* := \sup\{i : x_i < y_n\}$; then for all $K \in (y_n, x_m)$ $g_{\mathrm{y}}(K) = 0$ and $f_{\mathrm{x}}(K) = \sum_{i=i^*}^m p_i^{\mathrm{x}}(x_i - K)$ which yields the conclusion.

Now define a function $G : [0, y_n] \to \mathbb{R}$ as

$$\begin{aligned} G(K) &:= g_{\mathrm{y}}(K) - f_{\mathrm{x}}(K) = \sum_{\{j:y_j>K\}} p^{\mathrm{y}}_j (y_j - K) - \sum_{\{i:x_i>K\}} p^{\mathrm{x}}_i (x_i - K) \\ &= \left(\sum_{\{i:x_i>K\}} p^{\mathrm{x}}_i - \sum_{\{j:y_j>K\}} p^{\mathrm{y}}_j \right) K + \sum_{\{j:y_j>K\}} p^{\mathrm{y}}_j y_j - \sum_{\{i:x_i>K\}} p^{\mathrm{x}}_i x_i. \end{aligned} \tag{3.8}$$

The function G is piecewise linear on $[0, y_n]$, not differentiable at the points $\{x_i\}_{1\le i\le m}$ and $\{y_j\}_{1\le j\le n}$, attains its maximum and minimum on $[0, y_n]$ and $G(0) = G(y_n) = 0$. The lemma then follows if $G(K) \ge 0$ for all $K \in [0, y_n]$. Let $\partial_+ G$ denote its right derivative on $[0, y_n]$, with convention $\partial_+ G(y_n) = 0$. The positivity of G together with the boundary conditions are equivalent to the following:

$$\partial_+ G(K) \ge 0, \text{ for all } K \le K^*, \quad \text{and} \quad \partial_+ G(K) \le 0, \text{ for all } K > K^*,$$

where $K^* := \mathrm{argmax}_K G(K)$. From Equation (3.8) we can write for any $K \in [0, y_n]$

$$\partial_+ G(K) = \sum_{\{i:x_i>K\}} p^{\mathrm{x}}_i - \sum_{\{j:y_j>K\}} p^{\mathrm{y}}_j = \sum_{\{j:y_j\le K\}} p^{\mathrm{y}}_j - \sum_{\{i:x_i\le K\}} p^{\mathrm{x}}_i = \delta F(K),$$

where the function δF is defined above. Since δF is piecewise constant it remains to show that δF has a maximum and a minimum attained on unique subsets of $[0, y_n]$.

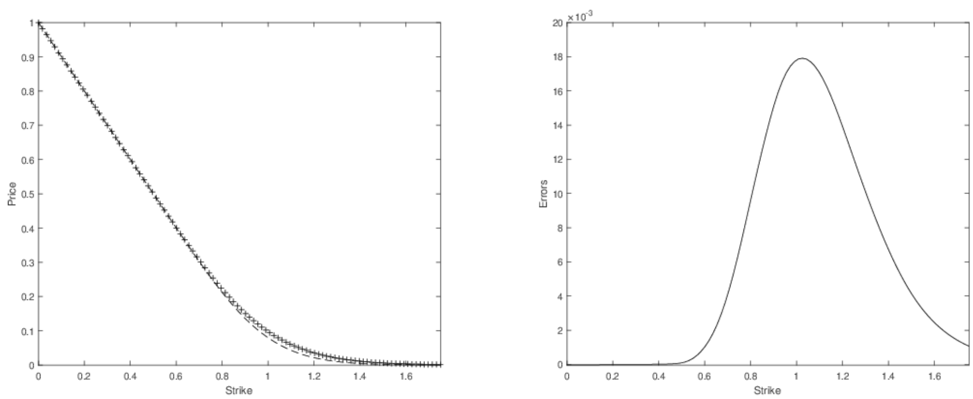
First note that $\delta F(0) = 0$ and $\delta F(y_n) = \mathrm{p}^{\mathrm{y}}([0, y_n]) - \mathrm{p}^{\mathrm{x}}([0, y_n]) = 0$. Moreover δF is bounded on $[0, y_n]$, i.e. $-1 \le \delta F \le 1$. For any $z \in [0, a) \cup (b, y_n]$ one has $\delta F(z) = \mathrm{p}^{\mathrm{y}}([0, z]) - \mathrm{p}^{\mathrm{x}}([0, z]) = \gamma([0, z])$ and therefore δF is an increasing function on $[0, a) \cup (b, y_n]$. On the other hand for any $z \in [a, b]$ one has

$$\begin{aligned} \delta F(z) &= \mathrm{p}^{\mathrm{y}}([0, z]) - \mathrm{p}^{\mathrm{x}}([0, z]) \\ &= \mathrm{p}^{\mathrm{y}}([a, z]) - \mathrm{p}^{\mathrm{x}}([a, z]) + \mathrm{p}^{\mathrm{y}}([0, a)) - \mathrm{p}^{\mathrm{x}}([0, a)) \\ &= -\eta([a, z]) + \gamma([0, a)). \end{aligned}$$

As $\gamma([0, a)) \ge 0$ and η is a measure on $[a, b]$, one has that δF is decreasing on $[a, b]$. Therefore there exist unique c, d, e and f such that $[c, d] \subset [0, y_n]$, $[e, f] \subset [0, y_n]$ and $[c, d] = \text{argmax}_z \delta F(z)$ and $[e, f] = \text{argmin}_z \delta F(z)$. It therefore follows that $\delta F(z) \geq 0$ for all $z \in [0, z^*)$ and $\delta F(z) \leq 0$ for all $z \in [z^*, y_n]$, where $z^* \in (d, e)$. $\square$

Dispersion Assumption 13 intuitively implies that the variance of the underlying price process increases with maturity, which is the case for all martingale models. Lemma 3.11 provides general conditions to ensure that convex order is preserved under discretization of continuous measures. These are however difficult to verify analytically for general distributions, or even for those obtained by Algorithm 3.8. In an example below we provide numerical evidence that the convex order is preserved for these distributions.

Example 3.13. *We consider discretization of the marginal distributions of the price process S generated by the Black-Scholes model. Let $t_1 = 1$ and $t_2 = 1.5$ years and the volatility parameter $\sigma = 0.2$. Figures 3.1a and 3.1b below present the functions g_{y}, f_{x} and G defined in Lemma 3.11 respectively. As a consistency check, the Call prices with maturity t_2 (crosses) are strictly greater than those with maturity t_1 (dashed line), and both functions are convex;*



(a) Functions f_{x} (dashed) and g_{y} (crosses). **(b)** Function $G(K) \geq 0$ for all $K \in [0, y_n]$.

Figure 3.1: Functions g_{y}, f_{x} and G defined in Lemma 3.11.

3.2.3 Discretization of the primal and dual problems

In this section we discuss discretizations of the primal (3.3), (3.5) and the dual (3.4), (3.6) problems. As discussed in the previous section we discretize the state space so that for each maturity $t \in \mathcal{T}$ there are $N(t) > 1$ points $0 < x_1^t < x_2^t < \cdots < x_{N(t)}^t$. Let us define discrete random variables $\widetilde{S}_t$ with $t \in \mathcal{T}$ such that $\mathrm{Supp}(\widetilde{S}_t) := \{x_1^t, x_2^t, \ldots, x_{N(t)}^t\}$. The random variables $\widetilde{S}_t$ approximate the price process S at each maturity $t \in \mathcal{T}$. Moreover each random variable $\widetilde{S}_t$ has distribution p^t obtained using Algorithm 3.8. Let us also define the set

$$\mathcal{X} := \left\{ (x_t)_{t\in\mathcal{T}} : x_t \in \mathrm{Supp}(\widetilde{S}_t) \text{ for each } t \in \mathcal{T} \right\}.$$

Recall that $d = \sum_{t\in\mathcal{T}} \kappa(t)$ is the total number of Call options available for static hedging. We then define the finite-dimensional approximation of the primal problem (3.3)

$$\min \left\{ \langle \mathrm{c}, \mathrm{w} \rangle + \lambda : (\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}} \right\}, \tag{3.9}$$

subject to constraints

$$A(\mathrm{w}, \lambda, \Theta; \mathrm{x}) - \Phi(\mathrm{x}) \geq 0, \text{ for all } \mathrm{x} = (x_t)_{t\in\mathcal{T}} \in \mathcal{X},$$

In order to define the finite-dimensional approximation of the dual problem (3.4), let us first define the set $\mathscr{M}_+^D$ of matrices of dimension $D := N(t_1) \times N(t_2) \times \cdots \times N(t_n)$ with non-negative entries. An element $\zeta := (\zeta_{\mathrm{x}})_{\mathrm{x}\in\mathcal{X}}$ in $\mathscr{M}_+^D$ represents a measure on $\mathcal{X}$. Let $\zeta_{\mathrm{x}} \in \mathbb{R}_+$ be a mass placed at $\mathrm{x} \in \mathcal{X}$ and we also denote by $\zeta_{\mathrm{x}^{i,j}} \in \mathbb{R}_+$ a mass placed at $\mathrm{x}^{i,j} \in \mathcal{X}$, where $i = 1, \ldots, n$ and $\mathrm{x}^{i,j} \in \mathbb{R}^n$ is a vector with the entry t_i fixed to be $x_j^{t_i}$ with $j = 1, \ldots, N(t_i)$. The finite-dimensional approximation is then written as

$$\max \left\{ \sum_{\mathrm{x}\in\mathcal{X}} \zeta_{\mathrm{x}} \Phi(\mathrm{x}) : \zeta \in \mathscr{M}_+^D \right\},$$

subject to constraints

$$\left(\sum_{\mathrm{x}^{i,j}\in\mathcal{X}} \zeta_{\mathrm{x}^{i,j}}\right)_{j=1,\dots,N(t_i)} = \mathrm{p}^{t_i}, \quad \left(\sum_{\mathrm{x}^{i,j}\in\mathcal{X}} (\mathrm{e}_{i-1}\bullet \mathrm{x})_T \zeta_{\mathrm{x}^{i,j}}\right)_{j=1,\dots,N(t_i)} = \mathbf{0} \in \mathbb{R}^{N(t_i)},$$

where e_{i-1} is a unit vector in $\mathbb{R}^n$ such that $(\mathrm{e}_{i-1}\bullet\mathrm{x})_T = (x_{t_i} - x_{t_{i-1}})$ for all $i = 1,\dots,n$. It must be noted that any feasible $\zeta \in \mathscr{M}_+^D$ is joint distribution of random variables $(\widetilde{S}_t)_{t\in\mathcal{T}}$ with marginal distributions equal to the distribution of $\widetilde{S}_t$ for each $t \in \mathcal{T}$ and satisfying the martingale condition for all $\Theta \in \tilde{\mathcal{H}}$.

Similarly the finite-dimensional approximation to the sub-hedging primal problem (3.5) can be stated as follows

$$\max\left\{\langle \mathrm{c}, \mathrm{w}\rangle + \lambda : (\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}\right\}, \tag{3.10}$$

subject to constraints

$$A(\mathrm{w}, \lambda, \Theta; \mathrm{x}) - \Phi(\mathrm{x}) \le 0, \text{ for all } \mathrm{x} = (x_t)_{t\in\mathcal{T}} \in \mathcal{X},$$

and the finite-dimensional approximation of its dual (3.6) is written as

$$\min\left\{\sum_{\mathrm{x}\in\mathcal{X}} \zeta_{\mathrm{x}} \Phi(\mathrm{x}) : \zeta \in \mathscr{M}_+^D\right\},$$

subject to constraints

$$\left(\sum_{\mathrm{x}^{i,j}\in\mathcal{X}} \zeta_{\mathrm{x}^{i,j}}\right)_{j=1,\dots,N(t_i)} = \mathrm{p}^{t_i}, \quad \left(\sum_{\mathrm{x}^{i,j}\in\mathcal{X}} (\mathrm{e}_{i-1}\bullet \mathrm{x})_T \zeta_{\mathrm{x}^{i,j}}\right)_{j=1,\dots,N(t_i)} = \mathbf{0} \in \mathbb{R}^{N(t_i)},$$

for all $i = 1,\dots,n$. The importance of incorporating the martingale conditions into the discretization is critical. This is easily seen in the following example for the dual problem, which also yields an issue for the primal problem.

Example 3.14. *Suppose that we are given only two maturities $t_1 < t_2$. Assume $\widetilde{S}_{t_1}$ can take value 0.75 or 1.25 each with 50% probability, whereas $\widetilde{S}_{t_2}$ can take value 0.5 or 1.5 each with 50% probability. Note that $\mathbb{E}\{\widetilde{S}_{t_1}\} = \mathbb{E}\{\widetilde{S}_{t_2}\} = 1$. The constraints of the dual problem $(\sum_{\mathrm{x}^{1,j}\in\mathcal{X}} \zeta_{\mathrm{x}^{1,j}})_{j=1,2} =$*

(0.5, 0.5) and $(\sum_{x^{2,j}\in\mathcal{X}}(e_1 \bullet x)_T \zeta_{x^{2,j}})_{j=1,2} = \mathbf{0}$ *fully determine the probabilities* $\zeta_{x^{1,1}} \in \{1/8, 3/8\}$ *and* $\zeta_{x^{1,2}} \in \{1/8, 3/8\}$.

The constraints $(\sum_{x^{2,j}\in\mathcal{X}} \zeta_{x^{2,j}})_{j=1,2} = p^{t_2}$ *are only true if* $p_1^{t_2} = p_2^{t_2} = 1/2$ *or* $\mathbb{E}\{\widetilde{S}_{t_2}\} = 1$. *Otherwise, there will be no solution to this LP. This stresses the importance of a consistent no-arbitrage discretization of the problem.*

3.3 Application to no-arbitrage bounds on forward volatility given marginals

We now illustrate the numerical methods developed in Sections 3.2.1, 3.2.2 and 3.2.3 on the Forward-Start options. We start by restricting the set of maturities $\mathcal{T}$ to only two maturities t and $t+\tau$, where $\tau > 0$; the state space Ω is then reduced to $\mathbb{R}_+^2$. The pay-off of a Forward-Start Straddle is $|S_{t+\tau} - \mathcal{K}S_t|$ for $\mathcal{K} > 0$. Note that $|S_{t+\tau} - \mathcal{K}S_t| = 2(S_{t+\tau} - \mathcal{K}S_t)_+ - S_{t+\tau} + \mathcal{K}S_t$, i.e. the Forward-Start Straddle can be perfectly hedged by a long position in 2 Forward-Start Call option with strike $\mathcal{K}$, a long position in $\mathcal{K}$ forward contracts maturing at time t and a short position in the one forward maturing at time $t+\tau$. Thus the problem of finding arbitrage-free bounds on the price of a Forward-Start Straddle is equivalent to finding arbitrage-free bounds on a price of a Forward-Start Call option with the same strike. Forward-Start option prices can be equated with forward implied volatility, which can be computed by a straightforward modification of the Black-Scholes model, an approach we now quickly recall. In the Black-Scholes model, the dynamics of the stock price process under the risk-neutral measure are given by $\mathrm{d}S_t = S_t\Sigma\mathrm{d}W_t$, $S_0 = 1$, where $\Sigma > 0$ represents the instantaneous volatility and W is a standard Brownian motion. Since the increments of the stock price process are stationary and independent, the Forward-Start option with pay-off $(S_{t+\tau} - KS_t)^+$ with $t, \tau > 0$ is worth $c_{BS}(\log(K), \Sigma\sqrt{\tau})$, where c_{BS} is defined in (2.1). For a given market or model price $C^{\mathrm{obs}}(t, \tau, K)$ of the option at strike K, Forward-Start date t and maturity τ, the forward implied volatility smile $\sigma_{t,\tau}(K)$ is then defined as the unique solution to $C^{\mathrm{obs}}(t, \tau, K) = c_{BS}(\log(K), \sigma_{t,\tau}(K)\sqrt{\tau})$.

We specialise our computations to the case when traded Call option prices are consistent with the Black-Scholes and the Heston models. In this case the marginal distributions of the price process S are known explicitly at times t and $t+\tau$ and we shall denote the marginal distributions as μ and ν respectively. Note that μ and ν are placed in convex order as a consequence of absence of arbitrage in the Black-Scholes and the Heston models. The set of martingale measures on $\mathbb{R}_+^2$ with marginal distributions equal to μ and ν is denoted by $\mathbb{M}(\mu,\nu)$ and is non-empty by above considerations. Then finding no-arbitrage bounds for the price of a Forward-Start Straddle is equivalent to solving infinite-dimensional LPs

$$\inf_{\zeta\in\mathbb{M}(\mu,\nu)} \int_{\mathbb{R}_+^2} |x-\mathcal{K}y|\zeta(\mathrm{d}x,\mathrm{d}y), \qquad \text{and} \qquad \sup_{\zeta\in\mathbb{M}(\mu,\nu)} \int_{\mathbb{R}_+^2} |x-\mathcal{K}y|\zeta(\mathrm{d}x,\mathrm{d}y). \tag{3.11}$$

3.3.1 Comparison with existing results

These problems have been studied by Hobson and Neuberger (2012) and Hobson and Klimmek (2014) in case of at-the-money Forward-Start Straddle ($\mathcal{K}=1$). The authors studied maps that optimally transported mass between marginal distributions μ and ν. Hobson and Klimmek (2014) found that in the lower bound case the solution is the optimal martingale measure $\zeta\in\mathbb{M}(\mu,\nu)$ that is concentrated on a trinomial map $\{p(x),x,q(x)\}$, where $p:[a,b]\to[0,a]$ and $q:[a,b]\to[b,\infty)$ (where a and b are such that μ and ν satisfy Assumption 13) are two decreasing functions that satisfy a system of coupled differential equations

$$\begin{aligned} p'(x) &= \frac{q(x)-x}{q(x)-p(x)}\frac{f_\mu(x)-f_\nu(x)}{f_\mu(p(x))-f_\nu(p(x))}, \\ q'(x) &= \frac{x-p(x)}{q(x)-p(x)}\frac{f_\mu(x)-f_\nu(x)}{f_\mu(q(x))-f_\nu(q(x))}, \end{aligned} \tag{3.12}$$

with boundary conditions

$$\begin{aligned}
p(b) &= \inf\left\{x \geq 0 : \gamma([0,x]) > 0\right\}, \\
q(b) &= \inf\left\{x \geq 0 : \gamma([0,x]) > \gamma([0,b])\right\}, \\
p(a) &= \sup\left\{x \geq 0 : \eta([0,x]) < \eta([0,a])\right\}, \\
q(a) &= \sup\left\{x \geq 0 : \gamma([0,x]) < 1\right\}.
\end{aligned}$$

By taking limits, we obtain that $p(a) = a$, $p(b) = 0$ $q(a) = +\infty$ and $q(b) = b$, see also Campi et al. (2017, Proposition 5.6).

Similarly Hobson and Neuberger (2012, Remark 5.4) showed that the solution for the upper bound is the optimal martingale measure $\zeta \in \mathbb{M}(\mu,\nu)$ and it is concentrated on a binomial map $\{g(x), f(x)\}$, where $f, g : \mathbb{R}_+ \to \mathbb{R}_+$ are two increasing, continuous and differentiable functions with $f(x) < x < g(x)$ for all $x \in \mathbb{R}_+$ that satisfy a system of integral equations

$$f^{-1}(y) = \int_{g^{-1}(y)}^{f^{-1}(y)} \frac{g(x)}{g(x) - f(x)} \mathrm{d}x, \qquad 1 = \int_{g^{-1}(y)}^{f^{-1}(y)} \frac{1}{g(x) - f(x)} \mathrm{d}x.$$

In both case one can compute the optimal values of the programmes (3.11).

On the other hand the problem of finding no-arbitrage bounds on the price of a Forward-Start Straddle with strike $\mathcal{K}$ can be approached using martingale optimal transport theory. This approach was studied in Beiglböck et al. (2013) by defining dual problems

$$\begin{aligned}
\int_{\mathbb{R}^2_+} |y - \mathcal{K}x| \zeta(\mathrm{d}x, \mathrm{d}y) &\geq \sup_{(\psi_0,\psi_1,\delta) \in \underline{Q}} \left\{ \int_{\mathbb{R}_+} \psi_0(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi_1(y)\nu(\mathrm{d}y) \right\}, \\
\int_{\mathbb{R}^2_+} |y - \mathcal{K}x| \zeta(\mathrm{d}x, \mathrm{d}y) &\leq \inf_{(\psi_0,\psi_1,\delta) \in \overline{Q}} \left\{ \int_{\mathbb{R}_+} \psi_0(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi_1(y)\nu(\mathrm{d}y) \right\},
\end{aligned} \tag{3.13}$$

where the sets of sub- and super-replicating portfolios are defined as follows

$$\begin{aligned}\underline{Q} :=&\{(\psi_0,\psi_1,\delta)\in L^1(\mu)\times L^1(\nu)\times B_b(\mathbb{R}_+):\\ &\psi_1(y)+\psi_0(x)+\delta(x)(y-x)\le|y-\mathcal{K}x|,\text{for all } x,y\in\mathbb{R}_+\},\end{aligned}$$

$$\begin{aligned}\overline{Q} :=&\{(\psi_0,\psi_1,\delta)\in L^1(\mu)\times L^1(\nu)\times B_b(\mathbb{R}_+):\\ &\psi_1(y)+\psi_0(x)+\delta(x)(y-x)\ge|y-\mathcal{K}x|,\text{for all } x,y\in\mathbb{R}_+\}.\end{aligned}$$

In fact the authors (Beiglböck et al., 2013, Section 1.5) restrict the set of the dual variables ψ_0 and ψ_1 to the set of linear combinations of finitely many European Call option pay-offs while restricting the set of trading strategies δ to $C_b(\mathbb{R}_+)$. It is then proved (Beiglböck et al., 2013, Theorem 1 and Corollary 1.1) (actually for a more general class of pay-off functions) that there is no duality gap between the primal and the dual problems hold. We briefly note that restricting the dual problems (3.13) in such manner is equivalent to the set-up of Section 2.2 and in this case the dual problems (2.27) and (2.29) coincide with the primal problems (3.11).

The optimal values may not be attained in the dual problems (3.13), as shown in Beiglböck et al. (2013, Proposition 4.1). Hobson and Neuberger (2012) and Hobson and Klimmek (2014) proved that if the marginal distributions μ and ν satisfy Assumption 13 then in addition to absence of duality gap between the primal (3.11) and the dual (3.13) problems the infimum and supremum in the dual problems (3.13) are attained in the at-the-money case, i.e. $\mathcal{K}=1$.

3.4 Numerical analysis of the no-arbitrage bounds

In the following subsections, we shall consider the observed vanilla Call option prices as computed from the Black-Scholes and the Heston model. We shall discretize the corresponding primal and the dual problems following Section 3.2.3, by considering the compact interval $[0,5]$ for the supports of the discretized random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$ with $m=n=500$ discretization points with $t=1$ and $\tau=0.5$ (the set of maturities is then

$\mathcal{T} := \{t, t+\tau\}$). The vectors of observed normalised strikes are taken to be $\mathfrak{K}_t = \mathfrak{K}_{t+\tau} = \{0.3, 0.4, 0.5, \ldots, 2\}$. We shall also use the monomial basis with $(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)(x) = (1, x, x^2, x^3, x^4)$ to discretize the hedge functions as in Equation (3.1).

3.4.1 Implementation of the Hobson-Klimmek transport plan

We now discuss implementation of the solution for the lower bound at-the-money Forward-Start Straddle discussed in Hobson and Klimmek (2014). The right-hand side of the equations in (3.12) are undefined at the boundary points. An application of L'Hôpital's rule shows that $\lim_{x\uparrow b} q'(x) = -1$ if $f'_\mu(b) \neq f'_\nu(b)$. On the other hand $\lim_{x\uparrow b} p'(x)$ depends on the marginal measures μ and ν. For instance, in the lognormal example in Section 3.4.2, we find that $p'(x) = \mathcal{O}\left(\mathrm{e}^{\alpha(\log p(x))^2}\right)$ for some $\alpha > 0$ as x tends to b from below, and $\lim_{x\uparrow b} p'(x) = -\infty$ (see for example Figure 3.2(b)). On the other hand if for example $f_\mu(0) \neq f_\nu(0)$ or $f'_\mu(0) \neq f'_\nu(0)$ then $\lim_{x\uparrow b} p'(x) = 0$.

In order to circumvent these issues so that we can apply the Runge-Kutta method to solve these ODEs, we introduce the following pre-processing step: fix a small $\varepsilon > 0$ (in our implementations, we choose $\varepsilon = 0.001$), integrate both sides of (3.12) over $[b-\varepsilon, b]$ and then approximate the right-hand side by using the rectangle rule for the integral with the unknown values $p^* := p(b-\varepsilon)$ and $q^* := q(b-\varepsilon)$. This yields the following simultaneous equations for p^* and q^* which we solve numerically:

$$\begin{aligned} p^* &= -\varepsilon \frac{q^* - b + \varepsilon}{q^* - p^*} \frac{f_\mu(b-\varepsilon) - f_\nu(b-\varepsilon)}{f_\mu(p^*) - f_\nu(p^*)}, \\ q^* &= b - \varepsilon \frac{b - \varepsilon - p^*}{q^* - p^*} \frac{f_\mu(b-\varepsilon) - f_\nu(b-\varepsilon)}{f_\mu(q^*) - f_\nu(q^*)}. \end{aligned}$$

These equations can easily be reduced into one root search; the first equation gives

$$q^* = \frac{(p^*)^2 \left[f_\mu(p^*) - f_\nu(p^*)\right] + \varepsilon(b-\varepsilon)\left[f_\mu(b-\varepsilon) - f_\nu(b-\varepsilon)\right]}{p^* \left[f_\mu(p^*) - f_\nu(p^*)\right] + \varepsilon\left[f_\mu(b-\varepsilon) - f_\nu(b-\varepsilon)\right]},$$

which in turn can be plugged into the second equation to solve for p^*. The pair (p, q) in (3.12) is then solved for $x \in [a, b-\varepsilon]$ using standard Runge-Kutta methods with the new boundary conditions $p(b-\varepsilon) = p^*, q(b-\varepsilon) = q^*$. The lower bound for the at-the-money Forward-Start Straddle price is then given using the optimal conditional density $\rho^*(Y = y|X = x)$ given in Hobson and Klimmek (2014, page 199):

$$\rho^*(Y = y|X = x) = \begin{cases} \dfrac{f_\eta(x)}{f_\mu(x)} \dfrac{q(x) - x}{q(x) - p(x)} \mathbb{1}_{\{y=p(x)\}}, & \text{if } y < x, \\ \left(1 - \dfrac{f_\eta(x)}{f_\mu(x)}\right), & \text{if } y = x, \\ \dfrac{f_\eta(x)}{f_\mu(x)} \dfrac{x - p(x)}{q(x) - p(x)} \mathbb{1}_{\{y=q(x)\}}, & \text{if } y > x, \end{cases}$$

and hence straightforward computations yield

$$\begin{aligned} \mathbb{E}(|Y - X|) &= \int_{\mathbb{R}} \mathbb{E}(|Y - X||X = x) f_\mu(x) \mathrm{d}x \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \rho^*(Y = y|X = x) f_\mu(x) |y - x| \mathrm{d}y \mathrm{d}x \\ &= \int_{\mathbb{R}} \int_{-\infty}^{x} f_\eta(x) \frac{q(x) - x}{q(x) - p(x)} \mathbb{1}_{\{y=p(x)\}} |y - x| \mathrm{d}y \mathrm{d}x \\ &+ \int_{\mathbb{R}} \int_{x}^{+\infty} f_\eta(x) \frac{x - p(x)}{q(x) - p(x)} \mathbb{1}_{\{y=q(x)\}} |y - x| \mathrm{d}y \mathrm{d}x \\ &= \int_a^b \frac{2(x - p(x))(q(x) - x)}{q(x) - p(x)} f_\eta(x) \mathrm{d}x, \end{aligned}$$

where we recall from Assumption 13 that $\eta = (\mu - \nu)_+$.

3.4.2 Application to the Black-Scholes model

Let $\mathcal{N}(\bar{m}, \Sigma^2)$ denote the Gaussian distribution with mean $\bar{m}$ and variance Σ^2, and assume that the random variables S_t and $S_{t+\tau}$ are distributed according to

$$\log(S_t) \sim \mathcal{N}\left(-\frac{1}{2}\Sigma^2 t, \Sigma^2 t\right) \text{ and } \log(S_{t+\tau}) \sim \mathcal{N}\left(-\frac{1}{2}\Sigma^2 (t+\tau), \Sigma^2 (t+\tau)\right).$$

In this case the forward volatility, i.e. the implied volatility computed from the Forward-Start option, is constant and also equal to Σ.

In Figure 3.2 we plot the lower and upper bounds for the forward implied volatility smile computed from the discretization of semi-infinite Problems (3.3) and (3.5) via (3.9) and (3.10) respectively. Figure 3.3a shows the pay-off of the static position with maturity 1 year and Figure 3.3b shows the pay-off of the static position with maturity 1.5 years. Note that the static position with the short maturity is short convexity whereas the static position with longer maturity is long convexity. Figure 3.4 shows the discretized delta hedge. The lower bound at-the-money case using Hobson-Klimmek solution and the LP solution are virtually identical (6.95% v.s. 6.98% respectively), illustrating consistency of the two approaches. Note that even in this simple case the range of possible forward smiles consistent with the two marginal laws is wide. The magnitude of the bounds produced is similar to bounds obtained in Henry-Labordère (2013) for cliquet options (Type-I Forward-Start option) which are very closely related to Forward-Start Straddles. This behaviour can be explained intuitively as no conditional instruments are used to hedge the Straddle, hence there is no restriction on the conditional probabilities between times t and $t+\tau$. The only restriction is that the two marginal laws at those times are placed in convex order, producing a very large class of feasible martingale measures.

3.4.3 Application to the Heston model

The marginal distributions for expiries $t = 1$ and $t + \tau = 1.5$ are now generated according to the Heston stochastic volatility model Heston (1993), in which the stock price process is the unique strong solution to the stochastic differential equation

$$\begin{aligned} \mathrm{d}S_t &= S_t\sqrt{V_t}\mathrm{d}W_t, & S_0 = 1, \\ \mathrm{d}V_t &= \kappa\left(\theta - V_t\right)\mathrm{d}t + \xi\sqrt{V_t}\mathrm{d}Z_t, & V_0 = v > 0, \end{aligned}$$

where W and Z are two one-dimensional standard Brownian motions with $\mathrm{d}\langle W, Z\rangle_t = \rho \mathrm{d}t$, and $\kappa, \theta, \xi > 0$ and $\rho \in [-1, 1]$. We consider here the follow-

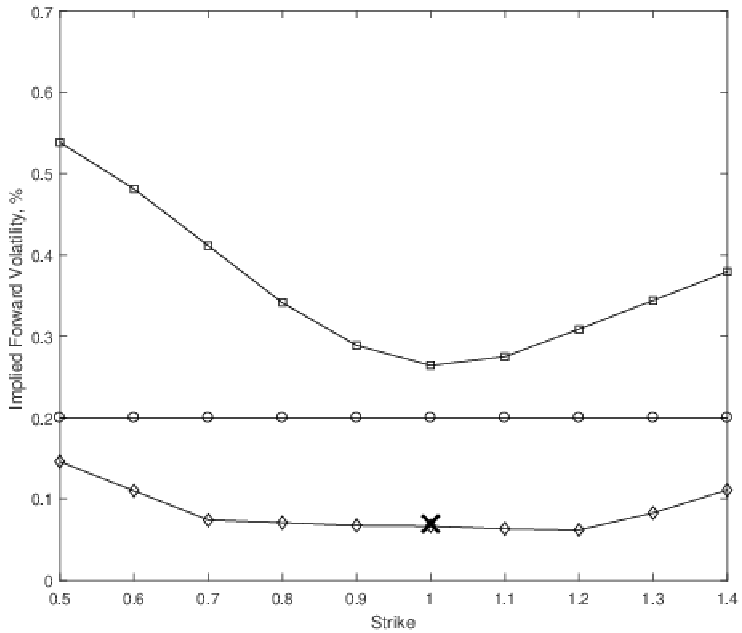
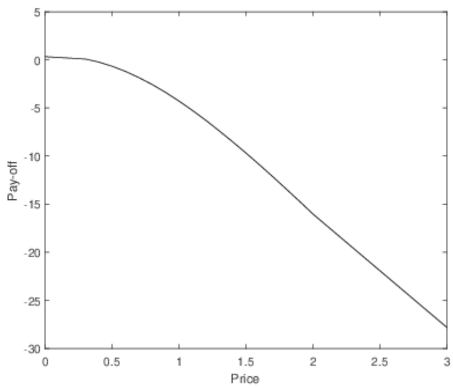
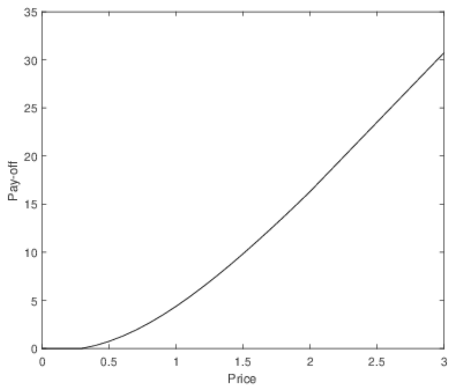


Figure 3.2: Optimal bounds in the Black-Scholes case. Upper bound (squares), lower bound (diamonds), the Black-Scholes forward volatility (circles), Hobson-Klimmek solution (cross).



(a) 1-year maturity.



(b) 1.5-year maturity.

Figure 3.3: Static position pay-offs in the Black-Scholes case.

ing values: $v = \theta = 0.07$, $\kappa = 1$, $\xi = 0.4$ and $\rho = -0.8$. Figure 3.5 shows

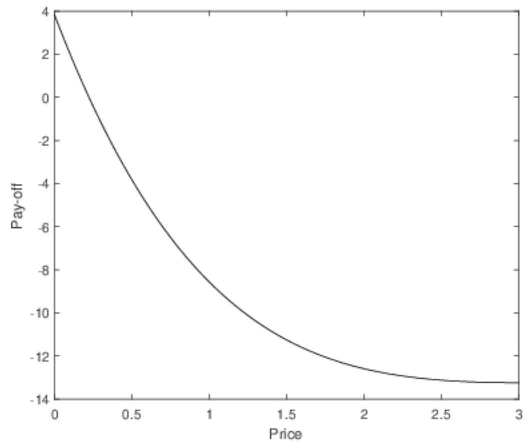


Figure 3.4: Optimal discretized delta hedge in the Black-Scholes case.

the Heston forward smile consistent with the marginals (computed using the inverse Fourier transform representation and a simple root search method) and the lower and upper bounds for the forward smile, from the discretized versions (3.9) and (3.10) of semi-infinite Problems (3.3) and (3.5) respectively. Figure 3.6 shows the pay-off of the static positions in the superhedge: one enters into positions with short convexity for the 1-year maturity (Figure 3.6a) and long convexity for the 1.5-year maturity (Figure 3.6b). Note that pay-offs of both static positions are remarkably similar to the pay-offs of the static positions in the Black-Scholes case, Figure 3.3, thus confirming the model-independent nature of the static superhedge. Figure 3.7 shows the discretized delta hedge. Note that oscillations are due to the fact that the basis of monomials was used to perform discretization and other possible choice might help address the problem. When compared with the delta hedge in the Black-Scholes model, Figure 3.4 it can be seen that both hedges are decreasing (bar some oscillations) in the price of the underlying. However the delta hedge in the Heston model seems to decrease linearly unlike the hedge in the Black-Scholes model. As in the Black-Scholes case, the Hobson-Klimmek solution (7.70%) and the LP solution (7.80%) are virtually the same for the lower bound at-the-money case.

In the examples explored in this section the range of forward smiles consis-

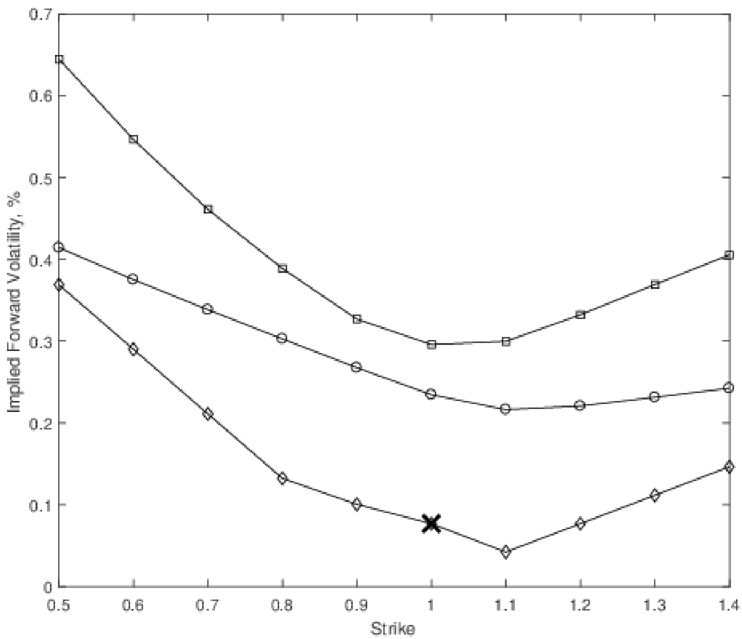
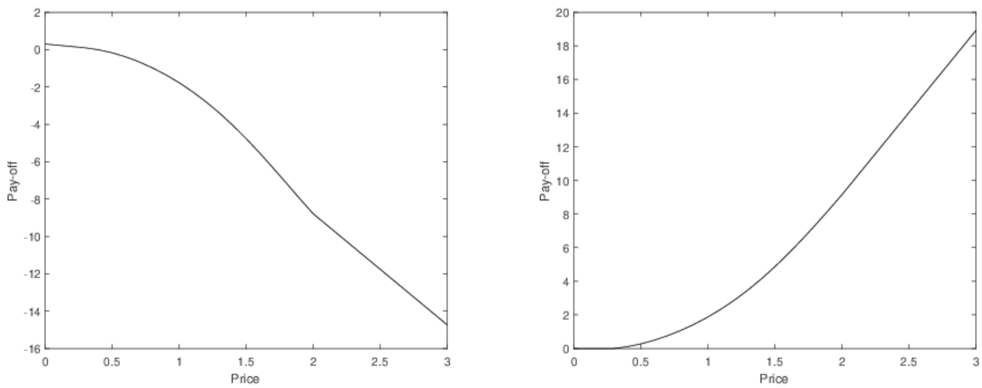


Figure 3.5: Optimal bounds in the Heston case. Upper bound (squares), lower bound (diamonds), the Heston forward volatility (circles), Hobson-Klimmek solution (cross).



(a) 1-year maturity. **(b)** 1.5-year maturity.

Figure 3.6: Static position pay-offs in the Heston case.

tent with the marginal laws is large, and of the same magnitude as in Henry-Labordère (2013, Section 5.5). This wide range of prices supports the claim

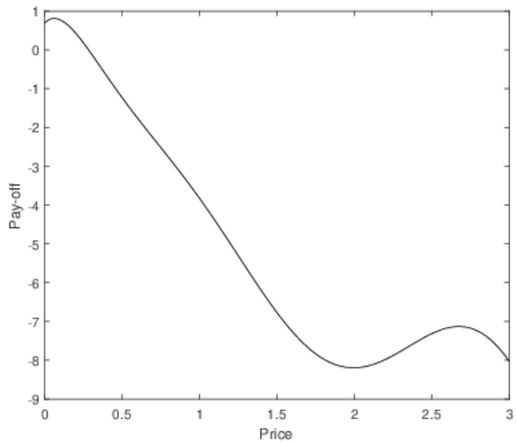


Figure 3.7: Optimal discretized delta hedge in the Heston case.

that using European vanilla options to replicate forward volatility-dependent claims seems illusory. Forward-Start options should therefore be seen as fundamental building blocks for exotic option pricing and not decomposable (or approximately decomposable) into European options. Forward-Start options are usually available from the stripping of cliquet options, on OTC markets. Cliquet options have been a popular instrument on Equity markets (Bühler, 2002), and also as a hedging tool for insurance companies (the so-called FLEX Index options, the details of which can be found on the CBOE website). Therefore models used for forward volatility-dependent exotics should be able to calibrate Forward-Start option prices and should produce realistic forward smiles consistent with trader expectations and observable prices.

3.4.4 On discretization of the state space

In both examples above, the discretion random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$ were supported on 500 points. This choice was arbitrary, and it is natural to question it. In the semi-infinite case, Shapiro (2009, Theorem 3.1) shows that the absence of duality gap between the primal problem and its dual guarantees the existence of a discretization, the value of which converges to that of the primal problem as the number of points increases. Rates of convergence have also been obtained for discretization schemes in semi-infinite programming, see Shapiro (2009); Still (2001). Our setting here is however of infinite-dimensional nature, and, to the best of our knowledge, no corresponding result exists. A general, theoretical, proof is outside the scope of our approach, and we now provide some numerical evidence about the convergence and stability of our discretization scheme. We choose two different discretization grids for the supports of the random variables $\widetilde{S}_t$ and $\widetilde{S}_{t+\tau}$, following Algorithm 3.8: a uniform grid and a grid consisting of roots of Legendre polynomials, both with the same number of points. Below, Tables 3.1 and 3.2 represent the optimal values of the sub-hedging dual problem as the number of discretization points increases, for different values of the Forward-Start Straddle strike $\mathcal{K}$. Clearly, refining the discretization grid produces only minor changes in the optimal values for the at-the-money case ($\mathcal{K} = 1$) and virtually none for the other cases. Tables 3.3 and 3.4 represent the optimal values of the super-hedging dual problem for different numbers of discretization points and different values of the Forward-Start Straddle strike $\mathcal{K}$. Refining the discretization grid yields only minor changes to the optimal values. In contrast to the sub-hedging case, the change in optimal values for the super-hedging problem can be observed for all strikes. Finally, both sub- and super-hedging dual problems seem to be very stable with respect to the discretization schemes and the number of points.

We would also like to comment on the rate of convergence of the optimal portfolios with respect to the partition refinements. Hobson and Neuberger (2012, Section 6.2) obtained explicit expressions for the dual variables (ψ_0, ψ_1, δ), defined in (3.13), in the lognormal case for the at-the-money Forward-Start Straddle ($\mathcal{K} = 1$):

$$\begin{aligned}
\psi_0(x) &= -\xi x \ln x + \xi x \ln\left(A \sinh(\xi^{-1})\right) + x \coth(\xi^{-1}), \\
\psi_1(y) &= \xi \left(y \ln y - y \ln(A/\xi) - y\right), \\
\delta(x) &= -\xi \ln\left(x/A \sinh(\xi^{-1})\right),
\end{aligned} \tag{3.14}$$

where the constants A and ξ are such that the expected value of the hedging portfolio $\psi_0(x) + \psi_1(y) + \delta(x)(y - x)$ is minimised under super-hedging constraints. They also showed that the bound achieved by the solution is tight (Hobson and Neuberger, 2012, Section 10). We compute the sup-norm error $\varepsilon(n)$ (n is the number of discretization points) between solutions of the discretized Primal (3.9) and the Hobson-Neuberger solution (3.14). We visually check the rate of convergence (i.e. the highest exponent r such that $\varepsilon(n) \sim \mathcal{O}(d_n^r)$) with respect to the discretization mesh size d_n by plotting $\log(\varepsilon(n))/\log(d_n)$ against $\log(d_n)$. The resulting plot is presented in Figure 3.8 below.

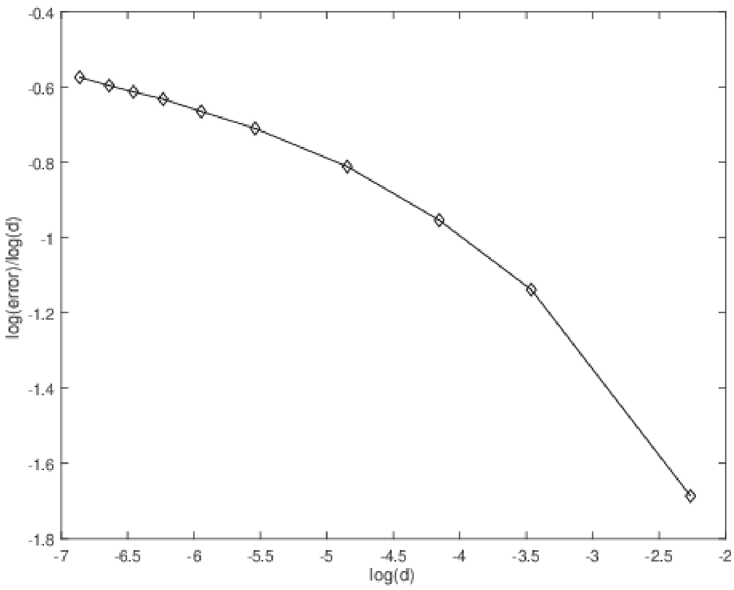


Figure 3.8: Plot of $\log(\varepsilon(n))/\log(d_n)$ against $\log(d_n)$, where $\varepsilon(n)$ is the sup-norm error.

Nb. of points / Strike	75	250	500	1000	2000
0.6	0.4	0.4	0.4	0.4	0.4
0.7	0.3	0.3	0.3	0.3	0.3
0.8	0.2	0.2	0.2	0.2	0.2
0.9	0.1001	0.1	0.1	0.1	0.1
1.0	0.0390	0.0384	0.0384	0.0384	0.0383
1.1	0.1004	0.1004	0.1004	0.1004	0.1004
1.2	0.2	0.2	0.2	0.2	0.2
1.3	0.3	0.3	0.3	0.3	0.3
1.4	0.4	0.4	0.4	0.4	0.4

Table 3.1: Optimal values of the sub-hedging primal problem (using roots of Legendre polynomials).

Nb. of points / Strike	75	250	500	1000	2000
0.6	0.4	0.4	0.4	0.4	0.4
0.7	0.3	0.3	0.3	0.3	0.3
0.8	0.2	0.2	0.2	0.2	0.2
0.9	0.10008	0.1	0.1	0.1	0.1
1.0	0.0388	0.0384	0.0384	0.0384	0.0383
1.1	0.1006	0.1004	0.1004	0.1004	0.1004
1.2	0.2	0.2	0.2	0.2	0.2
1.3	0.3	0.3	0.3	0.3	0.3
1.4	0.4	0.4	0.4	0.4	0.4

Table 3.2: Optimal values of the sub-hedging primal problem (using a uniform grid).

Strike \ Nb. of points	75	250	500	1000	2000
0.6	0.4147	0.4156	0.4157	0.4157	0.4157
0.7	0.3241	0.3255	0.3257	0.3257	0.3257
0.8	0.2394	0.2411	0.2413	0.2413	0.2414
0.9	0.1707	0.1743	0.1745	0.1746	0.1746
1.0	0.1453	0.1485	0.1489	0.1490	0.1490
1.1	0.1786	0.1815	0.1816	0.1817	0.1817
1.2	0.2513	0.2536	0.2538	0.2539	0.2539
1.3	0.3373	0.3394	0.3396	0.3397	0.3397
1.4	0.4292	0.4314	0.4316	0.4316	0.4316

Table 3.3: Optimal values of the super-hedging primal problem (using roots of Legendre polynomials).

Strike \ Nb. of points	75	250	500	1000	2000
0.6	0.4150	0.4157	0.4157	0.4157	0.4157
0.7	0.3245	0.3256	0.3257	0.3257	0.3257
0.8	0.2397	0.2411	0.2413	0.2413	0.2414
0.9	0.1716	0.1743	0.1746	0.1746	0.1746
1.0	0.1463	0.1487	0.1489	0.1490	0.1490
1.1	0.1795	0.1814	0.1817	0.1817	0.1817
1.2	0.2514	0.2538	0.2539	0.2539	0.2539
1.3	0.3376	0.3395	0.3396	0.3397	0.3397
1.4	0.4300	0.4315	0.4316	0.4316	0.4316

Table 3.4: Optimal values of the super-hedging primal problem (using a uniform grid).

4

Perturbation Analysis of Model-Independent Hedging Problems

Extrapolation of the total implied variance discussed in Section 2.2.4 restricts the feasible sets of the dual super- and sub-hedging problems (2.27) and (2.29) as well as the feasible set of the semi-infinite approximations of the dual problems defined in (3.4) and (3.6). On the other hand the feasible sets of primal problems (3.3) and (3.5) are enlarged by adding non-traded Call options with prices consistent with extrapolation. As this assumption is exogenous, in this chapter we discuss sensitivity of optimal values of the dual problems to extrapolation of the total implied variance.

The rest of the chapter is organised as follows. Section 4.1 reviews different notions of directional differentiability and how they apply to convex functions. We then embed the semi-infinite approximations to the primal problems (3.3) and (3.5) into a family of perturbed problems in Section 4.2. The dual problems are equivalently embedded in a family of perturbed problems, where the perturbations are changes in input Call option prices. We apply results of Section 4.1 to derive sensitivity of the dual problems that turn out to be the inner product of the perturbation vector and optimal portfolios weights obtained by solving unperturbed primal problems. When perturbations are itself parametrized we derive sensitivity of the values of the dual problems to the parameters via application of the chain rule. Section 4.3 presents the sensitivity analysis of arbitrage-free bounds of at-the-money Forward-Start Straddle building on results from Chapter 3.

4.1 Directional derivatives of convex functions

Let X be a normed topological vector space and let $g : X \to \overline{\mathbb{R}}$ be an extended real-valued function. It is convex if its epigraph

$$\text{epi } g := \{(x, \mu) \in X \times \mathbb{R} : \mu \geq g(x)\}$$

is a convex subset of the vector space $X \times \mathbb{R}^n$ (Rockafellar, 1970, Section 4). If the effective domain dom g is non-empty and the function g never takes value $-\infty$ then it is a proper convex function. Proper convex functions enjoy certain continuity properties.

Theorem 4.1. *(Zălinescu, 2002, Theorem 2.2.9) If a convex function $g : X \to \mathbb{R}$ is bounded above on a neighbourhood of a point in* dom g *then g is continuous on the interior of its domain. If g is not proper then it is identically $-\infty$ on* int (dom g)*.*

It is clear that if the convex function g is continuous on the interior of its domain then there exists a point $x_0 \in \text{dom } g$ such that it is bounded above on a neighbourhood of x_0. If in addition X is a Banach space endowed with the norm topology then continuity of g on the interior of its domain is equivalent to local Lipschitz continuity.

Proposition 4.2. *(Bonnans and Shapiro, 2000, Proposition 2.107 (v)) If g is the proper convex function and* int dom $g \neq \emptyset$ *then g is locally Lipschitz continuous on* int dom g*.*

If X is a finite-dimensional vector space, then any proper convex function is continuous on the relative interior of its effective domain.

Theorem 4.3. *(Aliprantis and Border, 2007, Theorem 7.24) If X is a finite-dimensional vector space then any proper convex function $g : X \to \mathbb{R}^n \cup \{+\infty\}$ is continuous on* ri dom g.

If a convex set $C \in \mathbb{R}^n$ is of dimension n then aff $C = \mathbb{R}^n$ and ri $C =$ int C. Therefore if dom g has the same dimension as the underlying finite-dimensional space X then Theorem 4.3 implies that g is continuous on int dom g.

Let X^* be topological dual of the space X. A point $x^* \in X^*$ is a subgradient of a convex function g at x_0 if the following holds

$$g(x) \geq g(x_0) + \langle x - x_0, x^* \rangle ,$$

for every $x \in X$. The subdifferential of g at x_0 is the set of all subgradients at x_0

$$\partial g(x_0) := \{x^* \in X^* : g(x) \geq g(x_0) + \langle x - x_0, x^* \rangle \text{ for all } x \in X\} ,$$

and if the subdifferential at x_0 is non-empty then the convex function g is said to be subdifferentiable at x_0 (Aliprantis and Border, 2007, Section 7.4).

The convex conjugate of the function g is the convex function $g^* : X^* \to \mathbb{R}^n$ defined to be

$$g^*(x^*) := \sup_{x \in X} \{\langle x, x^* \rangle - g(x)\} .$$

In turn, the convex conjugate of g^* is defined as

$$g^{**}(x) := \sup_{x^* \in X^*} \{\langle x, x^* \rangle - g^*(x^*)\} .$$

We summarise some properties of the convex conjugate of a convex function in the following proposition.

Proposition 4.4. *(Bonnans and Shapiro, 2000, Proposition 2.118) Let $g : X \to \mathbb{R}^n$ be a convex function. Then the following hold* *(i) If $g^{**}(x)$ is finite for some $x \in X$ then*

$$\partial g^{**}(x) = \operatorname{argmax} \{\langle x, x^* \rangle - g^*(x^*)\}.$$

(ii) If g is subdifferentiable at x, then $g^{**}(x) = g(x)$.

(iii) If $g^{**}(x) = g(x)$ and is finite then $\partial g^{**}(x) = \partial g(x)$.

Fenchel-Moreau Theorem states that the function g is convex and lower semi-continuous if and only if $g = g^{**}$, where we define the function g to be lower semi-continuous if $g = \operatorname{lsc} g$.

Definition 4.5. (Bonnans and Shapiro, 2000, Definition 2.44) A function g (not necessarily convex) is directionally differentiable at $x \in X$ in a direction h if the following limit exists in $\mathbb{R}$

$$g'(x, h) := \lim_{t \downarrow 0} \frac{g(x + th) - g(x)}{t}.$$

If g is directionally differentiable at $x \in X$ in any direction $h \in X$ we say it is directionally differentiable at x.

It is easily seen that for any $\alpha \geq 0$ one has $g'(x, \alpha h) = \alpha g'(x, h)$, i.e. g' is positively homogeneous in h. If g is a proper convex function then the directional derivative is also convex in x and h. We recall the definition of Gâteaux differentiability below.

Definition 4.6. (Benyamini and Lindenstrauss, 1998, Definition 4.1) If the map $g'(x, \cdot) : X \to Y$ is continuous and linear then g is Gâteaux differentiable at x. In particular if X and Y are normed Banach spaces then continuity of $g'(x, \cdot)$ is equivalent to boundedness in the operator norm.

Definition 4.7. (Bonnans and Shapiro, 2000, Definition 2.45) The mapping g is directionally differentiable at $x \in X$ in the Hadamard sense if it is directionally differentiable at x and for any sequences $(h_n)_{n \in \mathbb{N}} \subset X$ converging to h and $(t_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ converging to 0 the following equality holds:

$$g'(x, h) := \lim_{n \to \infty} \frac{g(x + t_n h_n) - g(x)}{t_n}.$$

In addition if $g'(x, \cdot)$ is linear in h then it is said to be Hadamard differentiable at x.

We now outline several properties of different directional derivatives.

Proposition 4.8. *(Bonnans and Shapiro, 2000, Proposition 2.46) If g is directionally differentiable at $x \in X$ in the Hadamard sense then $g'(x, \cdot)$ is continuous on X.*

We also contrast these two notions of directional differentiability with Fréchet differentiability, the definition of which we now recall.

Definition 4.9. (Andrews and Hopper, 2011, Definition A.1.) If $g'(x, \cdot)$ is the Gâteaux derivative of g at x and it exists and is uniformly continuous on $B_x(1) \subset X$ (a ball of radius 1 centred at x) then g is said to be Fréchet differentiable at x. In particular one can write the Fréchet derivative as a limit

$$Dg(x) := \lim_{h \to 0} \frac{g(x+h) - g(x)}{\|h\|_X}.$$

It can be equivalently expressed as

$$g(x+h) = g(x) + Dg(x)h + o(\|h\|_X),$$

for all $h \in X$ and $g'(x, h) = Dg(x)h$.

Unlike the Gâteaux derivative Fréchet derivative is not direction dependent, which makes it a very strong notion of directional differentiability. In particular Fréchet differentiability implies Hadamard differentiability and if underlying space X is finite-dimensional then Fréchet differentiability is equivalent to differentiability in the usual sense.

Hadamard derivative may lack continuity properties, which makes it the weakest notion of directional differentiability introduced so far. However if the function g is directionally differentiable and locally Lipschitz continuous then we can state some continuity properties of the Hadamard directional derivative.

Proposition 4.10. *(Bonnans and Shapiro, 2000, Proposition 2.49) Suppose that g is directionally differentiable at x and Lipschitz continuous (with a constant L) in a neighbourhood of x. Then g is directionally differentiable at x in the Hadamard sense and the directional derivative $g'(x, \cdot)$ is Lipschitz continuous (with the same constant L) on X.*

If X is a finite-dimensional vector space then the situation simplifies considerably. If g is also locally Lipschitz continuous at $x \in X$ then all different notions of directional differentiability are equivalent. It follows by Proposition 4.2 that for any proper convex function on a finite-dimensional vector space all notions of directional differentiability coincide.

We now state several properties of directional derivatives in general terms.

Proposition 4.11. *(Chain rule (Bonnans and Shapiro, 2000, Proposition 2.47)) If the function $g : X \to Y$ is directionally differentiable at a point x and the function $f : Y \to Z$ is Hadamard differentiable at $y = g(x)$ then the mapping $f \circ g$ is directionally differentiable at x and the following holds:*

$$(f \circ g)'(x, h) = f'(y, g'(x, h)).$$

Moreover if g is Fréchet differentiable at x and f is Fréchet differentiable at y then $f \circ g$ is Fréchet differentiable at x.

Outline of proof. Since g is directionally differentiable at x it follows that $g(x + th) = g(x) + tg'(x, h) + o(t)$. As f is Hadamard differentiable at y it is also directionally differentiable at y and hence

$$f(g(x + th)) = f(y + tg'(x, h) + o(t)) = f(y) + tf'(y, g'(x, h)) + o(t).$$

Hence $f \circ g$ is directionally differentiable.

Now, if f and g are Fréchet differentiable it implies that

$$\begin{aligned} f(g(x + h)) &= f(y + Dg(x)h + o(\|h\|_X)) \\ &= f(y) + f'(y, Dg(x)h)h + o(\|h\|_X) \\ &= f(y) + Dg(x)Df(y)h + o(\|h\|_X). \end{aligned}$$

As Dg and Df are uniformly continuous on $B_x(1)$ it implies that $f \circ g$ is Fréchet differentiable at x. $\square$

Proposition 4.12. *(Bonnans and Shapiro, 2000, Proposition 2.126 (iv-v)) Assume that X is a Banach space endowed with the norm topology. Assume also that the function $g : X \to \mathbb{R}$ is convex and continuous at a point $x \in X$. Then the following holds*

(i) *g is sub-differentiable at x;*

(ii) *$\partial g(x)$ is a non-empty, convex and weak* compact subset of X^*;*

(iii) *g is Hadamard directionally differentiable at x and the following holds for all directions $h \in X$*

$$g'(x, h) = \sup_{x^* \in \partial g(x)} \langle x^*, h \rangle .$$

It is clear that if in Proposition 4.12(iii) one has $\partial g(x) = \{a\}$ then $g'(x, h) = \langle a, h \rangle$ and g is Hadamard differentiable at x. It must be noted that similar result is proved in (Rockafellar, 1970, Theorem 23.4) when X is a finite-dimensional vector space.

4.2 Perturbation analysis

We embed the primal (3.3) and the dual (3.4) problems into a family of perturbed problems by introducing a vector $\mathrm{u} := (u_1^t, \dots, u_{\kappa(t)}^t)_{t \in \mathcal{T}} \in \mathbb{R}^d$ of price perturbations. Given an option with pay-off $\Phi \in U_h(\Omega)$ let $\overline{\vartheta}_p : \mathbb{R}^d \to \overline{\mathbb{R}}$ denote the value of the perturbed super-hedging primal problem

$$\overline{\vartheta}_p(\mathrm{u}) := \inf \left\{ \lambda + \langle \mathrm{c} + \mathrm{u}, \mathrm{w} \rangle : (\mathrm{w}, \lambda, \Theta) \in \overline{\mathcal{F}}_p \right\}, \tag{4.1}$$

where $\overline{\mathcal{F}}_p$ is the feasible set defined as

$$\overline{\mathcal{F}}_p := \left\{ (\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}} : A(\mathrm{w}, \lambda, \Theta; \omega) - \Phi(\omega) \geq 0 \text{ for all } \omega \in \Omega \right\}.$$

The explicit dependence on the pay-off Φ in the notations is dropped for simplicity. It must be noted that the value function $\overline{\vartheta}_p$ is convex and $\overline{\vartheta}_p(0)$ coincides with the value of the unperturbed primal problem (3.3).

We now define a Lagrangian function

$$L_\lambda^\Theta(\mathrm{w}, \mu) := \lambda + \langle \mathrm{c}, \mathrm{w}\rangle - \langle A(\mathrm{w}, \lambda, \Theta; \cdot) - \Phi, \mu\rangle \,. \tag{4.2}$$

We now take the supremum of $L_\lambda^\Theta(\mathrm{w}, \mu) + \langle \mathrm{u}, \mathrm{w}\rangle$ over $\mu \in (\mathcal{M}_h)_+(\Omega)$, the space of non-negative Borel measures of bounded variation that integrate h to a finite constant, to obtain

$$\sup_{\mu \in (\mathcal{M}_h)_+(\Omega)} \left\{ L_\lambda^\Theta(\mathrm{w}, \mu) + \langle \mathrm{u}, \mathrm{w}\rangle \right\} = \begin{cases} \lambda + \langle \mathrm{c} + \mathrm{u}, \mathrm{w}\rangle \,, & \text{if } (\mathrm{w}, \lambda, \Theta) \in \overline{\mathcal{F}}_p, \\ +\infty, & \text{otherwise.} \end{cases}$$

Therefore the primal problem is equivalent to

$$\inf_{(\mathrm{w},\lambda,\Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}} \sup_{\mu \in (\mathcal{M}_h)_+(\Omega)} \{ L_\lambda^\Theta(\mathrm{w}, \mu) + \langle \mathrm{u}, \mathrm{w}\rangle \}.$$

On the other hand if infimum is taken over $(\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}$ first, we obtain the following expression

$$\begin{aligned} &\inf_{(\mathrm{w},\lambda,\Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}} \left\{ L_\lambda^\Theta(\mathrm{w}, \mu) + \langle \mathrm{u}, \mathrm{w}\rangle \right\} \\ &= \inf_{(\mathrm{w},\lambda,\Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}} \{ \langle \Phi, \mu\rangle + \lambda + \langle \mathrm{c} + \mathrm{u}, \mathrm{w}\rangle - \langle A(\mathrm{w}, \lambda, \Theta; \cdot), \mu\rangle \} \,. \end{aligned}$$

The expression on the right is not equal to $-\infty$ if

$$\lambda + \langle \mathrm{c} + \mathrm{u}, \mathrm{w}\rangle = \langle A(\mathrm{w}, \lambda, \Theta; \cdot), \mu\rangle \,,$$

for all $(\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}$. Expanding the right-hand side according to Definition (3.2) and comparing the terms on the left and the right of the equality we see that it holds if

$$\langle \lambda, \mu\rangle = \lambda, \quad \langle (\Theta \bullet S)_T, \mu\rangle = 0, \quad \text{and} \quad \langle \mathrm{Cw}, \mu\rangle = \langle \mathrm{c} + \mathrm{u}, \mathrm{w}\rangle \,.$$

In particular the last equality can be re-written as follows

$$0 = \langle \mathrm{Cw}, \mu \rangle - \langle \mathrm{c} + \mathrm{u}, \mathrm{w} \rangle = \langle \mathrm{w}, \mathrm{C}^* \mu \rangle - \langle \mathrm{c} + \mathrm{u}, \mathrm{w} \rangle = \langle \mathrm{C}^* \mu - \mathrm{c} - \mathrm{u}, \mathrm{w} \rangle ,$$

where $\mathrm{C}^* : \mu \mapsto \mathrm{C}^* \mu \in \mathbb{R}^d$ defined as $\mathrm{C}^* \mu := \int_\Omega \mathrm{C}(\omega) \mu(\mathrm{d}\omega)$ is the adjoint map of $\mathrm{C} : \mathrm{w} \mapsto \mathrm{Cw} \in C_h(\Omega)$. As the inner product on the right-hand side is null for all $\mathrm{w} \in \mathbb{R}^d$ it follows that

$$\int_\Omega \mathrm{C}(\omega) \mu(\mathrm{d}\omega) = \mathrm{c} + \mathrm{u}.$$

Thus let the function $\overline{\vartheta}_d : \mathbb{R}^d \to \overline{\mathbb{R}}$ be the value of the perturbed dual problem defined as

$$\overline{\vartheta}_d(\mathrm{u}) := \sup \{ \langle \Phi, \mu \rangle : \mu \in \mathbb{M}_\mathrm{u} \} , \tag{4.3}$$

where $\mathbb{M}_\mathrm{u}$ is the feasible set of all non-negative Borel measures that integrate h to a finite constant

$$\mathbb{M}_\mathrm{u} := \{ \mu \in (\mathcal{M}_h)_+(\Omega) : \langle (\Theta \bullet S)_T, \mu \rangle = 0, \; \mathrm{C}^* \mu = \mathrm{c} + \mathrm{u} \} , \tag{4.4}$$

satisfying the martingale condition (2.19) for all $\Theta \in \tilde{\mathcal{H}}$ and is consistent with perturbed Call option prices. Note that due to presence of a riskless bond in the hedging portfolio, a feasible measure turns out to be a probability measure. The value $\overline{\vartheta}_d(0)$ is equal to the value of the unperturbed dual problem (3.4).

Similarly for a fixed $\Phi \in L_h(\Omega)$ we can embed the sub-hedging primal problem (3.5) in a family of perturbed problems as follows

$$\underline{\vartheta}_p(\mathrm{u}) := \sup \left\{ \lambda + \langle \mathrm{c} + \mathrm{u}, \mathrm{w} \rangle : (\mathrm{w}, \lambda, \Theta) \in \underline{\mathcal{F}}_p \right\} , \tag{4.5}$$

where we recall that the feasible set $\underline{\mathcal{F}}_p$ is defined as

$$\underline{\mathcal{F}}_p := \left\{ (\mathrm{w}, \lambda, \Theta) \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}} : \Phi(\omega) - A(\mathrm{w}, \lambda, \Theta; \omega) \geq 0 \text{ for all } \omega \in \Omega \right\} .$$

Similarly to the primal problem (4.1) the value $\underline{\vartheta}_p(0)$ corresponds to the value of the unperturbed primal problem (3.5). The perturbed dual problem reads

$$\underline{\vartheta}_d(\mathrm{u}) := \inf \{\langle \Phi, \mu \rangle : \mu \in \mathbb{M}_{\mathrm{u}}\}, \tag{4.6}$$

where $\mathbb{M}_{\mathrm{u}}$ is the feasible set defined in (4.4). Note, that as with the primal problem (4.5) the value $\underline{\vartheta}_d(0)$ is the value of the unperturbed dual problem (3.6).

Conditions for absence of duality gap between the primal (4.1) and the dual (4.3) problems can be expressed in terms of conditions on the adjoint map C^*. Let us first define the range of the adjoint map C^* as

$$\begin{aligned}\mathbf{M} := \big\{\mathrm{u} \in \mathbb{R}^d : \text{there exists } \mu \in (\mathcal{M}_h)_+(\Omega),\ \mathrm{u} = -\mathrm{c} + \mathrm{C}^*\mu, \\ \langle (\Theta \bullet S)_T, \mu \rangle = 0 \text{ for all } \Theta \in \tilde{\mathcal{H}}\big\}.\end{aligned} \tag{4.7}$$

The sufficient condition to guarantee absence of the duality gap between the primal and the dual problems at some $\mathrm{u}_0 \in \mathbb{R}^d$ is that $\mathrm{u}_0 \in \operatorname{int} \mathbf{M}$, which follows by arguments above. This condition goes back to Karlin and Studden (1966, Chapter XII, Theorem 2.1) in the context of generalised Tchebycheff inequalities. It can also be found in various forms in Anderson and Nash (1987, Theorem 4.4) and Bonnans and Shapiro (2000, Condition 5.259). The duality result can be stated as follows.

Theorem 4.13. *Bonnans and Shapiro (2000, Theorems 5.99) Assume that* $\mathrm{u}_0 \in \operatorname{int} \mathbf{M}$*, where* $\mathbf{M}$ *is the moment cone defined in* (4.7)*. Then there is no duality gap between the primal* (4.1) *and the dual* (4.3) *problems, i.e.* $\overline{\vartheta}_p(\mathrm{u}_0) = \overline{\vartheta}_d(\mathrm{u}_0)$*. Similarly the equality* $\underline{\vartheta}_p(\mathrm{u}_0) = \underline{\vartheta}_d(\mathrm{u}_0)$ *holds for programmes* (4.5) *and* (4.6)*.*

A similar result was used in Davis et al. (2014) to prove no duality gap result under condition of absence of weak arbitrage opportunities. We adapt their result here to show that absence of weak arbitrage opportunities implies no duality gap between the primal (4.1) and the dual (4.3) problems (and similarly for programmes (4.5) and (4.6)).

Proposition 4.14. *Suppose for some perturbation* $\mathrm{u}_0 \in \mathbb{R}^d$ *the prices* $\mathrm{u}_0 + \mathrm{c}$ *do not admit weak arbitrage opportunities. Then there is no duality gap be-*

tween the primal (4.1) *and the dual* (4.3) *super-hedging problems. Similarly there is no duality gap between the primal and the dual sub-hedging problems* (4.5) *and* (4.6) *respectively.*

Proof. We prove that absence of weak arbitrage opportunities imply that $\mathrm{u}_0 \in \operatorname{int} \mathbf{M}$. Absence of weak arbitrage opportunities is equivalent to existence of a model $\mu \in \mathbb{M}_{\mathrm{u}}$ for prices $\mathrm{c} + \mathrm{u}_0$ (Davis and Hobson, 2007, Theorem 4.2). Moreover following Davis et al. (2014, Proof of Proposition 3.1) in order to show $\mathrm{u}_0 \in \operatorname{int} \mathbf{M}$ it is sufficient to note that for any entry $c(K_i^t, t) + u_i^t$ of the vector $\mathrm{c} + \mathrm{u}_0$ the following inequalities hold

$$(1 - K_i^t)_+ < c(K_i^t, t) + u_i^t < 1,$$

for all $i = 1, \dots, \kappa(t)$ and $t \in \mathcal{T}$. As $\mu \mapsto \mathrm{C}^*\mu$ is a continuous function on $(\mathcal{M}_h)_+(\Omega)$ (as an application of Beiglböck, Henry-Labordère, and Penkner (2013, Lemma 2.2) to re-scaled $\mu \in (\mathcal{M}_h)_+(\Omega)$) one can also find a real positive constant $\varepsilon > 0$ such that any vector v in the open ball $\mathcal{B}_\varepsilon(\mathrm{c} + \mathrm{u})$ centred around $\mathrm{c} + \mathrm{u}$ satisfies Assumption 12 and therefore $\mathrm{u} \in \operatorname{int} \mathbf{M}$ so the theorem follows. □

Having established no duality gap result between the primal (4.1) and the dual (4.3) (and between programmes (4.5) and (4.6)) we can now discuss sensitivity of the programme values to the perturbation. In particular, a consequence of Theorem 4.13 is that $\overline{\vartheta}_d$ is continuous at u_0. Moreover if $\overline{\vartheta}_p$ is finite at u_0 we have the following result.

Proposition 4.15. *Assume that there is no duality gap between the primal* (4.1) *and the dual* (4.3) *problems at some* $\mathrm{u}_0 \in \mathbb{R}^d$*. Moreover assume that* $\overline{\vartheta}_p(\mathrm{u}_0) < \infty$*. Then* $\overline{\vartheta}_d$ *is Hadamard directionally differentiable at* u_0 *and the derivative in any direction* $h \in \mathbb{R}^d$ *is written as follows*

$$(\overline{\vartheta}_d)'(\mathrm{u}_0, h) = \inf_{x \in \mathfrak{S}_{\mathrm{u}_0}} \langle x, h \rangle ,$$

where $\mathfrak{S}_{\mathrm{u}_0} \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}$ *is the set of optimal solutions of the primal problem* (4.1) *at* u_0*.*

Proof. By a change of variables $\mu \mapsto -\mu$ we turn the dual problem into the minimization problem

$$w(\mathrm{u}) := \inf \{\langle \Phi, \mu \rangle : -\mu \in \mathbb{M}_{\mathrm{u}}\}.$$

One of course has $w(\mathrm{u}) = -\overline{\vartheta}_d(\mathrm{u})$. Let us now calculate the convex conjugate of w at $\mathrm{u}^* \in \mathbb{R}^d$

$$\begin{aligned}
w^*(\mathrm{u}^*) &= \sup_{\mathrm{u}\in\mathbb{R}^d} \{\langle \mathrm{u}, \mathrm{u}^* \rangle - w(\mathrm{u})\} \\
&= \sup_{\mu\in(\mathcal{M}_h)_+(\Omega)} \sup_{\mathrm{u}\in\mathbb{R}^d} \{\langle \mathrm{u}, \mathrm{u}^* \rangle - \langle \Phi, \mu \rangle - \chi_{\mathbb{M}_{\mathrm{u}}}(-\mu)\} \\
&= \sup_{\mu\in(\mathcal{M}_h)_+(\Omega)} \sup_{\mathrm{u}\in\mathbb{R}^d} \{\langle \mathrm{u}, \mathrm{u}^* \rangle - \langle \Phi, \mu \rangle - \langle \mathrm{u} + \mathrm{c} + \mathrm{C}^*\mu, \mathrm{u}^* \rangle + \langle \mathrm{u} + \mathrm{c} + \mathrm{C}^*\mu, \mathrm{u}^* \rangle \\
&+ \langle (\Theta \bullet S)_T, \mu \rangle - \langle (\Theta \bullet S)_T, \mu \rangle + \langle \lambda, \mu \rangle - \lambda - \langle \lambda, \mu \rangle + \lambda - \chi_{\mathbb{M}_{\mathrm{u}}}(-\mu)\} \\
&= \sup_{\mu\in(\mathcal{M}_h)_+(\Omega)} \{L^{\Theta}_{\lambda}(-\mathrm{u}^*, -\mu) + \sup_{\mathrm{u}\in\mathbb{R}^d} \{\langle \mathrm{u} + \mathrm{c} - \mathrm{C}^*(-\mu), \mathrm{u}^* \rangle - \lambda \\
&+ \langle \lambda + (\Theta \bullet S)_T, -\mu \rangle - \chi_{\mathbb{M}_{\mathrm{u}}}(-\mu)\}\},
\end{aligned}$$

where L is defined in (4.2). The second supremum in the above expression is finite if and only if $-\mu \in \mathbb{M}_u$ and hence the convex conjugate reads

$$w^*(\mathrm{u}^*) = \sup_{-\mu\in\mathbb{M}_{\mathrm{u}}} \left\{L^{\Theta}_{\lambda}(-\mathrm{u}^*, -\mu)\right\}.$$

The convex conjugate of w^* is written as follows

$$\begin{aligned}
w^{**}(\mathrm{u}) &= \sup_{\mathrm{u}^*\in\mathbb{R}^d} \{\langle \mathrm{u}, \mathrm{u}^* \rangle - w^*(\mathrm{u}^*)\} \\
&= \sup_{\mathrm{u}^*\in\mathbb{R}^d} \inf_{-\mu\in\mathbb{M}_{\mathrm{u}}} \left\{\langle \mathrm{u}, \mathrm{u}^* \rangle - L^{\Theta}_{\lambda}(-\mathrm{u}^*, -\mu)\right\} \\
&= -\inf_{\mathrm{u}^*\in\mathbb{R}^d} \sup_{\mu\in\mathbb{M}_{\mathrm{u}}} \left\{\langle \mathrm{u}, -\mathrm{u}^* \rangle + L^{\Theta}_{\lambda}(-\mathrm{u}^*, \mu)\right\} \\
&= -\inf_{\mathrm{u}^*\in\mathbb{R}^d} \sup_{\mu\in\mathbb{M}_{\mathrm{u}}} \left\{\langle \mathrm{u}, \mathrm{u}^* \rangle + L^{\Theta}_{\lambda}(\mathrm{u}^*, \mu)\right\} = -\overline{\vartheta}_p(\mathrm{u}),
\end{aligned}$$

where the third equality is a consequence of change of variables $-\mu \mapsto \mu$ and the last equality is a result of change of variables $-\mathrm{u}^* \mapsto \mathrm{u}^*$. The Young-Fenchel inequality implies that $w \geq w^{**}$ and hence we recover weak duality between the primal (4.1) and the dual (4.3) problems, i.e. $\overline{\vartheta}_p(\mathrm{u}) \geq \overline{\vartheta}_d(\mathrm{u})$.

By assumption there is no duality gap, i.e. $\overline{\vartheta}_d(\mathrm{u}_0) = \overline{\vartheta}_p(\mathrm{u}_0)$ and hence $w(\mathrm{u}_0) = w^{**}(\mathrm{u}_0)$, which implies that w is lower semi-continuous by Fenchel-Moreau Theorem. Moreover as $\mathrm{u}_0 \in \operatorname{int} \mathbf{M}$ it follows that w is continuous at u_0 by Theorem 4.1. By Proposition 4.12(i) the sub-differential $\partial w(\mathrm{u}_0)$ is non-empty and by Proposition 4.12(iii) the function w is Hadamard directionally differentiable at u_0 in any direction $h \in \mathbb{R}^d$ such that

$$w'(\mathrm{u}_0, h) = \sup_{\mathrm{u}^* \in \partial w(\mathrm{u}_0)} \langle \mathrm{u}^*, h \rangle .$$

We have by the Young-Fenchel inequality that

$$w(\mathrm{u}_0) = \langle \mathrm{u}_0, \mathrm{u}^* \rangle - w^*(\mathrm{u}^*),$$

if and only if $\mathrm{u}^* \in \partial w(\mathrm{u}_0)$ and hence it follows that $w^{**}(\mathrm{u}_0) = w(\mathrm{u}_0)$. The primal problem (4.1) can be expressed as $-w^{**}(\mathrm{u}_0)$ by discussion above and it is finite by assumption. Hence we have that $\partial w(\mathrm{u}_0) = -\mathfrak{S}_{\mathrm{u}_0}$ (the set of optimal solutions of the primal problem (4.1) at u_0) and it follows that

$$w'(\mathrm{u}_0, h) = \sup_{\mathrm{u}^* \in -\mathfrak{S}_{\mathrm{u}_0}} \langle \mathrm{u}^*, h \rangle = - \inf_{\mathrm{u}^* \in \mathfrak{S}_{\mathrm{u}_0}} \langle \mathrm{u}^*, h \rangle .$$

On the other hand as $w(\mathrm{u}_0) = -\overline{\vartheta}_d(\mathrm{u}_0)$ we have the following

$$\begin{aligned}
(\overline{\vartheta}_d)'(\mathrm{u}_0, h) &= \lim_{t \downarrow 0} \frac{\overline{\vartheta}_d(\mathrm{u}_0 + th) - \overline{\vartheta}_d(\mathrm{u}_0)}{t} \\
&= \lim_{t \downarrow 0} \frac{-w(\mathrm{u}_0 + th) + w(\mathrm{u}_0)}{t} \\
&= - \lim_{t \downarrow 0} \frac{w(\mathrm{u}_0 + th) - w(\mathrm{u}_0)}{t} \\
&= -w'(\mathrm{u}_0, d).
\end{aligned}$$

The statement of the proposition then follows. □

Corollary 4.16. *Assume that there is no duality gap between the primal* (4.5) *and the dual* (4.6) *problems at some* $\mathrm{u}_0 \in \mathbb{R}^d$. *Moreover assume that* $\underline{\vartheta}_p(\mathrm{u}_0) <$

∞. *Then* $\underline{\vartheta}_d$ *is Hadamard directionally differentiable at* u_0 *and the derivative in any direction* $h \in \mathbb{R}^d$ *is written as follows*

$$(\underline{\vartheta}_d)'(\mathrm{u}_0, h) = \sup_{x \in \Lambda_{\mathrm{u}_0}} \langle x, h \rangle \,,$$

where $\Lambda_{\mathrm{u}_0} \in \mathbb{R}^{d+1} \times \tilde{\mathcal{H}}$ *is the set of optimal solutions of the primal problem* (4.5) *at* u_0.

Proof. Noting that $\underline{\vartheta}_d(\mathrm{u}_0)$ is itself a minimization problem, the statement follows from the proof of Proposition 4.15 with $w = \underline{\vartheta}_d$. $\square$

If the perturbation u is itself parametrised by a vector $\mathrm{s} \in \mathbb{R}^n$ for some $n < \infty$ and it is continuously differentiable with respect to this parameter then we have the following application of the Chain Rule 4.11.

Corollary 4.17. *Let all assumptions in Proposition 4.15 hold. Assume also that* $\mathrm{u}_0 := \mathrm{u}(\mathrm{s}_0)$ *is continuously differentiable with respect to the parameter* $\mathrm{s}_0 \in \mathbb{R}^n$. *Then we have the following*

$$(\overline{\vartheta}_d \circ \mathrm{u})'(\mathrm{s}_0, h) = \inf_{\mathrm{u}^* \in \mathfrak{S}_{\mathrm{u}_0}} \langle \mathrm{u}^*, \nabla \mathrm{u}(\mathrm{s}_0) h \rangle \,,$$

where $\nabla \mathrm{u}(\mathrm{s}_0)$ *is the Jacobian matrix evaluated at* s_0.

Proof. As u is continuously differentiable it is Fréchet differentiable and $(\mathrm{u})'(\mathrm{s}_0, h) = \nabla \mathrm{u}(\mathrm{s}_0) h$. The function $\overline{\vartheta}_d$ is Hadamard directionally differentiable at u by Proposition 4.15. Hence we can apply the Chain Rule 4.11 to obtain the result. $\square$

Similar result holds for the sub-hedging problem and we state it without proof.

Corollary 4.18. *Let all assumptions in Corollary 4.16 hold. Assume also that* $\mathrm{u}_0 := \mathrm{u}(\mathrm{s}_0)$ *is continuously differentiable with respect to parameter* $\mathrm{s}_0 \in \mathbb{R}^n$. *Then we have the following*

$$(\underline{\vartheta}_d \circ \mathrm{u})'(\mathrm{s}_0, h) = \sup_{\mathrm{u}^* \in \Lambda_{\mathrm{u}_0}} \langle \mathrm{u}^*, \nabla \mathrm{u}(\mathrm{s}_0) h \rangle \,,$$

where $\nabla \mathrm{u}(\mathrm{s}_0)$ *is the Jacobian matrix.*

It must be noted that if there exists a unique solution to the super- or sub-hedging primal problem (4.1) (respectively (4.5)) at u_0 then the set of optimal solutions $\mathfrak{S}_{\mathrm{u}_0}$ (respectively Λ_{u_0}) is a singleton and the derivatives in Proposition 4.15 and Corollaries 4.16, 4.17 and 4.18 are all linear in h.

As a consequence of linearity of the derivatives Goberna and López (2014, Section 4.1) show that there exists a neighbourhood U of u_0 such that for all $\mathrm{u} \in U$ the value of the perturbed super-hedging dual problem (4.3) can be approximated at u as follows

$$\overline{\vartheta}_d(\mathrm{u}) = \overline{\vartheta}_d(\mathrm{u}_0) + \langle \mathrm{u}^*, \mathrm{u} - \mathrm{u}_0 \rangle + o(\mathrm{u} - \mathrm{u}_0),$$

where $\mathrm{u}^* \in \mathfrak{S}_{\mathrm{u}_0}$ and similarly for the perturbed sub-hedging problem (4.6). This approximation can be naturally extended to the case when the perturbation u is itself parametrised. In particular for all s in the neighbourhood S of s_0 the following approximation of the perturbed dual problem (4.3) holds

$$\overline{\vartheta}_d \circ \mathrm{u}(\mathrm{s}) = \overline{\vartheta}_d \circ \mathrm{u}(\mathrm{s}_0) + \langle \mathrm{u}^*, \nabla \mathrm{u}(\mathrm{s}_0)(\mathrm{s} - \mathrm{s}_0) \rangle + o(\mathrm{s} - \mathrm{s}_0), \tag{4.8}$$

with the similar expression for the perturbed sub-hedging problem (4.6).

4.3 Application to Forward-Start Straddle

We shall perform a sensitivity analysis of optimal values of robust hedging of the Forward-Start Straddle with respect to extrapolation of the total implied variance for each maturity $t \in \mathcal{T}$. In the sequel we assume that there is a unique solution to the primal perturbed problems (4.1) and (4.5) as the question of uniqueness of solutions is beyond the scope of this thesis. The super- and sub-hedges of the Forward-Start Straddle with pay-off $|S_{t+\tau} - \mathcal{K} S_t|$ (for some $t, \tau > 0$ and various strikes $\mathcal{K} > 0$) was discussed in Section 3.3. Two cases for input option prices were considered: Call option prices were consistent with either the Black-Scholes or the Heston models. The models were chosen as they produce very different profiles of the total implied variance. However in practice the input Call option prices are given by the market and arbitrage-free extrapolation of variance is subject to some form of parametrization.

Let us suppose that we are given $\kappa(t), \kappa(t+\tau) < \infty$ Call options maturing at time t with normalised strikes $K_1^t, \dots, K_{\kappa(t)}^t$ and Call options maturing at time $t+\tau$ with normalised strikes $K_1^{t+\tau}, \dots, K_{\kappa(t+\tau)}^{t+\tau}$. The vector of normalised Call prices then reads

$$\mathrm{c} = \big(c(K_1^t, t), \dots, c(K_{\kappa(t)}^t, t), c(K_1^{t+\tau}, t+\tau), \dots, c(K_{\kappa(t+\tau)}^{t+\tau}, t+\tau)\big).$$

We parametrise the total implied variance surface w by a vector of parameters $\mathrm{s} \in \mathbb{R}^l$ for some $l < \infty$ such that that the resulting surface is arbitrage free and grows at most linearly in the wings for each maturity t and $t+\tau$. Let us denote by $w(\cdot, \cdot; \mathrm{s})$ the parametrised surface.

Assumption 14. Parametrisation of the total implied variance surface $w(\cdot, \cdot; \mathrm{s})$ is continuously differentiable with respect to s.

We can then calculate the resulting parametrised implied volatility for each log-moneyness k and each $t \in \mathcal{T}$ as $I(k, t; \mathrm{s}) = \sqrt{w(k, t; \mathrm{s})}$, where $I(\cdot, \cdot; \mathrm{s})$ denotes parametrised total implied volatility. We then define the vector of perturbed prices as follows

$$\begin{aligned}\mathrm{c} + \mathrm{u}(\mathrm{s}_0) := \big(&c_{BS}(k_1^t, I(k_1^t, t; \mathrm{s}_0)), \dots, c_{BS}(k_{\kappa(t)}^t, I(k_{\kappa(t)}^t, t; \mathrm{s}_0)),\\ &c_{BS}(k_1^{t+\tau}, I(k_1^{t+\tau}, t+\tau; \mathrm{s}_0)), \dots, c_{BS}(k_{\kappa(t)}^{t+\tau}, I(k_{\kappa(t+\tau)}^{t+\tau}, t+\tau; \mathrm{s}_0))\big),\end{aligned}$$

where $k = \log(K)$ as before. We can compute sensitivities of perturbed prices with respect to parameters s.

Lemma 4.19. *Sensitivity of a perturbed price $c_{BS}(k_i^t, I(k_1^i, t; \mathrm{s}))$, with $i = 1, \dots, \kappa(t)$ and $t \in \mathcal{T}$ with respect to a parameter s_j for $j = 1, \dots, l$ reads*

$$\frac{\partial c_{BS}(k_i^t, I(k_i^t, t; \mathrm{s}))}{\partial s_j} = \frac{V(k_i^t, \mathrm{s})}{2I(k_i^t, t; \mathrm{s})\sqrt{t}} \frac{\partial w(k_i^t, t; \mathrm{s})}{\partial s_j},$$

where $V(k, \cdot) := \partial c_{BS}(k, \sigma\sqrt{t})/\partial\sigma|_{\sigma = I(k,t;\cdot)/\sqrt{t}}$ is the Black-Scholes Vega of *the Call option with log-moneyness k and the Black-Scholes implied volatility equal to $I(k,t;\cdot)/\sqrt{t}$.*

Proof. A simple application of the chain rule together with Assumption 14 yields the following

$$\begin{aligned}
\frac{\partial c_{BS}(k_i^t, I(k_i^t, t; \mathrm{s}))}{\partial s_j} &= \frac{\partial c_{BS}(k_i^t, I(k_i^t; \mathrm{s})\sqrt{t})}{\partial s_j} \\
&= V(k_i^t, \mathrm{s})\frac{\partial I(k_i^t; \mathrm{s})}{\partial s_j} \\
&= V(k_i^t, \mathrm{s})\frac{\partial w(k_i^t, t; \mathrm{s})}{\partial s_j}\frac{\mathrm{d}I(k_i^t; \mathrm{s})}{\mathrm{d}w(k_i^t, t; \mathrm{s})} \\
&= \frac{V(k_i^t, \mathrm{s})}{2I(k_i^t, t; \mathrm{s})\sqrt{t}}\frac{\partial w(k_i^t, t; \mathrm{s})}{\partial s_j}.
\end{aligned}$$

□

The Jacobian matrix $\nabla \mathrm{u}_0(s)$ of the perturbed Call prices then reads

$$\nabla \mathrm{u}_0(\mathrm{s}) := \begin{pmatrix}
\frac{\partial c_{BS}(k_1^t, I(k_1^t, t; \mathrm{s}))}{\partial s_1} & \cdots & \frac{\partial c_{BS}(k_1^t, I(k_1^t, t; \mathrm{s}))}{\partial s_l} \\
\vdots & \ddots & \vdots \\
\frac{\partial c_{BS}(k_{\kappa(t)}^t, I(k_{\kappa(t)}^t, t; \mathrm{s}))}{\partial s_1} & \cdots & \frac{\partial c_{BS}(k_{\kappa(t)}^t, I(k_{\kappa(t)}^t, t; \mathrm{s}))}{\partial s_l} \\
\frac{\partial c_{BS}(k_1^{t+\tau}, I(k_1^{t+\tau}, t+\tau; \mathrm{s}))}{\partial s_1} & \cdots & \frac{\partial c_{BS}(k_1^{t+\tau}, I(k_1^{t+\tau}, t+\tau; \mathrm{s}))}{\partial s_l} \\
\vdots & \vdots & \vdots \\
\frac{\partial c_{BS}(k_{\kappa(t+\tau)}^{t+\tau}, I(k_{\kappa(t+\tau)}^{t+\tau}, t+\tau; \mathrm{s}))}{\partial s_1} & \cdots & \frac{\partial c_{BS}(k_{\kappa(t+\tau)}^{t+\tau}, I(k_{\kappa(t+\tau)}^{t+\tau}, t+\tau; \mathrm{s}))}{\partial s_l}
\end{pmatrix}.$$

In the sequel we assume the same set-up as in Section 3.4. In particular we assume that $\mathcal{T} = \{1, 1.5\}$, the set of trading strategy is discretized employing a monomial basis of degree at most 4 and there are 18 options available for each maturity $t \in \mathcal{T}$ for static hedging with moneyness in $\{0.3, 0.4, 0.5, \ldots, 2.0\}$. However we shall assume that only a subset of those options has quotable market prices and the rest of Call options are priced by extrapolating the total implied variance. For numerical implementation the state space is taken to be $[0, 5] \times [0, 5]$ with 500 discretization points for both maturities.

4.3.1 Application to the Black-Scholes model

Let us now assume that only prices of at-the-money Call options are observable on the market for each maturity $t \in \mathcal{T}$. In this case it is not unreasonable to fit the Black-Scholes model to the observed prices. The only parameter in the Black-Scholes model that needs calibration is the diffusion coefficient Σ in the stochastic differential equation that governs the evolution of the stock price process S

$$\mathrm{d}S_t = S_t \Sigma \mathrm{d}W_t, \quad S_0 = 1,$$

where W is one-dimensional standard Brownian motion. We let $\Sigma = 0.2$ as in Section 3.4.2. The resulting total implied variance function $w : \mathbb{R} \times \mathcal{T} \to \mathbb{R}_+$ is constant in log-moneyness k for a fixed maturity and reads $w(k,t) = \Sigma^2 t$, with $t \in \mathcal{T}$.

Let us now suppose that the actual shape of the total implied variance for a fixed maturity $t \in \mathcal{T}$ reads

$$w(k,t;\mathrm{s}) = a_t |k| + \Sigma^2 t, \tag{4.9}$$

where $a_t \in \mathbb{R}$ is the slope of the right and the left wing of the total implied variance and $\mathrm{s} = (a_t, a_{t+\tau}) \in \mathbb{R}^2$ is the vector of parameters (where we also let $\mathcal{T} = \{t, t+\tau\}$ with $t = 1$ and $\tau = 0.5$). For a fixed maturity $t \in \mathcal{T}$ the function g defined in (2.13) must be non-negative for all $k > k^*$ and Proposition 2.7 implies that it is equivalent to $a_t \in [0,2]$ and existence of such $k^* \in \mathbb{R}_+$. As we propose extrapolation of the total implied variance to the right for all $k > 0$ and to the left for all $k < 0$ it implies that $k^* = 0$ (as $g(0) = \Sigma^2 t > 0$) which places further restrictions on a_t. In particular if $\Sigma^2 t \geq 2 - \sqrt{2 - a_t^2}$ then $g(k) \geq 0$ for all $k > 0$ by Proposition 2.7 and Corollary 2.8. This inequality places an upper bound on a_t for each maturity $t \in \mathcal{T}$ such that any extrapolation with the slope satisfying the bound is free of arbitrage. If $\Sigma^2 t < 2 - \sqrt{2 - a_t^2}$ then

$$g(k) > 0, \quad \text{for all} \quad k > \frac{a_t^2(\Sigma^2 t + 2) - 8\Sigma^2 t + 2a_t\sqrt{\Sigma^4 t^2 - 4\Sigma^2 t + a_t^2}}{a_t(4 - a_t^2)}.$$

It follows that the proposed extrapolation (4.9) is arbitrage-free if the expression on the right-hand side of the inequality above is equal to zero. The resulting quartic equation in a_t does not have real roots for either maturity $t \in \mathcal{T}$ when $\Sigma = 0.2$ and $\mathcal{T} = \{1, 1.5\}$. Hence the only viable values for a_t are between 0 and $\sqrt{4 - (2 - \Sigma^2 t)^2}$ for each $t \in \mathcal{T}$ (where the upper bound is obtained by solving the quadratic equation $\Sigma^2 t = 2 - \sqrt{2 - a_t^2}$).

Assumption 15. We assume the slopes for both maturities are equal to each other, i.e. $a_t = a_{t+\tau}$.

Remark 4.20. One can of course assume that the slopes are different for each maturity but then a condition $\partial_t w(k,t) \geq 0$ for all $t \in \mathcal{T}$ and $k \in \mathbb{R}$ must be satisfied in order to ensure absence of calendar spread arbitrage (Gatheral and Jacquier, 2014, Lemma 2.1). This condition in particular implies that the slices (i.e. functions $w(\cdot, t)$ for different $t \in \mathcal{T}$) of the implied total variance surface do not cross for any $k \in \mathbb{R}$. Therefore a potential choice for the slopes would be to increase the value of the slope for each wing as maturity increases.

Naturally the Jacobian $\nabla \mathrm{u}_0(\mathrm{s})$ now reads

$$\begin{pmatrix} \partial c_{BS}(k_1^t, I(k_1^t, t; \mathrm{s}))/\partial a_t & 0 \\ \vdots & \vdots \\ \partial c_{BS}(k_{\kappa(t)}^t, I(k_{\kappa(t)}^t, t; \mathrm{s}))/\partial a_t & 0 \\ 0 & \partial c_{BS}(k_1^{t+\tau}, I(k_1^{t+\tau}, t+\tau; \mathrm{s}))/\partial a_{t+\tau} \\ \vdots & \vdots \\ 0 & \partial c_{BS}(k_{\kappa(t+\tau)}^{t+\tau}, I(k_{\kappa(t+\tau)}^{t+\tau}, t+\tau; \mathrm{s}))/\partial a_{t+\tau} \end{pmatrix},$$

and as a simple consequence of Lemma 4.19 and Equation (4.9) we get that

$$\frac{\partial c_{BS}(k_i^t, I(k_i^t, t; \mathrm{s}))}{\partial a_t} = \frac{V(k_i^t, \mathrm{s})|k_i^t|}{2I(k_i^t, t; \mathrm{s})\sqrt{t}},$$

for all $i = 1, \ldots, \kappa(t)$ and $t \in \mathcal{T}$.

Below we present numerical results for the super- and sub-hedging primal programmes for the at-the-money Forward-Start Straddle, i.e. $\mathcal{K} = 1$. Tables 4.1 and 4.2 summarise result of the perturbation analysis for the hedging problems (4.1) and (4.5) respectively. The column "Perturbation" contains the values of the slopes of extrapolation of the total implied variance. As discussed above the slopes of the left and right wing are chosen to be the same and slopes for each maturity t and $t+\tau$ are chosen to be parallel to each other. As expected the optimal values of the perturbed problems converge to the optimal value of the unperturbed problem presented in the first row (i.e. the unperturbed problem is such that we take the Black-Scholes total implied variance for each $k \in \mathbb{R}$ for both maturities). The column "Est. Value" contains values that were computed via (4.8) as the value of the unperturbed problem plus values of derivatives and perturbations in the first two columns. The last column contains absolute difference between the optimal value of a perturbed problem obtained by solving the LP (4.1) and the value of the programme estimated via (4.8). As it can be seen the estimation becomes increasingly good the closer the value of the perturbation is to zero. It confirms that the perturbation results presented in Section 4.2 are local in nature.

Perturbation	Derivative	Optimal Value	Est. Value	Abs. Diff.
0	0	0.1490	0.1490	0
5.00E-05	6.51E-06	0.1490	0.1490	2.98E-10
1.00E-04	1.30E-05	0.1490	0.1490	1.19E-08
5.0E-03	6.0E-04	0.1496	0.1496	1.57E-06
0.0476	0.0062	0.1544	0.1552	7.75E-04
0.2020	0.0263	0.1563	0.1753	1.90E-02

Table 4.1: Perturbation of the super-hedging primal problem for the at-the-money Forward-Start Straddle in the Black-Scholes case.

4.3.2 Application to the Heston model

Let us now assume that for each maturity $t \in \mathcal{T}$ prices for Call options with moneyness in $\mathfrak{K}_{\text{market}} := \{0.8, 0.9, \ldots, 1.2\}$ are observed on the market. As in Section 3.4.3 we assume that the observed prices are consistent with the Heston stochastic volatility model (Heston, 1993), in which the stock price

Perturbation	Derivative	Optimal Value	Est. Value	Abs. Diff.
0	0	0.0385	0.0385	0
5.00E-05	-2.82E-06	0.0385	0.0385	2.88E-07
1.00E-04	-5.63E-06	0.0385	0.0385	3.42E-07
5.0E-03	-2.8E-04	0.0383	0.0383	1.16E-05
0.0476	-2.7E-03	0.0365	0.0359	6.10E-04
0.2020	-0.011	0.0357	0.0272	8.53E-03

Table 4.2: Perturbation of the sub-hedging primal problem for the at-the-money Forward-Start Straddle in the Black-Scholes case.

process is the unique strong solution to the stochastic differential equation

$$\begin{aligned} \mathrm{d}S_t &= S_t\sqrt{V_t}\mathrm{d}W_t, & S_0 = 1, \\ \mathrm{d}V_t &= \kappa\left(\theta - V_t\right)\mathrm{d}t + \xi\sqrt{V_t}\mathrm{d}Z_t, & V_0 = v > 0, \end{aligned} \tag{4.10}$$

where W and Z are two one-dimensional standard Brownian motions with $\mathrm{d}\langle W, Z\rangle_t = \rho \mathrm{d}t$, and $\kappa, \theta, \xi > 0$ and $\rho \in [-1, 1]$. We assume that $v = \theta = 0.07$, $\kappa = 1$, $\xi = 0.4$ and $\rho = -0.8$. In principle calibrating the Heston model provides an extrapolation of the total implied variance, however such parametrisation is not available in closed form and thus we make a simplifying assumption on the extrapolation of the implied variance beyond observable prices to perform a sensitivity analysis similar to the Black-Scholes case discussed above. Let us suppose that the total implied variance is extrapolated linearly to the left and the right of the observed Call prices for each maturity $t \in \mathcal{T}$. Let us define $L := \min\{i = 1, \dots, 18 : K_L = \min \mathfrak{K}_{\text{market}}\}$ and $R := \max\{i = 1, \dots, 18 : K_R = \max \mathfrak{K}_{\text{market}}\}$. Then for a vector of parameters $\mathrm{s} := (q_t, p_t, q_{t+\tau}, p_{t+\tau})$ we have the following extrapolation for the left wing

$$w(k, \cdot; \mathrm{s}) = \psi(q_\cdot)|k - k_L| + w(k_L, \cdot), \qquad \text{for all } k \leq k_L, \tag{4.11}$$

and the extrapolation for the right wing reads

$$w(k, \cdot; \mathrm{s}) = \psi(p_\cdot)|k - k_R| + w(k_R, \cdot), \qquad \text{for all } k \geq k_R, \tag{4.12}$$

where the function $\psi : \mathbb{R}_+ \to [0,2]$ is defined in (2.8) and we also let $\mathcal{T} = \{t, t+\tau\}$ with $t = 1$ and $\tau = 0.5$.

The Jacobian $\nabla \mathrm{u}_0(\mathrm{s})$ in this case reads

$$\begin{pmatrix} u_{1,1} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ u_{L-1,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \\ 0 & u_{R+1,2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & u_{m(t),2} & 0 & 0 \\ 0 & 0 & u_{1,3} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \partial u_{L-1,3} & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_{R+1,4} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & u_{m(t+\tau),4} \end{pmatrix},$$

where $u_{i,1} = \partial c_{BS}(k_i^t, I(k_i^t, t; \mathrm{s}))/\partial q_t$ for $i = 1, \dots, L-1$ are the sensitivities of the Call option prices to extrapolation of the left wing of the implied total variance at time t and $u_{i,2} = \partial c_{BS}(k_i^t, I(k_i^t, t; \mathrm{s}))/\partial p_t$ for $i = R+1, \dots, \kappa(t)$ are the sensitivities of the Call option prices to extrapolation of the right wing of the implied total variance at time t; $u_{i,3} = \partial c_{BS}(k_i^{t+\tau}, I(k_i^{t+\tau}, t+\tau; \mathrm{s}))/\partial q_{t+\tau}$ for $i = 1, \dots, L-1$ are the Call option price sensitivities with respect to extrapolation of the left wing of the implied total variance at time $t+\tau$ and $u_{i,4} = \partial c_{BS}(k_i^{t+\tau}, I(k_i^{t+\tau}, t+\tau; \mathrm{s}))/\partial p_{t+\tau}$ for $i = R+1, \dots, \kappa(t+\tau)$ are the sensitivities with respect to extrapolation of the right wing of the implied total variance at time $t+\tau$. Note that rows of zeros correspond to sensitivities of the traded Call option prices, which naturally do not depend on the extrapolation of the wings.

Lemma 4.21. *Let $w(k,t;\mathrm{s})$ be defined as in Equations* (4.11) *and* (4.12) *for each $t \in \mathcal{T}$ and $k \in \mathbb{R}$. Then the following expression for the partial derivative of $w(k,t;\mathrm{s})$ with respect to s_i, where $i = 1,\dots,4$ reads*

$$\frac{\partial w(k,t;\mathrm{s})}{\partial s_i} = -\frac{\left|k - \mathbb{1}_{\{i=1,3\}}(i)k_L - \mathbb{1}_{\{i=2,4\}}(i)k_R\right| \psi(s_i)}{\sqrt{s_i^2 + s_i}}.$$

Proof. We apply the chain rule to get

$$\frac{\partial w(k,t;\mathrm{s})}{\partial s_i} = \left|k - \mathbb{1}_{\{i=1,3\}}(i)k_L - \mathbb{1}_{\{i=2,4\}}(i)k_R\right| \frac{\partial \psi(s_i)}{\partial s_i},$$

and

$$\frac{\partial \psi(s_i)}{\partial s_i} = \frac{\partial}{\partial s_i}\left[2 - 4\left(\sqrt{s_i^2 + s_i} - s_i\right)\right]$$

$$= \frac{4\left(\sqrt{s_i^2 + s_i} - s_i\right) - 2}{\sqrt{s_i^2 + s_i}} = -\frac{\psi(s_i)}{\sqrt{s_i^2 + s_i}}.$$

□

Then by Lemmas 4.19 and 4.21 we have the following

$$\frac{\partial c_{BS}(k_i^t, I(k_i^t,t;\mathrm{s}))}{\partial s_j} = -\frac{V(k_i^t,\mathrm{s})\left|k - \mathbb{1}_{\{j=1,3\}}(i)k_L - \mathbb{1}_{\{j=2,4\}}(i)k_R\right| \psi(s_j)}{2I(k_i^t,t;\mathrm{s})\sqrt{t(s_j^2 + s_j)}},$$

for all $j = 1,\dots 4$, $i = 1,\dots,m(t)$ and $t \in \mathcal{T}$.

As discussed in Benaim and Friz (2008, Section 6.3) the MGF of $\log(S_t)$ in the Heston model satisfies Lemma 2.5 and therefore the slope of the total implied variance for a fixed maturity t as k tends to infinity becomes equal to $\psi(p)$ where the parameter p can be calculated as a root of a non-linear equation

$$
\begin{aligned}
&(\kappa-\rho\xi p)^2 \\
&\quad + \left(\xi^2(p^2-p)-(\kappa-\rho\xi p)^2\right)^{1/2}\cot\left(\frac{\left(\xi^2(p^2-p)-(\kappa-\rho\xi p)^2\right)^{1/2}t}{2}\right)=0.
\end{aligned}
\tag{4.13}
$$

We can also use the above equation to calculate the slope of the left wing of a slice of the total implied variance for a fixed maturity t and k tends to negative infinity. First we note that $1/S$ still follows the stochastic differential equation (4.10) with amended parameters. In particular let $X_t := \log(S_t)$ and $Y_t = -X_t$. Itô's lemma implies that

$$\mathrm{d}X_t = -\frac{V_t}{2}\mathrm{d}t + \sqrt{V_t}\mathrm{d}W_t,$$

and

$$\mathrm{d}Y_t = \frac{V_t}{2}\mathrm{d}t + \sqrt{V_t}\mathrm{d}B_t,$$

where $\mathrm{d}B_t := \sqrt{V_t}\mathrm{d}t - \mathrm{d}W_t$ is the Brownian motion with drift. Also note that $Z_t = \rho W_t + \sqrt{1-\rho^2}W_t^1$ where W and W^1 are two independent Brownian motions. It follows that

$$\mathrm{d}Z_t = \rho\left(\sqrt{V_t}\mathrm{d}t - \mathrm{d}B_t\right) + \sqrt{1-\rho^2}W_t^1 = \rho\sqrt{V_t}\mathrm{d}t + \mathrm{d}W_t^2,$$

where $W_t^2 = -\rho B_t + \sqrt{1-\rho^2}W_t^1$ and we have the following stochastic differential equation for the instantaneous variance V

$$\mathrm{d}V_t = \tilde{\kappa}\left(\tilde{\theta} - V_t\right)\mathrm{d}t + \xi\sqrt{V_t}\mathrm{d}W_t^2,$$

where $\tilde{\kappa} := \kappa - \rho\xi$ and $\tilde{\theta} := \kappa\theta/(\kappa-\rho\xi)$. Thus the inverse of S follows the Heston stochastic differential equation (4.10) with parameters $\tilde{\kappa}, \tilde{\theta}, \xi > 0$ and $\tilde{\rho} := -\rho \in [-1, 1]$ only if $\kappa > \rho\xi$ which is automatically satisfied as $\rho < 0$ in our case. As the higher moments of $1/S$ are the negative moments of S the parameter q of the slope $\psi(q)$ of the left wing can be calculated as a solution of the non-linear equation (4.13) with parameters $\tilde{\kappa}$ and $\tilde{\rho}$ substituted instead of κ and ρ (see Section 2.1.2 for discussion on connection between moments of the underlying stock price process and slopes of the total implied variance at extreme strikes). Thus we can calculate parameters s using Equation (4.13) and discussion above.

Table 4.3 presents the sets of slopes used to extrapolate the total implied variance for maturities 1 year and 1.5 years respectively. Perturbation sets are numbered for ease of reference and Set 1 corresponds to the unperturbed case. The parameters in this set are calculated by solving non-linear equation (4.13). As discussed in Remark 4.20, other perturbation sets were chosen so that the slices of the total implied variance do not cross. Figures 4.1 and 4.2 show implied volatility for maturities t and $t + \tau$ respectively plotted for different values of perturbation parameters. Circles represent the volatilities that correspond to prices available on the market and squares are the extrapolated values. The top left plot shows the unperturbed implied volatility profile in both figures and subsequent plots are numbered left to right and top to bottom with numbering corresponding to the numbering in Table 4.3.

Tables 4.4 and 4.5 present results of the perturbation analysis for the hedging problems (4.1) and (4.5) respectively. As in the Black-Scholes case discussed in Section 4.3.1 the results are in line with expectations, as the approximation becomes less accurate as value of perturbation parameters deviate from the unperturbed case (presented in the first row of each Table). It also confirms that the perturbation results obtained in Section 4.2 are local in nature.

It must be noted that results in the Black-Scholes and the Heston case imply that the at-the-money Forward-Start Straddle is not very sensitive to errors in extrapolation of the spot total implied variance. It implies that even if the extrapolation is very inaccurate, the price of Forward-Start options close to at-the-money will not vary significantly. These findings also confirm the results obtained in Sections 3.4.2 and 3.4.3, i.e. European options cannot effectively hedge forward volatility claims and instead Forward-Start options should be viewed as input data (when traded liquidly) into calibration of models used for pricing of forward volatility dependent exotics.

Perturbation Set	1	2	3	4	5	6
q_{t_1}	5.058	5.06	5.2	6	10	12
p_{t_1}	24.21	24.22	24.35	25.1	35	37
$\psi(q_{t_1})$	0.0901	0.09011	0.0879	0.077	0.0476	0.04
$\psi(p_{t_1})$	0.0202	0.02022	0.0201	0.0195	0.0141	0.0133
q_{t_2}	6.83	6.84	6.9	7.1	10	12
p_{t_2}	30.714	30.72	30.73	31.1	35	37
$\psi(q_{t_2})$	0.0683	0.0682	0.0676	0.0659	0.0476	0.04
$\psi(p_{t_2})$	0.016	0.01601	0.016	0.0158	0.0141	0.0133

Table 4.3: Perturbation parameters and corresponding total implied variance slopes.

Perturbation set	Derivative	Optimal Value	Est. Value	Abs. Diff.
1	0	0.1616	0.1616	0
2	-1.10E-05	0.1616	0.1616	2.77E-08
3	1.27E-04	0.1617	0.1617	5.16E-06
4	1.05E-03	0.1624	0.1627	2.38E-04
5	3.77E-03	0.1627	0.1654	2.65E-03
6	4.59E-03	0.1625	0.1662	3.69E-03

Table 4.4: Perturbation of the super-hedging primal problem for the at-the-money Forward-Start Straddle in the Heston case.

Perturbation set	Derivative	Optimal Value	Est. Value	Abs. Diff.
1	0	0.04455	0.04455	0
2	2.76E-06	0.04455	0.04455	6.78E-09
3	-3.41E-05	0.04452	0.04451	4.83E-06
4	-2.80E-04	0.04437	0.04427	1.06E-04
5	-1.02E-03	0.04432	0.04353	7.90E-04
6	-1.26E-03	0.04436	0.04329	1.07E-03

Table 4.5: Perturbation of the sub-hedging primal problem for the at-the-money Forward-Start Straddle in the Heston case.

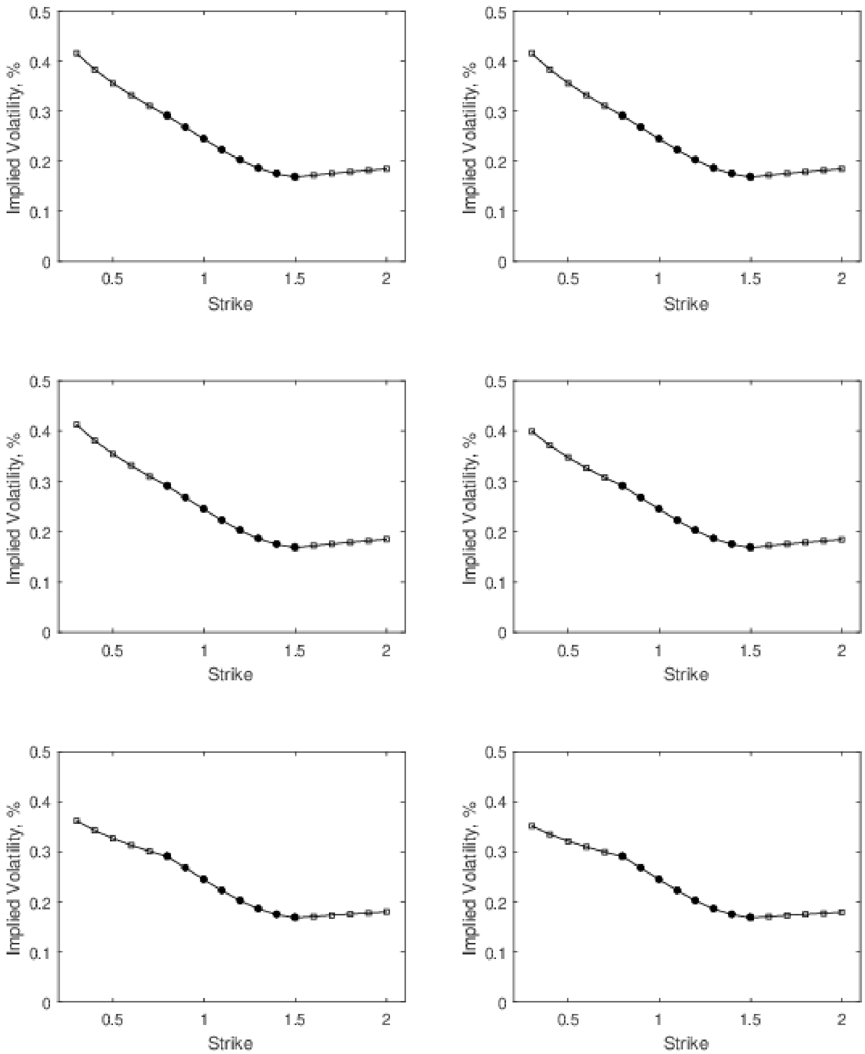


Figure 4.1: Perturbed spot implied volatility profiles for different sets of parameters, $t = 1.0$. Market data is marked with circles, extrapolated implied volatilities with squares. Top row corresponds to perturbation sets 1 (left) and 2 (right), second row - 3 (left) and 4 (right), third row - 5 (left) and 6 (right).

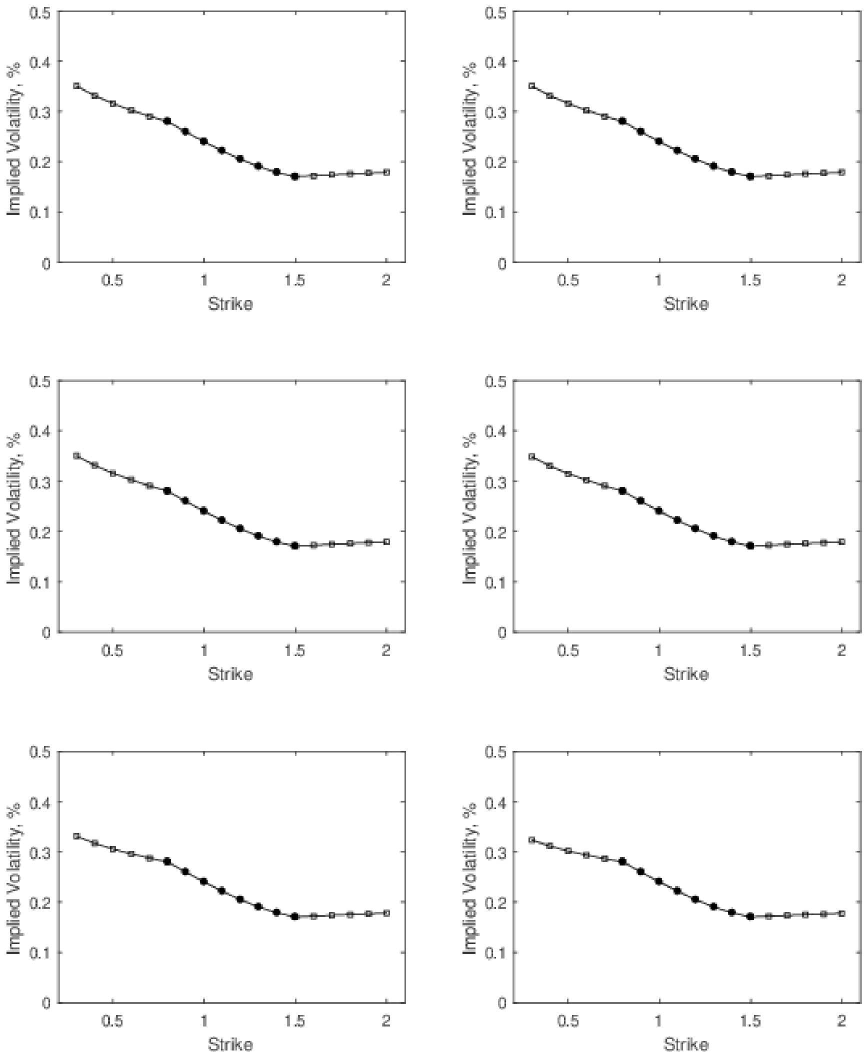


Figure 4.2: Perturbed spot implied volatility profiles for different sets of parameters, $t + \tau = 1.5$. Market data is marked with circles, extrapolated implied volatilities with squares. Top row corresponds to perturbation sets 1 (left) and 2 (right), second row - 3 (left) and 4 (right), third row - 5 (left) and 6 (right).

A

Banach lattices

In this appendix we outline properties of Banach lattices that are used throughout this thesis and we follow closely Aliprantis and Border (2007) in our account.

Definition A.1. (Aliprantis and Border, 2007, Section 8.1) Let X be a vector space endowed with a partial order "$\geq$", i.e. a transitive, reflexive and antisymmetric relation. If for any two elements $x_1, x_2 \in X$ their minimum $x_1 \wedge x_2$ and their maximum $x_1 \vee x_2$ also belong to X, then the space X is a *vector lattice* or a *Riesz space.*

For an element $x \in X$ of a Riesz space its positive part is defined as $x_+ := x \vee 0$ and its negative part is $x_- := (-x) \vee 0$. Then the absolute value of x is defined as $|x| := x_+ + x_-$. We also define a notion of an order unit.

Definition A.2. (Aliprantis and Border, 2007, Section 8.7) An element $u \in X_{++}$ is called an *order unit* if for all $x \in X$ there exists a positive constant $\lambda > 0$ such that $-\lambda u \leq x \leq \lambda u$.

A subset $E \subset X$ of a Riesz space is *solid* if for an element $e \in E$ the following inequality $-e \leq x \leq e$ implies that $x \in E$ (Aliprantis and Border, 2007, Definition 8.13).

Definition A.3. (Aliprantis and Border, 2007, Definition 8.44) A locally convex solid topology τ on a Riesz space X is a locally convex topology where τ has a base of neighbourhoods of zero that are solid.

If the locally convex topology is Hausdorff then we have the following result.

Theorem A.4. *(Aliprantis and Border, 2007, Theorem 8.43) In a locally solid Hausdorff space (X, τ), the positive cone X_+ is τ-closed.*

If the vector space X is norm-complete then it becomes a Banach lattice, an important subset of locally convex topological Riesz spaces.

Definition A.5. (Bichteler, 1998, Section IV.3, Definition 3.2) If a Riesz space X is endowed with a norm $\|\cdot\|_X$ that makes it complete then it is called a *Banach lattice*.

Theorem A.6. *(Aliprantis and Tourky, 2007, Theorem 1.36) Every positive linear functional $f : X \to \mathbb{R}$ on a Banach lattice X is continuous.*

Definition A.7. (Aliprantis and Border, 2007, Section 9.3) A linear operator $\Lambda : X \to Y$ between two Riesz spaces is a *lattice isomorphism* if $\Lambda(x_1 \vee x_2) = \Lambda(x_1) \vee \Lambda(x_2)$ for all $x_1, x_2 \in X$ and is one-to-one. Moreover if X and Y are normed spaces and a linear isomorphism Λ satisfies $\|\Lambda(x)\|_Y = \|x\|_X$ for all $x \in X$, then Λ is a *lattice isometry*.

We finish this section with an example of a Banach lattice that is used throughout this thesis.

Example A.8. *The space of continuous and bounded functions $C_b(X)$ on a topogical space X with the sup-norm $\|\cdot\|_\infty$ is a Banach lattice. The partial order "$\geq$" is defined pointwise on $C_b(X)$, i.e. $f \geq 0$ implies $f(x) \geq 0$ for all $x \in X$. For any two elements $f_1, f_2 \in C_b(X)$ the relation $f_1 \geq f_2$ holds only if $f_1 - f_2 \in (C_b)_+(X)$. The strict binary relation "$>$" is defined such that for any two elements $f_1, f_2 \in C_b(X)$, if $f_1 > f_2$ implies that $f_1 \geq f_2$ and $f_1 - f_2 \neq 0$. The lattice operations on $C_b(X)$ are defined pointwise as $f \vee g(x) := \max\{f(x), g(x)\}$ and $f \wedge g(x) := \min\{f(x), g(x)\}$ respectively. It can be seen that functions $f \vee g$ and $f \wedge g$ also belong to $C_b(X)$ and hence*

$C_b(X)$ is a vector lattice. Moreover $C_b(X)$ is a Banach space endowed with the usual sup-norm $\|f\|_\infty = \sup_{x\in X} |f(x)|$, hence it is a Banach lattice by Definition A.5.

B

Measures on topological spaces

In this appendix we present several results about measures on topological spaces. We also present the link between continuous linear functionals on the space of continuous and bounded functions $C_b(X)$ and measures on the underlying topological space X. We first summarise several topological concepts and results from measure theory.

Definition B.1. (Bogachev, 2007, Section 6.3) Given a topological space X a σ-algebra formed by sets of the form $\{x \in X : f(x) > 0\}$ for all $f \in C_b(X)$ is the *Baire σ-algebra* and is denoted by $\mathcal{B}a(X)$.

Some authors define Baire σ-algebra using $C_c(X)$, the space of continuous functions on X with compact support. However for locally compact metrizable spaces X the definitions agree (Aliprantis and Border, 2007, Section 4.12).

Definition B.2. (Bogachev, 2007, Definition 6.1.2) Let X be a topological space.

(i) X is a *Hausdorff* space if every two distinct points in X have disjoint neighbourhoods;

(ii) A Hausdorff space X is *regular* if, for every $x \in X$ and every closed set $Z \subseteq X$ such that $x \notin Z$ there exist disjoint open sets U and V such that $x \in U$ and $Z \subset V$;

(iii) A Hausdorff space X is *completely regular* if for every $x \in X$ and every closed $Z \subseteq X$ such that $x \notin Z$ there exists a continuous function $f : X \to [0, 1]$ such that $f(x) = 1$ and $f(z) = 0$ for all $z \in Z$;

(iv) A Hausdorff space X is called *normal* if for every disjoint closed sets Z_1 and Z_2 in X there exist disjoint open sets U and V such that $Z_1 \subset U$ and $Z_2 \subset V$;

(v) A Hausdorff space X is called *perfectly normal* if every closed set $Z \subseteq X$ is the pre-image of zero for some continuous function f on X, i.e. $Z = \{x \in X : x = f^{-1}(0) \text{ where } f \in C(X)\}$.

In particular all metric spaces are perfectly normal. Indeed, if we take $x \neq y \in X$, a metric space with a metric d, then there exists two open balls $B_\delta(x)$ and $B_\delta(y)$ where $\delta := 1/2d(x, y)$ such that $B_\delta(x) \cap B_\delta(y) = \emptyset$ and hence the space is Hausdorff. If two A and B are two disjoint closed sets in X then $A_1 := \{x \in X : \inf_{y\in A} d(x, A) < \inf_{y\in B} d(x, B)\}$ and $B_1 := \{x \in X : \inf_{y\in B} d(x, B) < \inf_{y\in A} d(x, A)\}$ are their open disjoint neighbourhoods. Hence (X, d) is normal. Then it follows from Urysohn's Lemma (Dunford and Schwartz, 1957, Theorem 2, Chapter 1, Section 5) that it is a perfectly normal space.

Proposition B.3. *(Bogachev, 2007, Proposition 6.3.4) Let X be a perfectly normal space. Then the Borel σ-algebra $\mathcal{B}(X)$ over X is equal to $\mathcal{B}a(X)$.*

Proof. Since $\mathcal{B}a(X)$ is the smallest σ-algebra such that all continuous functions are measurable with respect to it implies that $\mathcal{B}a(X) \subseteq \mathcal{B}(X)$ as continuous functions between topological spaces are Borel measurable.

On the other hand let $B \in \mathcal{B}(X)$ be a closed subset of X. As X is perfectly normal $B = \{x \in X : g(x) = 0, \text{ for some } g \in C(X)\}$. In particular, let the continuous function $g := \phi \circ f : X \to \mathbb{R}$ be such that $g(x) := \inf_{y \in f(B)} |f(x) - y|$ for all $x \in X$ and $f : X \to \mathbb{R}$ is a continuous function. Then $g(x) = 0$ for all $x \in B$ and $g(x) > 0$ for all $x \in X \setminus B$. Then it follows that $B \in \mathcal{B}a(X)$ and hence $\mathcal{B}(X) \subseteq \mathcal{B}a(X)$. $\square$

In particular all Baire measures on a perfectly normal space X are Borel. Before we proceed let us define the total variation of a measure.

Definition B.4. (Aliprantis and Border, 2007, Section 10.10) Let $\mathcal{A}$ be a algebra of subsets of X. For any set $A \subset X$ we define a *total variation* of a measure μ on $(X, \mathcal{A})$ as

$$|\mu|(A) := \sup \left\{ \sum_{i=1}^{n} |\mu(A_i)| : (A_i)_{i=1,\dots,n} \subset \mathcal{A} \text{ and } A = \bigcup_{i=1}^{n} A_i \right\}, \tag{B.1}$$

where $(A_i)_{i=1,\dots,n}$ is a finite partition of A. We say that μ is of bounded variation if $|\mu|(X) < \infty$ and we may also refer to a measure of bounded variation as finite. Moreover denote by $\|\mu\| := |\mu|(X)$ the total variation norm of μ.

The above definition can be equivalently stated for a σ-algebra, in which case a finite partition of a set A in (B.1) is replaced by a countable partition. It follows from the definition that a measure μ is finite when its total variation $\|\mu\|$ is finite.

Definition B.5. A σ-additive measure on a σ-algebra $\mathcal{A}$ is called *tight* if for every $A \in \mathcal{A}$ and $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset A$ such that $|\mu|(A \setminus K_\varepsilon) < \varepsilon$.

If we now restrict the space X further Borel measures posses further properties as presented in the next result.

Theorem B.6. *(Bogachev, 2007, Theorem 7.1.7) Let X be a metric space. Then every Borel measure μ is on X is regular. If X is complete and separable, then every Borel measure μ is tight.*

We now present the result that shows correspondence between linear functionals on $C_b(X)$ and Baire measures.

Theorem B.7. *(Bogachev, 2007, Theorem 7.10.1) Let X be a topological space. The formula*

$$\Lambda(f) = \int_X f(x)\mu(\mathrm{d}x) \tag{B.2}$$

establishes a one-to-one correspondence between signed Baire measures μ of bounded variation on X and continuous linear functionals Λ on $C_b(X)$ with the following property:

$$\lim_{n\to\infty} \Lambda(f_n) = 0, \tag{B.3}$$

for every sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(X)$ pointwise decreasing to zero.

Corollary B.8. *Let X be a Polish space. The formula* (B.2) *establishes a one-to-one correspondence between tight signed Borel measures μ of bounded variation and continuous linear functionals Λ on $C_b(X)$ possessing property* (B.3)*.*

Proof. By Theorem B.7 we have a one-to-one correspondence between signed Baire measures of bounded variation and linear functionals on $C_b(X)$ that satisfy property (B.3). As X is Polish, it is a metric space and hence by discussion above it is perfectly normal. Therefore all Baire measures on X are Borel by Proposition B.3. Moreover X is complete and separable, hence all Borel measures are tight by Theorem B.6. $\square$

In particular positive continuous linear functionals on $C_b(X)$ possessing property (B.3) can be represented by non-negative Baire measures of bounded variation. When X is a Polish space, the representation holds with non-negative tight Borel measures of bounded variation by above. Moreover rescaling a non-negative tight Borel measure μ by its total variation $\|\mu\|$ yields a Borel probability measure on X and the space of Borel probability measures spans the space of signed Borel measures of bounded variation (Aliprantis and Border, 2007, Chapter 15).

We now restrict our attention to X being a Polish space. It is well-known that the space of signed Borel measures $\mathcal{M}(X)$ endowed with the weak* topology has the topological dual that can be identified with $C_b(X)$ (Deuschel and Stroock, 2001, Lemma 3.2.3). In fact it can be shown that the topological dual of $\mathcal{M}_h(X)$, the space of signed Borel measures that integrate a positive function h (where h satisfies Assumption 1) to a finite constant, is the weighted space $C_h(X)$, see Guo, Tan, and Touzi (2016, Theorem A.2) for the case when h is linear. Consider the positive cone $(\mathcal{M}_h)_+(X)$. Its negative polar is by definition a set in $C_h(X)$

$$(\mathcal{M}_h)^*_+(X) := \{f \in C_h(X) : \langle f, \mu \rangle \geq 0 \text{ for all } \mu \in (\mathcal{M}_h)_+(X)\}, \tag{B.4}$$

where the bilinear form $\langle \cdot, \cdot \rangle : C_h(X) \times \mathcal{M}_h(X) \to \mathbb{R}$ is defined

$$\langle f, \mu \rangle := \int_X f(x)\mu(\mathrm{d}x),$$

for all $f \in C_h(X)$ and $\mu \in \mathcal{M}_h(X)$.

Lemma B.9. *The negative polar defined in* (B.4) *is equal to the positive cone* $(C_h)_+(X)$.

Proof. It is obvious that $(C_h)_+(X) \subset (\mathcal{M}_h)^*_+(X)$. For the reverse inclusion, note that Dirac measures δ_x are elements of $(\mathcal{M}_h)_+(X)$ for all $x \in X$. □

We now show that the positive cone $(\mathcal{M}_h)_+(X)$ is closed in the weak* topology and therefore by Bipolar Theorem (Aliprantis and Border, 2007, Theorem 5.103) the negative polar of the cone $(C_h)_+(X)$ is precisely $(\mathcal{M}_h)_+(X)$. Also note that re-scaling each $\mu \in (\mathcal{M}_h)_+(X)$ by its total variation $\|\mu\|$ yields $\mathcal{P}_h(X)$, the space of Borel probability measures that integrate h to a finite constant.

Lemma B.10. *The positive cone* $(\mathcal{M}_h)_+(X)$ *is closed in the weak* topology.*

Proof. For any $\mu \in (\mathcal{M}_h)_+(X)$ let us define a non-negative Borel measure $\tilde{\mu}$ as

$$\frac{\mathrm{d}\mu}{\mathrm{d}\tilde{\mu}} := \frac{\mathcal{C}}{h}, \tag{B.5}$$

where $\mathcal{C} := \left(\int_\Omega 1/h(\omega)\tilde{\mu}(\mathrm{d}\omega)\right)^{-1} < \infty$ as h satisfies Assumption 1. Then for any $f \in C_h(X)$ we have

$$\int_X f(x)\mu(\mathrm{d}x) = \int_X \frac{f}{h}(x)\mathcal{C}\tilde{\mu}(\mathrm{d}x) = \int_X g(x)\tilde{\mu}(\mathrm{d}x),$$

where $g := \mathcal{C}f/h$ and $g \in C_b(X)$.

Define a norm on the space of signed Borel measures of bounded variation $\mathcal{M}(X)$ as (Bogachev, 2007, Section 8.3)

$$\|\mu\|_0 = \sup\left\{\int_X f(x)\mu(\mathrm{d}x) : f \in \mathrm{Lip}_1(X),\ \sup_{x\in X}|f(x)| \leq 1\right\},$$

where

$$\mathrm{Lip}_1(X) := \{f : X \to \mathbb{R},\ |f(x) - f(y)| \leq d(x,y),\ \text{for all } x, y \in X\}.$$

As X is a Polish space, this norm generates a topology that coincides with the weak* topology on $\mathcal{M}_+(X)$ (Bogachev, 2007, Theorem 8.3.2). Moreover the space $\mathcal{M}_+(X)$ is Polish (Bogachev, 2007, Theorem 8.9.4) and hence it is closed in the weak* topology. Therefore for any sequence $(\tilde{\mu}_n)_{n\in\mathbb{N}} \subset \mathcal{M}_+(X)$ converging in the weak* topology to $\tilde{\mu}_0 \in \mathcal{M}_+(X)$ we have

$$\int_X g(x)\tilde{\mu}_0(\mathrm{d}x) = \lim_{n\to\infty}\int_X g(x)\tilde{\mu}_n(\mathrm{d}x),$$

and using (B.5) we can construct a sequence $(\mu_n)_{n\in\mathbb{N}} \subset (\mathcal{M}_h)_+(X)$ converging in the weak* topology to $\mu_0 \in (\mathcal{M}_h)_+(X)$. $\square$

C

The Stone-Čech Compactification

In this appendix we summarise the Stone-Čech compactification of a completely regular Hausdorff space X. Note that throughout the thesis we apply the results of this chapter to locally compact Polish spaces which are metrizable and perfectly normal by (Aliprantis and Border, 2007, Lemma 3.21), hence completely regular. We follow closely arguments presented in (Kelley, 1975, Section 5.24) and (Aliprantis and Border, 2007, Section 2.17).

Let X be a completely regular Hausdorff space. Let $C_b(X)$ be the space of continuous and bounded function on X. For each $f \in C_b(X)$ there exists a real constant $K_f > 0$ such that $|f(x)| \leq K_f$ for all $x \in X$. We can define the evaluation mapping $e : X \to \mathbb{R}^{C_b(X)}$ by $e(x) = (f(x))_{f \in C_b(X)}$ as the evaluation of each $f \in C_b(X)$ at $x \in X$. The image of X via the evaluation map $e(X)$ can then be represented as a subspace of $\prod_{f \in C_b(X)}[-K_f, K_f]$ and by Tychonoff's Theorem (Aliprantis and Border, 2007, Theorem 2.61) this product is a compact subset of $\mathbb{R}^{C_b(X)}$. Thus $\overline{e(X)}$ is itself a compact subset of $\mathbb{R}^{C_b(X)}$ and is called the Stone-Čech compactification of X which we denote by $\breve{X}$. The following result summarises this construction.

Theorem C.1. *(Aliprantis and Border, 2007, Theorem 2.79) Let X be as above. If Y is a compact Hausdorff space and $g : X \to Y$ is a continuous mapping, then g extends uniquely to a continuous mapping from the Stone-Čech compactification $\breve{X}$ to Y.*

The Stone-Čech compactification is unique as shown in the next result.

Corollary C.2. *(Aliprantis and Border, 2007, Corollary 2.80) Let C be a compactification of X and suppose that if Y is a compact Hausdorff space and $g : X \to Y$ is continuous, then g has a unique continuous extension from C to Y. Then C is homeomorphic to $\breve{X}$.*

Stone-Čech compactification $\breve{X}$ can be thought of as the union of the space X and the set of all limit points of all the nets in X such that every $f \in C_b(X)$ has a continuous extension to $\breve{X}$ as presented in the next result.

Corollary C.3. *(Aliprantis and Border, 2007, Corollary 2.81) Let C be a compactification of X and suppose that every $f \in C_b(X)$ has a unique continuous extension from X to C. Then C is homeomorphic to $\breve{X}$.*

Let us denote by $\breve{f}$ the unique extension of any $f \in C_b(X)$ to $\breve{X}$. As we assume throughout this appendix that X is a completely regular Hausdorff space we have the following result.

Corollary C.4. *(Aliprantis and Border, 2007, Corollary 2.84) Let X be as above. The mapping $f \mapsto \breve{f}$ is a lattice isometry (see Definition A.7) from $C_b(X)$ onto $C(\breve{X})$, i.e. $C_b(X) = C(\breve{X})$.*

It must be noted that the compactification $\breve{X}$ is not metrizable in general (Willard, 1970, Theorem 24.13).

References

Acciaio, B., M. Beiglböck, F. Penkner, and W. Schachermayer (2016). A Model-Free Version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem. *Mathematical Finance 26*(2), 233–251.

Aksamit, A., Z. Hou, and J. Obłój (2016, May). Robust Framework for Quantifying the Value of Information in Pricing and Hedging. *arXiv:1605.02539v1*.

Albrecher, H., P. A. Mayer, and W. Schoutens (2008). General Lower Bounds for Arithmetic Asian Option Prices. *Applied Mathematical Finance 15*(2), 123–149.

Aliprantis, C. D. and K. C. Border (2007). *Infinite-Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Aliprantis, C. D. and R. Tourky (2007). *Cones and Duality*. Graduate studies in mathematics. American Mathematical Society.

Alvarez, A., S. Ferrando, and P. Olivares (2013). Arbitrage and Hedging in a Non-Probabilistic Framework. *Mathematics and Financial Economics 7*(1), 1–28.

Anderson, E. and P. Nash (1987). *Linear Programming in Infinite-Dimensional Spaces*. New York: John Wiley & Sons.

Andrews, B. and C. Hopper (2011). *The Ricci Flow in Riemannian Geometry*. Lecture Notes in Mathematics. Springer Berlin Heidelberg.

Avellaneda, M., A. Levy, and A. Parás (1995). Pricing And Hedging Derivative Securities In Markets With Uncertain Volatilities. *Applied Mathematical Finance 2*, 73–88.

Badikov, S., A. Jacquier, D. Q. Liu, and P. Roome (2017). No-Arbitrage Bounds for the Forward Smile Given Marginals. *Quantitative Finance 17*(8), 1243–1256.

Baker, D. (2012). *Martingales with Specified Marginals*. Ph. D. thesis, Paris VI.

Bartl, D., M. Kupper, D. J. Prömel, and L. Tangpi (2017, May). Duality for Pathwise Superhedging in Continuous Time. *arXiv:1705.02933*.

Beiglböck, M., P. Henry-Labordère, and F. Penkner (2013, July). Model-Independent Bounds for Option Prices: a Mass Transport Approach. *Finance and Stochastics 17*(3), 477–501.

Beiglböck, M., M. Nutz, and N. Touzi (2016). Complete Duality for Martingale Optimal Transport on the Line. *arXiv:1507.00671v3*.

Benaim, S. and P. Friz (2008, 03). Smile Asymptotics II: Models with Known Moment Generating Functions. *J. Appl. Probab. 45*(1), 16–32.

Benaim, S. and P. Friz (2009). Regular Variation and Smile Asymptotics. *Mathematical Finance 19*(1), 1–12.

Benaim, S., P. Friz, and R. Lee (2008). *On Black-Scholes Implied Volatility at Extreme Strikes*, pp. 19–45. John Wiley & Sons, Inc.

Benyamini, Y. and J. Lindenstrauss (1998). *Geometric Nonlinear Functional Analysis*, Volume 1 of *American Mathematical Society colloquium publications*. American Mathematical Soc.

Bertsimas, D. and I. Popescu (2002). On the Relation Between Option and Stock Prices: A Convex Optimization Approach. *Operations Research 50*(2), 358–374.

Bichteler, K. (1998). *Integration - A Functional Approach*. Birkhäuser Advanced Texts Basler Lehrbücher. Springer Basel.

Bogachev, V. I. (2007). *Measure Theory*. Springer Berlin Heidelberg.

Bonnans, J. F. and A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research and Financial Engineering. Springer New York.

Borwein, J., R. Choksi, and P. Maréchal (2003). Probability Distributions of Assets Inferred from Option Prices via the Principle of Maximum Entropy. *SIAM Journal on Optimization 14*(2), 464–478.

Borwein, J. and A. S. Lewis (1991). Duality Relationships for Entropy-Like Minimization Problems. *SIAM Journal of Control and Optimization 29*(2), 325–338.

Bouchard, B. and M. Nutz (2015). Arbitrage and Duality in Nondominated Discrete-Time Models. *Annals of Applied Probability 25*(2), 823–859.

Boyle, P. P. and X. S. Lin (1997). Bounds on Contingent Claims Based on Several Assets. *Journal of Financial Economics 46*(3), 383–400.

Breeden, D. T. and R. H. Litzenberger (1978, October). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business 51*(4), 621–651.

Brown, H., D. Hobson, and L. Rogers (2001). Robust Hedging of Barrier Options. *Mathematical Finance 11*(3), 285–314.

Bühler, H. (2002). Applying Stochastic Volatility Models for Pricing and Hedging Derivatives. *www.quantitative-research.de/dl/021118SV.pdf*. published online.

Burzoni, M., M. Frittelli, and M. Maggis (2016). Universal Arbitrage Aggregator in Discrete-Time Markets under Uncertainty. *Finance and Stochastics 20*(1), 1–50.

Campi, L., I. Laachir, and C. Martini (2017, April). Change of Numeraire in the Two-Marginals Martingale Transport Problem. *Finance and Stochastics 21*(2), 471–486.

Cassese, G. (2008). Asset Pricing with no Exogenous Probability Measure.

Chen, X., G. Deelstra, J. Dhaene, and M. Vanmaele (2008). Static Super-Replication for a Class of Exotic Options. *Insurance: Mathematics and Insurance 42*, 1067–1085.

Clark, S. A. (1993). The Valuation Problem in Arbitrage Price Theory. *Journal of Mathematical Economics 22*(5), 463–478.

Clark, S. A. (2000). Arbitrage Approximation Theory. *Journal of Mathematical Economics 33*(2), 167 – 181.

Clark, S. A. (2003, March). An Infinite-Dimensional LP Duality Theorem. *Mathematics of operations research 28*(2), 233–245.

Clark, S. A. (2006). Necessary and Sufficient Conditions for Solving Infinite-Dimensional Linear Inequalities. *Positivity 10*(3), 475–489.

Cousot, L. (2007). Conditions on Option Prices for Absence of Arbitrage and Exact Calibration. *Journal of Banking & Finance 31*(11), 3377–3397.

Cox, A. M. G., Z. Hou, and J. Obłój (2016, Jul). Robust Pricing and Hedging under Trading Restrictions and the Emergence of Local Martingale Models. *Finance and Stochastics 20*(3), 669–704.

Cox, A. M. G. and S. M. Kinsley (2017, February). Robust Hedging of Options on a Leveraged Exchange Traded Fund. *arXiv:1702.07169v1*.

Cox, A. M. G. and J. Obłój (2011a). Robust Hedging of Double Touch Barrier Options. *SIAM J. Financial Math. 2*, 141–182.

Cox, A. M. G. and J. Obłój (2011b). Robust Pricing and Hedging of Double No-Touch Options. *Finance and Stochastics 15*(3), 573–605.

Cox, A. M. G. and J. Wang (2013, 06). Root's Barrier: Construction, Optimality and Applications to Variance Options. *Ann. Appl. Probab. 23*(3), 859–894.

Curran, M. (1994). Valuing Asian and Portfolio Options by Conditioning on the Geometric Mean Price. *Management Science 40*(12), 1705–1711.

d'Aspremont, A. and L. E. Ghaoui (2006). Static Arbitrage Bounds on Basket Option Prices. *Math. Program. 106*(3), 467–489.

Davis, M. H. A. and D. Hobson (2007). The Range of Traded Option Prices. *Mathematical Finance 17*, 1–14.

Davis, M. H. A. and J. Obłój (2008). Market Completion Using Options. *Advances in Mathematics of Finance. Banach Center Publications 83*, 49–60.

Davis, M. H. A., J. Obłój, and V. Raval (2014). Arbitrage Bounds for Prices of Weighted Variance Swaps. *Mathematical Finance 24*(4), 821–854.

De Marco, S. and P. Henry-Labordère (2015). Linking Vanillas and VIX Options: A Constrained Martingale Optimal Transport Problem. *SIAM Journal on Financial Mathematics 6*(1), 1171–1194.

Deelstra, G., J. Liinev, and M. Vanmaele (2004). Pricing of Arithmetic Basket Options by Conditioning. *Insurance: Mathematics and Economics* (34), 55–77.

Delbaen, F. (1992). Representing Martingale Measures when Asset Prices are Continuous and Bounded. *Mathematical Finance 2*(2), 107–130.

Delbaen, F. and W. Schachermayer (1994). A General Version of the Fundamental Theorem of Asset Pricing. *Mathematische Annalen 300*(3), 463–520.

Delbaen, F. and W. Schachermayer (2006). *The Mathematics of Arbitrage.* Springer Berlin Heidelberg.

Denis, L. and C. Martini (2006, 05). A Theoretical Framework for the Pricing of Contingent Claims in the Presence of Model Uncertainty. *Ann. Appl. Probab. 16*(2), 827–852.

Deuschel, J. D. and D. M. Stroock (2001). *Large Deviations.* American Mathematical Society.

Dolinsky, Y. and H. M. Soner (2014). Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields 160*(1-2), 391–427.

Dunford, N. and J. Schwartz (1957). *Linear operators, part 1 : general theory*. Pure and applied mathematics. London, New York, Sydney: Academic press.

Föllmer, H. and Y. Kabanov (1997). Optional Decomposition and Lagrange Multipliers. *Finance and Stochastics 2*, 69–81.

Fouque, J., G. Papanicolaou, R. Sircar, and K. Sølna (2011). *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*. Cambridge University Press.

Galichon, A., P. Henry-Labordère, and N. Touzi (2014). A Stochastic Control Approach to No-Arbitrage Bounds Given Marginals, with an Application to Lookback Options. *Annals of Applied Probability 24*(1), 312–336.

Gatheral, J. (2006). *The Volatility Surface: A Practitioner's Guide*. Wiley Finance. Wiley.

Gatheral, J. and A. Jacquier (2014). Arbitrage-free SVI volatility surface. *Quantitative Finance 14*(1), 59–71.

Goberna, M. and M. López (2014). *Post-Optimal Analysis in Linear Semi-Infinite Optimization*. SpringerBriefs in Optimization. Springer-Verlag New York.

Guo, G., A. Jacquier, C. Martini, and L. Neufcourt (2016). Generalized Arbitrage-Free SVI Volatility Surfaces. *SIAM Journal on Financial Mathematics 7*(1), 619–641.

Guo, G., X. Tan, and N. Touzi (2016). Tightness and Duality of Martingale Transport on the Skorokhod Space. *arXiv:1507.01125v2*.

Guyon, J., R. Menegaux, and M. Nutz (2017). Bounds for VIX Futures Given S&P 500 Smiles. *arXiv:1609.05832v2*.

Harrison, M. and S. Pliska (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications 11*(3), 215–260.

Henry-Labordère, P. (2013). Automated Option Pricing: Numerical Methods. *International Journal of Theoretical and Applied Finance 16*, 135–162.

Henry-Labordère, P. (2017). *Model-Free Hedging: A Martingale Optimal Transport Viewpoint*. Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis Group.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies 6*, 327–343.

Hobson, D. (1998). Robust Hedging of the Lookback Option. *Finance and Stochastics 2*, 329–347.

Hobson, D. (2010). The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices. *Paris-Princeton Lectures on Mathematical Finance*.

Hobson, D. and M. Klimmek (2012). Model-Independent Hedging Strategies for Variance Swaps. *Finance and Stochastics 16*(4), 611–649.

Hobson, D. and M. Klimmek (2014). Robust Price Bounds for the Forward Starting Straddle. *Finance and Stochastics 19*(1), 189–214.

Hobson, D., P. Laurence, and T.-H. Wang (2005). Static Arbitrage Upper Bounds for the Prices of Basket Options. *Quantitative Finance 5*, 329–342.

Hobson, D. and A. Neuberger (2012). Robust Bounds for Forward Start Options. *Mathematical Finance 22*(1), 31–56.

Hodges, S. D. and A. Neuberger (2000, July). Rational Bounds on the Prices of Exotic Options. *SSRN:239711*. EFA Working Paper No. 0280.

Hou, Z. and J. Obłój (2015). On Robust Pricing-Hedging Duality in Continuous Time. *arXiv:1503.02822v2*.

Ibraimi, M., M. Leippold, and F. Stang (2015, May). A Robust Fundamental Theorem of Asset Pricing with Discrete Martingale Measures. *SSRN:2507328*. Preprint.

Jacod, J. and A. N. Shirayev (1998). Local Martingales and the Fundamental Asset Pricing Theorems in the Discrete-Time Case. *Finance and Stochastics* (2), 259–273.

Kantorovich, L. (1942). On Translocation of Masses. *Dokl. Akad Nauk USSR*, 199–201.

Karlin, S. and W. Studden (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley Interscience.

Kelley, J. (1975). *General Topology*. Graduate Texts in Mathematics. Springer New York.

Kramkov, D. (1996). Optional Decomposition of Supermartingales and Hedging Contingent Claims in Incomplete Security Markets. *Probability Theory and Related Fields 105*, 459–479.

Kreps, D. M. (1981). Arbitrage and Equilibrium in Economies with Infinitely Many Commodities. *Journal of Mathematical Economics 8*(1), 15 – 35.

Laurence, P. and T. Wang (2005). Sharp Upper and Lower Bounds for Basket Options. *Applied Mathematical Finance 12*, 253–282.

Leblanc, M. and C. Martini (2000, November). Unbounded Volatility in the Uncertain Volatility Model. *SSRN:1544866*. published online.

Lee, R. W. (2004). The Moment Formula for Implied Volatility at Extreme Strikes. *Mathematical Finance 14*(3), 469–480.

Lo, A. W. (1987). Semi-Parametric Upper Bounds for Option Prices and Expected Payoffs. *Journal of Financial Economics 19*(2), 373–387.

Lyons, T. (1995). Uncertain Volatility and the Risk-Free Synthesis of Derivatives. *Applied Mathematical Finance 2*, 117–133.

Merton, R. C. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science 4*(1), 141–183.

Nadtochiy, S. and J. Obłój (2017). Robust Trading of Implied Skew. *International Journal of Theoretical and Applied Finance 20*(02), 1750008.

Obłój, J. (2004). The Skorokhod Embedding Problem and its Offsprings. *Probability Surveys 1*, 321–390.

Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications* (1st ed.). Springer Publishing Company, Incorporated.

Riedel, F. (2015, April). Financial Economics Without Probabilistic Prior Assumptions. *Decisions in Economics and Finance 38*(1), 75–91.

Rockafellar, R. (1970). *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton Univ. Press.

Rogers, L. C. G. and Z. Shi (1995). The Value of an Asian Option. *Journal of Applied Probability 32*(4), 1077–1088.

Romano, M. and N. Touzi (1997). Contingent Claims and Market Completeness in a Stochastic Volatility Model. *Mathematical Finance 7*(4), 399–412.

Shapiro, A. (2009). Semi-Infinite Programming, Duality, Discretization and Optimality Conditions: Invited Review. *Optimization 58*(2), 133–161.

Skorokhod, A. V. (1965). *Studies in the Theory of Random Processes*. Reading: Addison-Wesley. Translated from Russian by Scripta Technica, Inc.

Stebegg, F. (2014). Model-Independent Pricing of Asian Options via Optimal Martingale Transport. *arXiv:1412.1429*.

Still, G. (2001). Discretization in Semi-Infinite Programming: the Rate of Convergence. *Mathematical Programming 91*(1), 53–69.

Strassen, V. (1965). The Existence of Probability Measures with Given Marginals. *The Annals of Mathematical Statistics 36*(2), 423–439.

Tanaka, K. and A. Toda (2013). Discrete Approximations of Continuous Distributions by Maximum Entropy. *Economic Letters 118*, 445–450.

Willard, S. (1970). *General Topology*. Reading, Massachusetts: Addison Wesley.

Zaev, D. A. (2015). On the Monge-Kantorovich Problem with Additional Linear Constraints. *Mathematical Notes 98*(5), 725–741.

Zălinescu, C. (2002). *Convex Analysis in General Vector Spaces*. World Scientific.

Zălinescu, C. (2008). On Zero Duality Gap and the Farkas Lemma for Conic Programming. *Mathematics of Operations Research 33*(4), 991–1001.