Contributions to model-independent finance via martingale optimal transport

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Daniel Matthias Schmithals, M.Sc.
aus Köln

Tag der mündlichen Prüfung: 12.12.2018

Referentin: Prof. Dr. Nicole Bäuerle

Korreferentin: Prof. Dr. Eva Lütkebohmert-Holtz

Acknowledgement

This PhD thesis has been written during my employment as research and teaching assistant at the Institute for Stochastics at Karlsruhe Institute of Technology (KIT). I am very grateful to a lot of people who supported me on the way.

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Nicole Bäuerle for giving me the opportunity to write this thesis, for all her helpful advice, encouragement and excellent guidance and for always having an open ear. My sincere thanks further go to Prof. Dr. Eva Lütkebohmert-Holtz for acting as second referee for this thesis.

I would like to thank all members of the Institute for Stochastics for creating a great working atmosphere. I thank all the professors for constantly providing possibilities to broaden one's perspective on stochastics both in joint lecture supervision and the regular research seminars. I thank our two secretaries for their endless support regarding the every day administrative tasks. I thank all my former and current colleagues at the Institute for Stochastics. They made the last years an unforgettable and always enjoyable part of my life.

I thank the great people I am lucky to call my friends both in Karlsruhe and Köln for their steady helpfulness and for making bad times more tolerable and good times simply marvellous.

I am and always will be deeply thankful to my family, my parents and my sister, for their never-ending support and unconditional love over all the years.

Last but most certainly not least, I thank my wonderful wife Vanessa for making me happy every single day.

Contents

Chapter 1

Introduction

Ever since in 1973, Black & Scholes [12] and Merton [61] introduced what we call Black-Scholes model today, pricing of exotic options using financial market models is an active research area in mathematical finance. Over the years, researchers and practitioners developed a still growing set of increasingly sophisticated option pricing models, such as diffusion models, jump-diffusion models [56, 57, 62], pure jump models [20, 60], local volatility models [23, 27], stochastic volatility models with and without jumps in the underlying and the volatility [2, 36, 40], and Lévy models with stochastic time [16, 70].

All models of this era work similar in principle. They fix a certain underlying probability space and assume that the random future behaviour of the underlying asset price process is specified somehow, for example as the solution of a stochastic differential equation. Further assuming no-arbitrage and completeness of the considered financial market, a unique equivalent martingale measure, i.e. a probability measure such that the discounted asset price process is a martingale, exists by the fundamental theorem of asset pricing. Then using the law of one price, it is possible to derive the uniquely determined price of an exotic option written on the underlying by calculating either the expected discounted payoff of the exotic option with respect to the equivalent martingale measure or the price of a self-financing, replicating hedging strategy. Most of the models are of parametric form, where the parameters are determined by calibrating the model to observable market prices of certain options. We refer to this model-based approach as classical mathematical finance and remark that the connection of pricing and hedging of exotic options, which we call pricing-hedging duality, is a revolving observation throughout mathematical finance in general.

In this thesis however, we pursue a different approach that we refer to as robust or model-independent (mathematical) finance. This approach is of growing interest since the seminal article of Hobson [41] on the lookback option from 1998 and the pioneering work of Beiglböck, Henry-Labordère & Penkner [5] connecting model-independent finance and optimal transport from 2013. Before we introduce the model-independent approach and the techniques used in the two mentioned papers in detail, we shortly discuss reasons to complement the classical methods in the first place.

For this purpose, we follow a categorization of model input into three different categories by Obłój [66]. That is, any model input may be assigned to either of the categories beliefs, information and rules. Classical models mostly rely on strong assumptions on the financial market that have to be considered as beliefs, while observable market prices of liquidly traded options that count as information are used for calibration purposes only. This leads to two important drawbacks.

Stressing that the assumptions are simplifying and quite regularly unrealistic, we recognize that the model prices are very likely to be inaccurate and unreliable. Indeed, various studies, see for example Schoutens, Simons & Tistaert [71], observe a great range of option prices when calibrating several different models to the same underlying market.

Furthermore, not using the information of observable option prices for more than just calibration may lead to various inconsistencies of the model. The best known inconsistency, that serves as an example here, is the so-called volatility smile as observed in the Black-Scholes model. The inconsistency is that calibrating the Black-Scholes model to observable prices of call options with different strike prices, the associated model prices differ from each other and also from most of the market prices.

The lack of reliability and consistency that appears in classical mathematical finance is addressed by the model-independent approach. The general recipe in terms of Obłój's categories is rather simple: Fewer beliefs, more information.

The first demand of this recipe is satisfied by omitting all assumptions on the dynamics of the underlying asset price process as well as the completeness. In fact, the only remaining assumption is that of no-arbitrage such that the existence of martingale measures is ensured.

In order to satisfy the second demand, we have to use observable market prices of certain options and the desired consistency with these prices. Then we may derive restricting conditions on the martingale measures that could potentially be used for the pricing of exotic options. The historical development of financial markets favours this approach. Indeed, Black, Scholes and Merton had to consider European call options as exotic options, the value of which is derived from the value of the underlying asset. However, over the years, trading of such options became so liquid that Dupire [26, 27] and others argued that they should rather be considered as contingent claims with exogenously fixed prices. Thus, the prices of European call options became available as information for pricing other, more complicated exotic options. In model-independent finance, this is used under the idealizing assumption that for the maturities of interest, the call option prices are observable for a continuum of strike prices.

A very good intuition how the prices may be used is given by Hobson [42], who states that „ideally we should use a model which calibrates perfectly to the full spectrum of traded calls. However, in principle there are many such models, and associated with each model which is consistent with the market prices of liquidly traded options, there may be a different price for the exotic. Instead, one might attempt to characterise the class of models which are consistent with the market prices of options. This is a very challenging problem, and a less ambitious target is to characterise the extremal elements of this set, and especially those models for which the price of the exotic is maximised or minimised.“ An alternative motivation comes from Cox & Obłój [17] who “want to answer two questions. First, for a given exotic option, what is the range of prices that we can charge for it without introducing a model-independent arbitrage? Second, if we see a price outside this range, how do we exploit it to make a riskless profit?”

Restating these motivating citations, we observe that the target of model-independent finance is to price exotic options such that the prices satisfy no-arbitrage and are consistent with certain observable call option prices. Therefore, no specific martingale measure but a set of different consistent martingale measures emerges from the analysis. Thus, no unique option price but a range of possible option prices may be derived. In return, the risk of model misspecification is eliminated.

This intuition was first formalized by Hobson in his famous paper [41], where he used the no-arbitrage assumption and the knowledge about the call option prices to derive upper and lower price bounds for the lookback option in continuous time. For this, two tools are crucial.

The first tool is the lemma of Breeden & Litzenberger [13, Sec. 2] from 1978. It states that the distribution of the price of some asset at a certain time, i.e. the marginal distribution of the asset price process at that time, may be inferred from the prices of call options for different strike prices on the same underlying and the same maturity. Combining this marginal condition with the usual martingale condition, Hobson derived the set of all potential pricing measures over which then the expected discounted payoff of the exotic option is maximized and minimized.

The second tool is the theorem of Dambis, Dubins & Schwarz, see for example Revuz & Yor [69, Chap. V, Theorem 1.6], that enabled Hobson to translate the pricing problems into Skorokhod-type stopping problems. That is, Hobson translated the continuous-time martingale, which the underlying asset price process is, into a Brownian motion and optimized over all stopping times such that the stopped Brownian motion has the previously inferred marginal distribution.

Furthermore, Hobson derived sub- and super-replicating hedging strategies that help to exploit arbitrage possibilities, if there are any.

The connection of model-independent finance and the Skorokhod embedding problem is quite nice, as the latter is well-studied. A great survey on the Skorokhod embedding problem was published by Obłój [66], providing various solutions that were known at the time. Following the ideas of Hobson, a great variety of researchers investigated several different exotic options using similar methods, see for example [14, 15, 17, 18, 19, 43, 44, 45, 46, 47]. A comprehensive survey on the connection between model-independent finance and the

Skorokhod embedding problem is provided by Hobson [42].

All articles following this approach do not only use the same techniques but also provide the same type of results. That is, for a particular path dependent exotic option, in each article upper and lower price bounds as well as super- and sub-replicating hedging strategies are derived using previously known solutions to the Skorokhod embedding problem. We can think of such results as pricing-hedging dualities for very specific pairs of optimization problems. The fact that this duality results apply for a wide range of exotic options attracted the search for more general duality results.

Such results were finally established in 2013 by Galichon, Henry-Labordère & Touzi [29] using stochastic optimal control theory in the continuous-time case and by Beiglböck, Henry-Labordère & Penkner [5] using methods from optimal transport in the discrete-time case. While the former actually generalizes the results derived using the Skorokhod embedding problem, as both is done in continuous time, the latter is the approach that we further pursue in this thesis. Before we discuss the approach in more detail, we shortly discuss optimal transport.

In optimal transport, the problem is to minimize the cost that transportation of mass from one point to another generates in the sense that a cost-minimal transport allocation is aimed for. Mathematically, we may specify the mass at the origins and the destinations by measures. Then, minimizing the transport cost is equivalent to minimizing the integral over a usually two-variate function representing the cost of transporting a unit of mass from one point to another with respect to the set of all couplings or so-called transport plans which have the specified measures as marginals. The problem was originally introduced by Monge [63] in 1781 and then refined by Kantorovich [53, 54] in 1948. A great variety of researchers considered the optimal transport problem and in the course of their research many important results on optimal transport were established, see for example Rachev & Rüschendorf [67, 68] or Villani [77] for excellent monographes on the topic.

Observing that there is an analogy between model-independent finance and optimal transport, as in both areas the marginals of the distribution over which some function is optimized are specified, Beiglböck, Henry-Labordère & Penkner [5] introduced a new research field that we refer to as martingale optimal transport. Re-interpreting the transport cost function as the payoff function of an exotic option and implementing the usual martingale condition of mathematical finance, the minimization problem of optimal transport cost evolves to the lower price bound problem of model-independent finance. Also, properly implementing the martingale condition in the dual problem of optimal transport, a pricing-hedging duality is shown using only the usual assumptions of model-independent finance, no-arbitrage and consistency with call options prices.

Here, the authors use three key tools. The first tool is again the lemma of Breeden & Litzenberger enabling the authors to infer the marginal distributions of all potential pricing measures. The second tool, the theorem of Strassen [73], is necessary to guarantee the well-posedness of the lower price bound problem, as it implies assumptions on the marginals which guarantee that pricing measures satisfying both the martingale and the marginal conditions do exist. Finally, the third tool, which rather is a toolbox, is the set of all ideas, techniques and results from optimal transportation theory.

Using the connections made by Hobson [41] and Beiglböck, Henry-Labordère & Penkner [5], a stream of articles emerged building up an immediate connection between the Skorokhod embedding problem and martingale optimal transport that enabled the authors to bring great progress to both areas using the methods and results from each other, see for example [3, 4, 6, 31, 33, 34, 50, 52]. Following the continuous-time approach, a variety of results emerged in this setting, see for example [24, 25, 35, 49, 75].

Finally, many researchers closely followed the ideas of Beiglböck, Henry-Labordère & Penkner [5], answering questions brought up by them about the existence of dual optimizers, see for example Beiglböck, Nutz & Touzi [9], about the structure of solutions to the pricing problem, see for example Beiglböck & Juillet [7], and about the improvement of the price bounds, see for example Lütkebohmert & Sester [59]. Other articles posed and answered new questions and generalized earlier results, see for example [6, 8, 22, 30, 32, 38, 39, 48, 51, 58, 65] and the recent book on martingale optimal transport by Henry-Labordère [37].

The present thesis shall join this type of work, as we generalize various earlier results and introduce different algorithmic tools to derive and approximate solutions to the price bound and hedging problems.

This thesis is build up as follows. In Chapter 2, we setup the notation and introduce basic notions from analysis and measure theory that we need in this thesis. In Section 2.1, we discuss several continuity and differentiability properties of real-valued functions. Most importantly, we present a characterization result for bounded, semi-continuous functions. In Section 2.2, we introduce the measure theoretic terminology. Starting from the types of measures we consider, we recall the usual notions up to couplings and marginals. Finally, we discuss different convergence types for sequences of measures. In particular, we introduce the Wasserstein distance and state a representation result that is useful to calculate Wasserstein distances explicitly.

In Chapter 3, we recall the principles of model-independent finance discussed in the introduction. In Section 3.1, we introduce two financial markets incorporating the essentials of model-independent finance and thus provide the mathematical framework for this thesis. In Section 3.2, we take a closer look at European call options. In particular, we discuss several important no-arbitrage properties of the prices of European call options that we regularly apply throughout this thesis.

In Chapter 4, we detail the connection between model-independent finance and classic optimal transport, thus explaining the notion of martingale optimal transport. In Section 4.1, we discuss classic optimal transport in some detail. We present various well-known results of this area that motivate recent research questions in martingale optimal transport. In Section 4.2, we present the lemma of Breeden & Litzenberger. This lemma connects model-independent finance and classic optimal transport using the assumptions made on observable call option prices. Then we define the set of martingale transport plans as an adaption of classic transport plans. This set builds the foundation of martingale optimal transport. We complement the definition of martingale transport plans in Section 4.3, where we introduce the concepts of convex order, potential functions and call option price functions in order to guarantee the existence of martingale transport plans using the theorem of Strassen. Thus, in Section 4.4, we may finally introduce the primal problems of martingale optimal transport in a well-defined manner. These problems allow an interpretation as upper and lower price bound problems in model-independent finance. The dual problems of martingale optimal transport, which allow an interpretation as super and sub hedging problems, are introduced in Section 4.5. The first duality result of martingale optimal transport derived by Beiglböck, Henry-Labordère & Penkner [5] is presented in Section 4.6. This result allows an interpretation as a pricing-hedging duality for exotic options written on a single underlying asset.

In Chapter 5, we begin to derive new results on model-independent finance via martingale optimal transport. In Section 5.1, we use the duality result of Beiglböck, Henry-Labordère & Penkner [5] as a guideline to prove a similar pricing-hedging duality result for the general case of exotic options written on more than one underlying asset. We further discuss some drawbacks of the general theory of martingale optimal transport concerning the existence of optimizers to the dual problem and the extent of the gap between the lower and upper price bounds. In Section 5.2, we present some conditions under which dual optimizers, i.e. optimal hedging strategies, do exist. In particular, we present a result of Beiglböck, Lim & Obłój [8] that implies certain Lipschitz properties for the optimizers. In Section 5.3, we present improvements of the price bounds based on additional market information on the asset return variances in the single asset case as derived by Lütkebohmert & Sester [59]. Then we discuss possible generalizations to the multi-asset case using the entire asset return covariance structure.

In Chapter 6, we consider the optimization problems separately and in a simpler setting in order to establish structural conditions under which the problems may be solved explicitly. In Sections 6.1 and 6.2, we present results of Beiglböck & Juillet [7] and of Henry-Labordère & Touzi [38]. We introduce the notion of monotonicity, a structural property of martingale transport plans. One of the results of Beiglböck & Juillet [7] implies optimality of monotone martingale transport plans for the pricing problems considering general underlying marginals and a certain type of payoff functions. We further introduce the martingale Spence Mirrlees condition, a structural property of payoff functions. Similar to the afore mentioned optimality statement, a result of Henry-Labordère & Touzi [38] says that monotone martingale transport plans are optimal for the pricing problems considering continuous marginals and payoff functions satisfying the martingale Spence Mirrlees condition. In this case, we also present algorithmic methods of Henry-Labordère & Touzi [38] that provide solutions to both the pricing and the hedging problems, thus recovering a special case of the pricing-hedging duality. In Section 6.3, we generalize the afore mentioned optimality results proving that monotone martingale transport plans are optimal for the pricing problems considering general marginals and payoff functions that satisfy the martingale Spence Mirrlees condition, thus unifying the previous results. In Section 6.4, we specialize ourselves to the case of discrete marginals. In this case, we weaken the condition on the payoff functions for the optimality results and we introduce constructive algorithms that provide optimizers for both the pricing and the hedging problems.

In Chapter 7, we consider the results of model-independent finance via martingale optimal transport from an application-oriented point of view. The key motivation is to overcome the idealizing assumption that we could uniquely determine the true asset price process marginals from observable call option prices. This is crucial in application, as we can observe only finitely many call option prices in the market. The general idea to overcome this assumption is to approximate the price bounds. Thus, in this chapter, we discuss convergence issues. In Section 7.1, we evaluate properties of the true call option price function and the associated marginals, which we also refer to as theoretical call option price functions and theoretical marginals, in order to derive plausibility checks for observable call option prices. Then we discuss how the observable prices can be transformed into empirical call option price functions and associated marginals and how those can be used to approximate the true price bounds. In Section 7.2, we investigate the convergence of the approximating price bounds towards the true price bounds. This is done under the assumptions that the payoff function satisfies the martingale Spence Mirrlees condition and that the true marginals have bounded support. In Section 7.3, we introduce an explicit sequence of empirical marginals and quantify the speed of the convergence of the associated empirical price bounds towards the true price bounds in the situation of the previous section. Finally, in Section 7.4, we generalize the convergence and the convergence speed results of the previous sections. We get rid of the assumption that the payoff functions satisfy the martingale Spence Mirrlees condition and we provide similar results for general sequences of empirical marginals. Ultimately, we also overcome the assumption that the true marginals have bounded support, thus providing a rather general result allowing to a numerically approximate the true price bounds in application.

CHAPTER 2

Preliminaries

In this chapter, we introduce the basic notions of analysis and measure theory required in this thesis. We work in the n-dimensional Euclidean space $\mathbb{R}^n$, $n \in \mathbb{N}$, equipped with its usual topology induced by the standard scalar product $\langle \cdot, \cdot \rangle$ and the corresponding Euclidean norm $\| \cdot \|$. We denote by $\mathcal{B}(\mathbb{R}^n)$ the Borel σ-algebra on $\mathbb{R}^n$. We often restrict ourselves to $\mathbb{R}^n_+ := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_j \geq 0, j = 1, \ldots, n\}$ equipped with the same scalar product and norm. By $\mathcal{B}(\mathbb{R}^n_+)$ we denote the Borel σ-algebra on $\mathbb{R}^n_+$.

Let $\mathcal{X} \in \{\mathbb{R}^n, \mathbb{R}^n_+\}$. For $A \subseteq \mathcal{X}$, we denote the *closure* of A by $\bar{A}$ and the *interior* of A by A°. For $x, y \in \mathcal{X}$ and $n = 1$, we write $x \vee y := \max\{x, y\}$ for the *maximum* of x and y, $x \wedge y := \min\{x, y\}$ for the *minimum* of x and y, $x^+ := x \vee 0$ for the *positive part* of x, and $x^- := (-x) \vee 0$ for the *negative part* of x.

2.1. Function spaces

Let $f : \mathcal{X} \to \mathbb{R}$ be a function. We always assume measurability properties for the functions we consider. Therefore, we define

$$\mathbb{L}^0(\mathcal{X}) := \{f : \mathcal{X} \to \mathbb{R} \mid f \text{ is } (\mathcal{B}(\mathcal{X}), \mathcal{B}(\mathbb{R}))\text{-measurable}\},$$

the set of all measurable functions on $\mathcal{X}$. We often use the notion of semi-continuous functions. The function f is called *upper semi-continuous* in $x \in \mathcal{X}$, if for any sequence $(x_k)_{k \in \mathbb{N}}$ in $\mathcal{X}$ such that $\lim_{k \to \infty} x_k = x$, we have

$$\limsup_{k \to \infty} f(x_k) \leq f(x),$$

and it is called *lower semi-continuous* in $x \in \mathcal{X}$, if for any sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathcal{X}$ such that $\lim_{k\to\infty} x_k = x$, we have

$$\liminf_{k\to\infty} f(x_k) \geq f(x).$$

A function is continuous if and only if it is upper and lower semi-continuous. We denote by

$$\mathcal{C}(\mathcal{X}) := \{f : \mathcal{X} \to \mathbb{R} \mid f \text{ is continuous}\}$$

the set of all continuous functions on $\mathcal{X}$, by

$$\mathcal{C}_b(\mathcal{X}) := \{f \in \mathcal{C}(\mathcal{X}) \mid f \text{ is bounded}\}$$

the set of all bounded, continuous functions on $\mathcal{X}$, and by

$$\mathcal{C}_c(\mathcal{X}) := \{f \in \mathcal{C}(\mathcal{X}) \mid f \text{ has compact support}\}$$

the set of all continuous functions with compact support $\operatorname{supp}(f) := \{x \in \mathcal{X} \mid f(x) \neq 0\}$. The function f is called *Lipschitz continuous* with constant $L > 0$, if

$$|f(x) - f(y)| \leq L\|x - y\|, \quad x, y \in \mathcal{X}.$$

We denote the set of all Lipschitz continuous functions with constant $L > 0$ by

$$\mathcal{C}_L(\mathcal{X}) := \{f \in \mathcal{C}(\mathcal{X}) \mid f \text{ is Lipschitz continuous with constant } L\}.$$

For a sequence $(f_k)_{k\in\mathbb{N}}$ of functions $f_k : \mathcal{X} \to \mathbb{R}$, $k \in \mathbb{N}$, we write $f_k \nearrow f$, if $f_k \leq f_{k+1}$, $k \in \mathbb{N}$, and $f_k(x) \overset{k\to\infty}{\to} f(x)$, $x \in \mathcal{X}$. We write $f_k \searrow f$, if $-f_k \nearrow -f$.

Lemma 2.1 ([10, Lemma 7.14]). *Let $\mathcal{X}$ be a metrizable space and $f : \mathcal{X} \to \mathbb{R} \cup \{-\infty, \infty\}$.*

1. *The function f is lower semi-continuous and bounded from below if and only if there is a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathcal{C}_b(\mathcal{X})$ such that $f_k \nearrow f$.*

2. *The function f is upper semi-continuous and bounded from above if and only if there is a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathcal{C}_b(\mathcal{X})$ such that $f_k \searrow f$.*

Let us now proceed with differentiability considerations. We denote by

$$\mathcal{C}^k(\mathcal{X}) := \{f \in \mathcal{C}(\mathcal{X}) \mid f \text{ is } k \text{ times continuously differentiable}\}$$

the set of all k times continuously differentiable functions on $\mathcal{X}$. If $\mathcal{X} = \mathbb{R}$, then we denote the *derivatives* of $f \in \mathcal{C}^k(\mathbb{R})$ by f', f'', f''', $f^{(4)}, \dots, f^{(k)}$. By $f'(\cdot-)$ and $f'(\cdot+)$, we denote the *left* and the *right derivative* of f respectively. If $\mathcal{X} = \mathbb{R}^n$, then we denote the *total derivatives* of $f \in \mathcal{C}^k(\mathbb{R}^n)$ by f', f'', f''', $f^{(4)}, \dots, f^{(k)}$. For the *partial derivatives*, we write

$$f_{x_i}(x_1, \dots, x_n) := \frac{\partial f(x_1, \dots, x_n)}{\partial x_i}, \quad i = 1, \dots, n.$$

Remark 2.2. By Rademacher's theorem, see for example [28, Theorem 3.1.6], the left and right derivatives exist for any convex function and are then equal almost everywhere. ◊

2.2. Measure theory

After discussing several properties of functions, we now consider measure theoretic aspects. Let $\mathcal{X} \in \{\mathbb{R}^n, \mathbb{R}^n_+\}$. We define

$$\mathcal{P}_\alpha(\mathcal{X}) \coloneqq \left\{\pi \;\middle|\; \pi \text{ is a Borel measure on } (\mathcal{X}, \mathcal{B}(\mathcal{X})) : \pi(\mathcal{X}) < \infty \text{ and } \int_\mathcal{X} |x|\pi(\mathrm{d}x) < \infty\right\},$$

the set of all Borel measures π on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ with finite mass $\pi(\mathcal{X})$ and finite *barycentre* or *first moment* $\mathbb{B}(\pi) \coloneqq \frac{1}{\pi(\mathcal{X})} \int_\mathcal{X} x\pi(\mathrm{d}x)$. By

$$\mathcal{P}(\mathcal{X}) \coloneqq \{\pi \in \mathcal{P}_\alpha(\mathcal{X}) \mid \pi(\mathcal{X}) = 1\},$$

we denote the set of all Borel probability measures with finite first moment. A measure $\pi \in \mathcal{P}_\alpha(\mathcal{X})$ is called *discrete*, if its support

$$\operatorname{supp}(\pi) \coloneqq \{x \in \mathcal{X} \mid \text{For any neighborhood } N_x \text{ of } x, \text{ we have } \pi(N_x) > 0\}$$

is a countable set. An $x \in \mathcal{X}$ is called an *atom* of π, if $\pi(\{x\}) > 0$. For a discrete measure, the support is equal to the set of all atoms, i.e. $\operatorname{supp}(\pi) = \{x \in \mathcal{X} \mid \pi(\{x\}) > 0\}$. A measure $\pi \in \mathcal{P}_\alpha(\mathcal{X})$ is called *continuous*, if $\{x \in \mathcal{X} \mid \pi(\{x\}) > 0\} = \emptyset$, i.e. if it has no atoms.

Now let $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \pi)$ be a probability space, $\mu \in \mathcal{P}(\mathbb{R})$ and $X : \mathcal{X} \to \mathbb{R}$ a random variable on $\mathcal{X}$, i.e. $X \in \mathbb{L}^0(\mathcal{X})$. Then μ is called the *law* of X under π, if

$$\mu(B) = \pi(X \in B), \quad B \in \mathcal{B}(\mathbb{R}).$$

In this case, we write $X \sim \mu$ or $X \sim_\pi \mu$ and we denote the *distribution function* of (the law of) X by

$$F_\mu : \mathbb{R} \to [0, 1], \quad x \mapsto \pi(X \leq x) = \mu((-\infty, x]),$$

and the *quantile function* of (the law of) X by

$$F_\mu^{-1} : (0, 1) \to \mathbb{R}, \quad t \mapsto \inf\{x \in \mathbb{R} \mid F_\mu(x) \geq t\}.$$

For a measure $\pi \in \mathcal{P}_\alpha(\mathcal{X})$, we denote by

$$\mathbb{L}^1(\mathcal{X}, \pi) \coloneqq \left\{f \in \mathbb{L}^0(\mathcal{X}) \;\middle|\; \int_\mathcal{X} |f(x)|\pi(\mathrm{d}x) < \infty\right\}$$

the set of all π-*integrable* functions on $\mathcal{X}$. If $\pi \in \mathcal{P}(\mathcal{X})$ and $f \in \mathbb{L}^1(\mathcal{X}, \pi)$, then we write

$$\mathbb{E}_\pi[f(X)] \coloneqq \int_\mathcal{X} f(x)\pi(\mathrm{d}x)$$

for the *expected value* of $f(X)$ with respect to π.

For a sequence $(\pi_k)_{k\in\mathbb{N}}$ in $\mathcal{P}(\mathcal{X})$ and $\pi \in \mathcal{P}(\mathcal{X})$, we say π_k *converges weakly* to π, denoted by $\pi_k \stackrel{w}{\to} \pi$, if for all $f \in \mathcal{C}_b(\mathcal{X})$, we have

$$\int_{\mathcal{X}} f(x)\pi_k(\mathrm{d}x) \stackrel{k\to\infty}{\to} \int_{\mathcal{X}} f(x)\pi(\mathrm{d}x).$$

For two measurable spaces $(\mathcal{X}_1, \mathcal{B}(\mathcal{X}_1))$ and $(\mathcal{X}_2, \mathcal{B}(\mathcal{X}_2))$, $\mathcal{X}_1, \mathcal{X}_2 \in \{\mathbb{R}^n, \mathbb{R}^n_+\}$, a measurable mapping $f : \mathcal{X}_1 \to \mathcal{X}_2$ and a measure $\pi \in \mathcal{P}_\alpha(\mathcal{X}_1)$, we define the *pushforward measure* $\pi^f \in \mathcal{P}_\alpha(\mathcal{X}_2)$ by

$$\pi^f(B) := (f_\#\pi)(B) := \pi(f^{-1}(B)), \quad B \in \mathcal{B}(\mathcal{X}_2).$$

Using this notation, we may write $\mu = \pi^X = X_\#\pi$ for the law μ of a random variable X under the measure π.

Now let $\mathcal{X} = \mathcal{X}_1 \times \ldots \mathcal{X}_m$, $\mathcal{X}_1, \ldots, \mathcal{X}_m \in \{\mathbb{R}^n, \mathbb{R}^n_+\}$, $m \in \mathbb{N}$. Then we define the i-th *projection mapping*, $i = 1, \ldots, m$, by

$$\mathrm{proj}^{x_i} : \mathcal{X} \to \mathcal{X}_i, \quad x = (x_1, \ldots, x_m) \mapsto x_i,$$

and the i-th *marginal* of $\pi \in \mathcal{P}_\alpha(\mathcal{X})$, $i = 1, \ldots, m$, by

$$\mu_i(B) := \left(\mathrm{proj}^{x_i}_\#\pi\right)(B), \quad B \in \mathcal{B}(\mathcal{X}_i).$$

Let $(\mathcal{X}_i, \mathcal{B}(\mathcal{X}_i), \mu_i)$, $i = 1, \ldots, m$, be probability spaces. Then, coupling $\mu_1, \ldots, \mu_m$ means constructing random variables $X_1, \ldots, X_m$ on some probability space $(\Omega, \mathbb{P})$ such that $X_i \sim_{\mathbb{P}} \mu_i$, $i = 1, \ldots, m$. The couple $(X_1, \ldots, X_m)$ is called *coupling* of $(\mu_1, \ldots, \mu_m)$. We also say π is a *coupling* of $(\mu_1, \ldots, \mu_m)$, if $(X_1, \ldots, X_m) \sim_{\mathbb{P}} \pi$.

In a measure theoretic sense, coupling $\mu_1, \ldots, \mu_m$ means constructing a measure π such that π admits $\mu_1, \ldots, \mu_m$ as marginals. There are three equivalent ways of rephrasing the marginal condition for π:

1. For all $i = 1, \ldots, m$, π satisfies $\mathrm{proj}^{x_i}_\#\pi = \mu_i$.

2. For all $i = 1, \ldots, m$ and all measurable sets $B_i \in \mathcal{B}(\mathcal{X}_i)$, π satisfies

$$\mu_i(B_i) = \pi\left(\mathcal{X}_1 \times \ldots \times \mathcal{X}_{i-1} \times B_i \times \mathcal{X}_{i+1} \times \ldots \times \mathcal{X}_m\right).$$

3. For all $i = 1, \ldots, m$ and all π-integrable (or non-negative, measurable) functions $\varphi_i : \mathcal{X}_i \to \mathbb{R}$, π satisfies

$$\int_{\mathcal{X}_1\times\ldots\times\mathcal{X}_m} \sum_{i=1}^m \varphi_i(x_i)\pi(\mathrm{d}(x_1, \ldots, x_m)) = \sum_{i=1}^m \int_{\mathcal{X}_i} \varphi_i(x_i)\mu_i(\mathrm{d}x_i).$$

Let us consider two rather extreme couplings as examples for $m = 2$. The first one is the independence coupling. That is, the coupling (X_1, X_2) is distributed according to the product measure of μ_1 and μ_2, i.e. $(X_1, X_2) \sim \mu_1 \otimes \mu_2$, and X_1 and X_2 are independent.

The second one is a so-called deterministic coupling. That is, for a measurable function $T : \mathcal{X}_1 \to \mathcal{X}_2$ we have $X_2 = T(X_1)$. Hence, we say $(X_1, X_2) \sim \pi$ is a deterministic coupling, if one of the following equivalent conditions is satisfied.

1. π is concentrated on the graph of a measurable function $T : \mathcal{X}_1 \to \mathcal{X}_2$.

2. $X_1 \sim \mu_1$ and $X_2 = T(X_1)$, where $\mu_2 = T_{\#}\mu_1$.

3. For all μ_2-integrable (or non-negative, measurable) functions $\psi : \mathcal{X}_2 \to \mathbb{R}$, we have
$$\int_{\mathcal{X}_2} \psi(x_2)\mu_2(\mathrm{d}x_2) = \int_{\mathcal{X}_1} \psi(T(x_1))\mu_1(\mathrm{d}x_1).$$

4. $\pi = (\mathrm{Id}, T)_{\#}\mu_1$.

Let $\mu_1 \in \mathcal{P}(\mathcal{X}_1), \ldots, \mu_m \in \mathcal{P}(\mathcal{X}_m)$ and assume that $\mathcal{X}_i$ is n_i-dimensional, $i = 1, \ldots, m$. We denote by $d = \sum_{i=1}^m n_i$ the dimension of $\mathcal{X}_1 \times \ldots \times \mathcal{X}_m$. Then we denote the set of all couplings of $(\mu_1, \ldots, \mu_m)$ by

$$\begin{aligned}&\Pi_d(\mu_1, \ldots, \mu_m)\\ &\quad\coloneqq \{\pi \in \mathcal{P}(\mathcal{X}_1 \times \ldots \mathcal{X}_m) \mid \mu_i(B_i) = \pi(\mathcal{X}_1 \times \ldots \times \mathcal{X}_{i-1} \times B_i \times \mathcal{X}_{i+1} \times \ldots \times \mathcal{X}_m),\\ &\qquad\qquad B_i \in \mathcal{B}(\mathcal{X}_i), i = 1, \ldots, m\}.\end{aligned}$$

The elements of $\Pi_d(\mu_1, \ldots, \mu_m)$ are called *transport plans*. We have $\Pi_d(\mu_1, \ldots, \mu_m) \neq \emptyset$, as $\mu_1 \otimes \ldots \otimes \mu_m \in \Pi_d(\mu_1, \ldots, \mu_m)$. If the dimensions are unspecified and irrelevant, with a slight abuse of notation, we also write $\Pi_m(\mu_1, \ldots, \mu_m)$.

Clearly, considering the marginals of some measure and building a coupling of some measures serve as inverse operations.

On $\mathcal{P}(\mathcal{X})$ we define the *Wasserstein distance* of $\mu, \nu \in \mathcal{P}(\mathcal{X})$ by
$$W(\mu, \nu) \coloneqq \inf_{\pi \in \Pi_2(\mu,\nu)} \mathbb{E}_\pi \left[\|X - Y\|\right], \tag{2.1}$$
where $X \sim \mu$ and $Y \sim \nu$ are random variables.

We denote the topology induced by the metric W on the metric space $(\mathcal{P}(\mathcal{X}), W)$ by $\mathcal{T}_1(\mathcal{X})$. We further denote the weak topology induced by bounded, continuous functions by $\mathcal{T}_{cb}(\mathcal{X})$. Observe that in our case, see for example Villani [77, Theorem 6.9], the Wasserstein distance metrizes weak convergence, as we consider measures with finite first moments.

We may characterize the Wasserstein distance using Lipschitz continuous functions by
$$W(\mu, \nu) = \sup_{f \in \mathcal{C}_1(\mathcal{X})} \int_{\mathcal{X}} f(x)(\mu - \nu)(\mathrm{d}x). \tag{2.2}$$

In order to calculate the Wasserstein distance of two probability measures, we often need the following, more explicit lemma.

Lemma 2.3 ([64, Sec. 6.1]). *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ and $\pi \in \Pi_2(\mu, \nu)$. Let X and Y be two real-valued random variables with distribution functions $F_\mu : \mathbb{R} \to [0,1]$ and $F_\nu : \mathbb{R} \to [0,1]$. Let (X, Y) be a coupling of X and Y with distribution function $H : \mathbb{R}^2 \to [0,1]$, i.e. $\mathbb{E}_H[|X - Y|] = \int_{\mathbb{R}^2} |x - y| H(\mathrm{d}(x,y))$. Then we have*

$$\mathbb{E}_H\left[|X - Y|\right] = \int_{-\infty}^{\infty} F_\mu(t) + F_\nu(t) - 2H(t,t)\mathrm{d}t.$$

In particular, we have

$$W(\mu, \nu) = \int_{-\infty}^{\infty} F_\mu(t) + F_\nu(t) - 2\min\left\{F_\mu(t), F_\nu(t)\right\}\mathrm{d}t = \int_{-\infty}^{\infty} |F_\mu(t) - F_\nu(t)|\mathrm{d}t.$$

Remark 2.4. A proof of Lemma 2.3 is given in Dall'Aglio [21, Sec. 1]. In particular, the author shows that one may give up the assumption that the expected values exist. ♢

Remark 2.5. Let X be a non-negative random variable with distribution function F. Then

$$\mathbb{E}[X] < \infty \Rightarrow (1 - F(x)) \cdot x \overset{x \to \infty}{\to} 0. \tag{2.3}$$

In order to prove the claim, we assume that the convergence does not hold. Then there is an $\varepsilon > 0$ such that for all $x_0 \in \mathbb{R}_+$ there is an $x \geq x_0$ such that

$$(1 - F(x)) \cdot x > \varepsilon, \quad \text{or equivalently} \quad 1 - F(x) > \frac{\varepsilon}{x}.$$

Now let $(x_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}_+$ such that the above inequalities hold for every x_k, and that satisfies $x_{k+1} > 2x_k$, $k \in \mathbb{N}$. As the distribution function F is monotone non-decreasing, we have

$$1 - F(x) \geq 1 - F(x_k) > \frac{\varepsilon}{x_k}$$

for all $x \in (x_{k-1}, x_k]$. Using the representation formula for the expected value, we get

$$\begin{aligned}\mathbb{E}[X] &= \int_0^{\infty} (1 - F(x))\mathrm{d}x \geq \sum_{k=1}^{\infty} \int_{x_k}^{x_{k+1}} (1 - F(x))\mathrm{d}x > \sum_{k=1}^{\infty} \int_{x_k}^{x_{k+1}} \frac{\varepsilon}{x_{k+1}}\mathrm{d}x \\ &= \sum_{k=1}^{\infty} \frac{\varepsilon}{x_{k+1}} \cdot (x_{k+1} - x_k) = \varepsilon \cdot \sum_{k=1}^{\infty} \left(1 - \frac{x_k}{x_{k+1}}\right) > \varepsilon \cdot \sum_{k=1}^{\infty} \frac{1}{2} = \infty,\end{aligned}$$

a contradiction to the existence of the expected value.

The converse implication in (2.3) is not true, as for example $1 - F(x) := \frac{1}{x\ln(x)}$ yields a counterexample.

Let now $r \in \mathbb{N}$. Then the general version of the statement in (2.3) reads as

$$\mathbb{E}[X^r] < \infty \Rightarrow (1 - F(x)) \cdot x^r \overset{x \to \infty}{\to} 0 \wedge F(x) \cdot x^r \overset{x \to -\infty}{\to} 0,$$

where we allow the random variable X to take values in $\mathbb{R}$. We do not discuss the assertion in detail, as we do not need it in later chapters. ♢

Chapter 3

Principles of model-independent finance

In this chapter, we model the mathematical framework that we use in this thesis and present some of its elementary properties. In Section 3.1, we model two different financial markets that implement some of the core assumptions of model-independent finance mathematically. In Section 3.2, we analyze call options and their price functions closely, as they play an important role in model-independent finance and thus throughout this thesis.

3.1. The underlying financial markets

As this thesis is somehow twofold, we introduce two financial markets. Though one is a special case of the other, we introduce the markets separately, because two different parts of this thesis use the framework given by one of the subsequently defined markets respectively. We start by introducing the more general financial market used in a first part and then specialize and restate everything for the so-called standard market used in a second part. The main advantages of doing so are much simpler notation and referencing in the standard market case.

3.1.1. The general market

Let $(\Omega_g, \mathcal{F}_g) = (\mathbb{R}^{nd}, \mathcal{B}(\mathbb{R}^{nd}))$, $n, d \in \mathbb{N}$, be the underlying measurable space. We already know from the introduction that in model-independent finance, we do not specify a certain probability measure on $(\Omega_g, \mathcal{F}_g)$, as we consider all probability measures that are admissible in a certain sense at once. Based on $(\Omega_g, \mathcal{F}_g)$ we consider a frictionless financial market in which we assume no-arbitrage.

Let $0 = t_0 < \ldots < t_n = T < \infty$ be discrete trading times and T the final maturity date. We denote $\mathcal{T} \coloneqq \{t_1, \ldots, t_n\}$.

We assume that there is a risk-free asset with price process $(S_t^0)_{t\in\mathcal{T}}$ and $S_t^0 = S_0^0 = 1$ for all $t \in \mathcal{T}$. That is, we assume that the financial market pays no interest rates.

We further assume that there are d non-redundant risky assets with corresponding price processes $S^j = (S_t^j)_{t\in\mathcal{T}}$ with $S_{t_i}^j \in \mathbb{R}$, $i = 1, \ldots, n$, and $S_0^j = s_0^j \in \mathbb{R}$ for all $j = 1, \ldots, d$, where s_0^j is the observable market price of asset j at time $t = 0$. We denote $S = (S^1, \ldots, S^d)$. We assume that the risky assets pay no dividends. The risky assets $S^1, \ldots, S^d$ are the potential underlyings for the exotic options considered in this thesis.

As the future asset prices are unknown, we model the price processes as stochastic processes using the canonical definition. That is, for all $i = 1, \ldots, n$ and all $j = 1, \ldots, d$, we define the single asset prices by coordinate mappings

$$S_{t_i}^j : \Omega_g \to \mathbb{R}, \quad \left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right) \mapsto s_{t_i}^j.$$

Analogously, for all $j = 1, \ldots, d$, we define the price process of a single asset by

$$S^j : \Omega_g \to \mathbb{R}^n, \quad \left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right) \mapsto (s_{t_1}^j, \ldots, s_{t_n}^j).$$

Thus, whenever we specify a probability measure on $(\Omega_g, \mathcal{F}_g)$, we have indeed random variables, random vectors and stochastic processes.

The crucial novelty of model-independent finance is, as explained earlier, to assume that the prices of call options on the underlyings are determined by market mechanics rather than the underlying asset prices. Thus, the call option prices are observable in the market at time 0. That is, we assume that call options with payoff function

$$\Phi_{i,j} : \Omega_g \times \mathbb{R} \to \mathbb{R}_+, \quad \left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d, k\right) \mapsto (s_{t_i}^j - k)^+$$

are liquidly traded for all $i = 1, \ldots, n$ and all $j = 1, \ldots, d$, i.e. for all underlying assets at all trading times and for all strike prices $k \in \mathbb{R}$. We denote by $C_{i,j}(k)$ the price of the call option with payoff $\Phi_{i,j}\left(S_{t_1}^1, \ldots, S_{t_n}^1, \ldots, S_{t_1}^d, \ldots, S_{t_n}^d, k\right)$. As we assume no-arbitrage, we may deduce some properties of the call option price functions $k \mapsto C_{i,j}(k)$, $i = 1, \ldots, n$, $j = 1, \ldots, d$. We discuss these properties in Section 3.2.

Finally, we introduce a general exotic option depending on S. An exotic option is represented by its payoff function

$$c : \Omega_g \to \mathbb{R}, \quad \left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right) \mapsto c\left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right),$$

where we assume $c \in \mathbb{L}^0(\Omega_g)$. In this thesis, we derive results concerning the price at time $t = t_0$ of the exotic option with random payoff $c\left(S_{t_1}^1, \ldots, S_{t_n}^1, \ldots, S_{t_1}^d, \ldots, S_{t_n}^d\right)$.

We already know that in model-independent finance the price is in general not uniquely defined. Instead, there is an interval of possible no-arbitrage prices. Hence, whenever we speak about the price of an exotic option, we actually mean upper and lower price bounds.

In order to find these price bounds, we introduce several optimization problems, namely two pricing problems and two hedging problems. We formalize these problems in Sections 4.4 and 4.5. We use the framework of this general financial market in Chapter 5.

Remark 3.1. Though we generally allow the asset prices S_t^j to take values in $\mathbb{R}$, we often consider the case that the prices take values in $\mathbb{R}_+$. This case emerges from the above by suitably replacing $\mathbb{R}$ by $\mathbb{R}_+$. This also applies to the next section. Further note that whenever we choose $d = 1$, we remove the associated index j in the notation. ◊

3.1.2. The standard market

It is quite hard to deduce results in model-independent finance for the general market case. Thus, we specialize ourselves to a situation in which we can say far more.

In particular, let in the general market case of the previous section be $n = 2$ and $d = 1$. Then we have the underlying measurable space $(\Omega_s, \mathcal{F}_s) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. As the notation of the previous section is rather complicated for this simple case, we change it suitably. We denote the trading times by $0 < t < T$, the risk-free asset by $B = (B_t, B_T) = (1, 1)$ with $B_0 = 1$ and for the risky asset we write $S = (X, Y)$ with $S_0 = s_0 \in \mathbb{R}$. We use the same canonical definitions as before in order to introduce the stochastic to the underlying measurable space. Finally, we denote by $C_t(k)$ the price of $\Phi_t(X, k) = (X - k)^+$ and by $C_T(k)$ the price of $\Phi_T(Y, k) = (Y - k)^+$. The exotic option under consideration is then represented by the payoff function

$$c : \Omega_s \to \mathbb{R}, \quad (x, y) \mapsto c(x, y),$$

where we assume $c \in \mathbb{L}^0(\Omega_s)$.

As the notations in the standard market case stay simpler when introducing the optimization problems, we proceed to distinguish the two market cases in the following chapter.

We use the more special framework of the standard market in Chapters 6 and 7.

3.2. Call options

We know from the heuristic description in the introduction as well as from the market specification in the previous section that call options play an important role in model-independent finance. Though we do not yet formally know why, we take a closer look at this special (exotic) option.

We stress the assumption of no-arbitrage, which also applies to the call option prices. Hence, we may derive some properties of the price function $k \mapsto C(k)$, where in this section $C(k)$ shall represent the price of some call option with strike price $k \in \mathbb{R}$. The underlying, which we generically denote by S, and the maturity of the call option, which we denote by T, are of no particular interest at the moment.

Lemma 3.2. *The mapping $C : \mathbb{R} \to \mathbb{R}_+$, $k \mapsto C(k)$ is monotone non-increasing and convex.*

Proof. 1. Monotonicity of C: Let $k \leq \ell$. This implies $(S_T - k)^+ \geq (S_T - \ell)^+$. If we assume $C(k) < C(\ell)$ to get a contradiction, then we realize an arbitrage by buying a call option with strike price k and selling one with strike price ℓ. In Table 3.1,

Portfolio	Price in $t = 0$	Payoff in $t = T$
Call option with strike k long	$C(k)$	$(S_T - k)^+$
Call option with strike ℓ short	$-C(\ell)$	$-(S_T - \ell)^+$
In total	$C(k) - C(\ell)$	$(S_T - k)^+ - (S_T - \ell)^+$
Value	< 0	≥ 0

Table 3.1.: Arbitrage strategy, if C is not monotone non-increasing.

we illustrate that this arbitrage strategy has a non-negative payoff in $t = T$ and a negative price in $t = 0$. Thus, it realizes a free lunch, which is a contradiction to the no-arbitrage assumption. Consequently, $C(k) \geq C(\ell)$ and C is non-increasing.

2. Convexity of C: Let $s \in \mathbb{R}$, $\lambda \in (0, 1)$ and $k_1, k_2 \in \mathbb{R}$. Then we have

$$\begin{aligned}(s - (\lambda k_1 + (1-\lambda)k_2))^+ &= (\lambda(s - k_1) + (1-\lambda)(s - k_2))^+ \\ &\leq \lambda(s - k_1)^+ + (1-\lambda)(s - k_2)^+.\end{aligned}$$

The left hand side is the payoff of a call option with strike price $(\lambda k_1 + (1-\lambda)k_2)$ for $S_T = s$ and the right hand side is the payoff of a portfolio of λ call options with strike price k_1 and $(1-\lambda)$ call options with strike price k_2 for $S_T = s$.

By the same arguments as for the monotonicity, we immediately get

$$C(\lambda k_1 + (1-\lambda)k_2) \leq \lambda C(k_1) + (1-\lambda)C(k_2),$$

which is the convexity of C. □

Remark 3.3. 1. Clearly, the assertion of Lemma 3.2 also holds when S_T takes values in $\mathbb{R}_+$. Then the call option price function is a mapping $C : \mathbb{R}_+ \to \mathbb{R}_+$.

2. By Remark 2.2, the left and right derivatives $C'(\cdot-)$ and $C'(\cdot+)$ of C exist and are equal almost everywhere on $\mathbb{R}$. Hence, C is differentiable almost everywhere. ◊

Lemma 3.4. *The mapping $C : \mathbb{R} \to \mathbb{R}_+$, $k \mapsto C(k)$ has the following properties.*

1. $\lim_{k\to\infty} C(k) = 0$.

2. $C'(k+) \geq -1$, $k \in \mathbb{R}$.

3. $\lim_{k\to-\infty} C(k) + k = s_0$.

If S_T takes values in $\mathbb{R}_+$, i.e. if $C : \mathbb{R}_+ \to \mathbb{R}_+$, then we may replace 2. and 3. by the following properties.

2.' $C'(0+) \geq -1$.

3.' $C(0) = s_0$.

Proof. 1. $\lim_{k\to\infty} C(k) = 0$: We know that C is a monotone non-increasing function. As the payoff of a call option is always non-negative, it is bounded from below by 0. Hence, we know that the limit exists.

In order to argue formally that the limit is indeed equal to 0, we need tools from Chapter 4. This is detailed in Remark 4.15. Heuristically, the assertion is the reasonable assumption that for increasing strike prices it is virtually impossible for the option to generate a positive payoff.

2. $C'(k+) \geq -1, k \in \mathbb{R}$: Assume there is an $\ell \in \mathbb{R}$ such that $C'(\ell+) < -1$. Then there is an $\ell' > \ell$ such that $C'(\ell'+) \geq -1$ with

$$\frac{C(\ell') - C(\ell)}{\ell' - \ell} < -1.$$

Indeed, otherwise we have an immediate contradiction to $C(k) \geq 0, k \in \mathbb{R}$. The former inequality is equivalent to

$$C(\ell') - C(\ell) + \ell' - \ell < 0.$$

Using this, we may realize an arbitrage using the arbitrage strategy illustrated in Table 3.2. It has a non-negative payoff in $t = T$ and a negative price in $t = 0$, which

	Price in $t = 0$	Payoff in $t = T$, if		
Portfolio		$S_T > \ell'$	$\ell' \geq S_T > \ell$	$\ell > S_T$
Call with strike ℓ' long	$C(\ell')$	$S_T - \ell'$	0	0
Call with strike ℓ short	$-C(\ell)$	$\ell - S_T$	$\ell - S_T$	0
Bond investment of $\ell' - \ell$	$\ell' - \ell$	$\ell' - \ell$	$\ell' - \ell$	$\ell' - \ell$
In total	$C(\ell') - C(\ell) + \ell' - \ell$	$S_T - \ell' - (S_T - \ell) + \ell' - \ell$	$\ell' - S_T$	$\ell' - \ell$
Value	< 0	$= 0$	≥ 0	≥ 0

Table 3.2.: Arbitrage strategy, if $C'(\ell+) \geq -1$.

is a free lunch and a contradiction to the no-arbitrage assumption. Thus, we have $C'(k+) \geq -1, k \in \mathbb{R}$.

3. $\lim_{k\to-\infty} C(k) + k = s_0$: As for $\lim_{k\to\infty} C(k) = 0$, we have to wait for Remark 4.15 to formally prove this assertion in the general case. However, if we assume that S_T takes values in $[\underline{s}, \infty)$, then we may prove the claim using simple no-arbitrage arguments. Therefore, assume $k \leq \underline{s}$. Then we even have $C(k) = s_0 - k$.

Indeed, we have $C(k) \geq s_0 - k$, as for the payoff we have $(S_T - k)^+ \geq S_T - k$ for all $k \in \mathbb{R}$. (Actually, as also $(S_T - k)^+ \geq 0$, we even have $C(k) \geq (s_0 - k)^+$.) Now assuming $C(k) > s_0 - k$, we have an arbitrage strategy buying the underlying, lending money and selling the call option short as illustrated in Table 3.3. We observe that

Portfolio	Price in $t = 0$	Payoff in $t = T$
Underlying long	s_0	S_T
Bond investment of $-k$	$-k$	$-k$
Call option with strike k short	$-C(k)$	$-(S_T - k)^+$
In total	$s_0 - k - C(k)$	$(S_T - k) - (S_T - k)^+$
Value	< 0	$= 0$

Table 3.3.: Arbitrage strategy if $C(k) > s_0 - k$.

$(S_T - k)^+ = S_T - k$, as $k \leq \underline{s}$. This strategy has a zero payoff in $t = T$ and a negative price in $t = 0$, which is a free lunch and a contradiction to the no-arbitrage assumption. Thus, we have $C(k) = s_0 - k$ for all $k \leq \underline{s}$. Heuristically, this property should also hold if $\underline{s}$ tends to $-\infty$.

2.' $C'(0+) \geq -1$: Clearly, this holds by choosing $\ell = 0$ in the proof of the second property. By the convexity of C, we have $C'(k+) \geq -1$ for all $k \in \mathbb{R}_+$.

3.' $C(0) = s_0$: A call option with strike price $k = 0$ has a payoff of $(S_T - 0)^+ = S_T$, as $S_T \geq 0$ by assumption. Hence, any price unequal s_0 immediately gives rise to an arbitrage strategy. □

The properties discussed are mainly guaranteed by no-arbitrage considerations. We gain some deeper insight into the properties of call options and their price functions in Section 4.2, where we connect call option price functions and probability measures. We end this chapter by introducing the notion of candidate functions.

Definition 3.5. A function $C : \mathbb{R} \to \mathbb{R}_+$ is called a *candidate function for call option prices*, if it satisfies the following conditions.

1. C is monotone non-increasing and convex.
2. $\lim_{k\to\infty} C(k) = 0$, $C'(k+) \geq -1$, $k \in \mathbb{R}$ and $\lim_{k\to-\infty} C(k) + k = s_0$.

A function $C : \mathbb{R}_+ \to \mathbb{R}_+$ is called a *candidate function for call option prices*, if it satisfies the following conditions.

1. C is monotone non-increasing and convex.
2. $\lim_{k\to\infty} C(k) = 0$, $C'(0+) \geq -1$ and $C(0) = s_0$.

We denote the sets of all candidate functions by $\mathcal{K}^C_{\mathbb{R}}$ and $\mathcal{K}^C$ respectively.

CHAPTER 4

Martingale optimal transport

In this chapter, we present the optimization problems which are at the core of our interest. We understand the importance of call options formally and introduce the notion of martingale transport plans. We provide results from classic optimal transport that bring up presumptions worthwhile to investigate and basic ideas how to handle the presumptions in the martingale case. In Section 4.1, we recall several results from classic optimal transport that motivate the investigation of duality, optimality and several structural properties in martingale optimal transport. In Section 4.2, we transfer classic optimal transport to model-independent finance presenting the celebrated lemma of Breeden & Litzenberger. This finally explains the immense importance of the assumptions on call options. We also define the set of martingale transport plans. In Section 4.3, we introduce the notion of convex order and the theorem of Strassen in order to guarantee existence of martingale transport plans. In Section 4.4, we finally introduce the optimization problems that yield the upper and lower price bounds for exotic options in the general and the standard market cases. We complement the price bound problems in Section 4.5, where we introduce the super and sub hedging problems that prove to be the dual problems. In Section 4.6, we recall the pioneering duality result of Beiglböck, Henry-Labordère & Penkner [5], which serves as a guideline for our main theorem in the general market case.

4.1. Classic optimal transport

In this section, we introduce the idea of classic optimal transport in order to understand that an adaption may be useful in model-independent finance. We present several results of classic optimal transport, some of which we need explicitly or in an idea-generating way.

Before we start to rigorously introduce the mathematics, we shortly explain the original idea of optimal transport. In 1781, Gaspard Monge [63] introduced and studied the problem of what we call classic optimal transport: Assume you have a certain amount of soil at specified locations and have to transport it to other, not necessarily different locations. As transport is costly, the question is, how can we assign the locations to each other in a cost-minimizing and in this sense optimal way. We may assume that the locations on both ends of the transport are given by probability measures. Then the transport problem becomes the problem of coupling the measures with each other in a cost-minimizing way.

Based on this intuition, we formalize the problem and present several related results. Thereby, we mostly follow Villani [77, Chap. 4, 5], but add a useful result from Kellerer [55]. Finally, we motivate subsequent results by restating well-known assertions as it is done in Beiglböck & Juillet [7] and Henry-Labordère & Touzi [38]. Though we only need the results for probability measures on the Euclidean space, we state them in full generality.

Let $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$ be two Polish probability spaces and $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ a measurable function indicating the cost $\phi(x, y)$ of transporting mass from some $x \in \mathcal{X}$ to some $y \in \mathcal{Y}$. Then Monge's original problem of optimal transport is

$$\inf_T \int_{\mathcal{X}} \phi(x, T(x))\mu(\mathrm{d}x),$$

where the infimum is taken over all *transport maps* T from μ to ν, i.e. all measurable functions $T : \mathcal{X} \to \mathcal{Y}$ such that $T_{\#}\mu = \nu$. However, this problem may be unsolvable, as there is not always a transport map between two probability measures.

Therefore, in 1948, Kantorovich [53, 54] introduced the more general so-called Monge-Kantorovich problem of optimal transport

$$P_{MK}(\phi) := \inf_{\pi \in \Pi_2(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} \phi(x, y)\pi(\mathrm{d}(x, y)). \tag{4.1}$$

Theorem 4.1 ([77, Theorem 4.1]). *Let $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$ be two Polish probability spaces. Let $a : \mathcal{X} \to \mathbb{R}\cup\{-\infty\}$ and $b : \mathcal{Y} \to \mathbb{R}\cup\{-\infty\}$ be two upper semi-continuous functions such that $a \in \mathbb{L}^1(\mathcal{X}, \mu)$ and $b \in \mathbb{L}^1(\mathcal{Y}, \nu)$. Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\infty\}$ be a lower semi-continuous cost function such that*

$$\phi(x, y) \geq a(x) + b(y), \quad (x, y) \in \mathcal{X} \times \mathcal{Y}.$$

Then there is a coupling of (μ, ν) which minimizes the total cost $\int_{\mathcal{X}\times\mathcal{Y}} \phi(x, y)\pi(\mathrm{d}(x, y))$ among all possible couplings π of (μ, ν), i.e. there is a minimizer for $P_{MK}(\phi)$ in (4.1).

As the proof helps to understand some parts of the proof of Theorem 5.1, we provide it in detail. In advance, we present three results that we need in the proof and also in later chapters. We start with the so-called theorem of Prohorov.

Theorem 4.2 ([1, Theorem 2.3]). *Let $(\mathcal{X}, d)$ be a Polish space with metric d. Then a family $\mathcal{K} \subseteq \mathcal{P}(\mathcal{X})$ is relatively compact with respect to the weak topology $\mathcal{T}_{cb}(\mathcal{X})$ if and only* *if it is tight, i.e. if for all $\varepsilon > 0$ there is a compact set $K_\varepsilon \subseteq \mathcal{X}$ such that $\pi(\mathcal{X} \setminus K_\varepsilon) \leq \varepsilon$ for all $\pi \in \mathcal{K}$.*

We now present two lemmata. The first lemma states the lower semi-continuity of the cost functional $\pi \mapsto \int \phi \mathrm{d}\pi$ and the second lemma states the tightness of certain sets of transport plans. Both lemmata are useful in Chapter 5.

Lemma 4.3 ([77, Lemma 4.3])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces, $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\infty\}$ a lower semi-continuous cost function and $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{-\infty\}$ an upper semi-continuous function such that $h \leq \phi$ on $\mathcal{X} \times \mathcal{Y}$. Let $(\pi_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ converging weakly to some $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ and such that $h \in \mathbb{L}^1(\mathcal{X} \times \mathcal{Y}, \pi_k)$, $h \in \mathbb{L}^1(\mathcal{X} \times \mathcal{Y}, \pi)$ and*

$$\int_{\mathcal{X}\times\mathcal{Y}} h(x, y)\pi_k(\mathrm{d}(x, y)) \overset{k\to\infty}{\to} \int_{\mathcal{X}\times\mathcal{Y}} h(x, y)\pi(\mathrm{d}(x, y)).$$

Then

$$\int_{\mathcal{X}\times\mathcal{Y}} \phi(x, y)\pi(\mathrm{d}(x, y)) \leq \liminf_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi(x, y)\pi_k(\mathrm{d}(x, y)).$$

In particular, if ϕ is non-negative, then

$$P_\phi : \begin{cases} \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \to \mathbb{R} \\ \pi \mapsto \int_{\mathcal{X}\times\mathcal{Y}} \phi(x, y)\pi(\mathrm{d}(x, y)) \end{cases}$$

is lower semi-continuous on $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$ equipped with the topology of weak convergence $\mathcal{T}_{cb}(\mathcal{X} \times \mathcal{Y})$.

Lemma 4.4 ([77, Lemma 4.4])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces. Let $\mathcal{P} \subseteq \mathcal{P}(\mathcal{X})$ and $\mathcal{Q} \subseteq \mathcal{P}(\mathcal{Y})$ be tight subsets of $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(\mathcal{Y})$ respectively. Then the set*

$$\Pi_2(\mathcal{P}, \mathcal{Q}) := \{\pi \in \Pi_2(\mu, \nu) \mid \mu \in \mathcal{P}, \nu \in \mathcal{Q}\}$$

is tight in $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$.

Proof of Theorem 4.1. We first prove that $\Pi_2(\mu, \nu)$ is compact. Since $\mathcal{X}$ and $\mathcal{Y}$ are Polish, $\{\mu\}$ and $\{\nu\}$ are tight in $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(\mathcal{Y})$ respectively. By Lemma 4.4, $\Pi_2(\mu, \nu)$ is tight in $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$. Thus, by Theorem 4.2, $\Pi_2(\mu, \nu)$ is relatively compact.

Let now $(\pi_k)_{k\in\mathbb{N}}$ be a sequence in $\Pi_2(\mu, \nu)$. By the relative compactness, there is a subsequence $(\pi_{k_n})_{n\in\mathbb{N}}$ that is weakly convergent to some π, i.e. for all $f \in \mathcal{C}_b(\mathcal{X} \times \mathcal{Y})$,

$$\int_{\mathcal{X}\times\mathcal{Y}} f(x, y)\pi_{k_n}(\mathrm{d}(x, y)) \overset{n\to\infty}{\to} \int_{\mathcal{X}\times\mathcal{Y}} f(x, y)\pi(\mathrm{d}(x, y)).$$

It remains to show that $\pi \in \Pi_2(\mu, \nu)$. For this purpose, choose $f(x, y)$ independent of y, i.e. $f(x, y) = g(x) \in \mathcal{C}_b(\mathcal{X})$. Then

$$\begin{aligned} \int_{\mathcal{X}} g(x)\mu(\mathrm{d}x) &= \int_{\mathcal{X}\times\mathcal{Y}} f(x, y)\pi_{k_n}(\mathrm{d}(x, y)) \\ &\overset{n\to\infty}{\to} \int_{\mathcal{X}\times\mathcal{Y}} f(x, y)\pi(\mathrm{d}(x, y)) = \int_{\mathcal{X}} g(x)(\mathrm{proj}^x_\# \pi)(\mathrm{d}x), \end{aligned}$$

which implies $\mathrm{proj}^x_{\#}\pi = \mu$. Analogously, we obtain $\mathrm{proj}^y_{\#}\pi = \nu$ and thus $\pi \in \Pi_2(\mu,\nu)$. Therefore, $\Pi_2(\mu,\nu)$ is closed and thus also compact.

Now we show that a minimizer does indeed exist. In order to do so, let $(\pi_k)_{k\in\mathbb{N}}$ a minimizing sequence in $\Pi_2(\mu,\nu)$. By the compactness we may assume it is weakly converging to some $\pi \in \Pi_2(\mu,\nu)$. Choosing $h(x,y) := a(x) + b(y)$ the conditions of Lemma 4.3 are satisfied and thus

$$\int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi(\mathrm{d}(x,y)) \le \liminf_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi_k(\mathrm{d}(x,y)).$$

That is, π is a minimizer. □

Remark 4.5. The proof of Theorem 4.1 and the proofs of Lemma 4.3 and Lemma 4.4, as presented in Appendix A.1, do not rely on aspects of dimensionality. Thus, the assertions hold true if, for $n \in \mathbb{N}$, we replace $\mathcal{X}$ and $\mathcal{Y}$ by $\mathcal{X}_1,\ldots,\mathcal{X}_n$, $\mathcal{X}\times\mathcal{Y}$ by $\mathcal{X}_1\times\ldots\times\mathcal{X}_n$, etc. ◊

Definition 4.6. Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary sets and $\phi : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}\cup\{\infty\}$ a cost function. A set $\Gamma \subseteq \mathcal{X}\times\mathcal{Y}$ is said to be *ϕ-cyclically monotone*, if

$$\sum_{i=1}^{N} \phi(x_i, y_i) \le \sum_{i=1}^{N} \phi(x_i, y_{i+1})$$

holds for any $N \in \mathbb{N}$ and any family $(x_1,y_1),\ldots,(x_N,y_N)$ of points in Γ, where $y_{N+1} := y_1$. A transport plan is said to be *ϕ-cyclically monotone*, if it is concentrated on a ϕ-cyclically monotone set.

A ϕ-cyclically monotone transport plan can not be improved by simply rerouting the transported masses along a cycle. That is, the total cost is not reduced by transporting mass for example from x_1 to y_2 instead of y_1, from x_2 to y_3 instead of y_2, etc. We can think of this property as a local optimality criterion. Clearly, an optimal transport plan is ϕ-cyclically monotone. The converse is less obvious but still true by Theorem 4.7.

Before stating this result, we introduce the concept of the dual problem. While the primal problem aims at minimizing the transport cost, the dual problem aims at maximizing a certain profit. Indeed, assume a transport company may buy soil at location x for the price $\varphi(x)$ and sell it at location y for the price $\psi(y)$. Then the profit of this transport company is $\psi(y) - \varphi(x)$. Since the customer may transport the soil herself for the cost $\phi(x,y)$, in order to be competitive the transport company's profit should satisfy $\psi(y) - \varphi(x) \le \phi(x,y)$. The natural dual problem to the Monge-Kantorovich problem of optimal transport thus is

$$D_{MK}(\phi) := \sup\left\{\int_{\mathcal{Y}} \psi(y)\nu(\mathrm{d}y) - \int_{\mathcal{X}} \varphi(x)\mu(\mathrm{d}x)\right\},$$

where the supremum is taken over all functions $\psi : \mathcal{Y} \to \mathbb{R}$ and $\varphi : \mathcal{X} \to \mathbb{R}$ such that

$$\psi(y) - \varphi(x) \le \phi(x,y), \quad (x,y) \in \mathcal{X}\times\mathcal{Y}.$$

From a mathematical point of view, we should only consider functions $\varphi \in \mathbb{L}^1(\mathcal{X},\mu)$ and $\psi \in \mathbb{L}^1(\mathcal{Y},\nu)$.

We are now able to state several well-known results on classic optimal transport. All of the following results have counterparts in model-independent finance, some of which are proved in this thesis.

Theorem 4.7 ([77, Theorem 5.10])**.** *Let $(\mathcal{X},\mu)$ and $(\mathcal{Y},\nu)$ be two Polish probability spaces and let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\infty\}$ be a lower semi-continuous cost function such that*

$$\phi(x,y) \geq a(x) + b(y), \quad (x,y) \in \mathcal{X} \times \mathcal{Y},$$

for upper semi-continuous functions $a : \mathcal{X} \to \mathbb{R}$ and $b : \mathcal{Y} \to \mathbb{R}$ with $a \in \mathbb{L}^1(\mathcal{X},\mu)$ and $b \in \mathbb{L}^1(\mathcal{Y},\nu)$. Then:

1. *Strong duality holds, i.e.*

$$\begin{aligned}
&\min_{\pi \in \Pi_2(\mu,\nu)} \left\{ \int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi(\mathrm{d}(x,y)) \right\} = P_{MK}(\phi) = D_{MK}(\phi) \\
&= \sup_{(\varphi,\psi)\in \mathbb{L}^1(\mathcal{X},\mu)\times\mathbb{L}^1(\mathcal{Y},\nu)} \left\{ \int_{\mathcal{Y}} \psi(y)\nu(\mathrm{d}y) - \int_{\mathcal{X}} \varphi(x)\mu(\mathrm{d}x) \,\middle|\, \psi(y) - \varphi(x) \leq \phi(x,y) \right\} \\
&= \sup_{(\varphi,\psi)\in \mathcal{C}_b(\mathcal{X})\times\mathcal{C}_b(\mathcal{Y})} \left\{ \int_{\mathcal{Y}} \psi(y)\nu(\mathrm{d}y) - \int_{\mathcal{X}} \varphi(x)\mu(\mathrm{d}x) \,\middle|\, \psi(y) - \varphi(x) \leq \phi(x,y) \right\}.
\end{aligned}$$

2. *If ϕ is real-valued and the optimal transport cost $P_{MK}(\phi)$ is finite, then there is a measurable ϕ-cyclically monotone set $\Gamma \subseteq \mathcal{X} \times \mathcal{Y}$ such that for any $\pi \in \Pi_2(\mu,\nu)$ the following are equivalent.*
   a) *π is optimal for $P_{MK}(\phi)$.*
   b) *π is ϕ-cyclically monotone.*
   c) *π is concentrated on Γ.*
   d) *There exist functions $\varphi : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ and $\psi : \mathcal{Y} \to \mathbb{R} \cup \{-\infty\}$ such that $\psi(y) - \varphi(x) \leq \phi(x,y)$ for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$ with equality π-almost surely.*

Before stating a more general duality result, let us revisit some further results in the two marginal case that motivate the work of Beiglböck & Juillet [7] and Henry-Labordère & Touzi [38], which we present in Sections 6.1 and 6.2.

Theorem 4.8 ([7, Theorem 1.1])**.** *Let $\mu,\nu \in \mathcal{P}(\mathbb{R})$ and $\phi : \mathbb{R}^2 \to \mathbb{R}$ be a cost function defined by $\phi(x,y) = h(y-x)$, where $h : \mathbb{R} \to \mathbb{R}$ is a strictly convex function. Assume that there are functions $a \in \mathbb{L}^1(\mathbb{R},\mu)$ and $b \in \mathbb{L}^1(\mathbb{R},\nu)$ such that $\phi(x,y) \geq a(x) + b(y)$ for all $(x,y) \in \mathbb{R}^2$. If $P_{MK}(\phi)$ is finite, then for $\pi \in \Pi_2(\mu,\nu)$, the following are equivalent.*

1. *π is optimal for $P_{MK}(\phi)$.*

2. *π preserves the order, i.e. there is a set $\Gamma \subseteq \mathbb{R}^2$ with $\pi(\Gamma) = 1$ such that for $(x,y),(x',y') \in \Gamma$ if $x < x'$, then $y \leq y'$.*

We define for $\mu, \nu \in \mathcal{P}(\mathbb{R})$ the so-called Hoeffding-Fréchet transport plan

$$\pi_{HF}(B) := \left(F_\mu^{-1} \otimes F_\nu^{-1}\right)_\# \lambda_{[0,1]}(B), \quad B \in \mathcal{B}\left(\mathbb{R}^2\right),$$

and the increasing mapping $T_{HF} : \mathbb{R} \to \mathbb{R}, x \mapsto F_\nu^{-1} \circ F_\mu(x)$. Furthermore, we define functions $\varphi_{HF} : \mathbb{R} \to \mathbb{R}$ and $\psi_{HF} : \mathbb{R} \to \mathbb{R}$ by

$$\varphi_{HF}(x) := \phi(x, T_{HF}(x)) - \psi_{HF}(T_{HF}(x)) \quad \text{and} \quad \psi'_{HF}(y) := \phi_y(T_{HF}^{-1}(y), y).$$

Definition 4.9. A function $\phi : \mathbb{R}^2 \to \mathbb{R}$ satisfies the *Spence Mirrlees condition*, if the partial derivative ϕ_{xy} exists and satisfies $\phi_{xy} > 0$.

Theorem 4.10 ([38, Theorem 2.2])**.** *Let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be an upper semi-continuous cost function with linear growth, i.e. there exists a constant $K \in \mathbb{R}$ such that*

$$\phi(x, y) \leq K(1 + |x| + |y|), \quad (x, y) \in \mathbb{R}^2.$$

Assume that ϕ satisfies the Spence Mirrlees condition. Assume further that μ is continuous and that $\varphi_{HF} \in \mathbb{L}^1(\mathbb{R}, \mu)$ and $\psi_{HF} \in \mathbb{L}^1(\mathbb{R}, \nu)$. Finally, denote

$$\overline{P}_{MK}(\phi) := \sup_{\pi \in \Pi_2(\mu,\nu)} \int_{\mathcal{X} \times \mathcal{Y}} \phi(x, y) \pi(\mathrm{d}(x, y))$$

and by $\overline{D}_{MK}(\phi)$ the dual minimization problem. Then:

1. $\overline{P}_{MK}(\phi) = \overline{D}_{MK}(\phi) = \int_{\mathbb{R}} \phi\left(x, T_{HF}(x)\right) \mu(\mathrm{d}x)$.
2. $\varphi_{HF}(x) + \psi_{HF}(y) \geq \phi(x, y)$ *and* $(\varphi_{HF}, \psi_{HF})$ *is a solution for* $\overline{D}_{MK}(\phi)$.
3. $\pi_{HF}(\mathrm{d}x, \mathrm{d}y) = \mu(\mathrm{d}x)\delta_{T_{HF}(x)}(\mathrm{d}y)$ *is a solution for* $\overline{P}_{MK}(\phi)$ *and* π_{HF} *is the unique optimal transport plan.*

All results so far matched the standard market case. Now we shortly consider the general market case at least for $d = 1$. Therefore, let $\mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$. Then $\Pi_n(\mu_1, \ldots, \mu_n)$ is a convex and, by the same arguments as in the proof of Theorem 4.1, weakly compact subset of $\mathcal{P}(\mathbb{R}^n)$. We state a Kantorovich-type duality result of Kellerer [55] similar to Theorem 4.7 for this multi-marginal situation, as we need it in the proof of Theorem 5.1.

Theorem 4.11 ([55, Theorem 2.14])**.** *Let $\mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$. Then, for all lower semi-continuous functions $\phi : \mathbb{R}^n \to [0, \infty]$, we have*

$$\begin{aligned} P_{MK}^n(\phi) &:= \inf_{\pi \in \Pi_n(\mu_1,\ldots,\mu_n)} \left\{ \int_{\mathbb{R}^n} \phi(x_1, \ldots, x_n) \pi(\mathrm{d}(x_1, \ldots, x_n)) \right\} \\ &= \sup_{\varphi_j \in \mathbb{L}^1(\mathbb{R},\mu_j)} \left\{ \sum_{j=1}^n \int_{\mathbb{R}} \varphi_j(x_j) \mu_j(\mathrm{d}x_j) \;\middle|\; \varphi_1(x_1) + \ldots + \varphi_n(x_n) \leq \phi(x_1, \ldots, x_n) \right\} \\ &=: D_{MK}^n(\phi). \end{aligned}$$

4.2. Transfer to model-independent finance

In this section, we understand the connection of classic optimal transport and our studies. Therefore, let us consider the general market case for $d = 1$. Recall that we assume to know the price functions of call options with n different maturities.

The connection from model-independent finance to classic optimal transport is made by the so-called lemma of Breeden & Litzenberger. We first state it in a special case in which the assertion may be denoted compactly and hence the consequences are easy to fathom.

Lemma 4.12 ([13, Sec. 2])**.** *Let, in the situation of the general market case with $d = 1$, $\mathbb{Q} \in \mathcal{P}(\mathbb{R}^n)$ be consistent with the price functions of call options, i.e. for all $i = 1, \ldots, n$ and all $k \in \mathbb{R}$, we have*

$$C_i(k) = \int_{\mathbb{R}^n} \Phi_i(s_{t_1}, \ldots, s_{t_n}, k)\mathbb{Q}(\mathrm{d}(s_{t_1}, \ldots, s_{t_n})). \tag{4.2}$$

If $C_i \in \mathcal{C}^2(\mathbb{R})$, then we have

$$F_{S_{t_i}}(k) = \mathbb{Q}(S_{t_i} \leq k) = 1 + C_i'(k) \tag{4.3}$$

for the distribution function of S_{t_i} under $\mathbb{Q}$ and

$$f_{S_{t_i}}(k) = \mathbb{Q}(S_{t_i} \in dk) = C_i''(k) \tag{4.4}$$

for the associated density for all $i = 1, \ldots, n$ and all $k \in \mathbb{R}$.

Remark 4.13.

1. Without any formal proof of the above statement, we immediately see that the consistency condition in (4.2) implies a strong connection between the price function of the call option on S_{t_i} and its marginal distribution under the consistent measure $\mathbb{Q}$. Indeed, denoting the marginal of S_{t_i} under $\mathbb{Q}$ by μ_i, we may rewrite the right hand side of the consistency condition as

$$\int_{\mathbb{R}^n} \Phi_i(s_{t_1}, \ldots, s_{t_n}, k)\mathbb{Q}(\mathrm{d}(s_{t_1}, \ldots, s_{t_n})) = \int_{\mathbb{R}} \Phi_i(s_{t_1}, \ldots, s_{t_n}, k)\mu_i(\mathrm{d}s_{t_i}).$$

   Hence, the price function is determined by the associated marginal. This directly yields one part of the one-to-one connection between the two notions.

2. The differentiability assumption is only important for the exact representations in (4.3) and (4.4). The core assertion, namely the one-to-one connection between call option price functions and the marginal distributions of the underlying under any consistent measure, remains valid without it. ◊

Lemma 4.14 ([42, Lemma 2.2])**.** *In the situation of Lemma 4.12, drop the differentiability assumption on C_i. Then, for all $i = 1, \ldots, n$ and all $k \in \mathbb{R}$, we have*

$$F_{S_{t_i}}(k) = 1 + C_i'(k+).$$

Lemma 4.12 is implied by Lemma 4.14. The assertion of Lemma 4.14 is discussed in Hobson [42, Lemma 2.2]. It can also be derived from the following option price formula of Bick [11, Proposition 1]. Assume that an exotic option, that only depends on the underlying at some future time point T, has a payoff function $c : \mathbb{R} \to \mathbb{R}$ that is twice continuously differentiable except in a countable set of points. Then the price of the exotic option is

$$\int_{\mathbb{R}} c(x)C''(x)\mathrm{d}x + \sum_{a\in\mathcal{D}(C')} \left(C'(a+) - C'(a-)\right) c(a),$$

where C denotes the price function of a call option on the same underlying with maturity T and $\mathcal{D}(C')$ denotes the set of all points where the left and right derivatives of C differ.

As the assertion of Lemma 4.14 is well-known and the proof does not yield any insight, we do not report it here.

Remark 4.15. Before we make the connection of model-independent finance and classic optimal transport even clearer, let us formally prove the first and the third assertion of Lemma 3.4. Let us first assume that, for some call option price function $C : \mathbb{R} \to \mathbb{R}_+$, we have $C(k) \overset{k\to\infty}{\to} c > 0$. Then, for $k \in \mathbb{R}$, using μ as the marginal distribution of S_T under all consistent measures, we have

$$\begin{aligned} c \le C(k) &= \int_{\mathbb{R}} (s-k)^+ \mu(\mathrm{d}s) = \int_{[k,\infty)} s\mu(\mathrm{d}s) - \int_{[k,\infty)} k\mu(\mathrm{d}s) \\ &= \mathbb{E}_\mu \left[S_T \mathbb{1}_{\{S_T \ge k\}} \right] - k \cdot (1 - F_\mu(k)). \end{aligned}$$

The second summand tends to 0 by (2.3), as μ has a finite first moment. Thus, we have

$$\lim_{k\to\infty} \mathbb{E}_\mu \left[S_T \mathbb{1}_{\{S_T \ge k\}} \right] \ge c > 0,$$

a contradiction to the existence of the expected value of S_T. Similarly, for $k \in \mathbb{R}$, we have

$$\begin{aligned} C(k) + k &= \int_{\mathbb{R}} (s-k)^+ \mu(\mathrm{d}s) + k = \int_{[k,\infty)} (s-k)\mu(\mathrm{d}s) + \int_{\mathbb{R}} k\mu(\mathrm{d}s) \\ &= \int_{[k,\infty)} s\mu(\mathrm{d}s) + \int_{(-\infty,k)} k\mu(\mathrm{d}s) = \mathbb{E}_\mu \left[S_T \mathbb{1}_{\{S_T \ge k\}} \right] + k \cdot F_\mu(k). \end{aligned}$$

Clearly, the first summand tends to s_0 as $k \to -\infty$, while the second summand tends to 0 by the general version of (2.3) in the case $r = 1$. ♢

Let us now explain the connection of classic optimal transport and model-independent finance as well as the implications of Lemma 4.14 therein. The target is to price an exotic option characterized by its payoff function $c : \mathbb{R}^n \to \mathbb{R}$. In order to reach this, we search for potential pricing measures, i.e. measures that yield in a certain sense reasonable prices.

In order to obtain reasonable prices, such a measure has, in particular, to return observable market prices correctly. As we assume that call option prices are observable, potential pricing measures have to satisfy the consistency condition in (4.2). By Lemma 4.14, all such measures have the same marginals, that we denote by $\mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$.

This implies that the set of all potential pricing measures is a subset of $\Pi_n(\mu_1, \ldots, \mu_n)$, which formally connects model-independent finance with classic optimal transport. For obvious reasons, we call elements of $\Pi_n(\mu_1, \ldots, \mu_n)$ *consistent measures*.

However, $\Pi_n(\mu_1, \ldots, \mu_n)$ is not the set of all potential pricing measures, as such measures have to be martingale measures. That is, the (discounted) price process has to be a discrete-time martingale with respect to any potential pricing measure. As the measures in $\Pi_n(\mu_1, \ldots, \mu_n)$ do not satisfy this condition in general, we have to adapt classic optimal transport to the subset of all martingale measures.

Definition 4.16. Let $\mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$. The elements of the set

$$\begin{aligned}&\mathcal{M}_n(\mu_1, \ldots, \mu_n)\\ &\quad := \{\mathbb{Q} \in \Pi_n(\mu_1, \ldots, \mu_n) \mid \mathbb{E}_{\mathbb{Q}}[S_{t_{i+1}} \mid S_{t_i}, \ldots, S_{t_1}] = S_{t_i}\ \mathbb{Q} - \text{a.s.}, i = 1, \ldots, n-1\}\end{aligned}$$

are called *martingale transport plans* or *potential pricing measures*.

We may characterize the martingale property of a transport plan in several ways. For $\mathbb{Q} \in \Pi_n(\mu_1, \ldots, \mu_n)$, the following are equivalent.

1. $\mathbb{Q} \in \mathcal{M}_n(\mu_1, \ldots, \mu_n)$.
2. For all $1 \leq i \leq n-1$ and all $h \in \mathcal{C}_b(\mathbb{R}^i)$, we have

$$\mathbb{E}_{\mathbb{Q}}[h(S_{t_1}, \ldots, S_{t_i})(S_{t_{i+1}} - S_{t_i})] = 0.$$

Using the definition of martingale transport plans, we may consider the same problems as in classic optimal transport only with a restricted underlying set of probability measures. Clearly, we may hope that similar results as presented in the previous section apply for the martingale case as well.

4.3. Convex order and Strassen's theorem

Before we may state the adapted problems in a well-defined way, we need to consider the question if and under which conditions potential pricing measures do exist, i.e. if $\mathcal{M}_n(\mu_1, \ldots, \mu_n) \neq \emptyset$. This question shall be answered in this section.

Unfortunately, the answer is not as easy as for $\Pi_n(\mu_1, \ldots, \mu_n)$, as there is no obvious element. Indeed, the non-emptiness crucially depends on the measures $\mu_1, \ldots, \mu_n$ and only holds under additional assumptions. Therefore, we introduce the notion of convex order.

Definition 4.17. Two measures $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ are said to be in *convex order*, denoted by $\mu \leq_c \nu$, if for any convex function $f : \mathbb{R} \to \mathbb{R}$ such that the integrals exist,

$$\int_{\mathbb{R}} f(x)\mu(\mathrm{d}x) \leq \int_{\mathbb{R}} f(x)\nu(\mathrm{d}x).$$

Remark 4.18. Let $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ be two measures such that $\mu \leq_c \nu$. Then, choosing $f(x) = 1$, we have

$$\mu(\mathbb{R}) = \int_{\mathbb{R}} 1\mu(\mathrm{d}x) \leq \int_{\mathbb{R}} 1\nu(\mathrm{d}x) = \nu(\mathbb{R}).$$

With $f(x) = -1$, we also have $\mu(\mathbb{R}) \geq \nu(\mathbb{R})$. Hence, μ and ν have the same mass. Choosing $f(x) = x$, we get

$$\mathbb{B}(\mu) = \frac{1}{\mu(\mathbb{R})} \int_{\mathbb{R}} x\mu(\mathrm{d}x) \leq \frac{1}{\nu(\mathbb{R})} \int_{\mathbb{R}} x\nu(\mathrm{d}x) = \mathbb{B}(\nu).$$

With $f(x) = -x$, we also get $\mathbb{B}(\mu) \geq \mathbb{B}(\nu)$. Hence, μ and ν have the same barycentre. $\diamondsuit$

In order to state characterizations of convex order that are easier to be checked than the definition, we introduce the notions of potential functions and call option price functions.

Definition 4.19. Let $\mu \in \mathcal{P}_\alpha(\mathbb{R})$. Then the *potential function* of μ is defined by

$$u_\mu : \mathbb{R} \to \mathbb{R}_+, \quad x \mapsto \int_{\mathbb{R}} |t - x|\mu(\mathrm{d}t).$$

Proposition 4.20 ([7, Proposition 4.1])**.** *Let $\mu \in \mathcal{P}_\alpha(\mathbb{R})$ and say $\gamma = \mu(\mathbb{R})$ is the mass of μ and let $\beta = \frac{1}{\gamma} \int x\mu(\mathrm{d}x)$ be the barycentre of μ. Then:*

1. *u_μ is convex.*
2. $\lim_{x\to-\infty} u_\mu(x) - \gamma|x - \beta| = 0 = \lim_{x\to\infty} u_\mu(x) - \gamma|x - \beta|.$

Conversely, if f is a function with these properties for some $\beta \in \mathbb{R}$ and $\gamma \in (0, \infty)$, then there exists a unique measure $\mu \in \mathcal{P}_\alpha(\mathbb{R})$ such that $f = u_\mu$.

Proposition 4.21 ([7, Proposition 4.2])**.** *Let $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$. Then:*

1. *We have $\mu \leq \nu$ if and only if $u_\nu - u_\mu$ is convex.*
2. *A sequence $(\mu_k)_{k\in\mathbb{N}}$ of measures in $\mathcal{P}_\alpha(\mathbb{R})$ with mass γ and barycentre β converges in $\mathcal{P}_\alpha(\mathbb{R})$ weakly to some measure μ if and only if the sequence $(u_{\mu_k})_{k\in\mathbb{N}}$ of the potential functions converges pointwise to a function u that is the potential function of some measure $\mu' \in \mathcal{P}_\alpha(\mathbb{R})$. In that case, $\mu = \mu'$.*

Definition 4.22. Let $\mu \in \mathcal{P}_\alpha(\mathbb{R})$. Then the *call option price function* corresponding to μ is defined by

$$C_\mu : \mathbb{R} \to \mathbb{R}_+, \quad k \mapsto \int_{\mathbb{R}} (x - k)^+ \mu(\mathrm{d}x).$$

Remark 4.23. We used the notion of a call option price function with respect to some generic underlying $S = (S_{t_1}, \dots, S_{t_n})$ before and we denoted it by $C_{S_{t_i}}$, $i = 1, \dots, n$. If we now denote by μ_i the marginal distribution of S_{t_i} under any potential pricing measure, then by Lemma 4.14 it is obvious that we have $C_{\mu_i} \equiv C_{S_{t_i}}$. $\diamondsuit$

With these definitions, we may answer the question under which conditions the set of martingale transport plans is non-empty. The characterization using the convex order goes back to Strassen [73, Theorem 8].

Proposition 4.24. *Let $\mu_1, \dots, \mu_n \in \mathcal{P}_\alpha(\mathbb{R})$. Then the following are equivalent.*

1. $\mathcal{M}_n(\mu_1, \dots, \mu_n) \neq \emptyset$.
2. $\mu_1 \leq_c \dots \leq_c \mu_n$.

If we assume $\mu_1(\mathbb{R}) = \dots = \mu_n(\mathbb{R})$, then also the following is equivalent.

3. $u_{\mu_1} \leq \dots \leq u_{\mu_n}$.

If we additionally assume $\mathbb{B}(\mu_1) = \dots = \mathbb{B}(\mu_n)$, then also the following is equivalent.

4. $C_{\mu_1} \leq \dots \leq C_{\mu_n}$.

Remark 4.25. Further equivalent statements as well as general proofs can be found in Shaked & Shanthikumar [72, Chap. 3]. One of the additional equivalent statements in the case of equal total masses of the measures shall be mentioned here: There exist stochastic kernels $\kappa_t(x_1, \dots, x_{t-1}, \mathrm{d}x_t)$ such that for all $(x_1, \dots, x_{t-1}) \in \mathbb{R}^{t-1}$, $2 \leq t \leq n$,

$$\int_{\mathbb{R}} |x_t| \kappa_t(x_1, \dots, x_{t-1}, \mathrm{d}x_t) < \infty \quad \text{and} \quad \int_{\mathbb{R}} x_t \kappa_t(x_1, \dots, x_{t-1}, \mathrm{d}x_t) = x_{t-1},$$

and for all $1 \leq t \leq n$,

$$\mu_t = \mathrm{proj}_{\#}^{x_t}(\mu_1 \otimes \kappa_2 \otimes \dots \otimes \kappa_n). \qquad \Diamond$$

Remark 4.26. When it comes to application in finance, the characterization of the existence of martingale transport plans via call option price functions is important, as observable call option prices increase with the maturity, i.e. for $1 \leq i < j \leq n$, we have $C_{\mu_i} \leq C_{\mu_j}$.

Indeed, to get a contradiction assume that for $1 \leq i < j \leq n$ and some $k \in \mathbb{R}$, we have $C_{\mu_i}(k) > C_{\mu_j}(k)$. Recall that the measures μ_i and μ_j correspond to future underlying prices S_{t_i} and S_{t_j}. Denote by $C^{t_i}_{S_{t_j}}(k)$ the price of the call option on S_{t_j} with strike price k at time t_i. Then we have

$$C^{t_i}_{S_{t_j}}(k) \geq (S_{t_i} - k)^+$$

by usual no-arbitrage arguments. This gives rise to an arbitrage strategy. Indeed, in Table 4.1 we illustrate the payoffs of a strategy where at time $t = 0$ we buy a call option on S_{t_j} and sell a call option on S_{t_i} short and at time $t = t_i$ we resell the call option on S_{t_j}. This

Portfolio	Price in $t = 0$	Payoff in $t = t_i$
Call option on S_{t_j} with strike k long, sold in $t = t_i$	$C_{\mu_j}(k)$	$C^{t_i}_{S_{t_j}}(k)$
Call option on S_{t_i} with strike k short	-$C_{\mu_i}(k)$	$-(S_{t_i} - k)^+$
In total	$C_{\mu_j}(k) - C_{\mu_i}(k)$	$C^{t_i}_{S_{t_j}}(k) - (S_{t_i} - k)^+$
Value	< 0	≥ 0

Table 4.1.: Arbitrage strategy, if call option prices do not increase with maturity.

strategy yields a free lunch and hence a contradiction to no-arbitrage. Hence, we have $C_{\mu_{t_i}} \leq C_{\mu_{t_j}}$ for all $1 \leq i < j \leq n$.

Thus, we may assume $\mathcal{M}_n(\mu_1, \dots, \mu_n) \neq \emptyset$ without imposing too many restrictions. $\Diamond$

4.4. Pricing problems: The primal approach

Knowing that martingale transport plans exist under certain conditions, in this section we may formally introduce the price bound problems for the two market cases.

4.4.1. The general market case

Recall the general market from Section 3.1.1. Using the observable call option prices, we receive information about the marginal distributions of the vector of asset price processes S. We denote the marginal of $S^j_{t_i}$ with respect to any potential pricing measure by $\mu_{i,j} \in \mathcal{P}(\mathbb{R})$ and write $\mu = (\mu_{1,1}, \dots, \mu_{n,1}, \dots, \mu_{1,d}, \dots, \mu_{n,d})$.

Let us assume that marginals corresponding to the same underlying asset are in convex order as time increases. That is, we assume

$$\mu_{1,j} \leq_c \dots \leq_c \mu_{n,j} \tag{4.5}$$

for all $j = 1, \dots, d$. Then we denote by

$$\Pi^d_n(\mu) := \left\{ \pi \in \mathcal{P}\left(\mathbb{R}^{nd}\right) \,\middle|\, \pi\left(S^j_{t_i} \in B\right) = \mu_{i,j}(B),\ i = 1, \dots, n, j = 1, \dots, d, B \in \mathcal{B}(\mathbb{R}) \right\}$$

the set of all nd-dimensional transport plans with the desired marginals. We observe that $\Pi^d_n(\mu) = \Pi_{nd}(\mu_{1,1}, \dots, \mu_{n,1}, \dots, \mu_{1,d}, \dots, \mu_{n,d})$. Also, for all $j = 1, \dots, d$, we denote by

$${}_j\Pi_n(\mu) := \left\{ \pi \in \mathcal{P}\left(\mathbb{R}^{n}\right) \,\middle|\, \pi\left(S^j_{t_i} \in B\right) = \mu_{i,j}(B),\ i = 1, \dots, n, B \in \mathcal{B}(\mathbb{R}) \right\}$$

the set of all n-dimensional transport plans with the marginals of the asset price process S^j and observe that ${}_j\Pi_n(\mu) = \Pi_n(\mu_{1,j}, \dots, \mu_{n,j})$ for all $j = 1, \dots, d$.

Based on these classic transport plan sets, we define the sets of martingale transport plans. For $j = 1, \dots, d$, we denote by

$${}_j\mathcal{M}_n(\mu) := \left\{ \mathbb{Q} \in {}_j\Pi_n(\mu) \,\middle|\, \mathbb{E}_{\mathbb{Q}}\left[S^j_{t_{i+1}} \,\middle|\, S^j_{t_i}, \dots, S^j_{t_1} \right] = S^j_{t_i} \ \mathbb{Q}\text{-a.s.}, i = 1, \dots, n-1 \right\},$$

the set of all martingale transport plans with respect to the marginals of S^j. We observe that ${}_j\mathcal{M}_n(\mu) = \mathcal{M}_n(\mu_{1,j}, \dots, \mu_{n,j})$ for all $j = 1, \dots, d$.

Finally, we define the set of all martingale transport plans incorporating the marginals of the vector of asset price processes S

$$\begin{aligned} \mathcal{M} &:= \mathcal{M}^d_n(\mu) \\ &:= \left\{ \mathbb{Q} \in \Pi^d_n(\mu) \,\middle|\, \mathbb{E}_{\mathbb{Q}}\left[S^j_{t_{i+1}} \,\middle|\, S^j_{t_i}, \dots, S^j_{t_1} \right] = S^j_{t_i} \ \mathbb{Q}\text{-a.s.}, i = 1, \dots, n-1,\ j = 1, \dots, d \right\}, \end{aligned}$$

for which we may also write

$$\begin{aligned} \mathcal{M} &= \Pi_{nd}({}_1\mathcal{M}_n(\mu), \dots, {}_d\mathcal{M}_n(\mu)) \\ &:= \left\{ \pi \in \mathcal{P}\left(\mathbb{R}^{nd}\right) \,\middle|\, \pi \text{ has } n\text{-dimensional marginals } \pi_j \in {}_j\mathcal{M}_n(\mu), j = 1, \dots, d \right\}. \end{aligned}$$

By (4.5), we know that ${}_j\mathcal{M}_n(\mu) \neq \emptyset$ for all $j = 1, \ldots, d$. Hence, we have

$$\mathcal{M} \supseteq \bigotimes_{j=1}^{d} {}_j\mathcal{M}_n(\mu) \neq \emptyset.$$

Recall that $c : \mathbb{R}^{nd} \to \mathbb{R}$ is the payoff function of some exotic option with payoff

$$c\left(S^1_{t_1}, \ldots, S^1_{t_n}, \ldots, S^d_{t_1}, \ldots, S^d_{t_n}\right).$$

We define the primal problem of model-independent finance in the general market case, that we also call multi-asset and multi-marginal case. It is the problem of finding the lower bound for the price of the exotic option with payoff function c,

$$P(c) := \inf_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}_{\mathbb{Q}}\left[c(S^1_{t_1}, \ldots, S^1_{t_n}, \ldots, S^d_{t_1}, \ldots, S^d_{t_n})\right]. \tag{4.6}$$

As usual, we may also consider the problem of finding the upper price bound

$$\overline{P}(c) := \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}_{\mathbb{Q}}\left[c(S^1_{t_1}, \ldots, S^1_{t_n}, \ldots, S^d_{t_1}, \ldots, S^d_{t_n})\right].$$

4.4.2. The standard market case

Now recall the standard market. Let $\mu \in \mathcal{P}(\mathbb{R})$ be the marginal distribution corresponding to X and $\nu \in \mathcal{P}(\mathbb{R})$ the marginal distribution corresponding to Y. We assume $\mu \leq_c \nu$ and hence $\mathcal{M}_2(\mu, \nu) \neq \emptyset$. Then we define the upper price bound problem

$$P_2^c(\mu, \nu) := \sup_{\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X, Y)], \tag{4.7}$$

and the lower price bound problem

$$\underline{P}_2^c(\mu, \nu) := \inf_{\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X, Y)]. \tag{4.8}$$

Remark 4.27. We use different kinds of notations for the primal problems in the two cases, because in the chapter in which we deal with the general market case, the underlying marginals are always the same, while in the chapters in which we study the standard market case, we consider the problems for different marginals simultaneously. ◊

4.5. Hedging problems: The dual approach

From classical mathematical finance we know pricing-hedging dualities. We also know that duality results hold in classic optimal transport. Thus, it is natural to hope for similar results in martingale optimal transport and model-independent finance. In this section, we introduce the relevant hedging problems for the two market cases.

4.5.1. The general market case

Heuristically, the hedging problem that proves to be dual to the lower price bound problem in (4.6) is the problem of finding the most expensive hedging strategy that sub-replicates the payoff of an exotic option. Formally, this is the maximization problem

$$D(c) := \sup \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}\left(s_{t_i}^j\right) \mu_{i,j}\left(\mathrm{d}s_{t_i}^j\right) = \sup \sum_{j=1}^{d} \sum_{i=1}^{n} \mathbb{E}_{\mu_{i,j}}\left[\varphi_{i,j}(S_{t_i}^j)\right], \tag{4.9}$$

where the supremum is taken over functions

$$\varphi_{i,j} \in \mathcal{S} := \left\{ u : \mathbb{R} \to \mathbb{R} \;\middle|\; u(x) = a + bx + \sum_{\ell=1}^{m} c_\ell (x - k_\ell)^+,\ a, b, c_\ell, k_\ell \in \mathbb{R}, m \in \mathbb{N} \right\}$$

such that there are functions $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ with

$$\begin{aligned} &\Psi_{(\varphi_{i,j}),(h_i^j)}\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right) \\ &\quad := \sum_{j=1}^{d} \sum_{i=1}^{n} \varphi_{i,j}\left(s_{t_i}^j\right) + \sum_{j=1}^{d} \sum_{i=1}^{n-1} h_i^j\left(s_{t_1}^j, \dots, s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right) \\ &\quad \leq c\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right) \end{aligned}$$

for all $\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right) \in \mathbb{R}^{nd}$. From this sub hedging inequality, we directly derive the weak duality inequality $P(c) \geq D(c)$.

Analogously to the sub hedging problem and corresponding to the upper price bound problem, we define the super hedging problem

$$\overline{D}(c) := \inf \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}\left(s_{t_i}^j\right) \mu_{i,j}\left(\mathrm{d}s_{t_i}^j\right) = \inf \sum_{j=1}^{d} \sum_{i=1}^{n} \mathbb{E}_{\mu_{i,j}}\left[\varphi_{i,j}(S_{t_i}^j)\right],$$

where the infimum is again taken over functions $\varphi_{i,j} \in \mathcal{S}$ such that there are functions $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ with $\Psi_{(\varphi_{i,j}),(h_i^j)} \geq c$ on $\mathbb{R}^{nd}$.

Remark 4.28.
1. Analogously to Beiglböck, Henry-Labordère & Penkner [5], the dual problem can be formulated more general, considering $\mu_{i,j}$-integrable functions $\varphi_{i,j}$ and bounded, measurable functions h_i^j. Anyhow, we see later that $\varphi_{i,j} \in \mathcal{S}$ is sufficient to achieve duality. Additionally, $\mathcal{S}$ is the set of payoff functions that can be build up using only the risk-free and the risky assets, and the call options that we assume to be liquidly traded. This guarantees that hedging strategies are meaningful when it comes to application.

2. We could also consider functions $h_i^j \in \mathcal{C}_b(\mathbb{R}^{id})$ in the sense that the dynamic investment in any asset may depend on the history of all assets. As this enlarges the class of functions over which we optimize, $D(c)$ does only increase. In a setting different to ours, Lim [58] considers this exact approach. ◊

4.5.2. The standard market case

In the standard market case, we have dual problems that have an interpretation in the sense of hedging as well. Therefore, we define the sets of all admissible super and sub hedging strategies

$$\begin{aligned}\mathcal{D}_2^{\geq c} &:= \{(\varphi,\psi,h) \mid \varphi^+ \in \mathbb{L}^1(\mathbb{R},\mu), \psi^+ \in \mathbb{L}^1(\mathbb{R},\nu), h \in \mathbb{L}^0(\mathbb{R}),\\ &\qquad \varphi(x)+\psi(y)+h(x)(y-x) \geq c(x,y),\ (x,y)\in\mathbb{R}^2\},\\ \mathcal{D}_2^{\leq c} &:= \{(\varphi,\psi,h) \mid \varphi^+ \in \mathbb{L}^1(\mathbb{R},\mu), \psi^+ \in \mathbb{L}^1(\mathbb{R},\nu), h \in \mathbb{L}^0(\mathbb{R}),\\ &\qquad \varphi(x)+\psi(y)+h(x)(y-x) \leq c(x,y),\ (x,y)\in\mathbb{R}^2\}.\end{aligned}$$

Hedging strategies of this form are called semi-static hedging strategies, as φ and ψ may be interpreted as static investments in European options with maturity t and T respectively, while h may be understood as a dynamic investment in the underlying asset. Clearly, similar interpretations apply in the general market case.

Using such hedging strategies, we may define the super hedging problem

$$\begin{aligned}\overline{D}_2^c(\mu,\nu) &:= \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\int_{\mathbb{R}}\varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}}\psi(y)\nu(\mathrm{d}y)\right\} \\ &= \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\mathbb{E}_\mu[\varphi(X)] + \mathbb{E}_\nu[\psi(Y)]\right\},\end{aligned} \tag{4.10}$$

which is the dual problem to the upper price bound problem in (4.7). Analogously, we may define the sub hedging problem

$$\begin{aligned}\underline{D}_2^c(\mu,\nu) &:= \sup_{(\varphi,\psi,h)\in\mathcal{D}_2^{\leq c}} \left\{\int_{\mathbb{R}}\varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}}\psi(y)\nu(\mathrm{d}y)\right\} \\ &= \sup_{(\varphi,\psi,h)\in\mathcal{D}_2^{\leq c}} \left\{\mathbb{E}_\mu[\varphi(X)] + \mathbb{E}_\nu[\psi(Y)]\right\},\end{aligned} \tag{4.11}$$

which is the dual problem to the lower price bound problem in (4.8).

4.6. The first duality result

In this section, we discuss the pioneering duality result of martingale optimal transport by Beiglböck, Henry-Labordère & Penkner [5]. For this purpose, we describe the financial market and the assumptions considered by the authors. We do not go into too much detail, as they work in the general market case for $d = 1$.

Thus, let us consider the general market for $d = 1$. That is, we consider a discrete-time financial market liquidly trading a single risky asset $S = (S_t)_{t\in\mathcal{T}}$. Denote by $S_0 := s_0 \in \mathbb{R}$ the price of S at time $t = 0$. At the core of interest is a general exotic option with payoff function $c : \mathbb{R}^n \to \mathbb{R}$, the payoff of which is $c(S_{t_1},\dots,S_{t_n})$.

We assume that for all maturities $t_1,\dots,t_n$ and all strike prices $k \in \mathbb{R}$, call options on the underlying, i.e. (exotic) options with payoff function

$$\Phi_i : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}_+, \quad (s_{t_1}, \ldots, s_{t_n}, k) \mapsto (s_{t_i} - k)^+$$

are liquidly traded at time $t = 0$ with price $C_i(k)$. This has well-known implications for the marginals of any potential pricing measure.

The problem under consideration is to find the lower price bound for the exotic option

$$P_B(c) \coloneqq \inf_{\mathbb{Q} \in \mathcal{M}_n(\mu_1, \ldots, \mu_n)} \mathbb{E}_{\mathbb{Q}}[c(S_{t_1}, \ldots, S_{t_n})].$$

The dual problem considers semi-static sub hedging strategies, i.e. payoff functions of the form

$$\Psi_{(\varphi_i),(h_i)}(s_{t_1}, \ldots, s_{t_n}) \coloneqq \sum_{i=1}^{n} \varphi_i(s_{t_i}) + \sum_{i=1}^{n-1} h_i(s_{t_1}, \ldots, s_{t_i})(s_{t_{i+1}} - s_{t_i}),$$

where the functions $\varphi_i : \mathbb{R} \to \mathbb{R}$ are assumed to be μ_i-integrable and the functions $h_i : \mathbb{R}^i \to \mathbb{R}$ are bounded and measurable. If payoff functions of that kind and additionally sub-replicating the payoff of the exotic option in the sense $\Psi_{(\varphi_i),(h_i)} \leq c$ do exist, then for any martingale transport plan $\mathbb{Q} \in \mathcal{M}_n(\mu_1, \ldots, \mu_n)$ holds the inequality

$$\mathbb{E}_{\mathbb{Q}}[c(S_{t_1}, \ldots, S_{t_n})] \geq \mathbb{E}_{\mathbb{Q}}[\Psi_{(\varphi_i),(h_i)}(S_{t_1}, \ldots, S_{t_n})] = \mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^{n} \varphi_i(S_{t_i})\right] = \sum_{i=1}^{n} \mathbb{E}_{\mu_i}\left[\varphi_i(S_{t_i})\right].$$

This leads to the dual problem of finding the most expensive sub hedging strategy

$$D_B(c) \coloneqq \sup_{\varphi_i \in \mathcal{S}} \left\{ \sum_{i=1}^{n} \mathbb{E}_{\mu_i}[\varphi_i(S_{t_i})] \,\middle|\, \exists h_i \in \mathcal{C}_b(\mathbb{R}^i) : \Psi_{(\varphi_i),(h_i)} \leq c \text{ on } \mathbb{R}^n \right\},$$

for which we have $P_B(c) \geq D_B(c)$. Beiglböck, Henry-Labordère & Penkner [5] prove the following strong duality theorem for the above pair of optimization problems.

Theorem 4.29 ([5, Theorem 1.1])**.** *Let $\mu_1, \ldots \mu_n$ be Borel probability measures on $\mathbb{R}$ such that $\mathcal{M}_n(\mu_1, \ldots, \mu_n) \neq \emptyset$. Let further $c : \mathbb{R}^n \to (-\infty, \infty]$ be a lower semi-continuous function such that there exists a constant $K \in \mathbb{R}$ with*

$$c(s_{t_1}, \ldots, s_{t_n}) \geq -K(1 + |s_{t_1}| + \ldots + |s_{t_n}|) \tag{4.12}$$

on $\mathbb{R}^n$. Then $P_B(c) = D_B(c)$. Also $P_B(c)$ is attained, i.e. there is a martingale transport plan $\mathbb{Q}^* \in \mathcal{M}_n(\mu_1, \ldots, \mu_n)$ such that $P_B(c) = \mathbb{E}_{\mathbb{Q}^*}[c]$. In general, $D_B(c)$ is not attained.

Chapter 5

Duality in the general market

In this chapter, our central aim is to generalize the pioneering duality result of model-independent finance using martingale optimal transport, namely Theorem 4.29, to the general market case. This is necessary in order to be able to treat multi-asset options such as basket options. This generalization is presented in Section 5.1, where we also list and illustrate some drawbacks of the theory. These are the possible non-existence of optimizers for the dual problem and the potentially large deviations of the upper and lower price bounds. In Section 5.2, we discuss several recent results on conditions under which dual optimizers do exist, some of which are important in Chapters 6 and 7. Finally, in Section 5.3, we investigate the question, how it is possible to tighten the price bound gap. We discuss recent progress using information on asset return variances in the situation considered in Theorem 4.29. Then we transfer some of the results to the general market case additionally using information on asset return covariances.

5.1. The general duality result

In this section, we generalize Theorem 4.29. Before we state and prove the generalized result, we discuss the general market and the associated optimization problems and detail the differences to the situation and the problems considered in Section 4.6.

As notations are similar, we indicate by an indexed B whenever we consider objects from that section, as the results are from Beiglböck, Henry-Labordère & Penkner [5]. That is, we write $\mathcal{M}_n^B(\mu)$ instead of $\mathcal{M}_n(\mu_1, \dots, \mu_n)$, etc. We discuss the differences that appear in the primal and dual problems respectively. For this purpose, we adapt the single asset situation from dimension n to dimension nd. That is, then there are nd instead of n trading times and the dimensions of the compared problems are equal. We stress that the differences result from the introduction of multiple underlying assets.

1. Primal problem: The formal difference is induced by the different sets from which the optimizing measures can come. Indeed, while $\Pi^B_{nd}(\mu)$ and $\Pi^d_n(\mu)$ are equal, the sets $\mathcal{M}^B_{nd}(\mu)$ and $\mathcal{M}$ differ even for the same dimension.

   Indeed, in $\mathcal{M}^B_{nd}(\mu)$ we have martingale conditions connecting all marginals, as the marginals correspond to the same underlying risky asset. Counting the conditions leads to $nd-1$ conditions of the form

   $$\mathbb{E}_{\mathbb{Q}}[S_{t_{i+1}} \mid S_{t_i} \dots S_{t_1}] = S_{t_i} \ \mathbb{Q}\text{-a.s.}, \quad i = 1, \dots, nd-1.$$

   Conversely, in $\mathcal{M}$ we have martingale conditions only connecting the marginals corresponding to the same underlying asset. Marginals of different assets are in no specific relation at all. Formally, for all $j = 1, \dots, d$, we have the $n-1$ conditions

   $$\mathbb{E}_{\mathbb{Q}}\left[S^j_{t_{i+1}} \mid S^j_{t_i}, \dots, S^j_{t_1}\right] = S^j_{t_i} \ \mathbb{Q}\text{-a.s.}, \quad i = 1, \dots, n-1.$$

   Hence, we have $d(n-1)$ conditions, i.e. $d-1$ fewer than in the single asset case.

2. Dual problem: As in the setting of Section 4.6 the nd marginals correspond to the same risky asset, the dynamic part of the hedging strategies contains $nd-1$ single timestep investments in the underlying. In the general market case, the hedging strategies only consist of $n-1$ such investments for every risky asset.

   In fact, the $d-1$ missing dynamic investment possibilities would allow trading of the form $h^j_n(\cdot)\left(S^{j+1}_{t_1} - S^j_{t_n}\right)$, $j = 1, \dots, d-1$, in the general market case. Though such swap-type strategies may be considered somehow, we only allow for classic dynamic hedging strategies corresponding to one single asset.

The mentioned differences of our problems to those introduced by Beiglböck, Henry-Labordère & Penkner [5] induce value changes for $P(c)$ and $D(c)$. As for the sets of considered martingale transport plans we have $\mathcal{M} \supseteq \mathcal{M}^B_{nd}(\mu)$, $P(c)$ decreases from the single asset case to the multi-asset case, i.e. $P(c) \leq P_B(c)$. However, in the multi-asset case we have fewer possible dynamic hedging strategies such that $D(c)$ decreases as well, i.e. $D(c) \leq D_B(c)$. Thus, we may still hope to achieve duality.

The key result in this chapter is the desired duality result for the optimization problems introduced in (4.6) and (4.9). It guarantees that the lower bound for the price of the exotic option and the price of the most expensive hedging strategy that sub-replicates the payoff of the exotic option in a pointwise sense are equal. It generalizes Theorem 4.29 as it allows the exotic option to depend on more than only one asset. In particular, it allows us to incorporate basket options.

Theorem 5.1. *Let $\mathcal{M} \neq \emptyset$ and $c : \mathbb{R}^{nd} \to (-\infty, \infty]$ be a lower semi-continuous payoff function such that there is a constant $K \in \mathbb{R}$ with*

$$c\left(s^1_{t_1}, \ldots, s^1_{t_n}, \ldots, s^d_{t_1}, \ldots, s^d_{t_n}\right) \geq -K \left(1 + \sum_{j=1}^{d} \sum_{i=1}^{n} \left|s^j_{t_i}\right|\right) \tag{5.1}$$

for all $(s^1_{t_1}, \ldots, s^1_{t_n}, \ldots, s^d_{t_1}, \ldots, s^d_{t_n}) \in \mathbb{R}^{nd}$. Then $P(c) = D(c)$ and there is a $\mathbb{Q}^* \in \mathcal{M}$ such that $P(c) = \mathbb{E}_{\mathbb{Q}^*}[c]$.

Considering $\tilde{c} \coloneqq -c$, we get the corresponding duality result for the upper price bound and the cheapest super-replicating hedging strategy.

Corollary 5.2. *Let $\mathcal{M} \neq \emptyset$ and $c : \mathbb{R}^{nd} \to [-\infty, \infty)$ be an upper semi-continuous payoff function such that there is a constant $K \in \mathbb{R}$ with*

$$c\left(s^1_{t_1}, \ldots, s^1_{t_n}, \ldots, s^d_{t_1}, \ldots, s^d_{t_n}\right) \leq K \left(1 + \sum_{j=1}^{d} \sum_{i=1}^{n} \left|s^j_{t_i}\right|\right)$$

for all $(s^1_{t_1}, \ldots, s^1_{t_n}, \ldots, s^d_{t_1}, \ldots, s^d_{t_n}) \in \mathbb{R}^{nd}$. Then $\overline{P}(c) = \overline{D}(c)$ and there is a $\mathbb{Q}^* \in \mathcal{M}$ such that $\overline{P}(c) = \mathbb{E}_{\mathbb{Q}^*}[c]$.

Before we prove Theorem 5.1 on page 44, we collect several auxiliary results. First, we introduce a duality result for classic optimal transport, that we use in the proof of Theorem 5.1 with the correct dimensionality, $\pi \in \Pi^d_n(\mu)$ and a certain choice for the function ϕ.

Proposition 5.3 ([5, Proposition 2.1])**.** *Let $\phi : \mathbb{R}^n \to (-\infty, \infty]$ be a lower semi-continuous function such that there is a constant $K \in \mathbb{R}$ with*

$$\phi(x_1, \ldots, x_n) \geq -K(1 + |x_1| + \ldots + |x_n|), \quad (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

and let $\mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$. Then

$$\begin{aligned} P^n_{MK}(\phi) &= \inf_{\pi \in \Pi_n(\mu_1, \ldots, \mu_n)} \left\{ \int_{\mathbb{R}^n} \phi(x_1, \ldots, x_n) \pi(\mathrm{d}(x_1, \ldots, x_n)) \right\} \\ &= \sup_{\varphi_i \in \mathcal{S}} \left\{ \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_i(x_i) \mu_i(\mathrm{d}x_i) \;\middle|\; \varphi_1(x_1) + \ldots + \varphi_n(x_n) \leq \phi(x_1, \ldots, x_n) \right\} \\ &= D^n_{MK}(\phi). \end{aligned}$$

By Theorem 4.11, the main task in the proof is to show that it suffices to consider functions $\varphi_i \in \mathcal{S}$ instead of functions $\varphi_i \in \mathbb{L}^1(\mathbb{R}, \mu_i)$. In order to do so, we need Lemma 2.1 and the following approximation lemma, the proof of which is reported in Appendix A.2.

Lemma 5.4. *Let $f \in \mathcal{C}_b(\mathbb{R})$, $\mu_1, \ldots, \mu_n \in \mathcal{P}(\mathbb{R})$ and $\varepsilon > 0$. Then there is a function $u \in \mathcal{S}$ such that $u \leq f$ on $\mathbb{R}$ and, for all $i = 1, \ldots, n$, we have*

$$\int_{\mathbb{R}} f(x) - u(x) \mu_i(\mathrm{d}x) < \varepsilon.$$

Proof of Proposition 5.3. In the proof of Theorem 5.1, we show that without loss of generality we may assume $\phi \geq 0$. By Lemma 5.4, we may expand the class of admissible sub hedging functions from $\mathcal{S}$ to $\mathcal{C}_b(\mathbb{R})$. Hence, we have to show

$$P^n_{MK}(\phi) = \sup_{\varphi_i \in \mathcal{C}_b(\mathbb{R})} \left\{ \sum_{i=1}^n \int \varphi_i(x_i)\mu_i(\mathrm{d}x_i) \;\middle|\; \varphi_1(x_1) + \ldots \varphi_n(x_n) \leq \phi(x_1, \ldots, x_n) \right\}. \quad (5.2)$$

We start the proof of (5.2) assuming $\phi \in \mathcal{C}_c(\mathbb{R}^n)$. Then in particular, ϕ is lower semi-continuous. By Theorem 4.11, for all $\eta > 0$ there are $\varphi_i \in \mathbb{L}^1(\mathbb{R}, \mu_i)$, $i = 1, \ldots, n$, such that $\varphi_1 + \ldots + \varphi_n \leq \phi$ and

$$P^n_{MK}(\phi) - \sum_{i=1}^n \int_{\mathbb{R}} \varphi_i(x_i)\mu_i(\mathrm{d}x_i) \leq \eta.$$

As $\mathcal{C}_c(\mathbb{R})$ is dense in $\mathbb{L}^1(\mathbb{R}, \mu_i)$, $i = 1, \ldots, n$, see for example [78, Lemma V.1.10], we may as well assume $\varphi_1, \ldots, \varphi_n \in \mathcal{C}_c(\mathbb{R})$. Hence, ϕ, $\varphi_1, \ldots, \varphi_n$ are uniformly bounded in particular.

Now we iteratively replace $\varphi_1, \ldots, \varphi_n \in \mathcal{C}_c(\mathbb{R})$ by $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n \in \mathcal{C}_b(\mathbb{R})$. We define

$$\begin{aligned} \tilde{\varphi}_1(x_1) &:= \varphi_1(x_1) + \inf_{x_2, \ldots, x_n \in \mathbb{R}} \left\{ \phi(x_1, \ldots, x_n) - \sum_{i=1}^n \varphi_i(x_i) \right\} \\ &= \inf_{x_2, \ldots, x_n \in \mathbb{R}} \left\{ \phi(x_1, \ldots, x_n) - \sum_{i=2}^n \varphi_i(x_i) \right\} =: \inf_{x_2, \ldots, x_n \in \mathbb{R}} H(x_1, \ldots, x_n). \end{aligned} \quad (5.3)$$

Let us prove $\tilde{\varphi}_1 \in \mathcal{C}_b(\mathbb{R})$. Clearly, the boundedness of $\tilde{\varphi}_1$ holds by definition as a sum of bounded functions.

Now let us consider the continuity of $\tilde{\varphi}_1$. For all $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x, x' \in \mathbb{R}$ with $|x - x'| < \delta$, we have

$$|H(x, x_2, \ldots, x_n) - H(x', x_2, \ldots, x_n)| = |\phi(x, x_2, \ldots, x_n) - \phi(x', x_2, \ldots, x_n)| < \varepsilon \quad (5.4)$$

for all $x_2, \ldots, x_n \in \mathbb{R}$, as ϕ is uniformly continuous. Hence, we also have

$$|\tilde{\varphi}_1(x) - \tilde{\varphi}_1(x')| = \left| \inf_{x_2, \ldots, x_n \in \mathbb{R}} H(x, x_2, \ldots, x_n) - \inf_{x_2, \ldots, x_n \in \mathbb{R}} H(x', x_2, \ldots, x_n) \right| \leq \varepsilon.$$

In order to prove the last inequality by contradiction, let us assume without loss of generality that for some $\varepsilon_0 > 0$ and all $\delta > 0$ there are $x, x' \in \mathbb{R}$ with $|x - x'| < \delta$ such that

$$\tilde{\varphi}_1(x) - \tilde{\varphi}_1(x') > \varepsilon_0. \quad (5.5)$$

Now let $\delta_0 = \delta_0(\varepsilon_0)$ be such that (5.4) is satisfied and $x_0 = x_0(\delta_0), x'_0 = x'_0(\delta_0) \in \mathbb{R}$ be such that (5.5) is satisfied. Let also $(x^{(k)})_{k \in \mathbb{N}}$ with $x^{(k)} = (x_2^{(k)}, \ldots, x_n^{(k)}) \in \mathbb{R}^{n-1}$ be a minimizing sequence for $\tilde{\varphi}_1(x'_0)$, i.e. $\tilde{\varphi}_1(x'_0) = \lim_{k \to \infty} H(x'_0, x^{(k)})$. As $|x_0 - x'_0| < \delta_0$, by (5.4) we have $-\varepsilon_0 < H(x_0, x^{(k)}) - H(x'_0, x^{(k)}) < \varepsilon_0$ for all $k \in \mathbb{N}$. Taking the limit we get

$$-\varepsilon_0 \leq \lim_{k \to \infty} H\left(x_0, x^{(k)}\right) - \tilde{\varphi}_1(x'_0) \leq \varepsilon_0.$$

However, contradicting the definition of $\tilde{\varphi}_1$, this implies by (5.5) that we have

$$\lim_{k\to\infty} H\left(x_0, x^{(k)}\right) < \tilde{\varphi}_1(x_0).$$

Hence, the assumption in (5.5) is wrong and thus the continuity of $\tilde{\varphi}_1$ holds.

Finally, we prove $\tilde{\varphi}_1$ to be a reasonable replacement function for φ_1. By (5.3), we have

$$\begin{aligned} &\tilde{\varphi}_1(x_1) + \varphi_2(x_2) + \ldots + \varphi_n(x_n) \leq \phi(x_1, \ldots, x_n) \\ \iff &\inf_{x_2,\ldots,x_n} \left\{ \phi(x_1, \ldots, x_n) - \sum_{i=2}^{n} \varphi_i(x_i) \right\} \leq \phi(x_1, \ldots, x_n) - \sum_{i=2}^{n} \varphi_i(x_i), \end{aligned}$$

where the latter inequality is trivially satisfied. As $\tilde{\varphi}_1 \in \mathcal{C}_b(\mathbb{R})$, we have $\tilde{\varphi}_1 \in \mathbb{L}^1(\mathbb{R}, \mu_1)$ as well. Finally, as $\phi \geq \varphi_1 + \ldots + \varphi_n$, we have $\tilde{\varphi}_1 \geq \varphi_1$ by definition. Thus, $\tilde{\varphi}_1$ is suitable for the optimization problem and improves the value of $D^n_{MK}(\phi)$.

Iterating this procedure for $i = 2, \ldots, n$, we replace $\varphi_i(x_i)$ by

$$\tilde{\varphi}_i(x_i) \coloneqq \inf_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n \in \mathbb{R}} \left\{ \phi(x_1, \ldots, x_n) - \sum_{j=1}^{i-1} \tilde{\varphi}_j(x_j) - \sum_{j=i+1}^{n} \varphi_j(x_j) \right\},$$

and thus arrive at the duality in (5.2) for $\phi \in \mathcal{C}_c(\mathbb{R}^n)$.

Now let $\phi \in \mathcal{C}_b(\mathbb{R}^n)$ be non-negative. If we manage to prove the duality in (5.2) in this case, using Lemma 2.1 and the exact same arguments as in the proof of Theorem 5.1, we may transfer the result to general lower semi-continuous functions $\phi : \mathbb{R}^n \to [0, \infty]$.

Recall that by definition we have $P^n_{MK}(\phi) \geq D^n_{MK}(\phi)$. To get the opposite inequality, we observe that $\Pi_n(\mu_1, \ldots, \mu_n)$ is tight by Lemma 4.4 and Remark 4.5. Thus, for every $m \in \mathbb{N}$ there is a compact set $\tilde{K}_m \subseteq \mathbb{R}^n$ such that for all $\pi \in \Pi_n(\mu_1, \ldots, \mu_n)$ we have $\pi(\mathbb{R}^n \setminus \tilde{K}_m) \leq \frac{1}{m}$. Furthermore, we define $K_m \coloneqq \tilde{K}_m \cup [-m, m]^n$. Then we clearly have

$$\pi\left(\mathbb{R}^n \setminus K_m\right) \leq \frac{1}{m}$$

for all $\pi \in \Pi_n(\mu_1, \ldots, \mu_n)$. For $m \in \mathbb{N}$, we define

$$\tilde{\phi}_m(x_1, \ldots, x_n) \coloneqq \begin{cases} \phi(x_1, \ldots, x_n), & (x_1, \ldots, x_n) \in K_m \\ 0, & (x_1, \ldots, x_n) \notin K_m. \end{cases}$$

Clearly, then $\tilde{\phi}_m$ has compact support and satisfies $0 \leq \tilde{\phi}_m \leq \phi$ for all $m \in \mathbb{N}$. We also have $\tilde{\phi}_m(x_1, \ldots, x_n) \nearrow \phi(x_1, \ldots, x_n)$ for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$ as $m \to \infty$. Smoothing $\tilde{\phi}_m$, we achieve a continuous function ϕ_m that satisfies the following conditions.

1. $\phi_m = \phi$ on K_m.

2. $0 \leq \tilde{\phi}_m \leq \phi_m \leq \phi$ on $\mathbb{R}^n$.

3. ϕ_m has compact support.

We thus have a sequence $(\phi_m)_{m\in\mathbb{N}}$ in $\mathcal{C}_c(\mathbb{R}^n)$ such that $\phi_m(x_1,\dots,x_n) \nearrow \phi(x_1,\dots,x_n)$ for all $(x_1,\dots,x_n) \in \mathbb{R}^n$ as $m \to \infty$. Let $\pi \in \Pi_n(\mu_1,\dots,\mu_n)$. Then we have

$$\begin{aligned}\int_{\mathbb{R}^n} &\phi(x_1,\dots,x_n) - \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))\\ &= \int_{K_m} \phi(x_1,\dots,x_n) - \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))\\ &\quad+ \int_{\mathbb{R}^n\setminus K_m} \phi(x_1,\dots,x_n) - \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))\\ &= \int_{\mathbb{R}^n\setminus K_m} \phi(x_1,\dots,x_n) - \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))\\ &\le B\cdot\pi\left(\mathbb{R}^n\setminus K_m\right) \le \frac{B}{m},\end{aligned}$$

where $B \ge 0$ is the smallest bound for the bounded function ϕ. Hence, we have

$$\begin{aligned}\int_{\mathbb{R}^n} \phi(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n)) &= \int_{\mathbb{R}^n} \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))\\ &\quad+ \int_{\mathbb{R}^n} \phi(x_1,\dots,x_n) - \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))\\ &\le \int_{\mathbb{R}^n} \phi_m(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n)) + \frac{B}{m}.\end{aligned}$$

Applying the infimum over all $\pi \in \Pi_n(\mu_1,\dots,\mu_n)$ on both sides, we get the first step of

$$P^n_{MK}(\phi) \le P^n_{MK}(\phi_m) + \frac{B}{m} = D^n_{MK}(\phi_m) + \frac{B}{m} \le D^n_{MK}(\phi) + \frac{B}{m},$$

where we use the duality result for $\phi_m \in \mathcal{C}_c(\mathbb{R}^n)$ and $\phi_m \le \phi$ in the second and the third step. With $m \to \infty$ we get the desired duality for $\phi \in \mathcal{C}_b(\mathbb{R}^n)$. □

With the adapted multi-marginal duality result of classic optimal transport proved, we proceed to collect further useful facts. We recall that $\Pi^d_n(\mu)$ is a compact and convex set. Consider $\mathcal{M} \subseteq \Pi^d_n(\mu)$ and assume $\mathcal{M} \neq \emptyset$. We show that $\mathcal{M}$ is as well compact with respect to the weak topology $\mathcal{T}_{cb}(\mathbb{R}^{nd})$. For this purpose, we need two lemmata. While the second lemma has to be adapted to our more general situation, the first lemma could be used as stated by Beiglböck, Henry-Labordère & Penkner in [5, Lemma 2.2]. However, we provide a more general version and report the proof in Appendix A.2.

Lemma 5.5. *Let $f \in \mathcal{C}(\mathbb{R}^n)$ and $\mu = (\mu_1,\dots,\mu_n) \in \mathcal{P}(\mathbb{R})^n$ be such that there are functions $f_i \in \mathbb{L}(\mathbb{R},\mu_i)$, $i = 1,\dots,n$, and a constant $K \in \mathbb{R}$ with*

$$|f(x_1,\dots,x_n)| \le K\left(1 + \sum_{i=1}^n f_i(x_i)\right), \quad (x_1,\dots,x_n) \in \mathbb{R}^n. \tag{5.6}$$

Then the mapping

$$\pi \mapsto \int_{\mathbb{R}^n} f(x_1,\dots,x_n)\pi(\mathrm{d}(x_1,\dots,x_n))$$

is continuous on $\Pi_n(\mu_1,\dots,\mu_n)$ with respect to the weak topology $\mathcal{T}_{cb}(\mathbb{R}^n)$.

Remark 5.6. Observe that the assertion of Lemma 5.5 holds in particular for the choice $f_i(x_i) = |x_i|$. That is, what is shown by Beiglböck, Henry-Labordère & Penkner [5] and it is all we need in the proof of Theorem 5.1. $\Diamond$

Lemma 5.7. *Let* $\mu = (\mu_{1,1}, \dots, \mu_{n,1}, \dots, \mu_{1,d}, \dots, \mu_{n,d}) \in \mathcal{P}(\mathbb{R})^{nd}$ *and* $\pi \in \Pi_n^d(\mu)$. *Then the following are equivalent:*

1. $\pi \in \mathcal{M} = \mathcal{M}_n^d(\mu)$.

2. *For all* $j = 1, \dots, d$, *all* $i = 1, \dots, n-1$ *and all* $h \in \mathcal{C}_b(\mathbb{R}^i)$, *we have*

$$\int_{\mathbb{R}^{nd}} h\left(s_{t_1}^j, \dots, s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right)\pi\left(\mathrm{d}\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right)\right) = 0.$$

Proof. 1. $\Rightarrow$ 2. Assume $\pi \in \mathcal{M}$. Then, by the definition of $\mathcal{M}$, we have

$$\begin{aligned}&\int_{\mathbb{R}^{nd}} h\left(s_{t_1}^j, \dots, s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right)\pi\left(\mathrm{d}\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right)\right)\\ &\quad= \mathbb{E}_\pi\left[\mathbb{E}_\pi\left[h(S_{t_1}^j, \dots, S_{t_i}^j)(S_{t_{i+1}}^j - S_{t_i}^j)\;\middle|\; S_{t_1}^j, \dots, S_{t_i}^j\right]\right] = 0\end{aligned}$$

for all $h \in \mathcal{C}_b(\mathbb{R}^i)$, all $j = 1, \dots, d$ and all $i = 1, \dots, n-1$.

2. $\Rightarrow$ 1. Assume $\pi \notin \mathcal{M}$. Then, by the properties of $\mathcal{M}$, there is a $j \in \{1, \dots, d\}$ and an $i \in \{1, \dots, n-1\}$ such that we have

$$\mathbb{E}_\pi\left[S_{t_{i+1}}^j - S_{t_i}^j \;\middle|\; S_{t_1}^j, \dots, S_{t_i}^j\right] \neq 0,$$

which is a contradiction to the second condition. $\square$

Proposition 5.8. *The set* $\mathcal{M}$ *is compact with respect to the weak topology* $\mathcal{T}_{cb}(\mathbb{R}^{nd})$.

Proof. As $\Pi_n^d(\mu)$ is compact and $\mathcal{M} \subseteq \Pi_n^d(\mu)$, it suffices to show that $\mathcal{M}$ is a closed set. Therefore, for all $i = 1, \dots, n-1$, all $j = 1, \dots, d$ and all $h \in \mathcal{C}_b(\mathbb{R}^i)$, we denote $h_i^j\left(s_{t_1}^j, \dots, s_{t_i}^j\right) := h\left(s_{t_1}^j, \dots, s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right)$. We also define

$$\mathcal{M}(h,i,j) := \left\{\pi \in \Pi_n^d(\mu) \;\middle|\; \int_{\mathbb{R}^{nd}} h_i^j\left(s_{t_1}^j, \dots, s_{t_i}^j\right)\pi\left(\mathrm{d}\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right)\right) = 0\right\}.$$

Then, by Lemma 5.7, we have $\mathcal{M} = \bigcap_{i=1}^{n-1}\bigcap_{j=1}^{d}\bigcap_{h\in\mathcal{C}_b(\mathbb{R}^i)}\mathcal{M}(h,i,j)$. The functions h_i^j are continuous by definition and they satisfy

$$\left|h_i^j\left(s_{t_1}^j, \dots, s_{t_i}^j\right)\right| \le B(h)\left|s_{t_{i+1}}^j - s_{t_i}^j\right| \le B(h)\left(\left|s_{t_{i+1}}^j\right| + \left|s_{t_i}^j\right|\right),$$

where $B(h) \ge 0$ is the smallest bound for h. Hence, the functions satisfy the conditions of Lemma 5.5 such that $\pi \mapsto \int h_i^j \mathrm{d}\pi$ is continuous. In particular, $\mathcal{M}(h,i,j)$ is closed for all $i = 1, \dots, n-1$, all $j = 1, \dots, d$ and all $h \in \mathcal{C}_b(\mathbb{R}^i)$. Thus, $\mathcal{M}$ is closed as an intersection of closed sets. $\square$

The last tool that we need to prove Theorem 5.1 is a minimax theorem as it is known from game theory, optimal control and related fields. We report such a theorem in Appendix A.2. In the proof of our general theorem, we use it together with the duality result in Proposition 5.3 and the approximation statement in Lemma 2.1.

Proof of Theorem 5.1. In order to improve readability, whenever the full nd-tuple appears, we write S for $\left(S^1_{t_1},\dots,S^1_{t_n},\dots,S^d_{t_1},\dots,S^d_{t_n}\right)$ and s for $\left(s^1_{t_1},\dots,s^1_{t_n},\dots,s^d_{t_1},\dots,s^d_{t_n}\right)$.

We start the proof showing that we may without loss of generality assume $c \geq 0$. Notice that if c satisfies the assertion of Theorem 5.1, then so does $\tilde{c} := c + \sum_{j=1}^d \sum_{i=1}^n \tilde{\varphi}_{i,j}$, where $\tilde{\varphi}_{i,j} \in \mathcal{S}$, $i = 1,\dots,n$, $j = 1,\dots,d$.

Indeed, on the one hand we have

$$\begin{aligned} P(\tilde{c}) = \inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_{\mathbb{Q}}[\tilde{c}(S)] &= \inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_{\mathbb{Q}}[c(S)] + \sum_{j=1}^d \sum_{i=1}^n \mathbb{E}_{\mu_{i,j}}\left[\tilde{\varphi}_{i,j}\left(S^j_{t_i}\right)\right] \\ &= P(c) + \sum_{j=1}^d \sum_{i=1}^n \mathbb{E}_{\mu_{i,j}}\left[\tilde{\varphi}_{i,j}\left(S^j_{t_i}\right)\right], \end{aligned}$$

where in the second equality we use that all $\mathbb{Q} \in \mathcal{M}$ have the same marginals and that the second summand does not depend on the dependence structure between the marginals.

On the other hand we have

$$D(\tilde{c}) = \sup \sum_{j=1}^d \sum_{i=1}^n \mathbb{E}_{\mu_{i,j}}\left[\varphi_{i,j}\left(S^j_{t_i}\right)\right],$$

where the supremum is taken over functions $\varphi_{i,j} \in \mathcal{S}$ such that there are $h^j_i \in \mathcal{C}_b(\mathbb{R}^i)$ with

$$\sum_{j=1}^d \sum_{i=1}^n \varphi_{i,j}\left(s^j_{t_i}\right) + \sum_{j=1}^d \sum_{i=1}^{n-1} h^j_i\left(s^j_{t_1},\dots,s^j_{t_i}\right)\left(s^j_{t_{i+1}} - s^j_{t_i}\right) \leq \tilde{c}(s) = c(s) + \sum_{j=1}^d \sum_{i=1}^n \tilde{\varphi}_{i,j}\left(s^j_{t_i}\right),$$

which is the case if and only if

$$\sum_{j=1}^d \sum_{i=1}^n (\varphi_{i,j} - \tilde{\varphi}_{i,j})\left(s^j_{t_i}\right) + \sum_{j=1}^d \sum_{i=1}^{n-1} h^j_i\left(s^j_{t_1},\dots,s^j_{t_i}\right)\left(s^j_{t_{i+1}} - s^j_{t_i}\right) \leq c(s). \tag{5.7}$$

We define $\psi_{i,j} := \varphi_{i,j} - \tilde{\varphi}_{i,j}$ and immediately have $\psi_{i,j} \in \mathcal{S}$. Thus, we have

$$D(\tilde{c}) = \sup \sum_{j=1}^d \sum_{i=1}^n \mathbb{E}_{\mu_{i,j}}\left[(\psi_{i,j} + \tilde{\varphi}_{i,j})\left(S^j_{t_i}\right)\right] = D(c) + \sum_{j=1}^d \sum_{i=1}^n \mathbb{E}_{\mu_{i,j}}\left[\tilde{\varphi}_{i,j}\left(S^j_{t_i}\right)\right],$$

as the $\psi_{i,j}$ are by (5.7) exactly the functions in $\mathcal{S}$ that are used for the optimization in $D(c)$. As $P(c) = D(c)$ holds, we have $P(\tilde{c}) = D(\tilde{c})$ as well. Hence, we may consider the function $\tilde{c}$, the non-negativity of which holds by (5.1), defined as

$$\tilde{c}(s) = c(s) + K\left(1 + \sum_{j=1}^d \sum_{i=1}^n \left|s^j_{t_i}\right|\right) \geq 0.$$

Thus, let us assume $c \geq 0$. Then the class of functions the theorem has to be shown for is the set of all lower semi-continuous functions $c : \mathbb{R}^{nd} \to [0, \infty]$. As those functions are in particular bounded from below by 0, we may even assume $c \in \mathcal{C}_b(\mathbb{R}^{nd})$ by Lemma 2.1.

For such payoff functions, we justify the applicability of Theorem A.1 on the sets

$$K \coloneqq \Pi_n^d(\mu),$$

the set of all transport plans on $\mathbb{R}^{nd}$ with suitable marginals, and

$$T \coloneqq \left(\mathcal{C}_b(\mathbb{R}) \times \mathcal{C}_b\left(\mathbb{R}^2\right) \ldots \times \mathcal{C}_b\left(\mathbb{R}^{n-1}\right)\right)^d,$$

the set of $d(n-1)$-tuples of bounded, continuous functions, and the function

$$f : K \times T \to \mathbb{R}, \quad (\pi, h) \mapsto \int_{\mathbb{R}^{nd}} \chi_{c,(h_i^j)}(s)\pi(\mathrm{d}s),$$

where we write $h = \left(h_i^j\right)_{i=1,\ldots,n-1}^{j=1,\ldots,d}$ and define $\chi_{c,(h_i^j)} : \mathbb{R}^{nd} \to \mathbb{R}$ by

$$\chi_{c,(h_i^j)}(s) \coloneqq c(s) - \sum_{j=1}^{d}\sum_{i=1}^{n-1} h_i^j\left(s_{t_1}^j, \ldots, s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right).$$

We already know that K is compact. The convexity of K and T are well-known and obvious respectively. Thus, we only have to show that f is indeed a suitable function.

Let us first check the second condition. Therefore, let $h \in T$ be arbitrary. We have to investigate the mapping $\pi \mapsto f(\pi, h)$, which is convex by definition of the integral. In order to check the continuity, recall that $c \in \mathcal{C}_b(\mathbb{R}^{nd})$ and $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ for all $i = 1, \ldots, n-1$ and all $j = 1, \ldots, d$. Hence, $\chi_{c,(h_i^j)}$ is continuous as a combination of continuous functions and it satisfies

$$\left|\chi_{c,(h_i^j)}(s)\right| \leq |c(s)| + \sum_{j=1}^{d}\sum_{i=1}^{n-1}\left|h_i^j\left(s_{t_1}^j, \ldots, s_{t_i}^j\right)\right|\left(\left|s_{t_{i+1}}^j\right| + \left|s_{t_i}^j\right|\right).$$

Denote by $B(c) \geq 0$ and $B(h_i^j) \geq 0$ the smallest bounds for the bounded functions c and h_i^j, $i = 1, \ldots, n-1$, $j = 1, \ldots, d$. Then we obtain

$$\begin{aligned}\left|\chi_{c,(h_i^j)}(s)\right| &\leq B(c) + \sum_{j=1}^{d}\sum_{i=1}^{n-1} B\left(h_i^j\right)\left(\left|s_{t_{i+1}}^j\right| + \left|s_{t_i}^j\right|\right) \\ &= B(c) + \sum_{j=1}^{d}\sum_{i=1}^{n}\tilde{B}_i^j\left|s_{t_i}^j\right| =: B(c)\left(1 + \sum_{j=1}^{d}\sum_{i=1}^{n} f_i^j\left(s_{t_i}^j\right)\right),\end{aligned}$$

where $\tilde{B}_i^j \geq 0$ and $f_i^j \in \mathbb{L}^1(\mathbb{R}, \mu_{i,j})$, $i = 1, \ldots, n$, $j = 1, \ldots, d$, are defined in such a way that the equalities hold. Clearly, the integrability holds, as the first moments of all marginals exist. Thus, (5.6) is satisfied and $\pi \mapsto f(\pi, h)$ is continuous by Lemma 5.5.

Now let us check the third condition. Therefore, let $\pi \in K$ be arbitrary. Then the mapping $h \mapsto f(\pi, h)$ is concave on T, as for all $h, \tilde{h} \in T$ and all $\lambda \in [0,1]$, we have

$$f\left(\pi, \left(\lambda h + (1-\lambda)\tilde{h}\right)\right) = \lambda f(\pi, h) + (1-\lambda) f\left(\pi, \tilde{h}\right).$$

Now let us prove the desired duality. By definition of the sub hedging problem, we have

$$D(c) = \sup_{\varphi_{i,j} \in \mathcal{S}} \left\{ \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}\left(s_{t_i}^j\right) \mu_{i,j}\left(\mathrm{d}s_{t_i}^j\right) \,\middle|\, \exists h \in T : \Psi_{(\varphi_{i,j}),(h_i^j)}(s) \le c(s) \right\}.$$

By the definitions of $\Psi_{(\varphi_{i,j}),(h_i^j)}$ and $\chi_{c,(h_i^j)}$, we have

$$\Psi_{(\varphi_{i,j}),(h_i^j)}(s) \le c(s) \iff \sum_{j=1}^{d} \sum_{i=1}^{n} \varphi_{i,j}\left(s_{t_i}^j\right) \le \chi_{c,(h_i^j)}(s) \tag{5.8}$$

for all $s = \left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right) \in \mathbb{R}^{nd}$. Decomposing the supremum, we have

$$\begin{aligned} D(c) &= \sup_{\varphi_{i,j} \in \mathcal{S}} \left\{ \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}\left(s_{t_i}^j\right) \mu_{i,j}\left(\mathrm{d}s_{t_i}^j\right) \,\middle|\, \exists h \in T : \Psi_{(\varphi_{i,j}),(h_i^j)}(s) \le c(s) \right\} \\ &\overset{(\diamond)}{=} \sup_{h \in T} \sup_{\varphi_{i,j} \in \mathcal{S}} \left\{ \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}\left(s_{t_i}^j\right) \mu_{i,j}\left(\mathrm{d}s_{t_i}^j\right) \,\middle|\, \sum_{j=1}^{d} \sum_{i=1}^{n} \varphi_{i,j}\left(s_{t_i}^j\right) \le \chi_{c,(h_i^j)}(s) \right\}. \end{aligned} \tag{5.9}$$

We justify $(\diamond)$ in the case that all the suprema are indeed maxima.

Let us first show that „$\le$“ holds. For this, let $\bar{\varphi}_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, d$, be maximizing functions for the supremum on the left hand side of $(\diamond)$ and let the value be denoted by $\mathcal{I}(\bar{\varphi}_{i,j})$. Then by definition there is an $\bar{h} \in T$ satisfying the left hand side of (5.8). By the equivalence, the functions $\bar{\varphi}_{i,j}$ are accessible for the inner supremum on the right hand side of $(\diamond)$ for fixed $h = \bar{h}$ in the outer supremum. Thus, we have

$$\mathcal{I}(\bar{\varphi}_{i,j}) \le \sup_{\varphi_{i,j} \in \mathcal{S}} \left\{ \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}\left(s_{t_i}^j\right) \mu_{i,j}\left(\mathrm{d}s_{t_i}^j\right) \,\middle|\, \sum_{j=1}^{d} \sum_{i=1}^{n} \varphi_{i,j}\left(s_{t_i}^j\right) \le \chi_{c,(\bar{h}_i^j)}(s) \right\}.$$

Clearly, maximizing over all possible $h \in T$ the right hand side gets only bigger.

Now let us show that „$\ge$“ is true. For this, let $\tilde{h}$ and $\tilde{\varphi}_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, d$, be maximizing functions for the double supremum on the right hand side of $(\diamond)$ with value $\mathcal{J}(\tilde{h}, \tilde{\varphi}_{i,j})$. Then the right hand side of (5.8) is satisfied by definition and thus the $\tilde{\varphi}_{i,j}$ are accessible for the supremum on the left hand side of $(\diamond)$. Thus, we get

$$\sup_{\varphi_{i,j} \in \mathcal{S}} \left\{ \sum_{j=1}^{d} \sum_{i=1}^{n} \int_{\mathbb{R}} \varphi_{i,j}(s_{t_i}^j) \mu_{i,j}(\mathrm{d}s_{t_i}^j) \,\middle|\, \exists h \in T : \Psi_{(\varphi_{i,j}),(h_i^j)}(s) \le c(s) \right\} \ge \mathcal{J}(\tilde{h}, \tilde{\varphi}_{i,j}).$$

In total, we have the desired equality in the case of existing maximizers. The same arguments work when we consider maximizing sequences instead of maximizers. Thus, $(\diamond)$ does indeed hold.

Now using Proposition 5.3 on the right hand side of (5.9) with the choice $\phi = \chi_{c,(h_i^j)}$, we get the first equality of

$$\begin{aligned} D(c) &= \sup_{h\in T}\inf_{\pi\in K}\int_{\mathbb{R}^{nd}} \chi_{c,(h_i^j)}(s)\pi(\mathrm{d}s) \\ &= \inf_{\pi\in K}\sup_{h\in T}\int_{\mathbb{R}^{n}} \chi_{c,(h_i^j)}(s)\pi(\mathrm{d}s) \\ &\overset{(\circ)}{=} \inf_{\mathbb{Q}\in\mathcal{M}}\int_{\mathbb{R}^{nd}} c(s)\mathbb{Q}(\mathrm{d}s) = P(c), \end{aligned}$$

while the second equality holds by application of Theorem A.1. The last equality holds by definition, but $(\circ)$ needs some more detailed justification.

For this, let $\pi \in \Pi_n^d(\mu) \setminus \mathcal{M}$ be fixed. Then, by Lemma 5.7, there are $i \in \{1, \ldots, n-1\}$, $j \in \{1, \ldots, d\}$ and $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ such that

$$\int_{\mathbb{R}^{nd}} h_i^j\left(s_{t_1}^j, \ldots, s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right)\pi(\mathrm{d}s) \neq 0.$$

Suitably scaling, we get this integral to be arbitrarily large. Hence, the supremum on the left hand side of $(\circ)$ becomes arbitrarily large for any $\pi \in \Pi_n^d(\mu) \setminus \mathcal{M}$. As we minimize the value over all $\pi \in K$, it suffices to consider $\mathbb{Q} \in \mathcal{M}$. Furthermore, for all $\mathbb{Q} \in \mathcal{M}$, we have

$$\int_{\mathbb{R}^{nd}} h_i^j(s_{t_1}^j, \ldots, s_{t_i}^j)(s_{t_{i+1}}^j - s_{t_i}^j)\mathbb{Q}(\mathrm{d}s) = 0,$$

for all $i = 1, \ldots, n-1$, $j = 1, \ldots, d$ and $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ by Lemma 5.7, Thus, by definition of $\chi_{c,(h_i^j)}$, we also have

$$\int_{\mathbb{R}^{nd}} \chi_{c,(h_i^j)}(s)\mathbb{Q}(\mathrm{d}s) = \int_{\mathbb{R}^{nd}} c(s)\mathbb{Q}(\mathrm{d}s),$$

which yields $(\circ)$. Thus, we have shown the desired duality for $c \in \mathcal{C}_b(\mathbb{R}^{nd})$.

In order to complete the proof, let us now assume that c is lower semi-continuous and as before $c \geq 0$. By Lemma 2.1, there is a sequence $(c_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{C}_b(\mathbb{R}^{nd})$ such that $0 \leq c_1 \leq c_2 \leq \ldots$ and $\sup_{k\in\mathbb{N}} c_k = c$.

Let us define a sequence $(\mathbb{Q}_k)_{k\in\mathbb{N}}$ of measures in $\mathcal{M}$ such that, for all $k \in \mathbb{N}$, we have

$$P(c_k) \geq \int_{\mathbb{R}^{nd}} c(s)\mathbb{Q}_k(\mathrm{d}s) - \frac{1}{k}. \tag{5.10}$$

Let $k \in \mathbb{N}$. Then there is a sequence $(\mathbb{Q}_\ell^k)_{\ell\in\mathbb{N}}$ of measures in $\mathcal{M}$ such that

$$P(c_k) = \inf_{\mathbb{Q}\in\mathcal{M}}\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}(\mathrm{d}s) = \lim_{\ell\to\infty}\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_\ell^k(\mathrm{d}s).$$

We may choose the sequence such that the convergence is monotone, i.e. such that we have $\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_\ell^k(\mathrm{d}s) \searrow P(c_k)$. Hence, for all $\varepsilon > 0$ there is an $\ell_0 \in \mathbb{N}$ such that

$$0 \leq \int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_\ell^k(\mathrm{d}s) - P(c_k) \leq \varepsilon$$

for all $\ell \geq \ell_0$. Choosing a suitable subsequence of $(\mathbb{Q}_\ell^k)_{\ell\in\mathbb{N}}$, we have

$$\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_\ell^k(\mathrm{d}s) - P(c_k) \leq \frac{1}{2k}$$

for all $\ell \geq k$ and in particular

$$\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_k^k(\mathrm{d}s) - P(c_k) \leq \frac{1}{2k}.$$

This yields a sequence $(\mathbb{Q}_k^k)_{k\in\mathbb{N}}$ of measures in $\mathcal{M}$ such that for all $k \in \mathbb{N}$, we have

$$P(c_k) \geq \int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_k^k(\mathrm{d}s) - \frac{1}{2k}.$$

Furthermore, we have $c_k \nearrow c$. That is, for all $\varepsilon > 0$ there is an $k_0 \in \mathbb{N}$ such that $0 \leq c - c_k \leq \varepsilon$ and, in particular, $c_k \geq c - \varepsilon$ for all $k \geq k_0$. Again choosing a suitable subsequence, we obtain $c_k \geq c - \frac{1}{2k}$ for all $k \in \mathbb{N}$. In total, we have

$$P(c_k) \geq \int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_k^k(\mathrm{d}s) - \frac{1}{2k} \geq \int_{\mathbb{R}^{nd}} \left(c(s) - \frac{1}{2k}\right)\mathbb{Q}_k^k(\mathrm{d}s) - \frac{1}{2k} = \int_{\mathbb{R}^{nd}} c(s)\mathbb{Q}_k^k(\mathrm{d}s) - \frac{1}{k}.$$

Choosing $(\mathbb{Q}_k)_{k\in\mathbb{N}} = (\mathbb{Q}_k^k)_{k\in\mathbb{N}}$, we have the desired sequence of measures in $\mathcal{M}$.

As $\mathcal{M}$ is compact by Proposition 5.8, we may assume that $(\mathbb{Q}_k)_{k\in\mathbb{N}}$ converges weakly to some $\tilde{\mathbb{Q}} \in \mathcal{M}$. Then we have

$$\begin{aligned}
P(c) &= \inf_{\mathbb{Q}\in\mathcal{M}} \int_{\mathbb{R}^{nd}} c(s)\mathbb{Q}(\mathrm{d}s) \leq \int_{\mathbb{R}^{nd}} c(s)\tilde{\mathbb{Q}}(\mathrm{d}s) \overset{(\diamond)}{=} \lim_{\ell\to\infty} \int_{\mathbb{R}^{nd}} c_\ell(s)\tilde{\mathbb{Q}}(\mathrm{d}s) \\
&\overset{(*)}{=} \lim_{\ell\to\infty}\left(\lim_{k\to\infty}\int_{\mathbb{R}^{nd}} c_\ell(s)\mathbb{Q}_k(\mathrm{d}s)\right) \overset{(\circ)}{\leq} \lim_{\ell\to\infty}\left(\lim_{k\to\infty}\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_k(\mathrm{d}s)\right) \\
&= \lim_{k\to\infty}\int_{\mathbb{R}^{nd}} c_k(s)\mathbb{Q}_k(\mathrm{d}s) \leq \lim_{k\to\infty}\int_{\mathbb{R}^{nd}} c(s)\mathbb{Q}_k(\mathrm{d}s) \overset{(\star)}{\leq} \lim_{k\to\infty}\left(P(c_k) + \frac{1}{k}\right) = \lim_{k\to\infty} P(c_k).
\end{aligned}$$

The equation marked with $(\diamond)$ holds by the monotone convergence theorem. The equation marked with $(*)$ holds by the definition of weak convergence, since $c_k \in \mathcal{C}_b(\mathbb{R}^{nd})$, $k \in \mathbb{N}$. The inequality marked with $(\circ)$ holds, since the sequence $(c_k)_{k\in\mathbb{N}}$ is monotone non-decreasing. Finally, the inequality marked with $(\star)$ holds by (5.10).

By definition of the sequence $(c_k)_{k\in\mathbb{N}}$, we have $D(c_k) \leq D(c)$ and $P(c_k) \leq P(c)$ for all $k \in \mathbb{N}$, which implies

$$D(c) \geq D(c_k) = P(c_k) \nearrow P(c).$$

Thus, we have $D(c) \geq P(c)$. Recalling the weak duality inequality $P(c) \geq D(c)$, we finally have $P(c) = D(c)$ for general lower semi-continuous functions $c : \mathbb{R}^{nd} \to [0, \infty)$.

It remains to show that the infimum in the primal problem is attained. If $P(c) = \infty$, then this is trivially satisfied. Thus, we assume $P(c) < \infty$. Then there is a sequence $(\mathbb{Q}_k)_{k\in\mathbb{N}}$ in $\mathcal{M} \subset \Pi_n^d(\mu)$ such that $P(c) = \lim_{k\to\infty} \int c\mathrm{d}\mathbb{Q}_k$. As $\mathcal{M}$ is compact, a subsequence of $(\mathbb{Q}_k)_{k\in\mathbb{N}}$ converges weakly to some $\mathbb{Q} \in \mathcal{M}$.

As the payoff function c is lower semi-continuous, for any sequence $(\pi_k)_{k\in\mathbb{N}}$ in $\Pi_n^d(\mu)$

weakly converging to some $\pi \in \Pi_n^d(\mu)$, we have by Lemma 4.3 and Remark 4.5 that $\pi \mapsto \int c \mathrm{d}\pi$ is lower semi-continuous, i.e.

$$\liminf_{k\to\infty} \int_{\mathbb{R}^{nd}} c(s)\pi_k(\mathrm{d}s) \geq \int_{\mathbb{R}^{nd}} c(s)\pi(\mathrm{d}s).$$

Thus, the limit measure $\mathbb{Q} \in \mathcal{M}$ is a minimizer of $P(c)$. □

Remark 5.9. In Section 4.1, we mentioned that in classic optimal transport an optimizer to the dual problem does not exist in general. The same applies to martingale optimal transport. In the following sections we discuss, among other topics, the existence of dual optimizers.

However, if we assume that a dual optimizer exists, i.e. if there are functions $\varphi_{i,j} \in \mathcal{S}$ and $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ such that $\Psi := \Psi_{(\varphi_{i,j}),(h_i^j)} \leq c$ and $\sum_{j=1}^d \sum_{i=1}^n \mathbb{E}_{\mu_{i,j}}\left[\varphi_{i,j}\left(S_{t_i}^j\right)\right] = D(c)$, and if further $\mathbb{Q}$ is a primal optimizer, i.e. if $\mathbb{E}_{\mathbb{Q}}[c] = P(c)$, then

$$0 \leq \mathbb{E}_{\mathbb{Q}}[c - \Psi] = P(c) - D(c) = 0.$$

That is, the sub-replicating hedging strategy Ψ replicates the payoff function c at least $\mathbb{Q}$-almost surely. This property of a dual optimizer is also used in the following chapters. ◊

Remark 5.10. A result similar to Theorem 5.1 may be obtained applying a result from Zaev [79] on optimal transport with linear constraints. In order to understand how this is done, let us shortly discuss the work of Zaev [79].

For this purpose, let $\mu_1, \dots, \mu_n \in \mathcal{P}(\mathbb{R})$ and define for all $i = 1, \dots, n$,

$$\mathcal{C}_{\mathbb{L}}(\mu_i) := \{f \in \mathbb{L}^1(\mathbb{R}, \mu_i) \cap \mathcal{C}(\mathbb{R})\},$$

the set of all continuous functions that are absolutely integrable with respect to μ_i. Further define

$$\mathcal{C}_{\mathbb{L}}(\mu) := \mathcal{C}_{\mathbb{L}}(\mu_1, \dots, \mu_n) := \left\{ h \in \mathcal{C}(\mathbb{R}^n) \;\middle|\; \exists f = \sum_{i=1}^n f_i \in \bigoplus_{i=1}^n \mathcal{C}_{\mathbb{L}}(\mu_i) \text{ with } |h| \leq f \right\}$$

equipped with the seminorm

$$\|h\|_{\mathbb{L}} := \sup_{\pi \in \Pi_n(\mu_1,\dots,\mu_n)} \int_{\mathbb{R}^n} h(x_1, \dots, x_n)\pi(\mathrm{d}(x_1, \dots, x_n)).$$

Now let $W \subseteq \mathcal{C}_{\mathbb{L}}(\mu)$ be some subspace and $c \in \mathcal{C}_{\mathbb{L}}(\mu)$. Then the author considers the constrained Monge-Kantorovich problem of optimal transport

$$\inf_{\pi \in \Pi_W} \int_{\mathbb{R}^n} c(x_1, \dots, x_n)\pi(\mathrm{d}(x_1, \dots, x_n)),$$

where

$$\Pi_W := \left\{ \pi \in \Pi_n(\mu_1, \dots, \mu_n) \;\middle|\; \int_{\mathbb{R}^n} w(x_1, \dots, x_n)\pi(\mathrm{d}(x_1, \dots, x_n)) = 0 \text{ for all } w \in W \right\}.$$

Theorem 5.11 ([79, Theorem 2.1]). *Let $\mu_1, \dots, \mu_n \in \mathcal{P}(\mathbb{R})$, $W \subseteq \mathcal{C}_{\mathbb{L}}(\mu)$ a subspace and $c \in \mathcal{C}_{\mathbb{L}}(\mu)$. Then*

$$\inf_{\pi \in \Pi_W} \int_{\mathbb{R}^n} c(x_1, \dots, x_n)\pi(\mathrm{d}(x_1, \dots, x_n)) = \sup \sum_{i=1}^{n} \int_{\mathbb{R}} f_i(x_i)\mu_i(\mathrm{d}x_i),$$

where the supremum is taken over functions $f_i \in \mathcal{C}_{\mathbb{L}}(\mu_i)$ such that there is a function $w \in W$ with $\sum_{i=1}^{n} f_i(x_i) + w(x_1, \dots, x_n) \leq c(x_1, \dots, x_n)$ for all $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Using this theorem, we may state a result similar to our general duality theorem. In order to do so, consider measures $\mu_{1,1}, \dots, \mu_{n,1}, \dots, \mu_{1,d}, \dots, \mu_{n,d} \in \mathcal{P}(\mathbb{R})$ and choose

$$W = \left\{ \sum_{j=1}^{d} \sum_{i=1}^{n-1} h_i^j \left(s_{t_1}^j, \dots, s_{t_i}^j\right) \left(s_{t_{i+1}}^j - s_{t_i}^j\right) \;\middle|\; h_i^j \in \mathcal{C}_b(\mathbb{R}^i), i = 1, \dots, n-1, j = 1, \dots, d \right\}.$$

The functions in W are obviously contained in $\mathcal{C}(\mathbb{R}^{nd})$ and further we have

$$\begin{aligned}
\left| \sum_{j=1}^{d} \sum_{j=1}^{n-1} h_i^j \left(s_{t_1}^j, \dots, s_{t_i}^j\right) \left(s_{t_{i+1}}^j - s_{t_i}^j\right) \right| &\leq \sum_{j=1}^{d} \sum_{j=1}^{n-1} B\left(h_i^j\right) \left|s_{t_{i+1}}^j - s_{t_i}^j\right| \\
&\leq \sum_{j=1}^{d} \sum_{j=1}^{n-1} B\left(h_i^j\right) \left(\left|s_{t_{i+1}}^j\right| + \left|s_{t_i}^j\right|\right) = \sum_{j=1}^{d} \sum_{j=1}^{n} \tilde{B}_i^j \left|s_{t_i}^j\right| \\
&=: \sum_{j=1}^{d} \sum_{j=1}^{n} f_{i,j}\left(s_{t_i}^j\right) =: f\left(s_{t_1}^1, \dots, s_{t_n}^1, \dots, s_{t_1}^d, \dots, s_{t_n}^d\right),
\end{aligned}$$

where $B(h_i^j) \geq 0$ are the smallest bounds for the bounded functions h_i^j and $\tilde{B}_i^j$ are suitable constants such that the equality holds. Thus, we have indeed $W \subseteq \mathcal{C}_{\mathbb{L}}(\mu)$.

Using Lemma 5.7, we obtain $\Pi_W = \mathcal{M}$ and thus the similarity of Theorem 5.1 and the discussed special case of Theorem 5.11. However, we observe some differences in the assumptions on the payoff function and the hedging functions between the two results.

Indeed, our theorem allows for more general payoff functions. In particular, we do not need continuity such that payoff functions with jumps, as for example for binary type options, are covered. Furthermore, the class of functions that we use for static hedging is a subclass of the class considered by Zaev that allows nice interpretation when it comes to application, as we only use payoff functions of liquidly traded options for the hedging. ♢

Let us close this section discussing some drawbacks of the theory of martingale optimal transport.

A first drawback is, as we already mentioned, that dual optimizers do not exist in general. This was first shown by Beiglböck, Henry-Labordère & Penkner [5, Proposition 4.1]. A much simpler counterexample is given by Beiglböck, Nutz & Touzi [9, Example 8.2]. As we do not contribute anything to the progress on existence of dual optimizers, but only present some important proceedings in that direction, we do not discuss any counterexamples.

Instead, we proceed with a second drawback which is less technical but as important.

That is, upper and lower price bounds may deviate vastly. This can for example be seen in Lütkebohmert & Sester [59, Sec. 5], who numerically investigate the price differences in the standard market case and also suggest methods to improve the situation using additional information about the variance of asset returns. We complement their numerical studies by illustrative investigations in the general market case for $d = n = 2$.

Example 5.12. In order to illustrate the problem, we do not need to consider marginals too complicated. Thus, we restrict ourselves to the case of simple discrete measures. In particular, we consider the marginals

$$\mu_{1,1} = \frac{1}{3}(\delta_8 + \delta_{10} + \delta_{12}) \leq_c \frac{1}{4}(\delta_7 + \delta_9 + \delta_{11} + \delta_{13}) = \mu_{2,1},$$
$$\mu_{1,2} = \frac{1}{3}(\delta_8 + \delta_{10} + \delta_{12}) \leq_c \frac{1}{5}(\delta_4 + \delta_7 + \delta_{10} + \delta_{13} + \delta_{16}) = \mu_{2,2}.$$

In Table 5.1, we state the payoff functions that we consider in this example.

Exotic option type	Payoff function
Basket option	$c_B\left(S^1_{t_1}, S^1_{t_2}, S^2_{t_1}, S^2_{t_2}\right) = \left(\frac{1}{4}(S^1_{t_1} + S^1_{t_2} + S^2_{t_1} + S^2_{t_2}) - 10\right)^+$
Binary option	$c_1\left(S^1_{t_1}, S^1_{t_2}, S^2_{t_1}, S^2_{t_2}\right) = \mathbb{1}_{\left\{S^1_{t_2} > S^1_{t_1}\right\}} \cdot \mathbb{1}_{\left\{S^2_{t_2} > S^2_{t_1}\right\}}$
Asian option	$c_A\left(S^1_{t_1}, S^1_{t_2}, S^2_{t_1}, S^2_{t_2}\right) = \frac{1}{4}\left(S^1_{t_2} - S^1_{t_1}\right)^+ \cdot \left(S^2_{t_2} - S^2_{t_1}\right)^+$
Variance option	$c_V\left(S^1_{t_1}, S^1_{t_2}, S^2_{t_1}, S^2_{t_2}\right) = \left(\frac{S^1_{t_2} - S^1_{t_1}}{S^1_{t_1}}\right)^2 \cdot \left(\frac{S^2_{t_2} - S^2_{t_1}}{S^2_{t_1}}\right)^2$
Covariance option	$c_C\left(S^1_{t_1}, S^1_{t_2}, S^2_{t_1}, S^2_{t_2}\right) = \left(\frac{S^1_{t_2} - S^1_{t_1}}{S^1_{t_1}} \cdot \frac{S^2_{t_2} - S^2_{t_1}}{S^2_{t_1}}\right)^+$

Table 5.1.: Types of payoff functions

Let us now compare the resulting upper and lower price bounds for exotic options with the given payoff functions with respect to the sets of measures

$$\mathcal{M}^2_2(\mu) \subseteq \Pi^2_{2,\geq 0}(\mu) \subseteq \Pi^2_2(\mu),$$

where we define

$$\Pi^2_{2,\geq 0}(\mu) := \left\{\pi \in \Pi^2_2(\mu) \,\middle|\, \mathrm{Cov}_\pi(S^1_{t_1}, S^1_{t_2}) \geq 0, \mathrm{Cov}_\pi(S^2_{t_1}, S^2_{t_2}) \geq 0\right\}.$$

This set is of interest as the dependence structure of martingales ensures non-negative covariance in the above sense. However, it is easier to implement the above condition of non-negative covariance than the martingale property itself.

In Tables 5.2 and 5.3, we state the upper and lower price bounds for the different options and the different underlying sets of measures. The values are calculated numerically using classic methods of linear programming, as the price bound problems reduce to finite dimensional linear programs in this case.

Payoff function c	$\inf_{\pi\in\Pi_2^2(\mu)} \mathbb{E}_\pi[c]$	$\inf_{\pi\in\Pi_{2,\geq 0}^2(\mu)} \mathbb{E}_\pi[c]$	$\inf_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c]$
c_B	0.075	0.09	0.1464
c_1	0	0	0
c_A	0	0	0
c_V	0.0007	0.0007	0.001
c_C	0	0	0

Table 5.2.: Lower price bounds for different exotic options

Payoff function c	$\sup_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c]$	$\sup_{\pi\in\Pi_{2,\geq 0}^2(\mu)} \mathbb{E}_\pi[c]$	$\sup_{\pi\in\Pi_2^2(\mu)} \mathbb{E}_\pi[c]$
c_B	1.022	1.033	1.033
c_1	0.6	0.6	0.6
c_A	0.9944	2.1125	2.675
c_V	0.0218	0.0859	0.1092
c_C	0.072	0.171	0.241

Table 5.3.: Upper price bounds for different exotic options

Finally, in Table 5.4, we collect the relative price range

$$\frac{\sup_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c] - \inf_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c]}{\inf_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c]}$$

and the relative sharpening of the price ranges by the martingale property

$$1 - \frac{\sup_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c] - \inf_{\mathbb{Q}\in\mathcal{M}_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c]}{\sup_{\mathbb{Q}\in\Pi_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c] - \inf_{\mathbb{Q}\in\Pi_2^2(\mu)} \mathbb{E}_\mathbb{Q}[c]}$$

for the different payoff functions, both denoted in percent.

Payoff function c	Relative price range	Relative sharpening
c_B	1277.77	8.6013
c_1	-	0
c_A	-	62.8262
c_V	15500	80.8295
c_C	-	70.0124

Table 5.4.: Some key numbers to measure the price bound quality in percent.

We see that the price bound differences are rather great even in the martingale case, but we also see that often the martingale property leads to a major improvement of the price bounds. However, there is room for further improvement. We discuss some possible improvements of the price bounds in Section 5.3, where we consider an approach using additional information on the asset prices. △

5.2. Existence of dual optimizers

In this section, we present some results on the existence of dual optimizers for martingale optimal transport. In Section 5.2.1, we present the first result in that direction derived by Beiglböck, Nutz & Touzi [9]. The authors introduce the notion of irreducibility in order to guarantee the existence of dual optimizers. We present some more recent results by Beiglböck, Lim & Obłój [8] that we need in Chapter 7. Those results are achieved in the standard market case. In Section 5.2.2, we present some generalizations to the general market case provided by Nutz, Stebegg & Tan [65] and Lim [58]. We focus on the results and do not discuss any difficulties or subtleties that are not important for our work.

5.2.1. The standard market case

We immediately state the central notion of irreducibility that goes back to Beiglböck, Nutz & Touzi [9]. It plays an important role in Chapters 6 and 7, where we consider the standard market case. The irreducibility property intuitively stems from the observation that points in which the potential functions of the marginals μ and ν touch, i.e. points $x \in \mathbb{R}$ such that $u_\mu(x) = u_\nu(x)$ holds, somehow restrict martingale transport plans.

Definition 5.13 ([9, Definition 2.2])**.** Let $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ be such that $\mu \leq_c \nu$. The pair (μ, ν) is called *irreducible*, if the set $I := \{u_\mu < u_\nu\}$ is connected and $\mu(I) = \mu(\mathbb{R})$. In this situation, let J be the union of I and any endpoints of I that are atoms of ν. Then (I, J) is called the *domain* of (μ, ν).

While Henry-Labordère & Touzi [38] introduce an equivalent definition in terms of call option price functions, Beiglböck, Nutz & Touzi [9] also discuss several useful consequences of this definition. In order to understand the definition of irreducibility thoroughly, we recall some of these consequences here.

First of all, recall that $\mu \leq_c \nu$ implies $u_\mu \leq u_\nu$. Hence, outside of I we have $u_\mu = u_\nu$. As both measures have the same mass and barycentre, we have $\nu(J) = \nu(\mathbb{R})$. Indeed, J is the smallest superset of I such that ν is concentrated on J. A very important property proved by Beiglböck, Nutz & Touzi [9] implies that whenever discussing the standard market case, we may assume the pair (μ, ν) of marginals to be irreducible.

Proposition 5.14 ([9, Proposition 2.3])**.** *Let $\mu \leq_c \nu$ and let $(I_k)_{1 \leq k \leq N}$ be the (open) components of $\{u_\mu < u_\nu\}$, where $N \in \{1, \ldots, \infty\}$. Set $I_0 := \mathbb{R} \setminus \bigcup_{k \geq 1} I_k$ and $\mu_k := \mu|_{I_k}$ for $k \geq 0$, so that $\mu = \sum_{k \geq 0} \mu_k$. Then there is a unique decomposition $\nu = \sum_{k \geq 0} \nu_k$ such that*

1. *$\mu_0 = \nu_0$ and $\mu_k \leq_c \nu_k$ for all $k \geq 1$.*
2. *$I_k = \{u_{\mu_k} < u_{\nu_k}\}$ for all $k \geq 1$.*

Moreover, any $\mathbb{Q} \in \mathcal{M}_2(\mu, \nu)$ admits a unique decomposition $\mathbb{Q} = \sum_{k \geq 0} \mathbb{Q}_k$ such that $\mathbb{Q}_k \in \mathcal{M}_2(\mu_k, \nu_k)$ for all $k \geq 0$.

In order to show dual attainment using irreducibility, the set of functions over which the dual problem is optimized has to be relaxed. For this purpose, a generalized integrability notion is introduced. In the next two definitions, we assume the pair (μ, ν) with $\mu \leq_c \nu$ to be irreducible with domain (I, J).

Definition 5.15 ([9, Def. 4.7 & 4.9])**.** Let $\varphi : I \to \mathbb{R} \cup \{-\infty, \infty\}$ and $\psi : J \to \mathbb{R} \cup \{-\infty, \infty\}$ be Borel functions. If there exists a concave function $\chi : J \to \mathbb{R}$ such that $\psi - \chi \in \mathbb{L}^1(\mathbb{R}, \mu)$ and $\psi + \chi \in \mathbb{L}^1(\mathbb{R}, \nu)$, then we say χ is a *concave moderator* for (φ, ψ) and define

$$\begin{aligned}\int_I \varphi(x)\mu(\mathrm{d}x) + \int_J \psi(y)\nu(\mathrm{d}y) := &\int_J (\varphi - \chi)(x)\mu(\mathrm{d}x) + \int_J (\psi + \chi)(y)\nu(\mathrm{d}y) \\ &+ \frac{1}{2}\int_I (u_\mu - u_\nu)(x)\chi''(\mathrm{d}x) - \int_{J\setminus I} \chi(y)\nu(\mathrm{d}y).\end{aligned}$$

Denote by $\mathbb{L}^c(\mu, \nu)$ the set of all pairs (φ, ψ) of Borel functions that admit

$$\frac{1}{2}\int_I (u_\mu - u_\nu)(x)\chi''(\mathrm{d}x) - \int_{J\setminus I} \chi(y)\nu(\mathrm{d}y) < \infty.$$

With the notion of a concave moderator and the relaxed notion of integrability, we may introduce the relaxed set of super hedging strategies.

Definition 5.16 ([9, Definition 5.1])**.** Let $c : I \times J \to [0, \infty]$. Then we denote by $\mathcal{D}^{co,pw}_{\mu,\nu}(c)$ the set of all triples (φ, ψ, h) of Borel functions $\varphi : \mathbb{R} \to \mathbb{R} \cup \{-\infty, \infty\}$, $\psi : \mathbb{R} \to \mathbb{R} \cup \{-\infty, \infty\}$ and $h : \mathbb{R} \to \mathbb{R}$ such that $(\varphi, \psi) \in \mathbb{L}^c(\mu, \nu)$ and

$$\varphi(x) + \psi(y) + h(x)(y - x) \geq c(x, y), \quad (x, y) \in I \times J.$$

Theorem 5.17 ([9, Theorem 6.2])**.** *Let $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ be such that $\mu \leq_c \nu$ and (μ, ν) is irreducible. Further let $c : \mathbb{R}^2 \to [0, \infty]$ be a payoff function.*

1. *If c is upper semianalytic, i.e. if the set $\{c \geq a\}$ is the image of a Borel subset of a Polish space under a Borel mapping for all $a \in \mathbb{R}$, then*

$$\sup_{\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)} \mathbb{E}_\mathbb{Q}[c(X, Y)] = \inf_{(\varphi,\psi,h) \in \mathcal{D}^{co,pw}_{\mu,\nu}(c)} \left\{ \int_\mathbb{R} \varphi(x)\mu(\mathrm{d}x) + \int_\mathbb{R} \psi(y)\nu(\mathrm{d}y) \right\} \in [0, \infty].$$

2. *If $\inf_{(\varphi,\psi,h) \in \mathcal{D}^{co,pw}_{\mu,\nu}(c)} \{\int_\mathbb{R} \varphi(x)\mu(\mathrm{d}x) + \int_\mathbb{R} \psi(y)\nu(\mathrm{d}y)\} < \infty$, then there exists a dual minimizer $(\varphi, \psi, h) \in \mathcal{D}^{co,pw}_{\mu,\nu}(c)$.*

Now we proceed with the results of Beiglböck, Lim & Obłój [8]. We remark that the authors actually consider the lower price bound problem in (4.8) and associated to that the sub hedging problem in (4.11). On the contrary, in the standard market case, we mostly consider the upper price bound problem in (4.7) and the super hedging problem in (4.10). However, the results of Beiglböck, Lim & Obłój [8] are crucial for our work in Chapter 7. Therefore, we do not present the original results but adapted versions that we can apply later.

The authors investigate the existence of dual optimizers from an application-oriented point of view in the sense that they try to find conditions on the payoff function and the marginals such that optimizers do not only exist but also have some nice properties. For this purpose, they introduce the notion of a dual optimizer independently of the super hedging problem in (4.10) using the intuition that we developed in Remark 5.9.

Definition 5.18 ([8, Definition 2.1])**.** Let $\mu \leq_c \nu$ and $c : \mathbb{R}^2 \to \mathbb{R}$ be a payoff function. Then a triple (φ, ψ, h) of functions $\varphi : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, $\psi : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ and $h : \mathbb{R} \to \mathbb{R}$ is called *dual minimizer*, if φ is finite μ-almost surely, ψ is finite ν-almost surely and, for any maximizer $\mathbb{Q}^* \in \mathcal{M}_2(\mu, \nu)$ of the upper price bound problem in (4.7), we have

$$\begin{aligned} \varphi(x) + \psi(y) + h(x)(y - x) &\geq c(x, y), && \text{for all } (x, y) \in \mathbb{R}^2, \\ \varphi(x) + \psi(y) + h(x)(y - x) &= c(x, y), && \text{for } \mathbb{Q}^*\text{-almost every } (x, y). \end{aligned}$$

Definition 5.19 ([8, Definition 2.2])**.** Let J be an interval and $\mu \in \mathcal{P}_\alpha(\mathbb{R})$. We say that a function $c : \mathbb{R}^2 \to \mathbb{R}$ is *semi-concave* in $y \in J$ μ-uniformly, if there exists a Borel function $u : J \to \mathbb{R}$ such that for μ-almost every x, the mapping $y \mapsto c(x, y) + u(y)$ is continuous and concave on J. In this case, we say that u is a *y-concavifier* on J for c.

Theorem 5.20 ([8, Theorem 2.3])**.** *Let $\mu \leq_c \nu$, $J := \operatorname{conv}(\operatorname{supp}(\nu))$ and $c : \mathbb{R}^2 \to \mathbb{R}$. Suppose that there exists a y-concavifier u on J for c. If J is not compact, then further suppose that $y \mapsto c(x, y) + u(y)$ is of linear growth on J. Then there exists a dual minimizer in the sense of Definition 5.18.*

Theorem 5.21 ([8, Theorem 2.5])**.** *Suppose the assumptions of Theorem 5.20 are satisfied and that further c is Lipschitz continuous on $J \times J$ and u is Lipschitz continuous on J. Then there exists a dual minimizer (φ, ψ, h) such that φ and ψ are Lipschitz continuous on J and $|h|$ is bounded on J.*

Remark 5.22.

1. If c and u in Theorem 5.21 are Lipschitz continuous with constant Λ, then the dual minimizer may be chosen such that φ and ψ are Lipschitz continuous with constants 19Λ and 17Λ on J, and $|h|$ is bounded by 18Λ on J. This is computed in the proof of Beiglböck, Lim & Obłój [8, Theorem 2.5].
2. In a former version, the authors prove Theorem 5.21 for compact J. Then the proof yields that the dual minimizer may be chosen such that φ and ψ are Lipschitz continuous with constants 7Λ and 5Λ on J, and $|h|$ is bounded by 6Λ on J.
3. Analyzing the proof of [8, Theorem 2.5], the authors recognize that the global Lipschitz condition may be weakened. Instead, one demands that there is a $\Lambda > 0$ such that for the domain (I, J) of every irreducible component of (μ, ν), we have
   - $c_y(x, b-) + u'(b-) - c_y(x, a+) - u'(a+) \leq 4\Lambda$ for all $x \in I = (a, b)$.
   - $|c(x, y) - c(x', y)| \leq \Lambda |x - x'|$ for all $x, x', y \in J$. ◊

For us, Theorem 5.21 and Remark 5.22 are of major importance in Chapter 7.

5.2.2. The general market case

Nutz, Stebegg & Tan [65] generalize many aspects of the work of Beiglböck, Nutz & Touzi [9] to the general market case for $d = 1$. In particular, they prove a generalized version of Theorem 5.17 restating the well-known duality and proving the existence of dual optimizers whenever the value of the dual problem is finite. The theorem is the same as Theorem 5.17 only for a version of the dual problem with a different space of functions over which the hedging may be optimized. As the introduction of this so-called dual space takes several pages, we skip the exact formulation of the generalized version.

Lim [58] generalizes some of the results from Beiglböck, Lim & Obłój [8] for the general market in the case $n = 2$. Therefore, let $\mu_{1,1}, \mu_{2,1} \ldots, \mu_{1,d}, \mu_{2,d} \in \mathcal{P}(\mathbb{R})$ be such that $\mu_{1,j} \leq_c \mu_{2,j}$, $j = 1, \ldots, d$ and denote $\mu_1 = (\mu_{1,1}, \ldots, \mu_{1,d})$, $\mu_2 = (\mu_{2,1}, \ldots, \mu_{2,d})$, and $\mu = (\mu_1, \mu_2)$. Consider the set of martingale transport plans $\mathcal{M}_2^d(\mu)$, which can also be written as the set of all probability measures of $\mathbb{R}^d \times \mathbb{R}^d$ such that

1. If $\mathbb{Q}^1$ and $\mathbb{Q}^2$ are the d-copulas induced by an element $\mathbb{Q} \in \mathcal{M}_2^d(\mu)$, then they have marginals $\mu_{1,1}, \ldots, \mu_{1,d}$ and $\mu_{2,1}, \ldots, \mu_{2,d}$ respectively.

2. If $(\mathbb{Q}_x)_x$ is a disintegration of $\mathbb{Q}$ with respect to $\mathbb{Q}^1$, then for any convex function $\phi : \mathbb{R}^d \to \mathbb{R}$ it must satisfy $\phi(x) \leq \int_{\mathbb{R}^d} \phi(y) \mathrm{d}\mathbb{Q}_x(y)$ $\mathbb{Q}^1$-almost surely.

In this situation, Lim [58] considers the same lower price bound problem as we do in (4.6). However, the dual problem is changed slightly. Indeed, instead of assuming $\varphi_{i,j} \in \mathcal{S}$ and $h^j \in \mathcal{C}_b(\mathbb{R})$ for $i = 1, 2$, $j = 1, \ldots, d$, the author considers $\mu_{i,j}$-integrable functions $\varphi_{i,j} : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ and $h^j : \mathbb{R}^d \to \mathbb{R}$ bounded for $i = 1, 2$, $j = 1, \ldots, d$. Differently to our situation and as already discussed in the second part of Remark 4.28, this allows that the dynamic investment in the assets depends on the history of all assets. Clearly, on $\mathbb{R}^{2d}$ the functions have to satisfy the usual sub hedging property

$$\sum_{j=1}^{d} \left[\varphi_{1,j}\left(s_{t_1}^j\right) + \varphi_{2,j}\left(s_{t_2}^j\right) + h^j\left(s_{t_1}^1, \ldots, s_{t_1}^d\right)\left(s_{t_2}^j - s_{t_1}^j\right)\right] \leq c\left(s_{t_1}^1, \ldots, s_{t_1}^d, s_{t_2}^1, \ldots, s_{t_2}^d\right). \tag{5.11}$$

Using such functions, the definition of a dual maximizer is quite similar to the definition of a dual minimizer from Definition 5.18.

Definition 5.23 ([58, Definition 2.1])**.** We say that $\left(\varphi_{1,j}, \varphi_{2,j}, h^j\right)_{j=1,\ldots,d}$ is a *dual maximizer* for the sub hedging problem in (4.9), if

1. The functions $\varphi_{i,j}$ are finite $\mu_{i,j}$-almost surely and such that (5.11) holds.

2. For any minimizer $\mathbb{Q} \in \mathcal{M}_2^d(\mu)$ of the lower price bound problem in (4.6), we have equality in (5.11) $\mathbb{Q}$-almost surely.

Theorem 5.24 ([58, Theorem 2.2]). *Let $(\mu_{1,j}, \mu_{2,j})$ be irreducible with domain (I_j, J_j) for $j = 1, \ldots, d$. Let $c : \mathbb{R}^{2d} \to \mathbb{R}$ be a lower semi-continuous payoff function and suppose that*

$$\left| c\left(s_{t_1}^1, \ldots, s_{t_2}^1, s_{t_1}^d, \ldots, s_{t_2}^d\right) \right| \le \sum_{j=1}^{d} v_j\left(s_{t_1}^j\right) + w_j\left(s_{t_2}^j\right)$$

for some $v_j \in \mathbb{L}^1(\mathbb{R}, \mu_{1,j})$ and $w_j \in \mathbb{L}^1(\mathbb{R}, \mu_{2,j})$, $j = 1, \ldots, d$. Assume further that there is a function $q : J_1 \times \ldots \times J_d \to \mathbb{R}$ such that $q \in \mathbb{L}^1(J_1 \times \ldots \times J_d, \pi)$ for all $\pi \in \Pi_d(\mu_{2,1}, \ldots, \mu_{2,d})$, and for all $(s_{t_2}^1, \ldots, s_{t_2}^d) \in J_1 \times \ldots \times J_d$, we have

$$\sup_{\left(s_{t_1}^1, \ldots, s_{t_1}^d\right) \in I_1 \times \ldots \times I_d} c\left(s_{t_1}^1, \ldots, s_{t_2}^1, s_{t_1}^d, \ldots, s_{t_2}^d\right) \le q\left(s_{t_2}^1, \ldots, s_{t_2}^d\right).$$

Then a dual maximizer exists.

5.3. Improvements of the price bounds

In this section, we present some results on the improvement of the price bounds for martingale optimal transport. In Section 5.3.1, we present an approach introduced by Lütkebohmert & Sester [59], who use additional information on the variance of the asset returns in order to tighten the price bounds in the situation of Beiglböck, Henry-Labordère & Penkner [5] as presented in Section 4.6. In Section 5.3.2, we adapt this approach to the general market using the entire covariance structure of the asset returns.

5.3.1. Improvements using information on return variance

When pricing an exotic option with payoff function $c : \mathbb{R}^n \to \mathbb{R}$ on an underlying risky asset $S = (S_{t_1}, \ldots, S_{t_n})$ with marginals $\mu_1 \le_c \ldots \le_c \mu_n$, Lütkebohmert & Sester [59] aim to improve the price bounds $P_B(c)$ and $\overline{P}_B(c) := \sup_{\mathbb{Q} \in \mathcal{M}_n(\mu_1, \ldots, \mu_n)} \mathbb{E}_{\mathbb{Q}}[c(S_{t_1}, \ldots, S_{t_n})]$ by incorporating more than just the marginals and the martingale property into the set of potential pricing measures. For this purpose, they assume to have information on the variance of the asset returns

$$\operatorname{Var}\left(\frac{S_{t_j} - S_{t_i}}{S_{t_i}}\right)$$

for all $1 \le i < j \le n$ and motivate how this could be achieved in application. Clearly, in order to have a well-defined variance, one does assume that the underlying takes positive values, i.e. $S_{t_i} > 0$, $i = 1, \ldots, n$, which is equivalent to $\operatorname{supp}(\mu_i) \subseteq (0, \infty)$, $i = 1, \ldots, n$.

The authors restrict the set of potential pricing measures to the set of martingale transport plans that additionally satisfy the above variance condition, i.e. to the set

$$\mathcal{V}(\sigma, \mathcal{I}, \mu_1, \ldots, \mu_n) := \left\{ \mathbb{Q} \in \mathcal{M}(\mu_1, \ldots, \mu_n) \;\middle|\; \operatorname{Var}_{\mathbb{Q}}\left(\frac{S_{t_j} - S_{t_i}}{S_{t_i}}\right) = \sigma_{ij}^2 \text{ for } (i,j) \in \mathcal{I} \right\},$$

where $\mathcal{I} \subseteq \{(i,j) \in \{1, \ldots, n\}^2, i < j\}$ is an index set and $\sigma = (\sigma_{ij})_{(i,j) \in \mathcal{I}}$ is the matrix of standard deviations of the asset returns. Then

$$\underline{P}_B(c) \leq \inf_{\mathbb{Q}\in\mathcal{V}(\sigma,\mathcal{I},\mu_1,\dots,\mu_n)} \mathbb{E}_{\mathbb{Q}}[c(S_{t_1},\dots,S_{t_n})] \leq \sup_{\mathbb{Q}\in\mathcal{V}(\sigma,\mathcal{I},\mu_1,\dots,\mu_n)} \mathbb{E}_{\mathbb{Q}}[c(S_{t_1},\dots,S_{t_n})] \leq \overline{P}_B(c).$$

In this scenario, Lütkebohmert & Sester [59] prove several results. Those results are among others a characterization result for the set $\mathcal{V}(\sigma,\mathcal{I},\mu_1,\dots,\mu_n)$ similar to Lemma 5.7, a strong duality result for the problems

$$\inf_{\mathbb{Q}\in\mathcal{V}(\sigma,\mathcal{I},\mu_1,\dots,\mu_n)} \mathbb{E}_{\mathbb{Q}}[c(S_{t_1},\dots,S_{t_n})]$$

and

$$\sup_{\varphi_i\in\mathcal{S}} \left\{ \sum_{i=1}^n \mathbb{E}_{\mu_i}\left[\varphi_i\left(S_{t_i}\right)\right] \middle| \exists h_i \in \mathcal{C}_b(\mathbb{R}^i), \alpha_{ij} \geq 0 : \Psi_{(\varphi_i),(h_i),(\alpha_{ij},\sigma_{ij})} \geq c \text{ on } \mathbb{R}^n \right\},$$

where

$$\begin{aligned}\Psi_{(\varphi_i),(h_i),(\alpha_{ij},\sigma_{ij})}(s_{t_1},\dots,s_{t_n}) \coloneqq& \sum_{i=1}^n \varphi_i(s_{t_i}) + \sum_{i=1}^{n-1} h_i(s_{t_1},\dots,s_{t_i})(s_{t_{i+1}}-s_{t_i}) \\ &+ \sum_{(i,j)\in\mathcal{I}} \alpha_{ij}\left(\left(\frac{s_{t_j}}{s_{t_i}}\right)^2 - 1 - \sigma_{ij}^2\right),\end{aligned}$$

similar to Theorem 4.29, and several results concerning the price gap between the lower and upper price bounds with respect to $\mathcal{V}(\sigma,\mathcal{I},\mu_1,\dots,\mu_n)$. Lütkebohmert & Sester [59] also discuss the improvement of the price bounds by the additional information in an extensive numerical study.

5.3.2. Improvements using information on return covariance

Now let us consider the situation of the general market case. Then proving similar results as Lütkebohmert & Sester [59] considering return variances for the multi-asset case would be a strong generalization, as this gives rise to basket options among others. Anyhow, we intend to derive even stronger results by taking the complete covariance structure of asset returns into account.

For this purpose, let $\mu_{1,1},\dots,\mu_{n,1},\dots,\mu_{1,d},\dots,\mu_{n,d} \in \mathcal{P}((0,\infty))$ be the marginals of d different assets $S^1,\dots,S^d$ taking strictly positive values with $\mu_{1,j} \leq_c \dots \leq_c \mu_{n,j}$ for all $j=1,\dots,d$ and let $c:\mathbb{R}^{nd}\to\mathbb{R}$ be a payoff function. We incorporate information on the covariance of the returns to the set of potential pricing measures $\mathbb{Q}$ by posing conditions of the type

$$\mathrm{Cov}_{\mathbb{Q}}\left(\frac{S^{j_1}_{t_{i_2}} - S^{j_1}_{t_{i_1}}}{S^{j_1}_{t_{i_1}}}, \frac{S^{j_2}_{t_{i_4}} - S^{j_2}_{t_{i_3}}}{S^{j_2}_{t_{i_3}}}\right) = \sigma(j_1,j_2,i_1,\dots,i_4), \tag{5.12}$$

where $j_1,j_2 \in \{1,\dots,d\}$, $i_1,\dots,i_4 \in \{1,\dots,n\}$, $j_1 \leq j_2$, $i_1 < i_2$, and $i_3 < i_4$. That is, we introduce the set of all *covariance restricted martingale transport plans*

$$\mathcal{C}(\Sigma,\mathcal{I}) := \{\mathbb{Q} \in \mathcal{M} \mid \mathbb{Q} \text{ satisfies (5.12) for all } (j_1,j_2,i_1,\dots,i_4) \in \mathcal{I}\},$$

where $\mathcal{I}$ is a suitable index set such that

$$\mathcal{I} \subseteq \left\{(j_1,j_2,i_1,\dots,i_4) \in \{1,\dots,d\}^2 \times \{1,\dots,n\}^4 \;\middle|\; j_1 \le j_2, i_1 < i_2 \text{ and } i_3 < i_4\right\}$$

and $\Sigma = (\sigma(j_1,j_2,i_1,\dots,i_4))_{(j_1,j_2,i_1,\dots,i_4)\in\mathcal{I}}$ contains all return covariances. We shortly denote a generic index by $\mathbb{I}_6 := (j_1,j_2,i_1,\dots,i_4)$.

In this situation, we consider the upper and lower price bound problems

$$\begin{aligned} P^{\mathcal{C}}_{nd}(c) &:= \inf_{\mathbb{Q}\in\mathcal{C}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S^1_{t_1},\dots,S^1_{t_n},\dots,S^d_{t_1},\dots,S^d_{t_n}\right)\right] \\ &\le \sup_{\mathbb{Q}\in\mathcal{C}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S^1_{t_1},\dots,S^1_{t_n},\dots,S^d_{t_1},\dots,S^d_{t_n}\right)\right] =: \overline{P}^{\mathcal{C}}_{nd}(c). \end{aligned} \tag{5.13}$$

Remark 5.25.

1. When it comes to application, we often choose $i_1 = i_3 = i_2 - 1 = i_4 - 1$ in order to consider one step returns with either $j_1 = j_2$ regarding the variances of the returns or $j_1 \neq j_2$ regarding the covariances of the returns of two different assets at the same time point.

2. For $k = 1,\dots,d$, we denote by

$$\mathcal{V}^k(\Sigma,\mathcal{I}) := \left\{\mathbb{Q} \in \mathcal{M}_n(\mu_{1,k},\dots,\mu_{n,k}) \;\middle|\; \operatorname{Var}_{\mathbb{Q}}\left(\frac{S^k_{t_j} - S^k_{t_i}}{S^k_{t_i}}\right) = \sigma(k,k,i,j,i,j) \right.$$
$$\left. \text{for all } (k,k,i,j,i,j) \in \mathcal{I}\right\}$$

the set of all martingale transport plans restricted with respect to the return variances of the asset S^k as considered by Lütkebohmert & Sester [59]. Then we have

$$\begin{aligned} \mathcal{C}(\Sigma,\mathcal{I}) \subseteq \mathcal{V}(\Sigma,\mathcal{I}) &:= \Pi_{nd}\left(\mathcal{V}^1(\Sigma,\mathcal{I}),\dots,\mathcal{V}^d(\Sigma,\mathcal{I})\right) \\ &:= \left\{\pi \in \mathcal{P}(\mathbb{R}^{nd}) \mid \pi \text{ has } n\text{-marginals } \pi_j \in \mathcal{V}^j(\Sigma,\mathcal{I})\right\} \subseteq \mathcal{M}, \end{aligned}$$

where $\mathcal{V}(\Sigma,\mathcal{I})$ is the set of martingale transport plans only incorporating variances. We define $\mathcal{I}^= := \{(j_1,j_2,i_1,\dots,i_4) \in \mathcal{I} \mid j_1 = j_2, i_1 = i_3, i_2 = i_4\}$, the subset of $\mathcal{I}$ only containing indices that lead to variance conditions. Then $\mathcal{V}(\Sigma,\mathcal{I}) = \mathcal{C}(\Sigma,\mathcal{I}^=)$. We may then consider the price bound problems

$$\begin{aligned} P^{\mathcal{V}}_{nd}(c) &:= \inf_{\mathbb{Q}\in\mathcal{V}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S^1_{t_1},\dots,S^1_{t_n},\dots,S^d_{t_1},\dots,S^d_{t_n}\right)\right] \\ &\le \sup_{\mathbb{Q}\in\mathcal{V}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S^1_{t_1},\dots,S^1_{t_n},\dots,S^d_{t_1},\dots,S^d_{t_n}\right)\right] =: \overline{P}^{\mathcal{V}}_{nd}(c). \end{aligned}$$

Opposed to $\mathcal{V}(\Sigma,\mathcal{I})$, we may consider the set of martingale transport plans only incorporating true covariances. That is the set $(\mathcal{C} \setminus \mathcal{V})(\Sigma, \mathcal{I}) := \mathcal{C}(\Sigma, \mathcal{I} \setminus \mathcal{I}^{=})$. We may then also consider the price bound problems

$$\begin{aligned} P_{nd}^{\mathcal{C}\setminus\mathcal{V}}(c) &:= \inf_{\mathbb{Q}\in(\mathcal{C}\setminus\mathcal{V})(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S_{t_1}^1, \ldots, S_{t_n}^1, \ldots, S_{t_1}^d, \ldots, S_{t_n}^d\right)\right] \\ &\leq \sup_{\mathbb{Q}\in(\mathcal{C}\setminus\mathcal{V})(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S_{t_1}^1, \ldots, S_{t_n}^1, \ldots, S_{t_1}^d, \ldots, S_{t_n}^d\right)\right] =: \overline{P}_{nd}^{\mathcal{C}\setminus\mathcal{V}}(c). \end{aligned}$$

3. As the asset price processes S^k, $k = 1, \ldots, d$, are martingales under any potential pricing measure, we may derive an alternative representation of the covariances. Indeed, we have

$$\begin{aligned} &\mathrm{Cov}_{\mathbb{Q}}\left(\frac{S_{t_{i_2}}^{j_1} - S_{t_{i_1}}^{j_1}}{S_{t_{i_1}}^{j_1}}, \frac{S_{t_{i_4}}^{j_2} - S_{t_{i_3}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right) = \mathrm{Cov}_{\mathbb{Q}}\left(\frac{S_{t_{i_2}}^{j_1}}{S_{t_{i_1}}^{j_1}}, \frac{S_{t_{i_4}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right) \\ &= \mathbb{E}_{\mathbb{Q}}\left(\frac{S_{t_{i_2}}^{j_1}}{S_{t_{i_1}}^{j_1}} \cdot \frac{S_{t_{i_4}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right) - \mathbb{E}_{\mathbb{Q}}\left(\frac{S_{t_{i_2}}^{j_1}}{S_{t_{i_1}}^{j_1}}\right) \cdot \mathbb{E}_{\mathbb{Q}}\left(\frac{S_{t_{i_4}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right) = \mathbb{E}_{\mathbb{Q}}\left(\frac{S_{t_{i_2}}^{j_1}}{S_{t_{i_1}}^{j_1}} \cdot \frac{S_{t_{i_4}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right) - 1, \end{aligned}$$

where in the last step we use the tower property and the martingale property. If we assume $j_1 = j_2$ and $i_2 \neq i_4$ or $i_1 \neq i_3$, then again using the tower property multiple times, we obtain that the asset return covariances are equal to zero, as martingale increments are always uncorrelated. Therefore, we may reduce the index set under consideration to

$$\begin{aligned} \mathcal{I} \subseteq \{(j_1, j_2, i_1, \ldots, i_4) \in \{1, \ldots, d\}^2 \times \{1, \ldots, n\}^4 \mid j_1 \leq j_2, i_1 < i_2, i_3 < i_4, \\ \text{and if } j_1 = j_2, \text{ then } i_2 = i_4 \text{ and } i_1 = i_3\}. \quad \diamondsuit \end{aligned}$$

Let us now at least shortly motivate how we could gain the necessary information about the return covariances. By the third part of the previous remark, we have to derive information about the common distribution of the asset returns $S_{t_{i_2}}^{j_1}/S_{t_{i_1}}^{j_1}$ and $S_{t_{i_4}}^{j_2}/S_{t_{i_3}}^{j_2}$ for all $\mathbb{I}_6 \in \mathcal{I}$. For this purpose, we may use observable market prices of suitable rainbow options. While it is difficult to derive the full distribution and thus the exact covariances, we at least present an approach to derive upper and lower bounds for the covariances.

We assume that the prices of forward start options with payoff functions

$$\max\left(\frac{S_{t_{i_2}}^{j_1} - S_{t_{i_1}}^{j_1}}{S_{t_{i_1}}^{j_1}}, F\right) \quad \text{and} \quad \max\left(\frac{S_{t_{i_4}}^{j_2} - S_{t_{i_3}}^{j_2}}{S_{t_{i_3}}^{j_2}}, F\right),$$

and the prices of put on max options on the asset return with payoff function

$$\left(K - \max\left\{\frac{S_{t_{i_2}}^{j_1}}{S_{t_{i_1}}^{j_1}}, \frac{S_{t_{i_4}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right\}\right)^+$$

are observable. We denote the associated option prices by $S_{i_1,i_2,j_1}(F)$, $S_{i_3,i_4,j_2}(F)$ and $P_{i_1,i_2,i_3,i_4,j_1,j_2}^M(K)$. If we assume that the prices are differentiable and calculated under a martingale measure $\mathbb{Q} \in \mathcal{M}$, then we have

$$\frac{\partial}{\partial F} S_{i_1,i_2,j_1}(F-1) = \mathbb{Q}\left(\frac{S^{j_1}_{t_{i_2}}}{S^{j_1}_{t_{i_1}}} \leq F\right) \quad \text{and} \quad \frac{\partial}{\partial F} S_{i_3,i_4,j_2}(F-1) = \mathbb{Q}\left(\frac{S^{j_2}_{t_{i_4}}}{S^{j_2}_{t_{i_3}}} \leq F\right)$$

as shown in Lütkebohmert & Sester [59, Sec. 2.2.1]. Proceeding analogously, we obtain

$$\begin{aligned}\frac{\partial}{\partial K} P^M_{i_1,i_2,i_3,i_4,j_1,j_2}(K) &= \frac{\partial}{\partial K} \mathbb{E}_{\mathbb{Q}}\left[\left(K - \max\left\{\frac{S^{j_1}_{t_{i_2}}}{S^{j_1}_{t_{i_1}}}, \frac{S^{j_2}_{t_{i_4}}}{S^{j_2}_{t_{i_3}}}\right\}\right)^+\right] \\ &= \mathbb{Q}\left(\max\left\{\frac{S^{j_1}_{t_{i_2}}}{S^{j_1}_{t_{i_1}}}, \frac{S^{j_2}_{t_{i_4}}}{S^{j_2}_{t_{i_3}}}\right\} \leq K\right).\end{aligned}$$

Now let $\mathbb{I}_6 = (j_1, j_2, i_1, \ldots, i_4) \in \mathcal{I}$ be fixed, $K_1, K_2 \in \mathbb{R}_+$ and denote $R_1 := S^{j_1}_{t_{i_2}}/S^{j_1}_{t_{i_1}}$ and $R_2 := S^{j_2}_{t_{i_4}}/S^{j_2}_{t_{i_3}}$. In order to bound the covariance $\mathrm{Cov}_{\mathbb{Q}}(R_1, R_2)$, we estimate the probability

$$\mathbb{Q}(R_1 > K_1, R_2 > K_2) = \mathbb{Q}(R_1 > K_1) + \mathbb{Q}(R_2 > K_2) - \mathbb{Q}(\{R_1 > K_1\} \cup \{R_2 > K_2\}).$$

For the last expression, we have

$$\begin{aligned}\mathbb{Q}(\max\{R_1, R_2\} > \max\{K_1, K_2\}) &\leq \mathbb{Q}(\{R_1 > K_1\} \cup \{R_2 > K_2\}) \\ &\leq \mathbb{Q}(\max\{R_1, R_2\} > \min\{K_1, K_2\}).\end{aligned}$$

Thus, we may define random vectors $(\underline{R}_1, \underline{R}_2)$ and $(\overline{R}_1, \overline{R}_2)$ by

$$\begin{aligned}&\mathbb{Q}(\underline{R}_1 > K_1, \underline{R}_2 > K_2) \\ &:= \min\{\max\{\mathbb{Q}(R_1 > K_1) + \mathbb{Q}(R_2 > K_2) - \mathbb{Q}(\max\{R_1, R_2\} > \min\{K_1, K_2\}), 0\}, 1\}, \\ &\mathbb{Q}(\overline{R}_1 > K_1, \overline{R}_2 > K_2) \\ &:= \min\{\max\{\mathbb{Q}(R_1 > K_1) + \mathbb{Q}(R_2 > K_2) - \mathbb{Q}(\max\{R_1, R_2\} > \max\{K_1, K_2\}), 0\}, 1\}.\end{aligned}$$

All defining probabilities are determined by observable market prices and we have

$$\mathbb{Q}(\underline{R}_1 > K_1, \underline{R}_2 > K_2) \leq \mathbb{Q}(R_1 > K_1, R_2 > K_2) \leq \mathbb{Q}(\overline{R}_1 > K_1, \overline{R}_2 > K_2).$$

By definition, that is equivalent to $(\underline{R}_1, \underline{R}_2) \leq_{uo} (R_1, R_2) \leq_{uo} (\overline{R}_1, \overline{R}_2)$, where $\leq_{uo}$ is the *upper orthant order* as defined in Shaked & Shantikumar [72, Chap. 6.G]. Now applying [72, Theorem 6.G.1 (a)], we have

$$\mathbb{E}_{\mathbb{Q}}[\underline{R}_1 \cdot \underline{R}_2] - 1 \leq \mathrm{Cov}_{\mathbb{Q}}(R_1, R_2) \leq \mathbb{E}_{\mathbb{Q}}[\overline{R}_1 \cdot \overline{R}_2] - 1,$$

as $f(x, y) = x \cdot y$ is the product of two non-negative increasing functions.

These are the desired bounds for the covariance. Clearly, whenever $R_1 = R_2$, i.e. when we have in fact variances, then the bounds collapse to singletons just as derived by Lütkebohmert & Sester [59].

However, the third part of Remark 5.25 does not only help to bound the covariance but also implies a characterization lemma for $\mathcal{C}(\Sigma, \mathcal{I})$ similar to Lemma 5.7.

Lemma 5.26. *Let $\mathbb{Q} \in \mathcal{M}$. Then the following are equivalent.*

1. $\mathbb{Q} \in \mathcal{C}(\Sigma, \mathcal{I})$.
2. *For all $\mathbb{I}_6 \in \mathcal{I}$, we have*

$$\mathrm{Cov}_{\mathbb{Q}}\left(\frac{S_{t_{i_2}}^{j_1} - S_{t_{i_1}}^{j_1}}{S_{t_{i_1}}^{j_1}}, \frac{S_{t_{i_4}}^{j_2} - S_{t_{i_3}}^{j_2}}{S_{t_{i_3}}^{j_2}}\right) = \sigma(\mathbb{I}_6).$$

3. *For all $\alpha_{\mathbb{I}_6} \in \mathbb{R}$, we have*

$$\sum_{\mathbb{I}_6 \in \mathcal{I}} \int_{\mathbb{R}^{nd}} \alpha_{\mathbb{I}_6}\left(\left(\frac{s_{t_{i_2}}^{j_1}}{s_{t_{i_1}}^{j_1}} \cdot \frac{s_{t_{i_4}}^{j_2}}{s_{t_{i_3}}^{j_2}} - 1 - \sigma(\mathbb{I}_6)\right)\right) \mathbb{Q}\left(\mathrm{d}\left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right)\right) = 0.$$

Proof. 1. $\iff$ 2. Holds by definition of $\mathcal{C}(\Sigma, \mathcal{I})$.

2. $\Rightarrow$ 3. Clearly, the equation in the second condition implies

$$\int_{\mathbb{R}^{nd}} \alpha_{\mathbb{I}_6}\left(\left(\frac{s_{t_{i_2}}^{j_1}}{s_{t_{i_1}}^{j_1}} \cdot \frac{s_{t_{i_4}}^{j_2}}{s_{t_{i_3}}^{j_2}} - 1 - \sigma(\mathbb{I}_6)\right)\right) \mathbb{Q}\left(\mathrm{d}\left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right)\right) = 0$$

for any $\alpha_{\mathbb{I}_6} \in \mathbb{R}$. Summing over all $\mathbb{I}_6 \in \mathcal{I}$ implies the claim.

3. $\Rightarrow$ 2. In order to prove the equation in the second condition for a certain $\mathbb{I}_6 \in \mathcal{I}$, choose $\alpha_{\mathbb{I}_6} = 1$ and $\alpha_{\mathbb{J}_6} = 0$ for all $\mathbb{J}_6 \in \mathcal{I} \setminus \{\mathbb{I}_6\}$. ☐

This lemma implies a natural dual problem, which we introduce in the following corollary, where we provide a strong duality result.

Corollary 5.27. *Let $\mu_{1,1}, \ldots, \mu_{n,1}, \ldots, \mu_{1,d}, \ldots, \mu_{n,d} \in \mathcal{P}((0,\infty))$ be such that for all $i = 1, \ldots, n$ and all $j = 1, \ldots, d$, we have $\mu_{1,j} \leq_c \ldots \leq_c \mu_{n,j}$, $\int_{(0,\infty)} x^2 \mu_{i,j}(\mathrm{d}x) < \infty$ and that there exist $b_{i,j} > 0$ with $\mathrm{supp}(\mu_{i,j}) \subseteq [b_{i,j}, \infty)$. Assume further that $\mathcal{C}(\Sigma, \mathcal{I}) \neq \emptyset$ and let $c : \mathbb{R}^{nd} \to (-\infty, \infty]$ be a lower semi-continuous payoff function such that there is a constant $K \in \mathbb{R}$ with*

$$c\left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right) \geq -K\left(1 + \sum_{j=1}^{d} \sum_{i=1}^{n} \left|s_{t_i}^j\right|\right)$$

for all $\left(s_{t_1}^1, \ldots, s_{t_n}^1, \ldots, s_{t_1}^d, \ldots, s_{t_n}^d\right) \in (0,\infty)^{nd}$. Then

$$\inf_{\mathbb{Q} \in \mathcal{C}(\Sigma, \mathcal{I})} \mathbb{E}_{\mathbb{Q}}\left[c\left(S_{t_1}^1, \ldots, S_{t_n}^1, \ldots, S_{t_1}^d, \ldots, S_{t_n}^d\right)\right] = \sup \sum_{j=1}^{d} \sum_{i=1}^{n} \mathbb{E}_{\mu_{i,j}}\left[\varphi_{i,j}(S_{t_i}^j)\right],$$

where the supremum is taken over $\varphi_{i,j} \in \mathcal{S}$ such that there are $h_i^j \in \mathcal{C}_b(\mathbb{R}^i)$ and $\alpha_{\mathbb{I}_6} \in \mathbb{R}$ with

$$\begin{aligned}
&\Psi_{(\varphi_{i,j}),(h_i^j),(\alpha_{\mathbb{I}_6})}\left(s_{t_1}^1,\dots,s_{t_n}^1,\dots,s_{t_1}^d,\dots,s_{t_n}^d\right) \\
&\quad := \sum_{j=1}^{d}\sum_{i=1}^{n}\varphi_{i,j}\left(s_{t_i}^j\right) + \sum_{j=1}^{d}\sum_{i=1}^{n-1} h_i^j\left(s_{t_1}^j,\dots,s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right) \\
&\quad + \sum_{\mathbb{I}_6\in\mathcal{I}}\alpha_{\mathbb{I}_6}\left(\left(\frac{s_{t_{i_2}}^{j_1}}{s_{t_{i_1}}^{j_1}}\cdot\frac{s_{t_{i_4}}^{j_2}}{s_{t_{i_3}}^{j_2}} - 1 - \sigma(\mathbb{I}_6)\right)\right) \\
&\quad \le c\left(s_{t_1}^1,\dots,s_{t_n}^1,\dots,s_{t_1}^d,\dots,s_{t_n}^d\right)
\end{aligned}$$

for all $\left(s_{t_1}^1,\dots,s_{t_n}^1,\dots,s_{t_1}^d,\dots,s_{t_n}^d\right) \in (0,\infty)^{nd}$. Furthermore, there is a minimizer $\mathbb{Q}^* \in \mathcal{C}(\Sigma,\mathcal{I})$ for the lower price bound problem in (5.13)*.*

Remark 5.28. The proof of Corollary 5.27 goes in the exact same lines as the proofs of Theorem 5.1 and Lütkebohmert & Sester [59, Corollary 3.2]. Therefore, we do not report it in any detail.

However, let us stress that the additional assumptions on the marginals are necessary in order to be able to bound $\chi_{(\varphi_{i,j}),(h_i^j),(\alpha_{\mathbb{I}_6})} : \mathbb{R}^{nd} \to \mathbb{R}$, which is defined by

$$\begin{aligned}
&\chi_{(\varphi_{i,j}),(h_i^j),(\alpha_{\mathbb{I}_6})}\left(s_{t_1}^1,\dots,s_{t_n}^1,\dots,s_{t_1}^d,\dots,s_{t_n}^d\right) \\
&\quad := c\left(s_{t_1}^1,\dots,s_{t_n}^1,\dots,s_{t_1}^d,\dots,s_{t_n}^d\right) \\
&\quad - \sum_{j=1}^{d}\sum_{i=1}^{n-1} h_i^j\left(s_{t_1}^j,\dots,s_{t_i}^j\right)\left(s_{t_{i+1}}^j - s_{t_i}^j\right) \\
&\quad - \sum_{\mathbb{I}_6\in\mathcal{I}}\alpha_{\mathbb{I}_6}\left(\left(\frac{s_{t_{i_2}}^{j_1}}{s_{t_{i_1}}^{j_1}}\cdot\frac{s_{t_{i_4}}^{j_2}}{s_{t_{i_3}}^{j_2}} - 1 - \sigma(\mathbb{I}_6)\right)\right),
\end{aligned}$$

by the sum of integrable functions as in condition (5.6) of Lemma 5.5.

Indeed, for the only part of $\chi_{(\varphi_{i,j}),(h_i^j),(\alpha_{\mathbb{I}_6})}$ that is different to $\chi_{(\varphi_{i,j}),(h_i^j)}$ as defined in the proof of Theorem 5.1, we have

$$\begin{aligned}
\left(\frac{s_{t_{i_2}}^{j_1}}{s_{t_{i_1}}^{j_1}}\cdot\frac{s_{t_{i_4}}^{j_2}}{s_{t_{i_3}}^{j_2}} - 1 - \sigma(\mathbb{I}_6)\right) &\le \left(\frac{s_{t_{i_2}}^{j_1}}{b_{i_1}^{j_1}}\cdot\frac{s_{t_{i_4}}^{j_2}}{b_{i_3}^{j_2}} - 1 - \sigma(\mathbb{I}_6)\right) \\
&\le \left(\left(\frac{s_{t_{i_2}}^{j_1}}{b_{i_1}^{j_1}}\right)^2 + \left(\frac{s_{t_{i_4}}^{j_2}}{b_{i_3}^{j_2}}\right)^2 - 1 - \sigma(\mathbb{I}_6)\right),
\end{aligned}$$

where the latter is the sum of integrable functions, as the second moments of the marginals do exist. ◊

THE TWO MARGINAL & TWO ASSET CASE

Let $S^1 = (X_1, X_2)$ and $S^2 = (Y_1, Y_2)$ be two underlying asset price processes with marginals $\mu_1, \mu_2, \nu_1, \nu_2 \in \mathcal{P}((0,\infty))$ such that $\mu_1 \leq_c \mu_2$ and $\nu_1 \leq_c \nu_2$ respectively. We consider a payoff function $c : \mathbb{R}^4 \to \mathbb{R}$ and the corresponding optimization problems $\inf_{\mathbb{Q}\in\mathcal{C}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}[c]$ and $\sup_{\mathbb{Q}\in\mathcal{C}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}[c]$, where $\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}$ with

$$\sigma_{XY} = \mathrm{Cov}_{\mathbb{Q}}\left(\frac{X_2 - X_1}{X_1} \cdot \frac{Y_2 - Y_1}{Y_1}\right), \quad \sigma_X^2 = \mathrm{Var}_{\mathbb{Q}}\left(\frac{X_2 - X_1}{X_1}\right), \quad \sigma_Y^2 = \mathrm{Var}_{\mathbb{Q}}\left(\frac{Y_2 - Y_1}{Y_1}\right),$$

and $\mathcal{I} = \{(j_1, j_2, i_1, \ldots, i_4) \in \{1,2\}^6 \mid j_1 \leq j_2, i_1 = i_3 = 1, i_2 = i_4 = 2\}$. Let us characterize the non-emptiness of $\mathcal{C}(\Sigma, \mathcal{I})$ and thus the well-posedness of the price bound problems.

Proposition 5.29. *Suppose $\mu_1 \leq_c \mu_2$ and $\nu_1 \leq_c \nu_2$, i.e. $\mathcal{M}_2(\mu_1,\mu_2) \neq \emptyset$, $\mathcal{M}_2(\nu_1,\nu_2) \neq \emptyset$ and $\mathcal{M} \neq \emptyset$, and define the functions $\tilde{c}_X(x_1,x_2) := \left(\frac{x_2}{x_1}\right)^2 - 1$, $\tilde{c}_Y(y_1,y_2) := \left(\frac{y_2}{y_1}\right)^2 - 1$ and $\tilde{c}_{XY}(x_1,x_2,y_1,y_2) := \left(\frac{x_2}{x_1}\frac{y_2}{y_1}\right) - 1$. Then $\mathcal{C}(\Sigma,\mathcal{I}) \neq \emptyset$ if and only if*

1. $\sigma_X^2 \in I_{\sigma_X^2} := \left[\inf_{\mathbb{Q}\in\mathcal{M}_2(\mu_1,\mu_2)} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_X], \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_1,\mu_2)} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_X]\right].$

2. $\sigma_Y^2 \in I_{\sigma_Y^2} := \left[\inf_{\mathbb{Q}\in\mathcal{M}_2(\nu_1,\nu_2)} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_Y], \sup_{\mathbb{Q}\in\mathcal{M}_2(\nu_1,\nu_2)} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_Y]\right].$

3. $\sigma_{XY} \in I_{\sigma_{XY}} := \left[\inf_{\mathbb{Q}\in\mathcal{V}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_{XY}], \sup_{\mathbb{Q}\in\mathcal{V}(\Sigma,\mathcal{I})} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_{XY}]\right].$

We shortly write $I_\Sigma := I_{\sigma_X^2} \times I_{\sigma_Y^2} \times I_{\sigma_{XY}}$.

Remark 5.30. Similar assertions hold for $(\mathcal{C} \setminus \mathcal{V})(\Sigma,\mathcal{I})$, where the first and the second condition vanish and the third condition becomes

$$\sigma_{XY} \in I_{\sigma_{XY}}^{\mathcal{C}\setminus\mathcal{V}} := \left[\inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_{XY}], \sup_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_{\mathbb{Q}}[\tilde{c}_{XY}]\right],$$

and for $\mathcal{V}(\Sigma,\mathcal{I})$, where the third condition vanishes. ♢

Proof of Proposition 5.29. By the compactness of $\mathcal{M}_2(\mu_1,\mu_2)$ and the lower and upper semi-continuity of $\tilde{c}_X$, analogously to the proof of Theorem 5.1, there are measures $\underline{\mathbb{Q}}, \overline{\mathbb{Q}} \in \mathcal{M}_2(\mu_1,\mu_2)$ such that $I_{\sigma_X^2} = \left[\mathbb{E}_{\underline{\mathbb{Q}}}[\tilde{c}_X], \mathbb{E}_{\overline{\mathbb{Q}}}[\tilde{c}_X]\right]$. If we now assume

$$\sigma_X^2 \in \left[\underline{\sigma}_X^2, \overline{\sigma}_X^2\right] := \left[\mathbb{E}_{\underline{\mathbb{Q}}}[\tilde{c}_X], \mathbb{E}_{\overline{\mathbb{Q}}}[\tilde{c}_X]\right],$$

then there is a $\lambda \in [0,1]$ such that $\sigma_X^2 = \lambda\underline{\sigma}_X^2 + (1-\lambda)\overline{\sigma}_X^2$. We also know that then $\lambda\underline{\mathbb{Q}} + (1-\lambda)\overline{\mathbb{Q}} \in \mathcal{V}^1(\Sigma,\mathcal{I})$, as $\mathcal{M}_2(\mu_1,\mu_2)$ is convex and the convex combination is chosen such that, by (5.15), the measure has variance σ_X^2. Hence, $\mathcal{V}^1(\Sigma,\mathcal{I}) \neq \emptyset$.

Conversely, if $\mathbb{Q} \in \mathcal{V}^1(\Sigma,\mathcal{I})$, then in particular $\mathbb{Q} \in \mathcal{M}_2(\mu_1,\mu_2)$. Hence, $\mathbb{E}_{\mathbb{Q}}[\tilde{c}_X] \in I_{\sigma_X^2}$.

The same applies to σ_Y^2 such that we have $\mathcal{V}^1(\Sigma,\mathcal{I}) \neq \emptyset$ if and only if the first condition is satisfied and $\mathcal{V}^2(\Sigma,\mathcal{I}) \neq \emptyset$ if and only if the second condition is satisfied.

Clearly, this means that $\mathcal{V}(\Sigma,\mathcal{I}) \neq \emptyset$ if and only if the first and the second condition are satisfied. Hence, it remains to show that there is a measure $\mathbb{Q} \in \mathcal{V}(\Sigma,\mathcal{I})$ such that $\operatorname{Cov}_{\mathbb{Q}}\left(\frac{X_2}{X_1}, \frac{Y_2}{Y_1}\right) = \sigma_{XY}$ if and only if the third condition is satisfied.

For this purpose, denote by $\underline{\sigma_{XY}}$ and $\overline{\sigma_{XY}}$ the lower and the upper bound for σ_{XY} given by the third condition, i.e. $I_{\sigma_{XY}} = [\underline{\sigma_{XY}}, \overline{\sigma_{XY}}]$. Then there exists a $\lambda \in [0,1]$ such that $\sigma_{XY} = \lambda \underline{\sigma_{XY}} + (1-\lambda)\overline{\sigma_{XY}}$. We now denote by $\underline{\mathbb{Q}}$ and $\overline{\mathbb{Q}}$ measures in $\mathcal{V}(\Sigma,\mathcal{I})$ that realize $\underline{\sigma_{XY}}$ and $\overline{\sigma_{XY}}$. Then we define the measure $\mathbb{Q} := \lambda\underline{\mathbb{Q}} + (1-\lambda)\overline{\mathbb{Q}}$. If we can show that $\mathbb{Q} \in \mathcal{V}(\Sigma,\mathcal{I})$ and that $\mathbb{Q}$ realizes the desired covariance σ_{XY}, then we have $\mathbb{Q} \in \mathcal{C}(\Sigma,\mathcal{I})$.

As for all suitable functions $f : \mathbb{R}^4 \to \mathbb{R}$

$$\mathbb{E}_{\mathbb{Q}}[f(X_1,X_2,Y_1,Y_2)] = \lambda \mathbb{E}_{\underline{\mathbb{Q}}}[f(X_1,X_2,Y_1,Y_2)] + (1-\lambda)\mathbb{E}_{\overline{\mathbb{Q}}}[f(X_1,X_2,Y_1,Y_2)], \tag{5.14}$$

it is clear that the marginal and martingale conditions of $\mathcal{V}(\Sigma,\mathcal{I})$ are satisfied by $\mathbb{Q}$.

While in general variances and covariances may not be expressed as a simple expected value as in (5.14), by the martingale property we have

$$\begin{aligned}
\operatorname{Var}_{\mathbb{Q}}\left(\frac{X_2-X_1}{X_1}\right) &= \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{X_2}{X_1}\right)^2 - 1\right], \\
\operatorname{Var}_{\mathbb{Q}}\left(\frac{Y_2-Y_1}{Y_1}\right) &= \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{Y_2}{Y_1}\right)^2 - 1\right], \\
\operatorname{Cov}_{\mathbb{Q}}\left(\frac{X_2-X_1}{X_1}, \frac{Y_2-Y_1}{Y_1}\right) &= \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{X_2}{X_1}\cdot\frac{Y_2}{Y_1}\right) - 1\right].
\end{aligned} \tag{5.15}$$

Hence, $\mathbb{Q}$ has the correct marginal return variances and the correct return covariance and thus $\mathbb{Q} \in \mathcal{C}(\Sigma,\mathcal{I})$.

Conversely, let $\mathbb{Q} \in \mathcal{C}(\Sigma,\mathcal{I})$ and denote $\sigma_{XY} = \mathbb{E}_{\mathbb{Q}}\left[\frac{X_2}{X_1}\cdot\frac{Y_2}{Y_1}\right] - 1$. Then the subset property $\mathcal{C}(\Sigma,\mathcal{I}) \subseteq \mathcal{V}(\Sigma,\mathcal{I})$ immediately implies the third condition. □

Example 5.31. Let us discuss some numerical aspects. We stick to the same exotic options and marginals as in Example 5.12. As we consider the general market case for $n = d = 2$, the notation of the payoff functions simplifies as presented in Table 5.5.

Exotic option type	Payoff function
Basket option	$c_B(X_1,X_2,Y_1,Y_2) = \left(\frac{1}{4}(X_1+X_2+Y_1+Y_2) - 10\right)^+$
Binary option	$c_1(X_1,X_2,Y_1,Y_2) = \mathbb{1}_{\{X_2>X_1\}}\cdot\mathbb{1}_{\{Y_2>Y_1\}}$
Asian option	$c_A(X_1,X_2,Y_1,Y_2) = \frac{1}{4}(X_2-X_1)^+\cdot(Y_2-Y_1)^+$
Variance option	$c_V(X_1,X_2,Y_1,Y_2) = \left(\frac{X_2-X_1}{X_1}\right)^2\cdot\left(\frac{Y_2-Y_1}{Y_1}\right)^2$
Covariance option	$c_C(X_1,X_2,Y_1,Y_2) = \left(\frac{X_2-X_1}{X_1}\cdot\frac{Y_2-Y_1}{Y_1}\right)^+$

Table 5.5.: Payoff functions

Recall that we consider the discrete marginals

$$\mu_1 = \frac{1}{3}(\delta_8 + \delta_{10} + \delta_{12}) \leq_c \frac{1}{4}(\delta_7 + \delta_9 + \delta_{11} + \delta_{13}) = \mu_2,$$
$$\nu_1 = \frac{1}{3}(\delta_8 + \delta_{10} + \delta_{12}) \leq_c \frac{1}{5}(\delta_4 + \delta_7 + \delta_{10} + \delta_{13} + \delta_{16}) = \nu_2.$$

In this situation, we discuss the impact of additional information about asset return covariances to the upper and lower price bounds. For this purpose, recall that for fixed marginals and payoff function, we may interpret the lower price bounds as functions of the covariance parameters, i.e.

$$\underline{\mathcal{C}} : I_\Sigma \to \mathbb{R}, \quad (\sigma_X^2, \sigma_Y^2, \sigma_{XY}) \mapsto \inf_{\mathbb{Q}\in\mathcal{C}(\Sigma,\mathcal{I})} \mathbb{E}_\mathbb{Q}[c],$$
$$\underline{\mathcal{V}} : I_\Sigma \to \mathbb{R}, \quad (\sigma_X^2, \sigma_Y^2, \sigma_{XY}) \mapsto \inf_{\mathbb{Q}\in\mathcal{V}(\Sigma,\mathcal{I})} \mathbb{E}_\mathbb{Q}[c],$$
$$\underline{\mathcal{C}\setminus\mathcal{V}} : I_\Sigma \to \mathbb{R}, \quad (\sigma_X^2, \sigma_Y^2, \sigma_{XY}) \mapsto \inf_{\mathbb{Q}\in(\mathcal{C}\setminus\mathcal{V})(\Sigma,\mathcal{I})} \mathbb{E}_\mathbb{Q}[c],$$
$$\underline{\mathcal{M}} : I_\Sigma \to \mathbb{R}, \quad (\sigma_X^2, \sigma_Y^2, \sigma_{XY}) \mapsto \inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}_\mathbb{Q}[c],$$

Analogously, we may define the upper price bounds as functions of the covariance parameters. We denote the functions by $\overline{\mathcal{C}}$, $\overline{\mathcal{V}}$, $\overline{\mathcal{C}\setminus\mathcal{V}}$, $\overline{\mathcal{M}}$. We collect the associated pairs of price bounds denoting them by $\mathcal{C}, \mathcal{V}, \mathcal{C}\setminus\mathcal{V}, \mathcal{M}$, where clearly $\mathcal{V}$ is actually independent of σ_{XY}, $\mathcal{C}\setminus\mathcal{V}$ is independent of σ_X^2 and σ_Y^2 and $\mathcal{M}$ is independent of all three parameters.

As Proposition 5.29 suggests, the parameters have to lie in certain intervals in order not to contradict the existence of a suitable pricing measure. We partition these intervals uniformly and calculate the price bounds for all parameter combinations in the partition. By the structure of the mappings under consideration, we are not able to present the bounds in tables or figures depending on all parameters, as these objects are three dimensional. Instead, for certain fixed σ_X^2 and σ_Y^2, we present the bounds as a function of σ_{XY}.

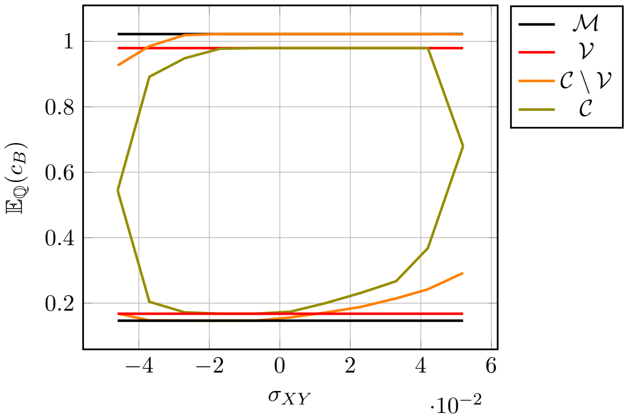
We first state the intervals from which the input parameters may come. We have

$$\sigma_X^2 \in [0.0211, 0.0298], \quad \sigma_Y^2 \in [0.1281, 0.2062], \quad \sigma_{XY} \in [-0.0656, 0.0718].$$

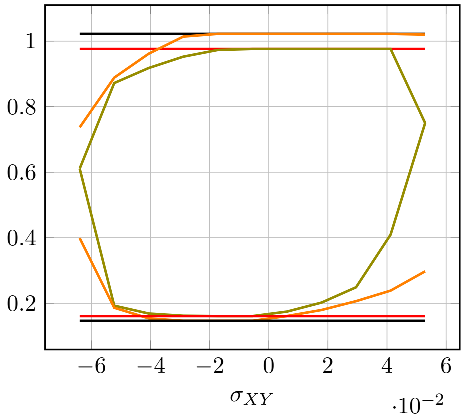
We stress that here the bounds for σ_{XY} result from the restriction by the martingale property only. If we assume that certain return variances are fixed, then the bounds become tighter.

Let us now consider the basket option price function c_B. In Figures 5.1a, 5.1b, 5.1c, 5.1d and 5.1e, we show the upper and lower price bounds for different return variance combinations. We combine high, medium and low variances from the possible intervals.

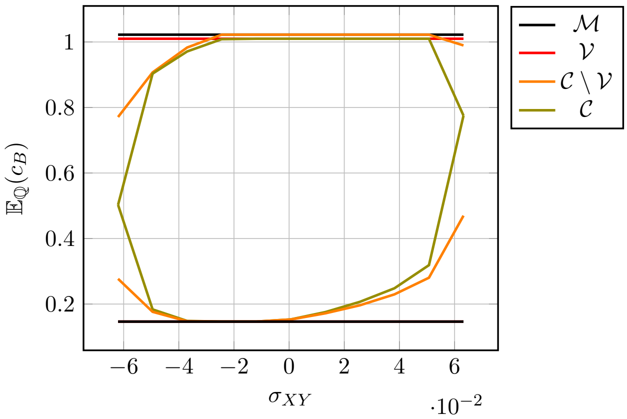
In black lines, we present the martingale optimal transport bounds that we already considered in Example 5.12. In red lines, we present the price bounds that result when incorporating the return variances σ_X^2 and σ_Y^2. In orange lines, we present the price bounds that only rely on the return covariance σ_{XY}. Finally, in olive lines, we present the price bounds that use the full information of σ_X^2, σ_Y^2 and σ_{XY}.



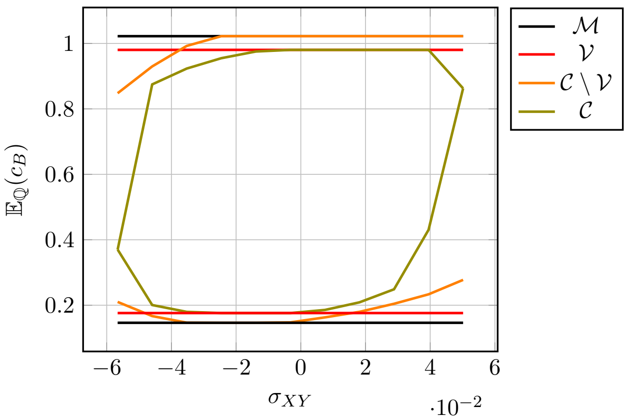
(a) $\sigma_X^2 = 0.022$, $\sigma_Y^2 = 0.1359$



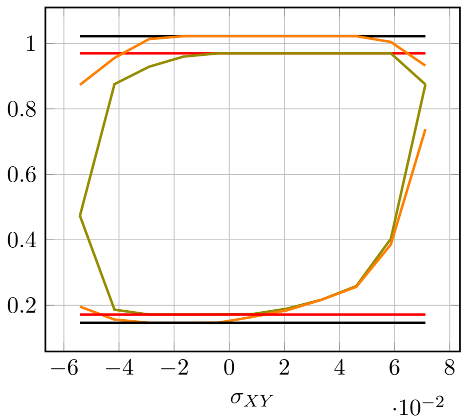
(b) $\sigma_X^2 = 0.022$, $\sigma_Y^2 = 0.1984$



(c) $\sigma_X^2 = 0.0255$, $\sigma_Y^2 = 0.1671$



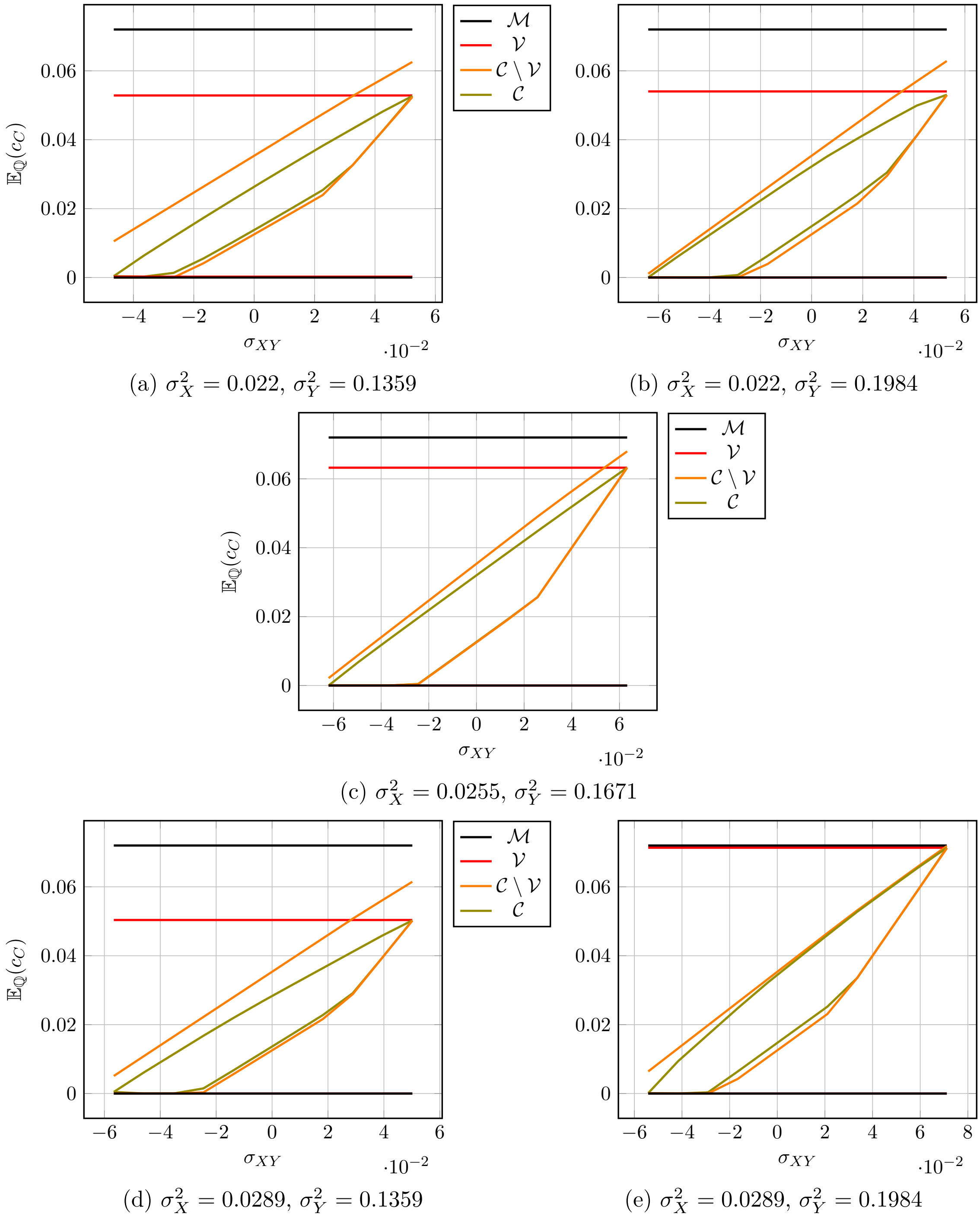
(d) $\sigma_X^2 = 0.0289$, $\sigma_Y^2 = 0.1359$



(e) $\sigma_X^2 = 0.0289$, $\sigma_Y^2 = 0.1984$

Figure 5.1.: Price bounds for basket options.

In Figures 5.2a, 5.2b, 5.2c, 5.2d and 5.2e, we show the same price bounds in the same variance situations for the covariance swap type payoff function c_C and in Figures 5.3a, 5.3b and 5.3c, we show price bounds in the case of medium variances $\sigma_X^2 = 0.0255$ and $\sigma_Y^2 = 0.1671$ for the binary type payoff function c_1, the asian type payoff function c_A and the variance swap type payoff function c_V respectively.



(a) $\sigma_X^2 = 0.022$, $\sigma_Y^2 = 0.1359$

(b) $\sigma_X^2 = 0.022$, $\sigma_Y^2 = 0.1984$

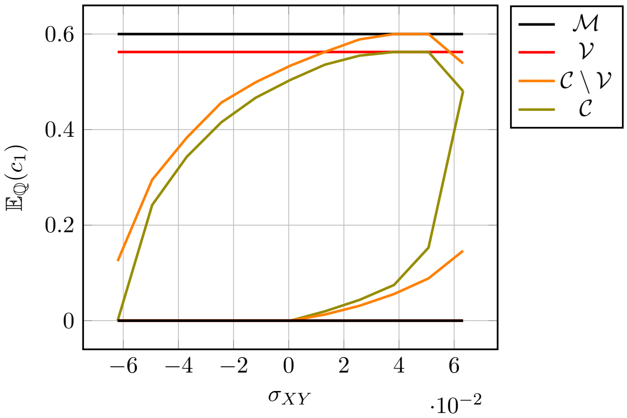
(c) $\sigma_X^2 = 0.0255$, $\sigma_Y^2 = 0.1671$

(d) $\sigma_X^2 = 0.0289$, $\sigma_Y^2 = 0.1359$

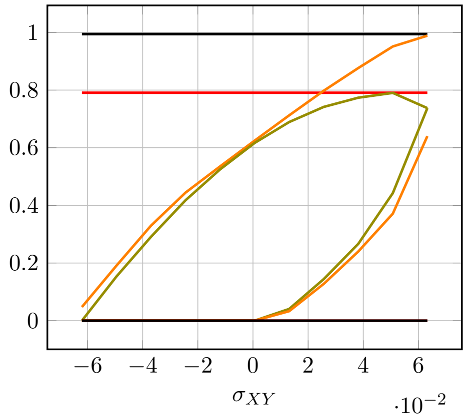
(e) $\sigma_X^2 = 0.0289$, $\sigma_Y^2 = 0.1984$

Figure 5.2.: Price bounds for covariance swaps.

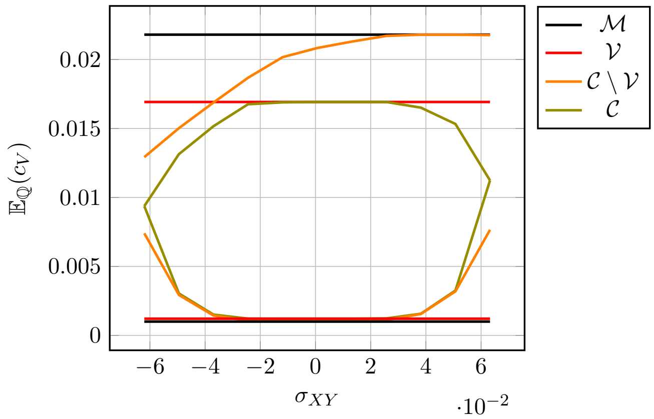
Analyzing the figures closely, several observations come up. First of all, the olive bounds are obviously the tightest and full covariance information yield quite strong improvements of the price bounds for some of the exotic options. This is the case for example for the covariance swap and the asian option. On the contrary, almost no improvement is observable for the basket option.



(a) Binary options



(b) Asian options



(c) Variance swaps

Figure 5.3.: Price bounds for $\sigma_X^2 = 0.0255$, $\sigma_Y^2 = 0.1671$.

The fact that the olive bounds $\underline{\mathcal{C}}$ and $\overline{\mathcal{C}}$ collapse for extreme choices of σ_{XY} for every choice of σ_X^2 and σ_Y^2 suggests that for any payoff function there are parameters $(\sigma_X^2, \sigma_Y^2, \sigma_{XY})$ such that $\underline{\mathcal{C}}(\sigma_X^2, \sigma_Y^2, \sigma_{XY}) = \overline{\mathcal{C}}(\sigma_X^2, \sigma_Y^2, \sigma_{XY})$. This observation is similar to Lütkebohmert & Sester [59, Proposition 3.4]. We also observe that for a certain covariance σ_{XY} these bounds touch $\underline{\mathcal{V}}$ and $\overline{\mathcal{V}}$ respectively.

While we are not able to formally prove the former, as no well-known structural results apply, the latter is virtually trivial, as the measures realizing the bounds in $\underline{\mathcal{V}}$ and $\overline{\mathcal{V}}$ realize a certain covariance.

Finally, we observe the convexity of $\underline{\mathcal{C}}$ and the concavity of $\overline{\mathcal{C}}$ at least with respect to σ_{XY}. However, looking at higher dimensional bound surfaces suggests that these properties hold indeed true for the bounds as functions of I_Σ. Thus, again similarly to Lütkebohmert & Sester [59, Proposition 3.6], we claim that for suitable c, the mapping $\overline{\mathcal{C}}$ is concave and the mapping $\underline{\mathcal{C}}$ is convex.

Structurally, the same observations hold true for the orange bounds $\underline{\mathcal{C} \setminus \mathcal{V}}$ and $\overline{\mathcal{C} \setminus \mathcal{V}}$, except they touch $\underline{\mathcal{M}}$ and $\overline{\mathcal{M}}$ respectively.

Analyzing the red bounds $\underline{\mathcal{V}}$ and $\overline{\mathcal{V}}$, we seem to lose the property that the bounds collapse to a unique price, as the gaps are quite large though we consider rather extreme variances σ_X^2 and σ_Y^2. Heuristically, this should be expected, as ultimately $\mathcal{V}(\Sigma, \mathcal{I})$ is just a set of transport plans of two specific two-dimensional marginals such that for any variance combination there is a great variety of potential pricing measures. △

CHAPTER 6

Monotonicity and optimality

In this chapter, we consider questions concerning explicit solutions to the price bound and hedging problems for certain classes of payoff functions. As this is rather difficult, we restrict ourselves to the standard market case and focus on the upper price bound problem in (4.7) and the associated super hedging problem in (4.10). We begin the chapter by defining the so-called monotone martingale transport plans and the martingale Spence Mirrlees condition. These definitions generalize the order preserving transport plan of Theorem 4.8 and the Spence Mirrlees condition of Definition 4.9 to the martingale case. They are relevant throughout this and the following chapter. In Section 6.1, we present some results of Beiglböck & Juillet [7] about existence, uniqueness and structural properties of monotone martingale transport plans. In Section 6.2, we discuss results of Henry-Labordère & Touzi [38] that build up on the results of the afore mentioned section. In particular, those are constructive procedures to find explicit solutions to both problems whenever the marginals are continuous. We illustrate one of the procedures with an extensive example. Finally, we point out major drawbacks of the illustrated procedure. In Section 6.3, we provide a generalization of some of the results presented in Sections 6.1 and 6.2, namely the optimality of the (left) monotone martingale transport plan for the upper price bound problem in (4.7) for payoff functions that satisfy the martingale Spence Mirrlees condition and arbitrary marginals. In Section 6.4, we further restrict ourselves to the case of discrete marginals. In this case, we discuss the structure of the pricing and hedging problems, improve the main result of Section 6.3 and provide an algorithm to determine the (left) monotone martingale transport plan. Complementing this, we introduce an algorithm to determine a solution to the hedging problem. Finally, we compare our determination techniques to those of other researchers and illustrate the algorithms.

Let us now directly introduce the two crucial definitions. We begin with the structural property of martingale transport plans that guarantees optimality for the upper price bound problem in (4.7) for a certain class of payoff functions.

Definition 6.1 ([7, Definition 1.4])**.** Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$. A martingale transport plan $\mathbb{Q} \in \mathcal{M}_2(\mu, \nu)$ is called *left monotone*, if there is a Borel set $\Gamma \subseteq \mathbb{R}^2$ with $\mathbb{Q}((X, Y) \in \Gamma) = 1$ and such that for $(x, y_1), (x, y_2), (x', y') \in \Gamma$ with $x < x'$, we have $y' \notin (y_1, y_2)$.
Respectively, $\mathbb{Q} \in \mathcal{M}_2(\mu, \nu)$ is called *right monotone*, if there is a Borel set $\Gamma \subseteq \mathbb{R}^2$ with $\mathbb{Q}((X, Y) \in \Gamma) = 1$ and such that for all $(x, y_1), (x, y_2), (x', y') \in \Gamma$ with $x > x'$, we have $y' \notin (y_1, y_2)$.
The set Γ is called *monotonicity set* of $\mathbb{Q}$ and we say it is left or right monotone respectively.

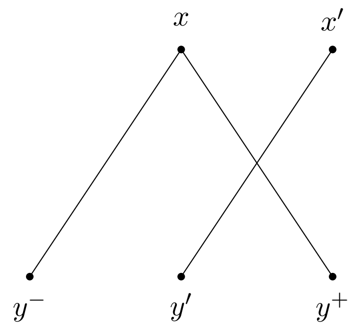


Figure 6.1.: Forbidden configuration for left monotonicity.

This definition was introduced by Beiglböck & Juillet [7] and picked up by Henry-Labordère & Touzi [38]. Both articles achieve impressive results about existence, uniqueness, optimality and the representation of monotone martingale transport plans. We recall those results as far as we generalize or complement them in this thesis. In order to do so, we need the structural property of payoff functions that guarantees that left monotone martingale transport plans are optimal for the upper price bound problem in (4.7).

Definition 6.2. A function $c : \mathbb{R}^2 \to \mathbb{R}$ satisfies the *martingale Spence Mirrlees condition*, if the partial derivative c_{xyy} exists and satisfies $c_{xyy} > 0$.

Remark 6.3. This definition goes back to Henry-Labordère & Touzi [38]. We may rephrase the definition. A function $c : \mathbb{R}^2 \to \mathbb{R}$ satisfies the martingale Spence Mirrlees condition, if one of the following conditions is satisfied.

1. The function c is measurable and the mapping $x \mapsto c(x, y)$ is continuously differentiable for all $y \in \mathbb{R}$ and such that $y \mapsto c_x(x, y)$ is strictly convex for all $x \in \mathbb{R}$.

2. The mapping $y \mapsto c(x', y) - c(x, y)$ is strictly convex for all $x' > x$.

3. For all $x' > x$, $y^+ > y' > y^-$, the function c satisfies

$$\frac{\frac{c(x',y^+)-c(x,y^+)}{x'-x} - \frac{c(x',y')-c(x,y')}{x'-x}}{y^+ - y'} - \frac{\frac{c(x',y')-c(x,y')}{x'-x} - \frac{c(x',y^-)-c(x,y^-)}{x'-x}}{y' - y^-} > 0.$$

The first alternative definition goes back to Beiglböck, Henry-Labordère & Touzi [6] and the second definition comes from Nutz, Stebegg & Tan [65]. The third alternative is based on our proof of Theorem 6.23. It is in particular useful in the case of discrete marginals. $\Diamond$

6.1. The first optimality results

In this section, we state the central results of Beiglböck & Juillet [7] concerning left monotone martingale transport plans in adapted versions such that the results apply to the upper price bound problem in (4.7).

For this purpose, let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ and $c : \mathbb{R}^2 \to \mathbb{R}$ be measurable and such that there exist $a \in \mathbb{L}^1(\mathbb{R}, \mu)$ and $b \in \mathbb{L}^1(\mathbb{R}, \nu)$ with

$$c(x, y) \leq a(x) + b(y), \quad x, y \in \mathbb{R}.$$

Then, for all $\pi \in \Pi_2(\mu, \nu)$, we have

$$\int_{\mathbb{R}^2} c(x, y)\pi(\mathrm{d}(x, y)) \in \left[-\infty, \int_{\mathbb{R}} a(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}} b(y)\nu(\mathrm{d}y)\right].$$

This property is called the *sufficient integrability condition.* Under these conditions, Beiglböck & Juillet [7] show the following results.

Theorem 6.4 ([7, Theorem 1.5])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$. Then there exists a unique left monotone martingale transport plan in $\mathcal{M}_2(\mu, \nu)$. We denote it by $\mathbb{Q}_{lc}(\mu, \nu)$.*

The next corollary is an analogue to the third part of Theorem 4.10, except we may not expect the martingale coupling $\mathbb{Q}_{lc}(\mu, \nu)$ to be deterministic, as it has to satisfy the martingale condition. This can only be true in the trivial case $\mu = \nu$.

Corollary 6.5 ([7, Corollary 1.6])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$ and suppose that μ is continuous. Then there exists a Borel set $S \subseteq \mathbb{R}$ and two measurable functions $T_d, T_u : S \to \mathbb{R}$ such that*

1. *$\mathbb{Q}_{lc}(\mu, \nu)$ is concentrated on the graphs of T_d and T_u.*
2. *For all $x \in S$, $T_d(x) \leq x \leq T_u(x)$.*
3. *For all $x < x' \in S$, $T_u(x) < T_u(x')$ and $T_d(x') \notin (T_d(x), T_u(x))$.*

This result is exploited by Henry-Labordère & Touzi [38], who determine the mappings T_d and T_u explicitly. We present their results in Section 6.2.

Beiglböck & Juillet [7] also show the optimality of $\mathbb{Q}_{lc}(\mu, \nu)$ for the upper price bound problem in (4.7) for certain types of payoff functions.

Theorem 6.6 ([7, Theorem 1.7])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$. Assume that for some differentiable function $h : \mathbb{R} \to \mathbb{R}$, the derivative of which is strictly concave, the function $c(x, y) = h(y - x)$ satisfies the sufficient integrability condition. If $P_2^c(\mu, \nu) > -\infty$, then $\mathbb{Q}_{lc}(\mu, \nu)$ is the unique maximizer of the upper price bound problem in (4.7).*

In the article, another type of payoff functions is proven to have $\mathbb{Q}_{lc}(\mu,\nu)$ as maximizer, namely functions of the form $c(x,y) = \varphi(x)\psi(y) \geq 0$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is non-negative and decreasing and $\psi : \mathbb{R} \to \mathbb{R}$ is non-negative and strictly concave.

We prove in Section 6.3 that there is a general class of payoff functions such that $\mathbb{Q}_{lc}(\mu,\nu)$ is optimal for the upper price bound problem in (4.7), which includes both types of payoff functions introduced by Beiglböck & Juillet [7].

Let now $t \in \mathbb{R}$ and $\pi \in \Pi_2(\mu,\nu)$. Then we define the measure ν_t^π by

$$\nu_t^\pi(B) := \mathrm{proj}_\#^y \left(\pi|_{(-\infty,t]\times\mathbb{R}}\right)(B), \quad B \in \mathcal{B}(\mathbb{R}).$$

Then π moves the mass of $\mu|_{(-\infty,t)}$ to ν_t^π. A transport plan $\pi \in \Pi_2(\mu,\nu)$ is uniquely defined by the family $(\nu_t^\pi)_{t\in\mathbb{R}}$. This yields an equivalent characterization of $\mathbb{Q}_{lc}(\mu,\nu) \in \mathcal{M}_2(\mu,\nu)$.

Theorem 6.7 ([7, Theorem 1.8])**.** *For every $t \in \mathbb{R}$, the measure $\nu_t^{\mathbb{Q}_{lc}(\mu,\nu)}$ is minimal with respect to the convex order in the family*

$$\left\{\nu_t^{\mathbb{Q}} \,\middle|\, \mathbb{Q} \in \mathcal{M}_2(\mu,\nu)\right\}.$$

The above results can be summarized as follows.

Theorem 6.8 ([7, Theorem 1.9])**.** *Let $\mu,\nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$. Let $h : \mathbb{R} \to \mathbb{R}$ be differentiable and such that h' is strictly concave. Assume that the payoff function $c(x,y) = h(y-x)$ satisfies the sufficient integrability condition. Moreover, assume that $P_2^c(\mu,\nu) > -\infty$ and let $\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)$. Then the following are equivalent.*

1. *$\mathbb{Q}$ is left monotone.*
2. *$\mathbb{Q}$ is optimal for the upper price bound problem in* (4.7)*.*
3. *$\mathbb{Q} = \mathbb{Q}_{lc}(\mu,\nu)$, i.e. for all $(\mathbb{Q}',t) \in \mathcal{M}_2(\mu,\nu)\times\mathbb{R}$ we have $\nu_t^{\mathbb{Q}} \leq_c \nu_t^{\mathbb{Q}'}$.*

Analogous results hold for the right monotone martingale transport plan and the lower price bound problem in (4.8). We detail the connection in Section 6.2.4.

6.2. The continuous optimality result

In this section, we state and discuss the already mentioned results of Henry-Labordère & Touzi [38] concerning the construction of the left monotone martingale transport plan and an associated super hedging strategy.

For this purpose, let $\mu,\nu \in \mathcal{P}(\mathbb{R})$ and define the *difference function* $\delta F : \mathbb{R} \to [-1,1]$, $x \mapsto F_\nu(x) - F_\mu(x)$. Let $c : \mathbb{R}^2 \to \mathbb{R}$ be upper semi-continuous and satisfying the sufficient integrability condition. Under the conditions of Corollary 6.5, Henry-Labordère & Touzi [38] construct the left monotone martingale transport plan explicitly. They also show that it is optimal for the upper price bound problem in (4.7) for a general class of payoff functions. For such payoff functions, optimizers for the super hedging problem in (4.10) are derived as well. The results are achieved under the following assumption.

Assumption 6.9 ([38, Asm. 3.5])**.** Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy $\mu \leq_c \nu$ and let μ be continuous.

Working under this assumption, we know by Theorem 6.4 and Corollary 6.5 that there exists a unique left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu) \in \mathcal{M}_2(\mu, \nu)$ such that $\mathbb{Q}_{lc}(\mu, \nu)$ is concentrated on the graphs of some measurable functions $T_d, T_u : S \to \mathbb{R}$ with $T_d(x) \leq x \leq T_u(x)$, $x \in S$. For the structure of $\mathbb{Q}_{lc}(\mu, \nu)$, we thus have

$$\mathbb{Q}_{lc}(\mu, \nu)(dx, dy) = \mu(dx) \otimes \left[q(x)\delta_{T_u(x)} + (1 - q(x))\delta_{T_d(x)}\right](dy), \tag{6.1}$$

where $q(x) := \frac{x - T_d(x)}{(T_u - T_d)(x)} \mathbb{1}_{\{(T_u - T_d)(x) > 0\}}$.

In order to present the results of Henry-Labordère & Touzi [38], we have to go into more detail than in the previous section, as the results are derived constructively. Thus, in order to understand and correctly state the results, we have to present certain parts of the construction itself.

Remark 6.10. Observe that by $\mu \leq_c \nu$, δF increases from zero at the left boundary and to zero at the right boundary of its support. As F_μ is continuous by Assumption 6.9, δF is upper semi-continuous. Thus, the local suprema of δF are attained in (l_μ, r_μ), where l_μ and r_μ are the left and the right boundary of the support of μ. Indeed, we have $\delta F = F_\nu$ on $(-\infty, l_\mu)$ and $\delta F = F_\nu - 1$ on (r_μ, ∞). Thus, δF is increasing in that area. Hence, there can not be any local suprema. Denoting by l_ν and r_ν the left and right boundary of the support of ν, by the convex order we have $l_\nu \leq l_\mu \leq r_\mu \leq r_\nu$. ♢

Let $M(\delta F)$ denote the set of all maximizers of δF and, for each $m \in M(\delta F)$, write $m_- := \sup\{x < m \mid \delta F(x) < \delta F(m)\}$ and $m_+ := \inf\{x > m \mid \delta F(x) < \delta F(m)\}$. Then the set

$$\mathbf{M}_0(\delta F) := \{m \in \mathbf{M}(\delta F) \mid m = m_+ \text{ and } \delta F \equiv \delta F(m) \text{ on } [m_-, m]\}$$

plays a crucial role in the construction. The construction works under the following additional assumption on μ and ν.

Assumption 6.11 ([38, Asm. 3.7])**.** Let ν be continuous and let $\mathbf{M}_0(\delta F)$ be finite.

It may also assumed that the pair (μ, ν) is irreducible, as the construction is done on separate irreducible components, which may be considered by Proposition 5.14. Then, by the continuity of μ and ν, we have $I = J$ for the domain of the irreducible pair (μ, ν).

6.2.1. Construction of the left monotone martingale transport plan

The construction relies on the representation of $\mathbb{Q}_{lc}(\mu, \nu)$ in (6.1). As $\mathbb{Q}_{lc}(\mu, \nu) \in \mathcal{M}_2(\mu, \nu)$, T_d and T_u have to be such that $\mathbb{Q}_{lc}(\mu, \nu)(X \in \mathrm{d}x) = \mu(\mathrm{d}x)$, $\mathbb{Q}_{lc}(\mu, \nu)(Y \in \mathrm{d}y) = \nu(\mathrm{d}y)$ and $\mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}[Y \mid X] = X$ hold. The first and the third condition are trivially satisfied. Hence, the construction is done in a fashion that guarantees $\mathbb{Q}_{lc}(\mu, \nu)(Y \in \mathrm{d}y) = \nu(\mathrm{d}y)$.

In the construction we need some more notations. We denote

$$g(x, y) := F_\nu^{-1}(F_\mu(x) + \delta F(y))$$

for all $x, y \in \mathbb{R}$, and for all $x \in \mathbb{R}$, we write

$$\gamma(x) \coloneqq \int_{-\infty}^{F_\nu^{-1}\circ F_\mu(x)} \xi F_\nu(\mathrm{d}\xi) - \int_{-\infty}^{x} \xi F_\mu(\mathrm{d}\xi).$$

Now let $A \in \mathcal{B}(\mathbb{R})$ be such that δF is increasing on A. Then, for $t \leq m \leq x$, we define

$$G_A^m(t,x) \coloneqq -\int_t^m [g(x,\zeta) - \zeta]\mathbb{1}_A(\zeta)\delta F(\mathrm{d}\zeta) + \int_m^x [g(\xi,m) - \xi]F_\mu(\mathrm{d}\xi).$$

As $\mathbf{M}_0(\delta F)$ is a finite set by assumption, let us write $\mathbf{M}_0(\delta F) = \{m_1^0, \ldots, m_n^0\}$, where $-\infty < m_1^0 < \ldots < m_n^0 < \infty$. Furthermore, we define

$$B_0 \coloneqq \{x \in \mathbb{R} \mid \delta F \text{ is increasing in a right neighborhood of } x.\}$$

and $x_0 \coloneqq \inf B_0$. We have $x \in B_0$, if for all $\varepsilon > 0$ there is an $x_\varepsilon \in (x, x+\varepsilon)$ such that $\delta F(x_\varepsilon) > \delta F(x)$. Observe that $x_0 < m_1^0$ and $\delta F \equiv 0$ on $(-\infty, x_0]$. The construction recursively relies on the following ingredients.

(I_1) $m_0 \in \{-\infty\} \cup \mathbf{M}_0(\delta F)$ and $A_0 \subset B_0 \cap (-\infty, m_0)$ with $\delta F > 0$ on A_0, satisfying $G_{A_0}^{m_0}(-\infty, \cdot) = \gamma(\cdot)$ and $\int_{-\infty}^{m_0} \mathbb{1}_{A_0}\phi(\mathrm{d}\delta F) = \int_{-\infty}^{m_0} \phi(\mathrm{d}\delta F)$ for all non-decreasing maps ϕ.

(I_2) $\bar{x}_0 \in B_0 \cap [m_0, m_n^0)$ and $t_0 \in A_0 \cup \{\infty\}$, satisfying $\delta F(t_0) = \delta F(\bar{x}_0) \geq 0$ and $G_{A_0}^{m_0}(t_0, \bar{x}_0) = 0$.

Lemma 6.12 ([38, Lemma 4.1])**.** *We define $m_1 \coloneqq \min\{\boldsymbol{M}_0(\delta F) \cap (\bar{x}_0, \infty)\}$ as well as $A_1 \coloneqq (A_0 \setminus [t_0, m_0]) \cup (\bar{x}_0, m_1)$. Then we have the following.*

1. *$\delta F > 0$ on A_1, $G_{A_1}^{m_1}(-\infty, \cdot) = \gamma(\cdot)$ and $\int_{-\infty}^{m_1} \mathbb{1}_{A_1}\phi(\mathrm{d}\delta F) = \int_{-\infty}^{m_1} \phi(\mathrm{d}\delta F)$ for all non-decreasing maps ϕ.*
2. *For all $x \geq m_1$ with $\delta F(x) \leq \delta F(m_1)$ there exists a unique scalar $t_{A_1}^{m_1}(x) \in A_1$ such that $G_{A_1}^{m_1}(t_{A_1}^{m_1}(x), x) = 0$.*
3. *The function $x \mapsto t_{A_1}^{m_1}(x)$ is decreasing μ-almost everywhere on $[m_1, x_1]$, where we define $x_1 \coloneqq \inf\{x > m_1 \mid g(x, t_{A_1}^{m_1}(x)) \leq x\}$.*
4. *If $x_1 < \infty$, then $x_1 \in B_0 \cap [m_1, m_n^0) \setminus \boldsymbol{M}_0(\delta F)$, and $\delta F(t_{A_1}^{m_1}(x_1)) = \delta F(x_1) \geq 0$.*

Using Lemma 6.12, the functions T_d and T_u may be constructed explicitly.

Algorithm 6.13 ([38, Sec. 4.2])**.** We start by defining $T_d(x) = T_u(x) = x$ for all $x \leq x_0$ and we continue the construction of T_d and T_u along the following steps.

1. Set $m_0 \coloneqq -\infty$, $A_0 \coloneqq \emptyset$, $\bar{x}_0 \coloneqq x_0$, $t_0 = -\infty$ and notice that (I_1) and (I_2) are obviously satisfied by these ingredients. We may then apply Lemma 6.12 and obtain

$$m_1 \coloneqq \min[\mathbf{M}_0(\delta F) \cap (x_0, \infty)] = \min \mathbf{M}_0(\delta F) = m_1^0,$$
$$A_1 \coloneqq (A_0 \setminus [t_0, m_0]) \cup (x_0, m_1) = (-\infty, m_1)\,.$$

We know that for all $x \geq m_1$ with $\delta F(x) \leq \delta F(m_1)$, there is a scalar $t^{m_1}_{A_1}(x) \in A_1$ and we choose $x_1 := \inf\left\{x > m_1 \;\middle|\; g\left(x, t^{m_1}_{A_1}(x)\right) \leq x\right\}$. Further define $t_1 := t^{m_1}_{A_1}(x_1)$. Then define the maps T_u and T_d on (x_0, x_1) by

$$T_d(x) = T_u(x) = x, \quad x_0 < x \leq m_1$$
$$T_d(x) = t^{m_1}_{A_1}(x), T_u(x) = g(x, T_d(x)), \quad m_1 \leq x < x_1.$$

If $x_1 = \infty$, this completes the construction and we set $m_j = x_j = \infty$ for all $j > 1$. Otherwise, Lemma 6.12 guarantees that the new ingredients (m_1, A_1, x_1, t_1) satisfy conditions (I_1) and (I_2) and we may continue with the next step.

i. Suppose that the maps T_d and T_u are defined on $(-\infty, x_{i-1})$ for some suitable quadruple $(m_{i-1}, A_{i-1}, x_{i-1}, t_{i-1})$ satisfying conditions (I_1) and (I_2). Using Lemma 6.12, we obtain $m_i := \min[M_0(\delta F) \cap (x_{i-1}, \infty)]$ and $A_i := (A_{i-1} \setminus [t_{i-1}, m_{i-1}]) \cup (x_{i-1}, m_i)$. By the existence of a scalar $t^{m_i}_{A_i}(x) \in A_i$, we find $x_i := \inf\left\{x > m_i \;\middle|\; g\left(x, t^{m_i}_{A_i}(x)\right) \leq x\right\}$ and $t_i := t^{m_i}_{A_i}(x_i)$. Then define the maps T_d and T_u on (x_{i-1}, x_i) by

$$T_d(x) = T_u(x) = x, \quad x_{i-1} < x \leq m_i$$
$$T_d(x) = t^{m_i}_{A_i}(x), T_u(x) = g(x, T_d(x)), \quad m_i \leq x < x_i.$$

If $x_i = \infty$, this completes the construction and we set $m_j = x_j = \infty$ for all $j > i$. Otherwise, Lemma 6.12 guarantees that the new ingredients (m_i, A_i, x_i, t_i) satisfy conditions (I_1) and (I_2) and we may continue with the next step.

As $\mathbf{M}_0(\delta F)$ is finite, the algorithm terminates after finitely many steps. The next theorem states that this yields the desired left monotone martingale transport plan using the probability kernel

$$\kappa_*(x, dy) := \mathbb{1}_D(x)\delta_{\{x\}}(dy) + \mathbb{1}_{D^c}(x)\left[q(x)\delta_{\{T_u(x)\}}(dy) + (1 - q(x))\delta_{\{T_d(x)\}(dy)}\right],$$

where $D := \bigcup_{i \geq 1}(x_{i-1}, m_i]$ and $q(x) = \frac{x - T_d(x)}{(T_u - T_d)(x)}\mathbb{1}_{\{(T_u - T_d)(x) > 0\}}$.

Theorem 6.14 ([38, Theorem 4.5]). *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$.*

1. *Assume the pair (μ, ν) is irreducible with domain (I, I) and satisfies Assumptions 6.9 and 6.11. Then we have*

$$\mathbb{Q}_{lc}(\mu, \nu)(\mathrm{d}x, \mathrm{d}y) = \mu(dx) \otimes \kappa_*(x, dy).$$

2. *Let $(\mu_k, \nu_k)_{k \geq 0}$ be the decomposition of (μ, ν) into irreducible components. Consider the decomposition $\mathbb{Q} = \sum_{k \geq 0} \mathbb{Q}_k \in \mathcal{M}_2(\mu, \nu)$ with $\mathbb{Q}_k \in \mathcal{M}_2(\mu_k, \nu_k)$, $k \geq 0$. Then $\mathbb{Q}$ is left monotone if and only if $\mathbb{Q}_k$ is left monotone for all $k \geq 1$.*

We discussed the construction of the left monotone martingale transport plan in the case of continuous marginals, which is the solution for the upper price bound problem in

(4.7) in the case that the payoff function satisfies the conditions of Theorem 6.6. We now present the construction of a solution to the super hedging problem in (4.10). Thereby, the duality of Theorem 5.1 in the standard market case is recovered. The irreducibility of (μ, ν) is important, as this implies the existence of a solution.

6.2.2. AN ASSOCIATED SUPER HEDGING STRATEGY

In this section, we present the construction of a triple $(\varphi_*, \psi_*, h_*) \in \mathcal{D}_2^{\geq c}$ such that

$$\int_{\mathbb{R}} \varphi_*(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}} \psi_*(y)\nu(\mathrm{d}y) = \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right]$$

as provided by Henry-Labordère & Touzi [38]. Such a triple is immediately optimal for the super hedging problem in (4.10), whenever $\mathbb{Q}_{lc}(\mu, \nu)$ is optimal for the upper price bound problem in (4.7), compare Definition 5.18. We only present the results and omit the construction process. The functions h_* and ψ_* are given by

$$\begin{aligned}
h_*'(x) &:= \frac{c_x\left(x, T_u(x)\right) - c_x\left(x, T_d(x)\right)}{T_u(x) - T_d(x)}, && x \in D^c,\\
h_*(x) &:= h_*\left(T_d^{-1}(x)\right) + c_y(x,x) - c_y\left(T_d^{-1}(x), x\right), && x \in D,\\
\psi_*'(x) &:= c_y\left(T_u^{-1}(x), x\right) - h_*\left(T_u^{-1}(x)\right), && x \in D^c,\\
\psi_*'(x) &:= c_y\left(T_d^{-1}(x), x\right) - h_*\left(T_d^{-1}(x)\right), && x \in D.
\end{aligned}$$

The corresponding function φ_* is given by

$$\begin{aligned}
\varphi_*(x) &:= \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}[c(X,Y) - \psi_*(Y) \mid X = x]\\
&= q(x)(c(x,T_u(x)) - \psi_*(T_u(x))) + (1-q(x))(c(x,T_d(x)) - \psi_*(T_d(x))), \quad x \in \mathbb{R}.
\end{aligned}$$

Finally, the exact values are chosen such that the function

$$x \mapsto c(x,T_u(x)) - \psi_*(T_u(x)) - [c(x,T_d(x)) - \psi_*(x)] - (T_u - T_d)(x)h(x)$$

is continuous.

Theorem 6.15 ([38, Theorem 5.1])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ with $\mu \leq_c \nu$ be such that (μ, ν) is irreducible and Assumptions 6.9 and 6.11 are satisfied. Assume further that $\varphi_*^+ \in \mathbb{L}^1(\mathbb{R}, \mu)$ and $\psi_*^+ \in \mathbb{L}^1(\mathbb{R}, \nu)$. Suppose the payoff function $c : \mathbb{R}^2 \to \mathbb{R}$ satisfies the martingale Spence Mirrlees condition. Then:*

1. *We have $(\varphi_*, \psi_*, h_*) \in \mathcal{D}_2^{\geq c}$.*

2. *We have*

$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)] = \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\int_{\mathbb{R}} \varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}} \psi(y)\nu(\mathrm{d}y)\right\}.$$

Also, $\mathbb{Q}_{lc}(\mu,\nu)$ is a maximizer for the left hand side and (φ_*,ψ_*,h_*) is a minimizer for the right hand side. That is,

$$\int_{\mathbb{R}}\int_{\mathbb{R}} c(x,y)\kappa_*(x,dy)\mu(dx) = \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}[c(X,Y)] = \int_{\mathbb{R}} \varphi_*(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}} \psi_*(y)\nu(\mathrm{d}y).$$

6.2.3. An exemplary construction

We now execute Algorithm 6.13 in the simplest possible situation, namely in the case of uniform marginals. We remark that this is one of the few situations, in which the algorithm yields analytical results for $\mathbb{Q}_{lc}(\mu,\nu)$. In almost all other cases, numerical procedures have to be used.

Example 6.16. We study the construction of the left monotone martingale transport plan in the case of continuous uniform marginal distributions. In this case, we achieve explicit results for the relevant expressions.

Therefore, let $a<c<d<b$ be such that $m := \frac{a+b}{2} = \frac{c+d}{2}$, and let $\mu \sim \mathcal{U}(c,d)$ and $\nu \sim \mathcal{U}(a,b)$. Then, for $x \in \mathbb{R}$, we have

$$\begin{aligned}
F_\mu(x) &= \frac{x-c}{d-c}\mathbb{1}_{(c,d)}(x) + \mathbb{1}_{[d,\infty)}(x),\\
f_\mu(x) &= \frac{1}{d-c}\mathbb{1}_{(c,d)}(x),\\
F_\nu(x) &= \frac{x-a}{b-a}\mathbb{1}_{(a,b)}(x) + \mathbb{1}_{[b,\infty)}(x),\\
f_\nu(x) &= \frac{1}{b-a}\mathbb{1}_{(a,b)}(x),\\
\delta F(x) &= \frac{x-a}{b-a}\mathbb{1}_{(a,c]}(x) + \left(\frac{x-a}{b-a} - \frac{x-c}{d-c}\right)\mathbb{1}_{(c,d)}(x) + \left(\frac{x-a}{b-a} - 1\right)\mathbb{1}_{[d,b)},\\
F_\nu^{-1}(x) &= \begin{cases} -\infty, & x \le 0\\ (b-a)x + a, & 0 < x < 1\\ \infty, & x \ge 1 \end{cases}
\end{aligned}$$

for the distribution, density and quantile functions that are relevant for the construction. The distributions satisfy the necessary conditions. Indeed, μ and ν are continuous with finite first moments and the unique maximum of δF is attained at c. In particular, there are only finitely many maxima.

Now let us calculate T_d and T_u and thereby observe that the second marginal of the resulting martingale transport plan is indeed ν. Afterwards, using certain choices of a,b,c,d, we deepen the understanding of left monotonicity providing an illustrative figure. Clearly, before we may determine the left monotone martingale transport plan associated to μ and ν, we have to show that $\mu \le_c \nu$. However, this is an easy consequence of the structure of the measures and we do not prove it in detail.

Thus, let us now determine T_d and T_u using Algorithm 6.13. For this purpose, observe that μ is concentrated right of the maximum of δF, which is attained in c. Hence, all of the mass of μ is split up by T_d and T_u. As the maximum is unique, we know that the algorithm terminates after one step. We start by calculating T_d. We choose

$$m_1 = \min M_0(\delta F) = c \quad \text{and} \quad A_1 = (x_0, m_1) = (-\infty, c)$$

such that we can disregard $\mathbb{1}_{A_1}$ in the defining equality $G^{m_1}_{A_1}(T_d(x), x) = 0$, which is thus equivalent to

$$\int_{T_d(x)}^{c} [g(x,\zeta) - \zeta]\delta F(\mathrm{d}\zeta) = \int_c^x [g(\xi, c) - \xi] F_\mu(\mathrm{d}\xi) \tag{6.2}$$

for $a \le T_d(x) \le c \le x \le d$. Recall $g(x,y) = F_\nu^{-1}(F_\mu(x) + \delta F(y))$.

We start by calculating the left hand side of the above equality. This yields

$$\begin{aligned}
\int_{T_d(x)}^{c} &\left(F_\nu^{-1}\left(F_\mu(x) + \delta F(\zeta)\right) - \zeta\right)\delta F(\mathrm{d}\zeta) = \int_{T_d(x)}^{c} \left(F_\nu^{-1}\left(\frac{x-c}{d-c} + \frac{\zeta - a}{b-a}\right) - \zeta\right)\delta F(\mathrm{d}\zeta) \\
&= \frac{1}{b-a}\int_{T_d(x)}^{c} \left(F_\nu^{-1}\left(\frac{x-c}{d-c} + \frac{\zeta - a}{b-a}\right) - \zeta\right)\mathrm{d}\zeta \\
&\overset{(\diamond)}{=} \frac{1}{b-a}\int_{T_d(x)}^{c} \left((b-a)\left(\frac{x-c}{d-c} + \frac{\zeta - a}{b-a}\right) + a - \zeta\right)\mathrm{d}\zeta \\
&= \int_{T_d(x)}^{c} \frac{x-c}{d-c}\mathrm{d}\zeta = (c - T_d(x))\frac{x-c}{d-c},
\end{aligned}$$

where in the first step we use $x \in (c,d)$ and in the second step we use $T_d(x) \ge a$. The equality under $(\diamond)$ is under closer review in Remark 6.17.

The right hand side of the equality in (6.2) is rearranged similarly. Skipping some intermediate steps, we get

$$\begin{aligned}
\int_c^x &\left(F_\nu^{-1}(F_\mu(\xi) + \delta F(c)) - \xi\right) F_\mu(\mathrm{d}\xi) = \frac{1}{d-c}\int_c^x \left(F_\nu^{-1}\left(\frac{\xi - c}{d-c} + \frac{c-a}{b-a}\right) - \xi\right)\mathrm{d}\xi \\
&\overset{(\circ)}{=} \frac{1}{d-c}\int_c^x \left(\frac{(b-a)(\xi - c)}{d-c} + c - \xi\right)\mathrm{d}\xi \\
&= \frac{1}{d-c}\left[\frac{b-a}{d-c}\int_c^x (\xi - c)\mathrm{d}\xi + \int_c^x (c-\xi)\mathrm{d}\xi\right] \\
&= \frac{1}{2}\frac{1}{d-c}\left[\frac{b-a}{d-c}(x-c)^2 - (x-c)^2\right],
\end{aligned}$$

where the equality under $(\circ)$ is as well discussed in Remark 6.17.

Plugging those results into the equation in (6.2), for all $x \in (c,d)$, we receive

$$T_d(x) = \frac{(d-c) - (b-a)}{2} F_\mu(x) + c.$$

Remark 6.17. Observe that solving equation (6.2) as above, we formally assume that the arguments, that we evaluate F_ν^{-1} in, are in $(0,1)$, as then the integrals are finite.

Clearly, these arguments are positive on both sides of the equation. Indeed, for the left hand side we have

$$F_\mu(x) + \delta F(\zeta) = F_\mu(x) - F_\mu(\zeta) + F_\nu(\zeta) \ge F_\nu(\zeta) > 0,$$

as $x > c > \zeta > a$. Similarly, for the right hand side we have

$$F_\mu(\xi) + \delta F(c) = F_\mu(\xi) - F_\mu(c) + F_\nu(c) \geq F_\nu(c) > 0,$$

as $\xi > c$. Thus, only the condition of being less than 1 may restrict the arguments.

Let $x \in \operatorname{supp}(\mu)$ and $c < \xi < x$. Then, for the right hand side, we have

$$F_\mu(\xi) + \delta F(c) = \frac{\xi - c}{d - c} + \frac{c - a}{b - a} \overset{!}{<} 1 \iff \xi < (d - c)\left(\frac{b - c}{b - a}\right) + c.$$

With $\xi \to x$ we have the condition $x \leq (d-c)\left(\frac{b-c}{b-a}\right) + c$. As we have $\frac{b-c}{b-a} < 1$, this upper bound is strictly less than d. Hence, the integral on the right hand side of (6.2) is infinite for all $x \in \left((d-c)\frac{b-c}{b-a} + c, d\right)$.

Now let $a < \zeta < c < x$. Then, for the left hand side, we analogously have

$$F_\mu(x) + \delta F(\zeta) = \frac{x - c}{d - c} + \frac{\zeta - a}{b - a} \overset{!}{<} 1 \iff \zeta < (b - a)\,\frac{d - x}{d - c} + a.$$

As long as the right hand side is greater than c, no problems occur. However, if

$$c > (b - a)\,\frac{d - x}{d - c} + a \iff x > (d - c)\frac{b - c}{b - a} + c,$$

then the integral on the left hand side of (6.2) is infinite. Thus, for $x \in \left((d-c)\frac{b-c}{b-a} + c, d\right)$, no uniquely defined scalar $t^{m_1}_{A_1}(x)$ exists.

Still, the analytic results that we derived for $T_d(x)$, $x < (d-c)\frac{b-c}{b-a} + c$, determine the function correctly for the complete support of μ by simply continuing T_d onto the area where it is not properly defined by the integral equation in (6.2). Thus, this procedure ultimately yields the left monotone martingale transport plan as desired.

However, if no closed form results may be derived, as it is the case for example for normal or log-normal marginals, then this gives rise to severe problems. Indeed, the necessary numerical considerations will fail in such cases. ◊

Now let us calculate the second characterizing mapping T_u. For $x \in (c, d)$, we have

$$\begin{aligned} T_u(x) = g(x, T_d(x)) &= F_\nu^{-1}\left(\frac{x - c}{d - c} + \frac{\frac{(d-c)-(b-a)}{2}F_\mu(x) + c - a}{b - a}\right) \\ &= (b - a)\left(\frac{x - c}{d - c} + \frac{\frac{(d-c)-(b-a)}{2}F_\mu(x) + c - a}{b - a}\right) + a = \frac{(d-c)+(b-a)}{2}F_\mu(x) + c. \end{aligned}$$

Remark 6.18. Considering the argument of F_ν^{-1} in the calculation of T_u and using the condition that the argument should be less than 1, similar to Remark 6.17 we obtain the additional restriction

$$x < 2(d - c)\frac{b - c}{(d - c) + (b - a)} + c,$$

where the right hand side is strictly less than d. As for T_d, continuation of the calculated function for such $x \in \operatorname{supp}(\mu)$ that do not satisfy the above condition is meaningful. ◊

Finally, let us calculate $q(x)$ in order to write down the left monotone martingale transport plan in closed form. For $x \in (c,d)$, we have $T_u(x) > T_d(x)$ and thus

$$\begin{aligned} q(x) &= \frac{x - \left(\frac{(d-c)-(b-a)}{2} F_\mu(x) + c\right)}{\frac{(d-c)+(b-a)}{2} F_\mu(x) + c - \left(\frac{(d-c)-(b-a)}{2} F_\mu(x) + c\right)} \\ &= \frac{2(d-c)x - ((d-c)-(b-a))(x-c) - 2(d-c)c}{2(d-c)(b-a)F_\mu(x)} \\ &= \frac{((d-c)+(b-a))F_\mu(x)}{2(b-a)F_\mu(x)} = \frac{1}{2}\left(1 + \frac{d-c}{b-a}\right) \in \left(\frac{1}{2}, 1\right]. \end{aligned}$$

Thus, we get

$$\mathbb{Q}_{lc}(\mu,\nu)(\mathrm{d}x, \mathrm{d}y) = \mu(\mathrm{d}x) \otimes \left(\frac{1}{2}\left(1 + \frac{d-c}{b-a}\right)\delta_{T_u(x)} + \frac{1}{2}\left(1 - \frac{d-c}{b-a}\right)\delta_{T_d(x)}\right)(\mathrm{d}y).$$

Let us now check that it has indeed the correct marginals. Therefore we show that $F_\nu(y) = \mathbb{Q}_{lc}(\mu,\nu)(Y \leq y)$. We immediately have

$$\begin{aligned} \mathbb{Q}_{lc}(\mu,\nu)(Y \leq y) &= \int_{\mathbb{R}} \int_{-\infty}^{y} \left(q(x)\,\delta_{T_u(x)} + (1-q(x))\,\delta_{T_d(x)}\right)(\mathrm{d}s)\,\mu(\mathrm{d}x) \\ &= \frac{1}{d-c} \int_c^d \left(q(x) \int_{-\infty}^{y} \delta_{T_u(x)}(\mathrm{d}s) + (1-q(x)) \int_{-\infty}^{y} \delta_{T_d(x)}(\mathrm{d}s)\right) \mathrm{d}x \\ &=: \frac{1}{d-c} \int_c^d \left(q(x)\,\mathcal{I}_1(x,y) + (1-q(x))\,\mathcal{I}_2(x,y)\right) \mathrm{d}x =: \mathcal{I}_3(y) \end{aligned}$$

In order to compute $\mathcal{I}_1(x,y)$, $\mathcal{I}_2(x,y)$ and $\mathcal{I}_3(y)$, we have to distinguish two cases.

1. Let $y < c$. Then $\mathcal{I}_1(x,y) \equiv 0$, as $T_u(x) \geq c$, and $\mathcal{I}_2(x,y) = \mathbb{1}_{\{T_d(x) \leq y\}} = \mathbb{1}_{\{x \geq T_d^{-1}(y)\}}$, where $T_d^{-1}(y) = 2\frac{(y-c)(d-c)}{(d-c)-(b-a)} + c$.

   Let now $a < y < c$. Using the above, we receive

$$\begin{aligned} \mathcal{I}_3(y) &= \frac{1}{d-c}\left(\frac{1}{2}\left(1 - \frac{d-c}{b-a}\right)\right) \int_c^d \mathbb{1}_{\{x \geq T_d^{-1}(y)\}} \mathrm{d}x \\ &= \frac{(b-a)-(d-c)}{2(b-a)(d-c)} \left[d - \max\left\{c, T_d^{-1}(y)\right\}\right]. \end{aligned}$$

   Now observe that $\frac{(y-c)(d-c)}{(d-c)-(b-a)} \geq 0$, as $y - c \leq 0$ and $(d-c)-(b-a) \leq 0$, such that $T_d^{-1}(y) \geq c$. Using this and $b + a = d + c$, we have

$$\begin{aligned} \mathcal{I}_3(y) &= \frac{(b-a)-(d-c)}{2(b-a)(d-c)} \left[d - c - 2\frac{(y-c)(d-c)}{(d-c)-(b-a)}\right] \\ &= \frac{(b-a)-(d-c)}{2(b-a)} + \frac{2(y-c)}{2(b-a)} \\ &= \frac{y-a}{b-a} + \frac{(b+a)-(d+c)}{2(b-a)} = \frac{y-a}{b-a} = F_\nu(y). \end{aligned}$$

   If $y < a$, then the marginal condition is trivially satisfied.

2. Let $y > c$. Then $\mathcal{I}_1(x,y) = \mathbb{1}_{\{c \le x \le T_u^{-1}(y)\}}$ and $\mathcal{I}_2(x,y) = \mathbb{1}_{\{T_d(x) \le c\}} = \mathbb{1}_{\{x \ge c\}}$, where $T_u^{-1}(y) = 2\frac{(y-c)(d-c)}{(d-c)+(b-a)} + c$.

   Let now $c < y < b$. Using the above, we receive

$$\begin{aligned}
\mathcal{I}_3(y) &= \frac{1}{d-c}\left(\int_c^d q(x)\mathbb{1}_{\{c \le x \le T_u^{-1}(y)\}}\mathrm{d}x + \int_c^d (1-q(x))\mathrm{d}x\right) \\
&= \frac{(d-c)+(b-a)}{2(d-c)(b-a)}\int_c^d \mathbb{1}_{\{x \le F_u^{-1}(y)\}}\mathrm{d}x + \frac{(b-a)-(d-c)}{2(b-a)(d-c)}\int_c^d 1\mathrm{d}x \\
&= \frac{(d-c)+(b-a)}{2(d-c)(b-a)}\min\left\{d-c, 2\frac{(y-c)(d-c)}{(d-c)+(b-a)}\right\} + \frac{(b-a)-(d-c)}{2(b-a)(d-c)}(d-c) \\
&= \min\left\{\frac{(d-c)+(b-a)}{2(b-a)}, \frac{y-c}{b-a}\right\} + \frac{1}{2} - \frac{d-c}{2(b-a)}.
\end{aligned}$$

   Here, the minimum is attained by the second term as $y \le b$. Thus, we get

$$\mathcal{I}_3(y) = \frac{1}{2(b-a)}((b-a)-(d-c)+2(y-c)) = \frac{y-a}{b-a} = F_\nu(y),$$

   where in the last step we use again $b + a = d + c$.

   If $b < y$, then the marginal condition is trivially satisfied.

Finally, let us consider an explicit example. For this, let now $-a = b = 2$, $-c = d = 1$. Then we have $m = 0$, the mappings T_d and T_u are defined by

$$\begin{aligned}
T_d &: [-1,1] \to [-2,-1], \quad x \mapsto -\frac{1}{2}(x+3) \\
T_u &: [-1,1] \to [-1,2], \quad x \mapsto \frac{3}{2}x + \frac{1}{2},
\end{aligned}$$

and we have $q(x) \equiv \frac{3}{4}$. In order to illustrate the transport plan, we provide Figure 6.2, where we see that the derivation of T_d and T_u by the presented algorithm works and yields the left monotone martingale transport plan. Formally, the derivation of $T_d(x)$ is correct only for $x \le \frac{1}{2}$. With the help of Figure 6.2 and the previous theoretical investigations, we see that the continuation indeed provides the left monotone martingale transport plan. △

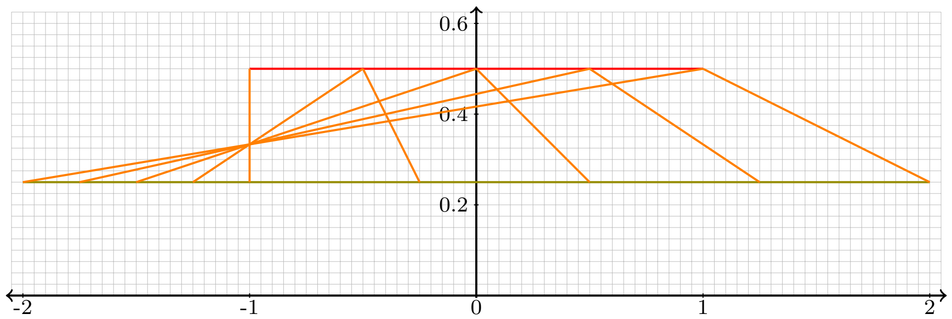


Figure 6.2.: f_μ (red), f_ν (olive) and pointwise mass transport between $\mu \sim \mathcal{U}(-1,1)$ and $\nu \sim \mathcal{U}(-2,2)$ for certain points in the support of μ (orange).

6.2.4. On connections between left and right monotonicity

We presented some properties of the left monotone martingale transport plan, such as uniqueness and optimality for the upper price bound problem in (4.7) with respect to a certain class of payoff functions. In order to find the optimal solution to the lower price bound problem in (4.8) for the same class of payoff functions, we have to find the right monotone martingale transport plan, that we denote by $\mathbb{Q}_{rc}(\mu, \nu)$.

The roles as maximizers and minimizers of those two martingale transport plans in the continuous marginal case exchange when choosing the opposite type of payoff function, i.e. replacing the condition $c_{xyy} > 0$ by $c_{xyy} < 0$. We detail these connections by transforming the problems, a method introduced in [38, Remark 5.2].

1. Let us assume $c_{xyy} < 0$. Then the upper price bound $\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right]$ is attained by the right monotone martingale transport plan $\mathbb{Q}_{rc}(\mu, \nu)$. Indeed, define $\bar{c}(x, y) := c(-x, -y)$. Then $\bar{c}_{xyy} > 0$ and thus, there is a stochastic kernel κ_* such that
$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[\bar{c}(X,Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \bar{c}(x,y)\kappa_*(x,dy)\mu(dx).$$
We define a pair of probability measures $(\bar{\mu}, \bar{\nu})$ by their distribution functions
$$F_{\bar{\mu}}(x) := 1 - F_\mu(-x) \quad \text{and} \quad F_{\bar{\nu}}(x) := 1 - F_\nu(-y).$$
Then we have
$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\bar{\mu},\bar{\nu})} \mathbb{E}_{\mathbb{Q}}[\bar{c}(X,Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \bar{c}(x,y)\bar{\kappa}_*(x,dy)\bar{\mu}(dx). \tag{6.3}$$
for some suitable stochastic kernel $\bar{\kappa}_*$. We now show that also
$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \bar{c}(x,y)\bar{\kappa}_*(x,dy)\bar{\mu}(dx).$$
In order to show this, replace X by $-X$ and Y by $-Y$. Then the martingale condition remains stable and for any $\bar{\mathbb{Q}} \in \mathcal{M}_2(\bar{\mu}, \bar{\nu})$, we have
$$\bar{\mathbb{Q}}(X \le x) = F_{\bar{\mu}}(x) = 1 - F_\mu(-x) \iff \bar{\mathbb{Q}}(-X \le x) = F_\mu(x)$$
and analogously $\bar{\mathbb{Q}}(-Y \le y) = F_\nu(y)$. That is, the random variables $-X$ and $-Y$ have marginals μ and ν under $\bar{\mathbb{Q}}$ respectively. Thus,
$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)] = \sup_{\mathbb{Q}\in\mathcal{M}_2(\bar{\mu},\bar{\nu})} \mathbb{E}_{\mathbb{Q}}[c(-X,-Y)] = \sup_{\mathbb{Q}\in\mathcal{M}_2(\bar{\mu},\bar{\nu})} \mathbb{E}_{\mathbb{Q}}[\bar{c}(X,Y)],$$
from which by (6.3), we deduce the desired equality
$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \bar{c}(x,y)\bar{\kappa}_*(x,dy)\bar{\mu}(dx).$$

We now know that for a payoff function such that $c_{xyy} < 0$, we find the maximizing martingale transport plan in $\mathcal{M}_2(\mu,\nu)$ as the maximizing one from $\mathcal{M}_2(\bar{\mu},\bar{\nu})$, when optimizing with respect to the payoff function $\bar{c}$. By the structure of $\bar{\mu}$ and $\bar{\nu}$, we see that this plan is right monotone with respect to μ and ν.

2. Let us now assume $c_{xyy} > 0$. Then the problem of finding the lower price bound is solved by the right monotone martingale transport plan as well. As $-c(x,y)$ again satisfies the condition $c_{xyy} < 0$, using the previous part, we observe

$$\inf_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[(c(X,Y))] = -\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[-c(X,Y)\right] = \int_{\mathbb{R}}\int_{\mathbb{R}} \bar{c}(x,y)\bar{\kappa}_*(x,dy)\bar{\mu}(dx).$$

The results of this section yield the overview in Table 6.1.

Optimizing martingale transport plan for the problem in	(4.7)	(4.8)
and payoff function s.t. $c_{xyy} > 0$	$\mathbb{Q}_{lc}(\mu,\nu)$	$\mathbb{Q}_{rc}(\mu,\nu)$
and payoff function s.t. $c_{xyy} < 0$	$\mathbb{Q}_{rc}(\mu,\nu)$	$\mathbb{Q}_{lc}(\mu,\nu)$

Table 6.1.: Optimality properties of left and right monotone martingale transport plans.

6.3. The general optimality result

In this section, we generalize Theorem 6.6 regarding the class of payoff functions and Theorem 6.15 regarding the class of marginals. For this purpose, we need some further definitions and results from Beiglböck & Juillet [7]. The idea is the same as for c-cyclical monotonicity in classic optimal transport. That basically is, any optimal transport plan should specify the best possible coupling target in the support of ν for any point in the support of μ. In the martingale case, this is rigourously formulated by the two following definitions.

Definition 6.19 ([7, Definition 1.10])**.** Let α be a measure on $\mathbb{R}^2$ with finite first moment in the second argument. A measure α' on the same space is called *competitor*, if α' has the same marginals as α and it satisfies

$$\int_{\mathbb{R}} y\alpha_x(\mathrm{d}y) = \int_{\mathbb{R}} y\alpha'_x(\mathrm{d}y)$$

for $\mathrm{proj}^x_\#\alpha$-almost every $x \in \mathbb{R}$, where $(\alpha_x)_{x\in\mathbb{R}}$ and $(\alpha'_x)_{x\in\mathbb{R}}$ are disintegrations of the measures α and α' with respect to $\mathrm{proj}^x_\#\alpha$, i.e. $\alpha = \mathrm{proj}^x_\#\alpha \otimes \alpha_x$ and $\alpha' = \mathrm{proj}^x_\#\alpha \otimes \alpha'_x$.

Definition 6.20 ([7, Definition A.1])**.** For a payoff function $c : \mathbb{R}^2 \to \mathbb{R}$, a Borel set $\Gamma \subset \mathbb{R}^2$ is called *finitely optimal* for c, if for every measure α on $\mathbb{R}^2$ such that $\mathrm{supp}(\alpha) \subseteq \Gamma$ and $|\,\mathrm{supp}(\alpha)| < \infty$ and every competitor α' of α, we have

$$\int_\Gamma c(x,y)\alpha(\mathrm{d}(x,y)) \geq \int_\Gamma c(x,y)\alpha'(\mathrm{d}(x,y)).$$

Using these definitions, we may present the so-called variational lemma that basically states an equivalence of the optimality of a transport plan and the finite optimality of the associated support set.

Lemma 6.21 ([7, Lemma 1.11])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$ and let $c \in \mathbb{L}^0(\mathbb{R}^2)$ be a payoff function that satisfies the sufficient integrability condition. Let $\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)$ be an optimal martingale transport plan leading to a finite price $P_2^c(\mu,\nu)$. Then there is a Borel set $\Gamma \subset \mathbb{R}^2$ such that $\mathbb{Q}(\Gamma) = 1$ and Γ is finitely optimal for c.*

Lemma 6.22 ([7, Lemma A.2])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$ and let $c \in \mathcal{C}_b(\mathbb{R}^2)$. Let $\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)$. If there is a finitely optimal set $\Gamma \subset \mathbb{R}^2$ for c with $\mathbb{Q}(\Gamma) = 1$, then $\mathbb{Q}$ is an optimal martingale transport plan for the upper price bound problem in* (4.7)*.*

Recall the notions of left monotonicity and the martingale Spence Mirrlees condition. With these notions and the above definitions and results, we may now prove the connection between the optimality of the left monotone martingale transport plan for the upper price bound problem in (4.7) and the satisfaction of the martingale Spence Mirrlees condition by the payoff function. In the proof, we proceed similar to the proof of [7, Theorem 6.1].

Theorem 6.23. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$. Suppose $c : \mathbb{R}^2 \to \mathbb{R}$ is a payoff function that satisfies the martingale Spence Mirrlees condition and the sufficient integrability assumption. Then the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu,\nu)$ is optimal for the upper price bound problem in* (4.7)*.*

Proof. Let $\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)$ be a finite optimizer for $P_2^c(\mu,\nu)$, which exists by Corollary 5.2. Let further $\Gamma \subset \mathbb{R}^2$ be a finitely optimal set for c with $\mathbb{Q}(\Gamma) = 1$, the existence of which is guaranteed by Lemma 6.21. We prove that Γ is a left monotone set.

In order to get a contradiction, we assume that $(x, y^-), (x, y^+), (x', y') \in \Gamma$ are such that they contradict the left monotonicity of Γ, i.e. $x < x'$ and $y^- < y' < y^+$. If we consider a measure α with finite support contained in Γ, then any competitor α' of α should satisfy

$$\int_\Gamma c(x,y)\alpha(\mathrm{d}(x,y)) \geq \int_\Gamma c(x,y)\alpha'(\mathrm{d}(x,y))$$

by the finite optimality of Γ.

We define such a measure α on Γ by

$$\alpha := \lambda\delta_{(x,y^-)} + (1-\lambda)\,\delta_{(x,y^+)} + \delta_{(x',y')},$$

where $\lambda \in [0,1]$ is such that $\lambda y^+ + (1-\lambda)y^- = y'$, i.e. $\lambda = \frac{y'-y^-}{y^+-y^-}$. A general competitor

$$\alpha' = \omega_1\delta_{(x,y^-)} + \omega_2\delta_{(x,y')} + \omega_3\delta_{(x,y^+)} + \omega_1'\delta_{(x',y^-)} + \omega_2'\delta_{(x',y')} + \omega_3'\delta_{(x',y^+)},$$

which we shortly denote by $\alpha' = \left(\omega_1, \omega_2, \omega_3, \omega_1', \omega_2', \omega_3'\right)^\top$, has to satisfy the conditions of Definition 6.19, which in our situation translate to the following conditions.

1. Marginal distribution $\mathrm{proj}^x_\# \alpha$:
   - Mass in x: $\omega_1 + \omega_2 + \omega_3 \overset{!}{=} 1$,
   - Mass in x': $\omega_1' + \omega_2' + \omega_3' \overset{!}{=} 1$.
2. Marginal distribution $\mathrm{proj}^y_\# \alpha$:
   - Mass in y^-: $\omega_1 + \omega_1' \overset{!}{=} \lambda$,
   - Mass in y': $\omega_2 + \omega_2' \overset{!}{=} 1$,
   - Mass in y^+: $\omega_3 + \omega_3' \overset{!}{=} 1 - \lambda$.
3. Conditional distributions α_x, α_x':
   - Integral equation for $x : \omega_1 y^- + \omega_2 y' + \omega_3 y^+ \overset{!}{=} \lambda y^- + (1-\lambda) y^+ =: \bar{y}$,
   - Integral equation for x': $\omega_1' y^- + \omega_2' y' + \omega_3' y^+ \overset{!}{=} y'$.

These conditions yield a linear equation system with variables $(\omega_1, \omega_2, \omega_3, \omega_1', \omega_2', \omega_3')$, which we may state in matrix form as

$$\left(\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & \lambda \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1-\lambda \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ y^- & y' & y^+ & 0 & 0 & 0 & \bar{y} \\ 0 & 0 & 0 & y^- & y' & y^+ & y' \end{array}\right)$$

We see that the fourth row is redundant by adding the first, second and third row and subtracting the fifth row, and that the sixth row is redundant by adding suitable multiples of the first, second and third row and subtracting the seventh row, where we use that $\bar{y} - \lambda y^- - y' - (1-\lambda) y^+ = -y'$.

We solve the above linear equation system in order to determine the form of a competitor a'. Clearly, this system may be solved using classic methods as follows.

$$\left(\begin{array}{cccccc|c} 1 & 0 & 0 & 1 & 0 & 0 & \lambda \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1-\lambda \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & y^- & y' & y^+ & y' \end{array}\right) \rightsquigarrow \left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & -1 & -1 & \lambda - 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1-\lambda \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & \frac{y^+ - y^-}{y' - y^-} & 1 \end{array}\right)$$

$$\overset{\text{Def.}\lambda}{\rightsquigarrow} \left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & \frac{1}{\lambda} - 1 & \lambda \\ 0 & 1 & 0 & 0 & 0 & -\frac{1}{\lambda} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1-\lambda \\ 0 & 0 & 0 & 1 & 0 & 1 - \frac{1}{\lambda} & 0 \\ 0 & 0 & 0 & 0 & 1 & \frac{1}{\lambda} & 1 \end{array}\right)$$

Thus, a general competitor α' of α has the form

$$\alpha' = \begin{pmatrix} \lambda \\ 0 \\ 1-\lambda \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} \frac{1}{\lambda} - 1 \\ -\frac{1}{\lambda} \\ 1 \\ 1 - \frac{1}{\lambda} \\ \frac{1}{\lambda} \\ -1 \end{pmatrix} = \begin{pmatrix} \lambda \\ 0 \\ 1-\lambda \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 1-\lambda \\ -1 \\ \lambda \\ \lambda - 1 \\ 1 \\ -\lambda \end{pmatrix} = \alpha + s \cdot \begin{pmatrix} 1-\lambda \\ -1 \\ \lambda \\ \lambda - 1 \\ 1 \\ -\lambda \end{pmatrix}.$$

We immediately see that $s \leq 0$ has to hold in order to have $\alpha' \geq 0$. For $s = 0$, we obtain α.

Now let us consider the integral difference

$$\int_\Gamma c(x,y)\alpha'(\mathrm{d}(x,y)) - \int_\Gamma c(x,y)\alpha(\mathrm{d}(x,y)) = \int_\Gamma c(x,y)(\alpha' - \alpha)(\mathrm{d}(x,y)).$$

As $\alpha' - \alpha = s \cdot \Big(1-\lambda, -1, \lambda, \lambda - 1, 1, -\lambda\Big)^\top$, we have

$$\begin{aligned} &\int_\Gamma c(x,y)(\alpha' - \alpha)(\mathrm{d}(x,y)) \\ &\quad = s\left[(1-\lambda)c(x,y^-) - c(x,y') + \lambda c(x,y^+) - (1-\lambda)c(x',y^-) + c(x',y') - \lambda c(x',y^+)\right]. \end{aligned}$$

Thus, we have $\int_\Gamma c(x,y)(\alpha' - \alpha)(\mathrm{d}(x,y)) > 0$ for all $s < 0$, if and only if

$$\lambda\left[c(x',y^+) - c(x,y^+)\right] + (1-\lambda)\left[c(x',y^-) - c(x,y^-)\right] - \left[c(x',y') - c(x,y')\right] > 0.$$

This condition however may be rewritten by the differentiability of the payoff function. We obtain the equivalent notion

$$\int_x^{x'} \lambda c_x(t,y^+) + (1-\lambda)c_x(t,y^-) - c_x(t,y')\mathrm{d}t > 0,$$

which is satisfied as c satisfies the martingale Spence Mirrlees condition. This however is a contradiction to the finite optimality of α. Hence, Γ is left monotone. □

Remark 6.24. The main assertion of Theorem 6.23 is independently reported in Beiglböck, Henry-Labordère & Touzi [6, Theorem 3.3]. The authors also use similar ideas as in the proof of [7, Theorem 6.1]. Thus, their proof is similar to our proof. However, our proof is beneficial, as it allows us to deduce a stronger result in the case of discrete marginals. ◊

6.4. Monotonicity and optimality in the discrete case

In this section, we consider the special case of discrete marginals. We improve the assertion of Theorem 6.23, present an algorithm to determine the left monotone martingale transport plan and an algorithm to find an optimal super hedging strategy. This complements the methods of Henry-Labordère & Touzi [38] in the continuous case. Finally, we discuss an extensive example in order to illustrate the results of this section.

For this purpose, we need several further results from Beiglböck & Juillet [7]. We start with a generalized version of the convex order of two measures.

Definition 6.25 ([7, Definition 4.3])**.** Two measures $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ are said to be in *extended convex order*, denoted by $\mu \leq_E \nu$, if for any non-negative convex function $f : \mathbb{R} \to \mathbb{R}_+$ such that the integrals exist,

$$\int f(x)\mu(\mathrm{d}x) \leq \int f(x)\nu(\mathrm{d}x).$$

Let us shortly discuss this notion. Since the class of test functions is reduced, we immediately have that $\mu \leq_c \nu$ implies $\mu \leq_E \nu$. The converse is not true. Indeed, if $\mu \leq \nu$, i.e. $\mu(B) \leq \nu(B)$ for all $B \in \mathcal{B}(\mathbb{R})$, then $\mu \leq_E \nu$. This is an easy consequence of an algebraic induction. As $\mu \leq \nu$ may hold for measures with different masses or barycentres, this implies that the extended convex order extends the convex order in the sense that new relations are introduced.

However, choosing $f \equiv 1$, $\mu \leq_E \nu$ implies $\mu(\mathbb{R}) \leq \nu(\mathbb{R})$. Assuming $\mu(\mathbb{R}) = \nu(\mathbb{R})$, the convex and the extended convex order are equivalent. This can easily be deduced using Proposition 4.24, as then we have

$$\mu \leq_c \nu \Rightarrow \mu \leq_E \nu \Rightarrow u_\mu \leq u_\nu \Rightarrow \mu \leq_c \nu.$$

An alternative proof can be found after Definition 4.3 in [7].

Beiglböck & Juillet [7] also show that $\mu \leq_E \nu$ if and only if there exists an $\eta \in \mathcal{P}_\alpha(\mathbb{R})$ such that $\mu \leq_c \eta$ and $\eta \leq \nu$. In the following, we are interested in a certain choice of η, namely the so-called shadow. In order to formally introduce this particular measure, assume $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ are such that $\mu \leq_E \nu$. Then we denote by F^ν_μ the set of all measures $\eta \in \mathcal{P}_\alpha(\mathbb{R})$ such that $\mu \leq_c \eta$ and $\eta \leq \nu$, i.e.

$$F^\nu_\mu := \{\eta \in \mathcal{P}_\alpha(\mathbb{R}) \mid \mu \leq_c \eta, \eta \leq \nu\}.$$

All measures in F^ν_μ have the same mass and the same barycentre as μ. Thus, it is meaningful to search for minimal and maximal elements in the partially ordered set $(F^\nu_\mu, \leq_c)$.

Lemma 6.26 ([7, Lemma 4.6])**.** *Let $\mu, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ be such that $\mu \leq_E \nu$. Then there exists a measure $S^\nu(\mu)$ such that*

1. $S^\nu(\mu) \leq \nu$.

2. $\mu \leq_c S^\nu(\mu)$.

3. If η is another measure satisfying 1. *and* 2.*, then $S^\nu(\mu) \leq_c \eta$.*

As a consequence of 3.*, the measure $S^\nu(\mu)$ is uniquely determined. Further it satisfies*

3.' If η is a measure such that $\eta \leq \nu$ and $\mu \leq_E \eta$, then $S^\nu(\mu) \leq_E \eta$.

There also is a measure $T^\nu(\mu)$ that is maximal in the convex order, i.e. in the third condition of Lemma 6.26, we have instead $\eta \leq_c T^\nu(\mu)$. This may be proven similar to the proof of Lemma 6.44.

Definition 6.27. The uniquely defined measure $S^{\nu}(\mu)$ from Lemma 6.26 is called the *shadow (measure)* of μ in ν.

As Beiglböck & Juillet [7] show, the shadow measure has several useful properties. First of all, it is associative in the sense of the following theorem.

Theorem 6.28 ([7, Theorem 4.8])**.** *Let $\gamma_1, \gamma_2, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ be such that $\gamma_1 + \gamma_2 \leq_E \nu$. Then $\gamma_2 \leq_E \nu - S^{\nu}(\gamma_1)$ and*

$$S^{\nu}(\gamma_1 + \gamma_2) = S^{\nu}(\gamma_1) + S^{\nu - S^{\nu}(\gamma_1)}(\gamma_2).$$

Corollary 6.29. *If $\gamma_1 + \gamma_2 \leq_c \nu$ holds in the situation of Theorem 6.28, then we even have $\gamma_2 \leq_c \nu - S^{\nu}(\gamma_1)$ and*

$$\nu = S^{\nu}(\gamma_1 + \gamma_2) = S^{\nu}(\gamma_1) + S^{\nu - S^{\nu}(\gamma_1)}(\gamma_2).$$

Furthermore, the shadow measure allows a general characterization of the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu)$.

Theorem 6.30 ([7, Theorem 4.18])**.** *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ and assume $\mu \leq_c \nu$. Then there is a unique probability measure $\mathbb{Q}$ on $\mathbb{R}^2$ that transports $\mu|_{(-\infty,x]}$ to $S^{\nu}\left(\mu|_{(-\infty,x]}\right)$. That is,*

$$\operatorname{proj}^x_{\#}\left(\mathbb{Q}|_{(-\infty,x]\times\mathbb{R}}\right) = \mu|_{(-\infty,x]} \quad \text{and} \quad \operatorname{proj}^y_{\#}\left(\mathbb{Q}|_{(-\infty,x]\times\mathbb{R}}\right) = S^{\nu}\left(\mu|_{(-\infty,x]}\right)$$

for all $x \in \mathbb{R}$. $\mathbb{Q}$ is a martingale transport plan, i.e. $\mathbb{Q} \in \mathcal{M}_2(\mu, \nu)$, and it is the left monotone martingale transport plan, i.e. $\mathbb{Q} = \mathbb{Q}_{lc}(\mu, \nu)$.

In the case of a discrete marginal μ, Beiglböck & Juillet [7] also provide an abstract description of the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu)$ using the shadow measures of the single atoms of μ.

Example 6.31 ([7, Example 4.20]). Let $\delta = \alpha\delta_x$ and $\delta \leq_E \nu$. Then $S^{\nu}(\delta)$ is the restriction of ν to a measure of the form $\nu' = (F_\nu^{-1})_{\#}\lambda_{[s,s']}$ with $s' = s + \alpha$ such that $\mathbb{B}(\nu') = x$.

If $\mu = \sum_{j=1}^N \delta_j := \sum_{j=1}^N \omega_j \delta_{x_j}$ with $x_1 < \ldots < x_N$, then

$$\mathbb{Q}_{lc}(\mu, \nu) = \sum_{j=1}^N \bar{\delta}_j \otimes S^{\nu - \nu_{j-1}}(\delta_j),$$

where $\bar{\delta}_j := \frac{\delta_j}{\delta_j(\{x_j\})}$ and $\nu_j := S^{\nu}(\mu_j)$ with $\mu_j := \sum_{i=1}^j \delta_i$. △

While the description of $\mathbb{Q}_{lc}(\mu, \nu)$ provided in the example is somehow useful when the quantile function F_ν^{-1} is well-defined and easy to determine, the description is not practical when the quantile function is defined ambiguously, which is the case when the marginals μ and ν are both discrete.

6.4.1. The discrete case

Assumption 6.32. Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be such that $\mu \leq_c \nu$ and (μ, ν) is irreducible. Let μ and ν be discrete, i.e. there are $N, M \in \mathbb{N} \cup \{\infty\}$ such that

$$\mu = \sum_{j=1}^{N} \omega_j \delta_{x_j} \quad \text{and} \quad \nu = \sum_{i=1}^{M} \vartheta_i \delta_{y_i},$$

where $\omega_j, \vartheta_i \geq 0$, $x_j, y_i \in \mathbb{R}$ for all $j = 1, \ldots, N$ and all $i = 1, \ldots, M$, $\sum_{j=1}^{N} \omega_j = \sum_{i=1}^{M} \vartheta_i = 1$ and $m := \sum_{j=1}^{N} \omega_j x_j = \sum_{i=1}^{M} \vartheta_i y_i < \infty$ hold.

Under this assumption, martingale transport plans $\mathbb{Q} \in \mathcal{M}_2(\mu, \nu)$ are of the form

$$\mathbb{Q} = \sum_{j=1}^{N} \sum_{i=1}^{M} q_{j,i} \delta_{(x_j, y_i)},$$

where the following additional constraints have to be satisfied.

1. The masses of $\mathbb{Q}$ are non-negative, i.e. we have $q_{j,i} \geq 0$, for all $j = 1, \ldots, N$ and all $i = 1, \ldots, M$.

2. The marginal distributions of $\mathbb{Q}$ are μ and ν, i.e. we have

$$\sum_{i=1}^{M} q_{j,i} = \omega_j, \quad j = 1, \ldots, N,$$
$$\sum_{j=1}^{N} q_{j,i} = \vartheta_i, \quad i = 1, \ldots, M.$$

   This clearly implies $\sum_{j=1}^{N} \sum_{i=1}^{M} q_{j,i} = 1$ such that $\mathbb{Q}$ is indeed a probability measure.

3. The measure $\mathbb{Q}$ satisfies the martingale condition. We know that the martingale condition may be characterized in different ways. Transferring the classic condition $\mathbb{E}_{\mathbb{Q}}[Y \mid X] = X$ to the discrete situation, we have

$$\sum_{i=1}^{M} \frac{q_{j,i}}{\omega_j} y_i = x_j, \quad j = 1, \ldots, N,$$

   as $\frac{q_{j,i}}{\omega_j} = \frac{\mathbb{Q}((X,Y)=(x_j,y_i))}{\mathbb{Q}(X=x_j)}$ is the correct conditional distribution. Rewriting this as

$$\sum_{i=1}^{M} q_{j,i} y_i = \omega_j x_j \iff \sum_{i=1}^{M} q_{j,i} y_i - \sum_{i=1}^{M} q_{j,i} x_j = 0 \iff \sum_{i=1}^{M} q_{j,i}(y_i - x_j) = 0,$$

   we find an alternative condition, which resembles the alternative characterization $\mathbb{E}_{\mathbb{Q}}[h(X)(Y - X)] = 0$ for all suitable $h : \mathbb{R} \to \mathbb{R}$.

Altogether, the upper price bound problem in (4.7) reduces to a possibly infinite dimensional linear programm in the discrete case, namely

$$\begin{aligned}
\max \quad & \sum_{j=1}^{N}\sum_{i=1}^{M} q_{j,i}c(x_j,y_i) \coloneqq \sum_{j=1}^{N}\sum_{i=1}^{M} q_{j,i}c_{j,i} \\
\text{s.t.} \quad & \sum_{i=1}^{M} q_{j,i} = \omega_j, \quad j=1,\dots,N, \\
& \sum_{j=1}^{N} q_{j,i} = \vartheta_i, \quad i=1,\dots,M, \\
& \sum_{i=1}^{M} q_{j,i}(y_i - x_j) = 0, \quad j=1,\dots,N, \\
& q_{j,i} \geq 0, \quad j=1,\dots,N, i=1,\dots,M.
\end{aligned} \tag{6.4}$$

Formulating a discrete version of the super hedging problem in (4.10) yields the problem

$$\begin{aligned}
\min \quad & \sum_{j=1}^{N} \omega_j p_j + \sum_{i=1}^{M} \vartheta_i r_i \\
\text{s.t.} \quad & p_j + r_i + h_j(y_i - x_j) \geq c_{j,i}, \quad j=1,\dots,N, i=1,\dots,M, \\
& p_j, h_j, q_i \in \mathbb{R}, \quad j=1,\dots,N, i=1,\dots,M.
\end{aligned} \tag{6.5}$$

Remark 6.33. If we assume $N, M < \infty$, then using classic methods of linear optimization, we may state some observations.

1. In linear optimization, the dual problem for an optimization problem of the form

   $$\max q^T c \text{ such that } Aq = b, q \geq 0$$

   is given by

   $$\min h^T b \text{ such that } h^T A \geq c.$$

   It is easy to check that the upper price bound problem in (6.4) and the super hedging problem in (6.5) indeed satisfy this connection.

2. As $\mu \leq_c \nu$, by Corollary 5.2 we know that there is an admissible solution to the upper price bound problem in (6.4). As $\sum_{j=1}^{N}\sum_{i=1}^{M} q_{j,i} = 1$ holds, we also know that the maximum is finite. Thus, the upper price bound problem is admissible and solvable. Hence, by the strong duality theorem, see for example Vanderbei [76, Theorem 5.2], we know that there is a solution to the super hedging problem in (6.5) as well, even without assuming irreducibility. ◊

6.4.2. The discrete optimality result

From Theorem 6.23, we know for which payoff functions the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu,\nu)$ is optimal for the upper price bound problem in (4.7).

The discrete version in (6.4) is slightly different from a structural point of view, as only the discrete points $x_1, \dots, x_N$ and $y_1, \dots, y_M$ are important. Thus, we may prove the optimality criterion to be slightly more general.

Theorem 6.34. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.32. Let $c : \mathbb{R}^2 \to \mathbb{R}$ be a payoff function such that*

$$\frac{\frac{c(x',y^+)-c(x,y^+)}{x'-x} - \frac{c(x',y')-c(x,y')}{x'-x}}{y^+ - y'} - \frac{\frac{c(x',y')-c(x,y')}{x'-x} - \frac{c(x',y^-)-c(x,y^-)}{x'-x}}{y' - y^-} > 0 \tag{6.6}$$

for all $x' > x$ and $y^+ > y' > y^-$ with $x, x' \in \operatorname{supp}(\mu)$ and $y^-, y', y^+ \in \operatorname{supp}(\nu)$. Then $\mathbb{Q}_{lc}(\mu, \nu)$ is optimal for the discrete upper price bound problem in (6.4).

Proof. As in the proof of Theorem 6.23, we find the optimality condition

$$\lambda \left[c(x', y^+) - c(x, y^+)\right] + (1 - \lambda) \left[c(x', y^-) - c(x, y^-)\right] - \left[c(x', y') - c(x, y')\right] > 0,$$

which is then equivalent to

$$\begin{aligned} &\lambda \left[c(x', y^+) - c(x', y') - c(x, y^+) + c(x, y')\right] \\ &- (1 - \lambda) \left[c(x', y') - c(x', y^-) - c(x, y') + c(x, y^-)\right] > 0. \end{aligned}$$

If we plugin $\lambda = \frac{y'-y^-}{y^+-y^-}$, multiply by $y^+ - y^-$ and divide by $y^+ - y'$ and $y' - y^-$, then we obtain

$$\frac{[c(x', y^+) - c(x', y') - c(x, y^+) + c(x, y')]}{y^+ - y'} - \frac{[c(x', y') - c(x', y^-) - c(x, y') + c(x, y^-)]}{y' - y^-} > 0.$$

Finally, dividing by $x' - x$ and sorting the terms suitably, we have the desired condition

$$\frac{\frac{c(x',y^+)-c(x,y^+)}{x'-x} - \frac{c(x',y')-c(x,y')}{x'-x}}{y^+ - y'} - \frac{\frac{c(x',y')-c(x,y')}{x'-x} - \frac{c(x',y^-)-c(x,y^-)}{x'-x}}{y' - y^-} > 0.$$

□

The following lemma shows that all functions that satisfy the martingale Spence Mirrlees condition satisfy condition (6.6) as well. In this sense, condition (6.6) is indeed more general than the martingale Spence Mirrlees condition.

Lemma 6.35. *Let $c : \mathbb{R}^2 \to \mathbb{R}$ be a function satisfying the martingale Spence Mirrlees condition. Then (6.6) holds for all $x' > x$, $y^+ > y' > y^-$.*

Proof. In the following we prove an inequality from which we may deduce the claimed implication by dividing by $(y^+ - y')(y' - y^-)$ and $x' - x$ on both sides. In order to derive the desired inequality, we repeatedly apply the fundamental theorem of calculus, which is possible by the differentiability assumption on c. As $c_{xyy} > 0$ and $s \geq y' \geq u$ for all

$s \in [y', y^+]$ and $u \in [y^-, y']$, we have

$$\begin{aligned} 0 &< \int_x^{x'} \int_{y'}^{y^+} \int_{y^-}^{y'} \int_u^s c_{xyy}(t,v)\mathrm{d}v\mathrm{d}u\mathrm{d}s\mathrm{d}t = \int_x^{x'} \int_{y'}^{y^+} \int_{y^-}^{y'} c_{xy}(t,s) - c_{xy}(t,u)\mathrm{d}u\mathrm{d}s\mathrm{d}t \\ &= \int_x^{x'} \int_{y'}^{y^+} c_{xy}(t,s)(y'-y^-) - (c_x(t,y') - c_x(t,y^-))\mathrm{d}s\mathrm{d}t \\ &= \int_x^{x'} (c_x(t,y^+) - c_x(t,y'))(y'-y^-) - (c_x(t,y') - c_x(t,y^-))(y^+-y')\mathrm{d}t \\ &= [c(x',y^+) - c(x,y^+) - (c(x',y') - c(x,y'))](y'-y^-) \\ &\quad - [c(x',y') - c(x,y') - (c(x',y^-) - c(x,y^-))](y^+-y'). \end{aligned}$$
□

Let us now introduce a discrete case analogue of the explicit construction of $\mathbb{Q}_{lc}(\mu,\nu)$ in Section 6.2.1. The structure of $\mathbb{Q}_{lc}(\mu,\nu)$ in this situation is abstractly discussed in Example 6.31. We translate this abstract description to a constructive algorithm.

6.4.3. Construction of the left monotone martingale transport plan

Assumption 6.36. Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.32. Additionally assume that μ and ν satisfy the following conditions.

1. $x_1 < \ldots < x_N$, $y_1 < \ldots < y_M$.
2. $\mu \neq \nu$.
3. μ has at least two different atoms of positive mass.

In this section, we detail an algorithm to determine $\mathbb{Q}_{lc}(\mu,\nu)$ whenever μ and ν satisfy Assumption 6.36. Instead of using the notion of the shadow as in Example 6.31, we proceed intuitively. However, to prove the correctness of the algorithm, we prove that the construction indeed yields the desired shadow measure. In order to clarify the connection to the discrete upper price bound problem in (6.4), we denote the mass of the atom (x_j, y_i) under $\mathbb{Q}_{lc}(\mu,\nu)$, i.e. the mass transported from x_j to y_i, by $q_{j,i}$.

Remark 6.37. Let us shortly discuss the additional properties of Assumption 6.36. The first property is necessary to be able to work with left monotonicity, as this depends on the order of the atoms. The two further conditions simplify notation and formalisms. The two special cases ruled out yield formal difficulties in the algorithm though they are trivial to handle. Indeed, if the second condition is violated, then the unique martingale transport plan is the identity transport $\mu \otimes \delta_{\mathrm{id}(x)}(\mathrm{d}x, \mathrm{d}y)$. If the third condition is violated, then the single atom $\mu = \omega\delta_x$ is trivially transported to the atoms of ν by just splitting the mass properly. ◊

Let us now develop an intuition on how to proceed in a constructive algorithm. It is natural to couple the single atoms $\omega_j \delta_{x_j}$ of μ stepwise and one after another with suitable measures $\rho_j \leq \nu$ in order to determine a transport plan. As we construct the left monotone martingale transport plan, this approach has to comply with some conditions.

1. The marginal distributions of the coupling have to be μ and ν. As we couple the atoms of μ stepwise and one after another, it is obvious that μ results as a marginal. To ensure that ν results as a marginal, the measures ρ_j, $j = 1, \ldots, N$, have to satisfy $\sum_{j=1}^{N} \rho_j = \nu$. This is satisfied by coupling the atoms $\omega_j \delta_{x_j}$ of μ with suitable measures $\rho_j \leq \nu - \sum_{i=1}^{j-1} \rho_i$.

2. The coupling has to satisfy the martingale condition. As $\mu \leq_c \nu$, a martingale transport plan does exist. Starting by coupling an arbitrary atom $\omega_j \delta_{x_j}$ of μ with some measure $\rho_j \leq \nu$, two conditions have to be satisfied not to contradict the martingale property.

   First, $\mathbb{B}(\rho_j) = x_j$ is mandatory, as this is the martingale condition for the measure ρ_j. However, as $\rho_j(\mathbb{R}) = \omega_j$ naturally holds, we may also require $\omega_j \delta_{x_j} \leq_c \rho_j$, implying both properties. Secondly, $\mu - \omega_j \delta_{x_j} \leq_c \nu - \rho_j$ is necessary to ensure the existence of a martingale transport plan between the residual measures.

   Hence, coupling all atoms of μ ordered by $j_1, \ldots, j_N$ should admit

   $$\omega_{j_i} \delta_{x_{j_i}} \leq_c \rho_{j_i}$$

   for all $i = 1, \ldots, N$. Denoting $\mu^{(k)} = \mu - \sum_{i=1}^{k} \omega_{j_i} \delta_{x_{j_i}}$ and $\nu^{(k)} = \nu - \sum_{i=1}^{k} \rho_{j_i}$ for all $k = 0, \ldots, N$, it should also admit

   $$\mu^{(k)} \leq_c \nu^{(k)}.$$

   This demand reminds us of Corollary 6.29. Indeed, the corollary yields the desired convex orders, whenever we can show that, for all $i = 1, \ldots, N$, we have

   $$\rho_{j_i} = S^{\nu^{(i-1)}} \left(\omega_{j_i} \delta_{x_{j_i}} \right).$$

3. The coupling has to be left monotone. As the monotonicity is a property that relies on the order of the atoms, it is natural to couple the atoms in a certain order such that no contradictions are introduced.

   Assume we couple the atom $\omega_j \delta_{x_j}$ of μ first. Then the structure of left monotonicity yields conditions for $x_1, \ldots, x_{j-1}$ and for $x_{j+1}, \ldots, x_N$ respectively. Indeed, for all $i = 1, \ldots, j-1$ and all $y \in \operatorname{supp}(\rho_j)$, we have to satisfy

   $$y \notin \operatorname{conv}(\operatorname{supp}(\rho_i))^\circ.$$

   Analogously, for all $i = j+1, \ldots, N$ and all $y \in \operatorname{supp}(\rho_i)$, we have to satisfy

   $$y \notin \operatorname{conv}(\operatorname{supp}(\rho_j))^\circ.$$

   If in any step any of the above conditions is violated, this is a contradiction to the left monotonicity of the constructed transport plan.

As it is easier to control only one of the conditions, it seems useful to start the coupling with either the smallest atom x_1 or, if $N < \infty$, the largest atom x_N and to proceed in either increasing or decreasing order. As we shall see later, starting with the largest atom x_N, the procedure may fail. Thus, we start with the smallest atom x_1 and proceed in increasing order. Then we have to ensure that $\nu^{(k)} = \nu - \sum_{j=1}^k \rho_j$ has no mass between any atoms of ρ_k, i.e. it is crucial that, for all $k = 0, \dots, N-1$, we have

$$\operatorname{supp}\left(\nu^{(k)}\right) \cap (\operatorname{conv}(\operatorname{supp}(\rho_k)))^\circ = \emptyset.$$

As motivated, the algorithm describes detailed how to couple the smallest atom $\omega_1\delta_{x_1}$ of μ with some measure $\rho_1 \leq \nu$ such that the expected relations as discussed above are satisfied. Iteratively applying this coupling procedure yields sequences $(\rho_1, \dots, \rho_N)$, $(\mu^{(0)}, \dots, \mu^{(N)})$ and $(\nu^{(0)}, \dots, \nu^{(N)})$ of measures such that $\mu^{(k)} \leq_c \nu^{(k)}$ for all $k = 0, \dots, N$, $\mu^{(N)} = \nu^{(N)} = 0$ and $\rho_N = \nu^{(N-1)}$, as in the N-th step there only remains the coupling of one single atom $\omega_N\delta_{x_N}$, which has to be coupled with the remaining residual measure $\nu^{(N-1)}$ as discussed in Remark 6.37. The sequences $(\omega_1\delta_{x_1}, \dots, \omega_N\delta_{x_N})$ and $(\rho_1, \dots, \rho_N)$ then determine $\mathbb{Q}_{lc}(\mu, \nu)$.

The iterative construction immediately guarantees the marginal and martingale conditions. The left monotonicity is also directly implied by the iteration procedure. After presenting the algorithm, it only remains to prove that the iteration may be continued after each step. As announced in the intuition, we do this using Corollary 6.29.

However, before we provide the algorithm formally, we introduce two auxiliary lemmata.

Lemma 6.38. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.36 and suppose $x_1 \in \operatorname{supp}(\nu)$, say $x_1 = y_\ell$ for some $\ell \in \{1, \dots, M\}$.*

1. *Then (x_1, y_ℓ) is an atom of $\mathbb{Q}_{lc}(\mu, \nu)$.*
2. *Further suppose $\ell = 1$. Then $\omega_1 \leq \vartheta_1$.*

Proof. 1. In order to get a contradiction, assume that x_1 is not coupled with y_ℓ under $\mathbb{Q}_{lc}(\mu, \nu)$. Then there are $y^-, y^+ \in \operatorname{supp}(\nu)$ with $y^- < y_\ell < y^+$ and such that x_1 is coupled with y^- and y^+. Also, there is some $x' \in \operatorname{supp}(\mu)$ with $x' > x_1$ and such that x' is coupled with y_ℓ. This contradicts the left monotonicity.

2. In order to get a contradiction, assume that $\omega_1 > \vartheta_1$. After coupling x_1 with y_1 there is mass of at least $\omega_1 - \vartheta_1$ left in x_1. This has then to be coupled with some atoms from $\operatorname{supp}(\nu) \setminus \{y_1\}$. As $y_i > x_1$ for all $i = 2, \dots, M$, this contradicts the martingale property. □

Lemma 6.39. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.36 and suppose $y_\ell < x_1 < y_{\ell+1}$ for some $\ell \in \{1, \dots, M-1\}$. Then (x_1, y_ℓ) and $(x_1, y_{\ell+1})$ are atoms of $\mathbb{Q}_{lc}(\mu, \nu)$.*

Proof. In order to get a contradiction, assume without loss of generality that x_1 is not coupled with $y_{\ell+1}$. Then there is a $y^+ \in \operatorname{supp}(\nu)$ with $y^+ > y_{\ell+1}$ and such that x_1 is coupled with at least y_ℓ and y^+. Also, there is an $x' \in \operatorname{supp}(\mu)$ with $x' > x_1$ and such that x' is coupled with $y_{\ell+1}$. This contradicts the left monotonicity. □

Algorithm 6.40. Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.36 and let $x_1 \in \operatorname{supp}(\mu)$ be the smallest atom of μ. The atom x_1 has mass ω_1 and we denote $\delta := \omega_1 \delta_{x_1}$. We distinguish two cases.

1. $x_1 \in \operatorname{supp}(\nu)$.

   Then there is an $\ell \in \{1, \dots, M\}$ with $x_1 = y_\ell$. By the first part of Lemma 6.38, x_1 has to be coupled with y_ℓ.

   1.1. $\ell = 1$.

   By the second part of Lemma 6.38, we have $\omega_1 \leq \vartheta_1$. We define $q_{1,1} := \omega_1$, $\rho_1 := q_{1,1}\delta_{y_1}$,

   $$\mu^{(1)} := \sum_{j=2}^{N} \omega_j \delta_{x_j} \quad \text{and} \quad \nu^{(1)} := \nu - \rho_1.$$

   By coupling δ with ρ_1, the total mass ω_1 of x_1 is transported to ν and we have $\mu^{(1)} \leq_c \nu^{(1)}$. We may thus apply the iteration step to these measures for the smallest atom x_2 of $\mu^{(1)}$.

   1.2. $\ell = 2, \dots, M-1$.

   We have $y_{\ell-1} < x_1 < y_{\ell+1}$.

   1.2.1. $\omega_1 \leq \vartheta_\ell$.

   As in 1.1, we define $q_{1,\ell} := \omega_1$, $\rho_1 := q_{1,\ell}\delta_{y_\ell}$,

   $$\mu^{(1)} := \sum_{j=2}^{N} \omega_j \delta_{x_j} \quad \text{and} \quad \nu^{(1)} := \nu - \rho_1.$$

   By coupling δ with ρ_1, the total mass ω_1 of x_1 is transported to ν and we have $\mu^{(1)} \leq_c \nu^{(1)}$. We may thus apply the iteration step to these measures for the smallest atom x_2 of $\mu^{(1)}$.

   1.2.2. $\omega_1 > \vartheta_\ell$.

   We define $q_{1,\ell} := \vartheta_\ell$ and $\rho_1' := q_{1,\ell}\delta_{y_\ell}$. That is, we do not yet transport the total mass ω_1 of x_1 to ν. Hence, we couple x_1 with further atoms of ν. By Lemma 6.39, we have to couple with $y_{\ell-1}$ and $y_{\ell+1}$ first. As $y_{\ell-1} < x_1 < y_{\ell+1}$, there are numbers $q_{1,\ell-1}', q_{1,\ell+1}' \geq 0$ such that

   $$q_{1,\ell-1}' + q_{1,\ell+1}' = \omega_1 - \rho_1'(\mathbb{R}),$$

   $$q_{1,\ell-1}' y_{\ell-1} + q_{1,\ell+1}' y_{\ell+1} + \rho_1'(\mathbb{R})\,\mathbb{B}(\rho_1') = \omega_1 x_1.$$

   Considering $q_{1,\ell-1}'$ and $q_{1,\ell+1}'$, four cases may appear.

1.2.2.1. $q'_{1,\ell-1} > \vartheta_{\ell-1}$ and $q'_{1,\ell+1} > \vartheta_{\ell+1}$. Proceed as in 2.1.$[[\ell-1],[\ell+1]]$.

1.2.2.2. $q'_{1,\ell-1} > \vartheta_{\ell-1}$ and $q'_{1,\ell+1} \leq \vartheta_{\ell+1}$. Proceed as in 2.2.$[[\ell-1],[\ell+1]]$.

1.2.2.3. $q'_{1,\ell-1} \leq \vartheta_{\ell-1}$ and $q'_{1,\ell+1} > \vartheta_{\ell+1}$. Proceed as in 2.3.$[[\ell-1],[\ell+1]]$.

1.2.2.4. $q'_{1,\ell-1} \leq \vartheta_{\ell-1}$ and $q'_{1,\ell+1} \leq \vartheta_{\ell+1}$. Proceed as in 2.4.$[[\ell-1],[\ell+1]]$.

1.3. $\ell = M$.

Then $x_1 = y_M$, which is possible only if $N = M = 1$ and $\omega_1 = \vartheta_M$. However, that is $\mu = \nu$, a contradiction to Assumption 6.36.

2. $x_1 \notin \operatorname{supp}(\nu)$.

Then there is an $\ell \in \{1, \ldots, M-1\}$ such that $y_\ell < x_1 < y_{l+1}$. Thus, there are numbers $q'_{1,\ell}, q'_{1,\ell+1} \geq 0$ such that

$$q'_{1,\ell} + q'_{1,\ell+1} = \omega_1,$$

$$q'_{1,\ell} y_\ell + q'_{1,\ell+1} y_{\ell+1} = \omega_1 x_1.$$

Considering $q'_{1\ell}$ and $q'_{1\ell+1}$, four cases may appear.

2.1. $[[\ell],[\ell+1]]$: $q'_{1,[\ell]} > \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]} > \vartheta_{[\ell+1]}$.

Then it is not possible to (additionally) couple x_1 with only $y_{[\ell]}$ and $y_{[\ell+1]}$. Thus, we define $q_{1,[\ell]} := \vartheta_{[\ell]}$ and $q_{1,[\ell+1]} := \vartheta_{[\ell+1]}$ as well as

$$\rho'_1 := \rho'_1 + q_{1,[\ell]} \delta_{y_{[\ell]}} + q_{1,[\ell+1]} \delta_{y_{[\ell+1]}}.$$

That is, we do not yet transport the total mass ω_1 of x_1 to ν. Hence, the remaining mass has to be coupled with $y_{[\ell]-1}$ and $y_{[\ell+1]+1}$. As $y_{[\ell]-1} < x_1 < y_{[\ell+1]+1}$, there are numbers $q'_{1,[\ell]-1}, q'_{1,[\ell+1]+1} \geq 0$ such that

$$q'_{1,[\ell]-1} + q'_{1,[\ell+1]+1} = \omega_1 - \rho'_1(\mathbb{R}),$$

$$q'_{1,[\ell]-1} y_{[\ell]-1} + q'_{1,[\ell+1]+1} y_{[\ell+1]+1} + \rho'_1(\mathbb{R}) \mathbb{B}(\rho'_1) = \omega_1 x_1.$$

Considering $q'_{1,[\ell]-1}$ and $q'_{1,[\ell+1]+1}$, four cases may appear.

2.1.1. $q'_{1,[\ell]-1} > \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]+1} > \vartheta_{[\ell+1]+1}$. Proceed as in 2.1.$[[\ell-1],[\ell+2]]$.

2.1.2. $q'_{1,[\ell]-1} > \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]+1} \leq \vartheta_{[\ell+1]+1}$. Proceed as in 2.2.$[[\ell-1],[\ell+2]]$.

2.1.3. $q'_{1,[\ell]-1} \leq \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]+1} > \vartheta_{[\ell+1]+1}$. Proceed as in 2.3.$[[\ell-1],[\ell+2]]$.

2.1.4. $q'_{1,[\ell]-1} \leq \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]+1} \leq \vartheta_{[\ell+1]+1}$. Proceed as in 2.4.$[[\ell-1],[\ell+2]]$.

2.2. $[[\ell],[\ell+1]]$: $q'_{1,[\ell]} > \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]} \leq \vartheta_{[\ell+1]}$.

Then it is not possible to (additionally) couple x_1 with only $y_{[\ell]}$ and $y_{[\ell+1]}$. Thus, we define $q_{1,[\ell]} := \vartheta_{[\ell]}$ and $\rho'_1 := \rho'_1 + \vartheta_{1[\ell]} \delta_{y_{[\ell]}}$. That is, we do not yet transport the total mass ω_1 of x_1 to ν. Hence, the remaining mass has to be coupled with

$y_{[\ell]-1}$ and $y_{[\ell+1]}$. As $y_{[\ell]-1} < x_1 < y_{[\ell+1]}$, there are numbers $q'_{1,[\ell]-1}, q'_{1,[\ell+1]} \geq 0$ such that

$$q'_{1,[\ell]-1} + q'_{1,[\ell+1]} = \omega_1 - \rho'_1(\mathbb{R}),$$

$$q'_{1,[\ell]-1} y_{[\ell]-1} + q_{1,[\ell+1]} y'_{[\ell+1]} + \rho'_1(\mathbb{R}) \mathbb{B}(\rho'_1) = \omega_1 x_1.$$

Considering $q'_{1,[\ell]-1}$ and $q'_{1,[\ell+1]}$, four cases may appear.

2.2.1. $q'_{1,[\ell]-1} > \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]} > \vartheta_{[\ell+1]}$. Proceed as in 2.1.[[$\ell-1$], [$\ell+1$]].

2.2.2. $q'_{1,[\ell]-1} > \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]} \leq \vartheta_{[\ell+1]}$. Proceed as in 2.2.[[$\ell-1$], [$\ell+1$]].

2.2.3. $q'_{1,[\ell]-1} \leq \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]} > \vartheta_{[\ell+1]}$. Proceed as in 2.3.[[$\ell-1$], [$\ell+1$]].

2.2.4. $q'_{1,[\ell]-1} \leq \vartheta_{[\ell]-1}$ and $q'_{1,[\ell+1]} \leq \vartheta_{[\ell+1]}$. Proceed as in 2.4.[[$\ell-1$], [$\ell+1$]].

2.3. [[ℓ], [$\ell+1$]]: $q'_{1,[\ell]} \leq \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]} > \vartheta_{[\ell+1]}$.

Then it is not possible to (additionally) couple x_1 with only $y_{[\ell]}$ and $y_{[\ell+1]}$. Thus, we define $q_{1,[\ell+1]} \coloneqq \vartheta_{[\ell+1]}$ and $\rho'_1 \coloneqq \rho'_1 + \vartheta_{1[\ell+1]} \delta_{y_{[\ell+1]}}$. That is, we do not yet transport the total mass ω_1 of x_1 to ν. Hence, the remaining mass has to be coupled with $y_{[\ell]}$ and $y_{[\ell+1]+1}$. As $y_{[\ell]} < x_1 < y_{[\ell+1]+1}$, there are numbers $q'_{1,[\ell]}, q'_{1,[\ell+1]+1} \geq 0$ such that

$$q'_{1,[\ell]} + q'_{1,[\ell+1]+1} = \omega_1 - \rho'_1(\mathbb{R}),$$

$$q'_{1,[\ell]} y_{[\ell]} + q'_{1,[\ell+1]+1} y_{[\ell+1]+1} + \rho'_1(\mathbb{R}) \mathbb{B}(\rho'_1) = \omega_1 x_1.$$

Considering $q'_{1,[\ell]}$ and $q'_{1,[\ell+1]+1}$, four cases may appear.

2.3.1. $q'_{1,[\ell]} > \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]+1} > \vartheta_{[\ell+1]+1}$. Proceed as in 2.1.[[$\ell$], [$\ell+2$]].

2.3.2. $q'_{1,[\ell]} > \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]+1} \leq \vartheta_{[\ell+1]+1}$. Proceed as in 2.2.[[$\ell$], [$\ell+2$]].

2.3.3. $q'_{1,[\ell]} \leq \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]+1} > \vartheta_{[\ell+1]+1}$. Proceed as in 2.3.[[$\ell$], [$\ell+2$]].

2.3.4. $q'_{1,[\ell]} \leq \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]+1} \leq \vartheta_{[\ell+1]+1}$. Proceed as in 2.4.[[$\ell$], [$\ell+2$]].

2.4. [[ℓ], [$\ell+1$]] : $q'_{1,[\ell]} \leq \vartheta_{[\ell]}$ and $q'_{1,[\ell+1]} \leq \vartheta_{[\ell+1]}$.

We define $q_{1,[\ell]} \coloneqq q'_{1,[\ell]}$ and $q_{1,[\ell+1]} \coloneqq q'_{1,[\ell+1]}$ as well as

$$\rho_1 \coloneqq \rho'_1 + q_{1,[\ell]} \delta_{y_{[\ell]}} + q_{1,[\ell+1]} \delta_{y_{[\ell+1]}}.$$

Further we define

$$\mu^{(1)} \coloneqq \sum_{j=2}^{N} \omega_j \delta_{x_j} \text{ and } \nu^{(1)} \coloneqq \nu - \rho_1.$$

By coupling δ with ρ_1, the total mass ω_1 of x_1 is transported to ν and we have $\mu^{(1)} \leq_c \nu^{(1)}$. Hence, we may apply the iteration step to these measures for smallest atom x_2 of $\mu^{(1)}$.

Let us now show that Corollary 6.29 applies in the cases 1.1, 1.2 and 2.4 in the sense that indeed $\mu^{(1)} \leq_c \nu^{(1)}$ holds. As this is only useful, if we ensure that in case 2, case 2.4 is reached somehow, we begin by showing that.

Lemma 6.41. *In Algorithm 6.40, if $x_1 \notin \operatorname{supp}(\nu)$ or $x_1 = y_\ell \in \operatorname{supp}(\nu)$ such that $\omega_1 > \vartheta_\ell$, then we reach case 2.4 after a finite number of recursive steps through cases 2.1, 2.2 and 2.3 for all $\mu, \nu \in \mathcal{P}(\mathbb{R})$ that satisfy Assumption 6.36.*

Proof. By assumption, x_1 is the smallest of at least two atoms of μ. Thus, the total mass $\omega_1 < \mu(\mathbb{R})$ of x_1 is coupled with a set of subsequent atoms from the countably many atoms of ν, say with $\{y_n, \ldots, y_m\}$. The algorithm guarantees that the masses $\vartheta_{n+1}, \ldots, \vartheta_{m-1}$ are used up entirely. In particular, after the coupling there is no remaining mass in (y_n, y_m) and we have $\omega_1 \leq \sum_{j=n}^{m} \vartheta_j$. We distinguish two cases.

1. $m < \infty$. Then x_1 is coupled with finitely many atoms. Hence, we clearly reach case 2.4 after a finite number of iterated applications of the cases 2.1, 2.2 and 2.3, as in each of this cases the mass of at least one of the atoms in $\{y_n, \ldots, y_m\}$ is used up.

2. $m = \infty$. Then x_1 is coupled with infinitely many atoms. However, this contradicts the martingale and the left monotonicity properties. Indeed, only the atoms $y_1, \ldots, y_n$ remain to be coupled with some atom x_2 of $\mu - \omega_1 \delta_{x_1}$. By assumption we have $x_2 > x_1$ and by construction we have $x_1 > y_n$. Thus, x_2 is greater than the greatest atom remaining in ν with which it could possibly be coupled. Thus $m = \infty$ is possible only if x_1 is the last remaining atom of μ. This however is ruled out by Assumption 6.36.

We conclude that we do reach case 2.4 after a finite number of case applications. □

In order to prove that Corollary 6.29 applies and implies $\mu^{(1)} \leq_c \nu^{(1)}$ in all relevant cases, we show that $\rho_1 = S_\nu(\delta)$, where in analogy to the algorithm we now write

$$\rho_1 = q_{1,n}\delta_{y_n} + \sum_{j=n+1}^{m-1} \vartheta_j \delta_{y_j} + q_{1,m}\delta_{y_m}.$$

Lemma 6.42. *Applying Algorithm 6.40, we have $\rho_1 = S^\nu(\delta)$.*

Proof.

1. By construction, we have $\rho_1 \leq \nu$.

2. As δ is an atom, we have $\delta \leq_c \eta$ for all η with the same mass and barycentre. By construction, ρ_1 satisfies this property.

3. Let η be a measure with $\delta \leq_c \eta$ and $\eta \leq \nu$. Then $\eta = \sum_{j=1}^{M} \gamma_j \delta_{y_j}$, where $0 \leq \gamma_j \leq \vartheta_j$, $j = 1, \ldots, M$. By construction, we have $\rho_1 = q_{1,n}\delta_{y_n} + \sum_{j=n+1}^{m-1} \vartheta_j \delta_{y_j} + q_{1,m}\delta_{y_m}$.

   Let us now show that $\rho_1 \leq_c \eta$. For this purpose, we write

$$\rho_1 = \rho_1 - (\rho_1 \wedge \eta) + (\rho_1 \wedge \eta) =: \rho_1' + (\rho_1 \wedge \eta),$$
$$\eta = \eta - (\rho_1 \wedge \eta) + (\rho_1 \wedge \eta) =: \eta' + (\rho_1 \wedge \eta).$$

Now it is sufficient to show that $\rho_1' \leq_c \eta'$, as adding $\rho_1 \wedge \eta$ does not change the convex order. Let us investigate $\rho_1 \wedge \eta$. By construction, ρ_1 is concentrated on $[y_n, y_m]$ and we also have $\rho_1 = \nu$ on $[y_{n+1}, y_{m-1}]$. Thus,

$$(\rho_1 \wedge \eta) = \begin{cases} \eta, & \text{on } (y_n, y_m) \\ \rho_1, & \text{on } (-\infty, y_n) \cup (y_m, \infty). \end{cases}$$

Concerning y_n and y_m we have to consider several cases.

3.1. $\rho_1 \geq \eta$ on $\{y_n, y_m\}$. Then $\rho_1 \wedge \eta = \eta$ on $\{y_n, y_m\}$ and thus

$$\rho_1' = \rho_1 - (\rho_1 \wedge \eta) = \begin{cases} \rho_1 - \eta, & \text{on } [y_n, y_m] \\ 0, & \text{on } [y_n, y_m]^c, \end{cases}$$

$$\eta' = \eta - (\rho_1 \wedge \eta) = \begin{cases} 0, & \text{on } [y_n, y_m] \\ \eta - \rho_1, & \text{on } [y_n, y_m]^c. \end{cases}$$

Now let $f : \mathbb{R} \to \mathbb{R}$ be a convex function and $g : \mathbb{R} \to \mathbb{R}$ the linear function with $f(y_n) = g(y_n)$ and $f(y_m) = g(y_m)$ as illustrated in Figure 6.3.

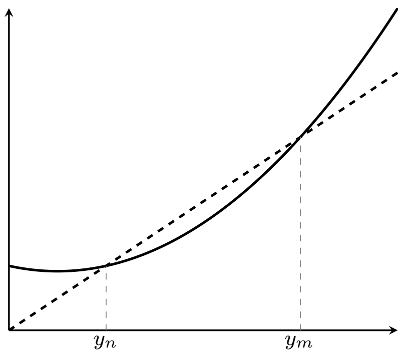


Figure 6.3.: A convex function (solid) and the intersecting linear function (dashed).

Then $f \leq g$ on $[y_n, y_m] = \operatorname{supp}(\rho_1')$ and $f \geq g$ on $[y_n, y_m]^c = \operatorname{supp}(\eta')$. Hence,

$$\int_{\mathbb{R}} f(y)\rho_1'(\mathrm{d}y) \leq \int_{\mathbb{R}} g(y)\rho_1'(\mathrm{d}y) = \int_{\mathbb{R}} g(y)\eta'(\mathrm{d}y) \leq \int_{\mathbb{R}} f(y)\eta'(\mathrm{d}y),$$

where in the equality we use that ρ_1 and η, and thus ρ_1' and η' have the same mass and the same barycentre, and that g is a linear function. Hence, $\rho_1' \leq_c \eta'$.

3.2. $\rho_1 \geq \eta$ in y_n and $\rho_1 \leq \eta$ in y_m. Then $\rho_1 \wedge \eta = \eta$ in y_n and $\rho_1 \wedge \eta = \rho_1$ in y_m. The argumentation is done replacing y_m by y_{m-1} in 3.1.

3.3. $\rho_1 \leq \eta$ in y_n and $\rho_1 \geq \eta$ in y_m. Then $\rho_1 \wedge \eta = \rho_1$ in y_n and $\rho_1 \wedge \eta = \eta$ in y_m. The argumentation is done replacing y_n by y_{n+1} in 3.1.

3.4. $\rho_1 \leq \eta$ on $\{y_n, y_m\}$. Then $\rho_1 \wedge \eta = \rho_1$ on $\{y_n, y_m\}$ The argumentation is done replacing y_n by y_{n+1} and y_m by y_{m-1} in 3.1. □

Remark 6.43. 1. Based on Example 6.31 and Algorithm 6.40, in the discrete marginal case, we now know how to choose s in $S_\nu(\delta) = (F_\nu^{-1})_\# \lambda_{[s,s']}$ with $s' = s + \omega_1$. Obviously, $S_\nu(\delta)$ equals ν in between the quantiles s and $s+\omega_1$ such that we have to choose

$$s = \sum_{j=1}^{n} \vartheta_j - q_{1,n}.$$

Analogously, we get the alternative representation $s' = \sum_{j=1}^{m-1} \vartheta_j + q_{1,m}$.

Similarly, the algorithm gives an intuition how the shadow $S^\nu(\delta) = (F_\nu^{-1})_\# \lambda_{[s,s']}$ looks like when ν is continuous. It is the measure that we receive by moving the correct mass along the distribution function of ν until the correct barycentre results. That is,

$$F_\nu(s') - F_\nu(s) = \omega_1 \quad \text{and} \quad \int_s^{s'} x F_\nu(\mathrm{d}x) = x_1.$$

2. We could expect that we also obtain the left monotone martingale transport plan, when we proceed the other way around, i.e. when we start with the greatest atom of μ and iteratively determine the greatest measure in convex order, $T_\nu(\delta)$, for each atom δ. However, this is not the case as we see in the following example.

   Consider the discrete probability measures

$$\mu = \frac{1}{6}\delta_1 + \frac{1}{2}\delta_3 + \frac{1}{3}\delta_6 \quad \text{and} \quad \nu = \frac{1}{15}\delta_{-2} + \frac{1}{2}\delta_2 + \frac{1}{6}\delta_4 + \frac{4}{15}\delta_8.$$

   They satisfy $\mu \leq_c \nu$, as we are able to construct a martingale transport plan between the two measures. Let us try to determine the left monotone martingale transport plan using $T^\nu(\cdot)$. Obviously, we have $T^\nu(\frac{1}{3}\delta_6) = \frac{1}{15}\delta_{-2} + \frac{4}{15}\delta_8$. Then, using the algorithm in consideration, we get

$$\mu^{(1)} := \mu - \frac{1}{3}\delta_6 = \frac{1}{6}\delta_1 + \frac{1}{2}\delta_3 \quad \text{and} \quad \nu^{(1)} := \nu - T^\nu(\delta_6) = \frac{1}{2}\delta_2 + \frac{1}{6}\delta_4.$$

   For those measure we immediately have $\mu^{(1)} \not\leq_c \nu^{(1)}$, as $l_{\mu^{(1)}} < l_{\nu^{(1)}}$. Hence, there is no martingale transport plan between those two measures and we are in particular not able to determine $T^{\nu - T^\nu(\frac{1}{3}\delta_6)}\left(\frac{1}{2}\delta_3\right)$.

   If we instead apply Algorithm 6.40 to the measures μ and ν, then we obtain

$$S^\nu\left(\frac{1}{6}\delta_1\right) = \frac{1}{24}\delta_{-2} + \frac{1}{8}\delta_2,$$
$$S^{\nu - S^\nu(\frac{1}{6}\delta_1)}\left(\frac{1}{2}\delta_3\right) = \frac{11}{36}\delta_2 + \frac{1}{6}\delta_4 + \frac{1}{36}\delta_8,$$

   and finally

$$S^{\nu - S^\nu(\frac{1}{6}\delta_1) - S^{\nu - S^\nu(\frac{1}{6}\delta_1)}(\frac{1}{2}\delta_3)}\left(\frac{1}{3}\delta_6\right) = \frac{1}{40}\delta_{-2} + \frac{5}{72}\delta_2 + \frac{43}{180}\delta_8.$$

   This indeed determines the left monotone martingale transport plan. ◊

In order to overcome this drawback of $T^\nu(\cdot)$, we now consider measures $\gamma_1, \gamma_2, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ such that $\gamma_1 + \gamma_2 \leq_c \nu$ and the set

$$F^{\nu,\leq_c}_{\gamma_1+\gamma_2} \coloneqq \{\eta \in \mathcal{P}_\alpha(\mathbb{R}) \mid \gamma_1 \leq_c \eta, \eta \leq \nu \text{ and } \gamma_2 \leq_c \nu - \eta\}.$$

It is clear that $S^\nu(\gamma_1) \in F^{\nu,\leq_c}_{\gamma_1+\gamma_2}$ holds and that $S^\nu(\gamma_1)$ is the minimal element with respect to the convex order in $F^{\nu,\leq_c}_{\gamma_1+\gamma_2}$. Indeed, this follows from Lemma 6.26. Now let us consider the maximal element with respect to the convex order in $F^{\nu,\leq_c}_{\gamma_1+\gamma_2}$.

Lemma 6.44. *Let $\gamma_1, \gamma_2, \nu \in \mathcal{P}_\alpha(\mathbb{R})$ with $\gamma_1 + \gamma_2 \leq_c \nu$. Then there is a measure $C^\nu(\gamma_1)$ such that*

1. *$C^\nu(\gamma_1) \leq \nu$,*
2. *$\gamma_1 \leq_c C^\nu(\gamma_1)$,*
3. *$\gamma_2 \leq_c \nu - C^\nu(\gamma_1)$,*
4. *If η is another measure with the above properties, then $\eta \leq_c C^\nu(\gamma_1)$.*

Proof. The proof goes essentially as the proof of [7, Lemma 4.6]. Thus, we translate the problem in the language of potential functions. We aim to find a suitable convex function $h : \mathbb{R} \to \mathbb{R}_+$ that is the potential function of $C^\nu(\gamma_1)$. In order to understand the conditions for a function to be suitable, we translate the first, the second and the third condition of the lemma. Here, we write $\mu = \gamma_1 + \gamma_2$ and denote by $k_1 = \gamma_1(\mathbb{R})$ the mass of γ_1 and by

$$m_1 = \frac{1}{k_1} \int_{\mathbb{R}} x \gamma_1(\mathrm{d}x)$$

the barycentre of γ_1. We further denote by $k_2 = \nu(\mathbb{R}) - \gamma_2(\mathbb{R}) = \mu(\mathbb{R}) - \gamma_2(\mathbb{R}) = \gamma_1(\mathbb{R}) = k_1$ the mass of $\nu - \gamma_2$ and by

$$\begin{aligned} m_2 &= \frac{1}{k_2} \int_{\mathbb{R}} x(\nu - \gamma_2)(\mathrm{d}x) = \frac{1}{k_1} \left(\int_{\mathbb{R}} x \nu(\mathrm{d}x) - \int_{\mathbb{R}} x \gamma_2(\mathrm{d}x) \right) \\ &= \frac{1}{k_1} \left(\int_{\mathbb{R}} x \mu(\mathrm{d}x) - \int_{\mathbb{R}} x \gamma_2(\mathrm{d}x) \right) = \frac{1}{k_1} \int_{\mathbb{R}} x \gamma_1(\mathrm{d}x) = m_1 \end{aligned}$$

the barycentre of $\nu - \gamma_2$, where in the above calculations we use $\mu \leq_c \nu$. By Propositions 4.20 and 4.21, the translations of the three defining properties are the following.

1. $u_\nu - h$ is convex.
2. $u_{\gamma_1} \leq h$ and $\lim_{-x \to \infty} h(x) - k_1|x - m_1| = 0 = \lim_{x \to \infty} h(x) - k_1|x - m_1|$.
3. $u_{\gamma_2} \leq u_\nu - h$, i.e $h \leq u_\nu - u_{\gamma_2}$.

We remark that the second part of the second condition in connection with the convexity already guarantees that h is indeed a potential function, see Proposition 4.20. Thus, the transformation in the third condition is unproblematic. Furthermore, we remark that in the third condition we omitted $\lim_{-x\to\infty} h(x) - k_2|x - m_2| = 0 = \lim_{-x\to\infty} h(x) - k_2|x - m_2|$, as this is implied by the second part of the second condition and the previous calculations.

Now we define the set of all relevant functions

$$U_F := \{h : \mathbb{R} \to \mathbb{R}_+ \mid h \text{ is convex and satisfies conditions } 1. - 3.\}$$

The set U_F contains all functions that could be the potential function of $C^\nu(\gamma_1)$. The actual potential function $\tilde{h}$ has to satisfy $\tilde{h} \geq h$ for all $h \in U_F$ by the fourth condition.

By the existence of $S^\nu(\gamma_1)$ and Corollary 6.29, we know that $U_F \neq \emptyset$. Thus, we define

$$\tilde{h}(x) := \sup_{h\in U_F} h(x), \quad x \in \mathbb{R}.$$

The function $\tilde{h}$ is the potential function of some measure, as it satisfies the conditions of Proposition 4.20. Furthermore, it satisfies the first, the second and the third condition and is by construction the greatest such function. The desired measure $C^\nu(\gamma_1)$ now is the unique measure $\rho \in \mathcal{P}_\alpha(\mathbb{R})$ such that $u_\rho \equiv \tilde{h}$. □

Remark 6.45. By the additional defining property of $C^\nu(\cdot)$ compared to $S^\nu(\cdot)$ and $T^\nu(\cdot)$, which basically ensures the validity of an analogue to Corollary 6.29, it is possible to introduce an abstract procedure to determine the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu)$. Using the measure $C^\nu(\cdot)$, we get it as follows.

For $\mu = \sum_{j=1}^N \delta_j = \sum_{j=1}^N \omega_j \delta_{x_j}$ with $x_1 < \ldots < x_N$, we have

$$\mathbb{Q}_{lc}(\mu, \nu) = \sum_{j=0}^{N-1} \bar{\delta}_{N-j} \otimes C^{\nu_{N-j}}(\delta_{N-j}),$$

where $\bar{\delta}_{N-j} := \frac{\delta_{N-j}}{\delta_{N-j}(\{x_{N-j}\})}$, $\nu_N := \nu$, and $\nu_j := \nu_{j+1} - C^{\nu_{j+1}}(\delta_{j+1})$, $j = 0, \ldots, N-1$.

However, a counterpart to Algorithm 6.40 is hard to provide, as the determination of the measure $C^\nu(\cdot)$ is difficult. ♢

Remark 6.46. Independently of our work, Hobson & Norgilas [48] also considered the problem of determining the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu)$. The results are similar to those of Beiglböck & Juillet [7] and Henry-Labordère & Touzi [38] in the sense that two mappings are defined and constructed that fully characterize $\mathbb{Q}_{lc}(\mu, \nu)$. In the case of a continuous marginal μ, the mappings provided by Hobson & Norgilas [48] and those provided of Henry-Labordère & Touzi [38] are in a one-to-one connection.

However, differently to Henry-Labordère & Touzi [38] and similar to our work, Hobson & Norgilas [48] provide a procedure to determine $\mathbb{Q}_{lc}(\mu, \nu)$ in the presence of atoms. Indeed, whenever μ is a discrete measure the procedure yields the desired martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu)$ in an algorithmic fashion and independently of the structure of ν. For non-discrete μ, the determination of $\mathbb{Q}_{lc}(\mu, \nu)$ has to be done by approximation.

Though the work of Hobson & Norgilas [48] covers all possible cases, we suggest to prefer our approach and that of Henry-Labordère & Touzi [38] in the purely discrete and purely continuous cases respectively, as even the algorithmic part of the general procedure is highly non-intuitive and complicated to implement and execute. We combine the three complementary approaches altogether in the following example. ◊

Example 6.47 (Construction of $\mathbb{Q}_{lc}(\mu,\nu)$ in the case of mixed marginals). We completely presented techniques to determine the left monotone martingale transport plan in the cases of purely discrete and purely continuous marginals. Also, we remarked that there is a generally functioning approach that is rather difficult to apply.

Now we consider a special case not discussed so far, namely the case of mixed marginals. We combine the three mentioned techniques to gain a procedure as easy as possible.

In particular, we consider the case of marginals supported in $\mathbb{R}_+$ having exactly one atom located in 0. Heuristically, this could model the value of a company with a positive default probability. Formally, let $\mu,\nu \in \mathcal{P}_\alpha(\mathbb{R}_+)$ be continuous and $0 < \omega_0 < \vartheta_0 < 1$ be such that

$$\mu' := \mu + \omega_0\delta_0 \leq_c \nu + \vartheta_0\delta_0 =: \nu',$$

$\mu',\nu' \in \mathcal{P}(\mathbb{R}_+)$ and (μ',ν') is irreducible. Clearly, $\omega_0 \leq \vartheta_0$ is necessary in order not to contradict the convex order $\mu' \leq_c \nu'$. Also, $\omega_0 < \vartheta_0$ is necessary in order not to contradict the irreducibility of (μ',ν'). However, this assumption is weak, as default probabilities increase over time.

Let us now start to determine the left monotone transport plan $\mathbb{Q}_{lc}(\mu',\nu')$. Algorithm 6.40 immediately tells us that we have to couple the total mass ω_0 of δ_0 in μ' with δ_0 in ν', as otherwise the martingale property is violated. As $\omega_0\delta_0 = S^{\nu'}(\omega_0\delta_0)$, by Corollary 6.29 we have

$$\mu \leq_c \nu + (\vartheta_0 - \omega_0)\delta_0 =: \nu + \chi_0\delta_0 =: \bar{\nu}.$$

This coupling does not interfere with the assertions of Theorem 6.30, as $\omega_0\delta_0 = \mu|_{(-\infty,0]}$. Thus, we proceed to determine $\mathbb{Q}_{lc}(\mu,\bar{\nu})$ such that

$$\mathbb{Q}_{lc}(\mu',\nu') = \mathbb{Q}_{lc}(\mu,\bar{\nu}) + \omega_0\delta_{(0,0)}.$$

Therefore, we consider μ and $\bar{\nu}$. By assumption, μ is continuous and hence we are in the situation of Corollary 6.5. Thus, we have

$$\mathbb{Q}_{lc}(\mu,\bar{\nu}) = \mu \otimes \left(q(x)\delta_{T_u(x)} + (1-q(x))\delta_{T_d(x)}\right),$$

where $q(x) = \frac{x - T_d(x)}{T_u(x)-T_d(x)}\mathbb{1}_{\{T_u(x)>T_d(x)\}}$ and $T_u, T_d : \mathbb{R}_+ \to \mathbb{R}_+$ are suitable mappings.

As $\bar{\nu}$ is not continuous, we may not apply the techniques of Henry-Labordère & Touzi [38] directly. However, the structure having only one atom at the left boundary of the support of $\bar{\nu}$ allows us to proceed very similarly. Indeed, we proceed as the authors do to gain a first intuition.

Thus, let us assume that there is a unique maximum $m \in \mathbb{R}_+$ of the difference function $\delta_F = F_{\bar{\nu}} - F_\mu$ that also serves as a bifurcation point for T_u and T_d, i.e. $T_u \equiv T_d \equiv \mathrm{id}$ on

$(0, m)$, $T_u : (m, \infty) \to (m, \infty)$ is monotone non-decreasing and $T_d : (m, \infty) \to (0, m)$ is monotone non-increasing. We also define

$$m_0 := \inf \{x \in \operatorname{supp}(\mu) \mid T_d(x) = 0, x \geq m\},$$

the smallest point in the support of μ that is coupled with δ_0. While it is easy to calculate the bifurcation point m, it is not immediately clear how to determine m_0. We come back to this later.

First, let us split up the measures μ and $\bar{\nu}$. We write

$$\mu = \mu|_{(0,m)} + \mu|_{(m,m_0)} + \mu|_{(m_0,\infty)} =: \mu_1 + \mu_2 + \mu_3.$$

We do this, as then we may easier couple in a left monotone fashion.

Indeed, on $(0, m)$ we have $T_u \equiv T_d \equiv \mathrm{id}$, i.e. $(0, m)$ is the set of particles of μ that are coupled with themselves at least almost surely. We stress that this is necessary in order to maintain the left monotonicity. Indeed, if there is a μ-non-null set $\mathcal{N}_s$ of points $x' \in (0, m)$ such that $T_u(x') \neq T_d(x')$, then we have $T_d(x') < x' < T_u(x')$ by the martingale property. Then there must be a $\bar{\nu}$-non-null set of points x'' such that there exists $x' \in \mathcal{N}_s$ with $x'' > x'$ and $T_d(x'') = x'$, a contradiction to the left monotonicity.

Further, on (m, m_0) we have $0 < T_d < T_u$ and hence this is the set of particles of μ, the mass of which is split up but not coupled with δ_0. This however implies that the measure $\nu_2 \leq \nu$ with which μ_2 is coupled, is continuous.

Finally, (m_0, ∞) is the set of particles of μ, the mass of which is coupled with δ_0.

These thoughts also imply the splitting of $\bar{\nu}$, where we write

$$\bar{\nu} = \mu_1 + S^{\bar{\nu}-\mu_1}(\mu_2) + S^{\bar{\nu}-\left(\mu_1+S^{\bar{\nu}-\mu_1}(\mu_2)\right)}(\mu_3) =: \nu_1 + \nu_2 + \nu_3.$$

We already stated that μ_1 has to be coupled with ν_1 via an identity-mapping, i.e. $\mathbb{Q}_{lc}(\mu_1, \nu_1) = \mu_1 \otimes \delta_{\mathrm{id}}$. We have $\mu_1 = S^{\bar{\nu}}(\mu_1) = \nu_1$ and $\omega_0\delta_0 + \mu_1 = \mu'|_{(-\infty,m]}$, which implies

$$\omega_0\delta_0 + \nu_1 = S^{\nu'}(\omega_0\delta_0) + S^{\bar{\nu}}(\mu_1) = S^{\nu'}(\omega_0\delta_0) + S^{\nu'-S^{\nu'}(\omega_0\delta_0)}(\mu_1) = S^{\nu'}\left(\mu'|_{(-\infty,m]}\right)$$

Hence, we do not have any contradictions to Theorem 6.30. Observe that this also holds for $\mu'|_{(-\infty,x]}$ and $S^{\nu'}\left(\mu'|_{(-\infty,x]}\right)$ for all $x \in (-\infty, m)$ such that by Theorem 6.30, the resulting coupling is the correct one. By Corollary 6.29, we have

$$\mu'' := \mu_2 + \mu_3 \leq_c \nu_2 + \nu_3 =: \nu''.$$

Now let us come back to m_0. As μ_2 and ν_2 are by construction continuous, we may apply the theory of Henry-Labordère & Touzi [38]. Thus, m_0 has to satisfy their integral equation and we may choose

$$m_0 := \inf \left\{ x > m \;\middle|\; \int_{(0,m)} (g(x, \zeta) - \zeta)\delta F(\mathrm{d}\zeta) = \int_{(m,m_0)} (g(\xi, m) - \xi) F_\mu(\mathrm{d}\xi) \right\}.$$

Sometimes the easy bounds $m \le m_0 \le F_\mu^{-1}(1-\chi_0)$ may be useful.

Applying the theory presented in Section 6.2.1, we now determine the maps $T_d|_{(m,m_0)}$ and $T_u|_{(m,m_0)}$ as in Algorithm 6.13 and thus get $\mathbb{Q}_{lc}(\mu_2,\nu_2)$. Remark that, by our assumption that m is the unique maximum, the algorithm terminates after only one step and that it actually also covers the identity mapping between μ_1 and ν_1. However, by definition we have $v_2 = S^{\bar{\nu}-\mu_1}(\mu_2)$ such that we have $\mu_3 \le_c \nu_3$ by Corollary 6.29. As before, we understand that we have $\omega_0\delta_0 + \mu_1 + \mu_2 = \mu'|_{(-\infty,m_0]}$ and

$$\omega_0\delta_0 + \nu_1 + \nu_2 = S^{\nu'}(\omega_0\delta_0) + S^{\bar{\nu}}(\mu_1) + S^{\bar{\nu}-\mu_1}(\mu_2) = S^{\nu'}\left(\mu'|_{(-\infty,m_0]}\right)$$

and analogously that $\mu'|_{(-\infty,x]}$ is coupled with $S^{\nu'}\left(\mu'|_{(-\infty,x]}\right)$ for all $x \in (-\infty,m_0)$ by construction. Hence, there are again no contradictions to Theorem 6.30.

Finally, we use the techniques of Hobson & Norgilas [48] to determine $\mathbb{Q}_{lc}(\mu_3,\nu_3)$. Then the same arguments as before guarantee that indeed

$$\mathbb{Q}_{lc}(\mu',\nu') = \omega_0\delta_{(0,0)} + \mathbb{Q}_{lc}(\mu_1,\nu_1) + \mathbb{Q}_{lc}(\mu_2,\nu_2) + \mathbb{Q}_{lc}(\mu_3,\nu_3).$$

If we skip the assumption of a unique maximum of δF, then we have to proceed similarly to the general version of Algorithm 6.13. As in the simple case discussed above, we have to adapt the algorithm such that in each step we use the techniques of Hobson & Norgilas [48] in order to check for particles the mass of which is transported to δ_0. That is, in every step of the algorithm, three instead of two cases are distinguished. However, we do not discuss this case in detail. △

6.4.4. An associated super hedging strategy

Assumption 6.48. Let $\mu,\nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.32. Additionally assume that μ and ν satisfy the following conditions.

1. $x_1 < \ldots < x_N$, $y_1 < \ldots < y_M$.
2. $N, M \in \mathbb{N}$.

In this section, we complement Algorithm 6.40 in the case of finitely supported measures that satisfy the conditions of Assumption 6.48, in the sense that we provide a technique to determine a super hedging strategy associated to $\mathbb{Q}_{lc}(\mu,\nu)$.

For this purpose, we rewrite the discrete super hedging problem in (6.5) in the fashion of the general super hedging problem in (4.10). In particular, we change the notation from p_j, q_i and h_j to $\varphi(x_j), \psi(y_i)$ and $h(x_j)$. Though this makes the notation more complicated, it also yields a better intuition. The rewritten super hedging problem is

$$\begin{aligned} &\min \sum_{j=1}^{N} \omega_j\varphi(x_j) + \sum_{i=1}^{M} \vartheta_i\psi(y_i) \qquad (6.7)\\ &\text{s.t. } \varphi(x_j) + \psi(y_i) + h(x_j)(y_i - x_j) \ge c(x_j,y_i), \quad j=1,\ldots,N, i=1,\ldots,M. \end{aligned}$$

Let in the following $\Gamma \subseteq \{x_1, \dots, x_N\} \times \{y_1, \dots, y_M\}$ be the monotonicity set of $\mathbb{Q}_{lc}(\mu, \nu)$, where we assume that μ and ν satisfy Assumption 6.48. The basic idea is to use the characterization of a dual minimizer from Definition 5.18. That is, we search numbers $\varphi(x_1), \dots, \varphi(x_N)$, $\psi(y_1), \dots, \psi(y_M)$ and $h(x_1), \dots, h(x_N)$ such that for all $j = 1, \dots, N$ and all $i = 1, \dots, M$, we have

$$\varphi(x_j) + \psi(y_i) + h(x_j)(y_i - x_j) \geq c(x_j, y_i), \tag{6.8}$$

where equality holds for all $(x_j, y_i) \in \Gamma$, i.e. $\mathbb{Q}_{lc}(\mu, \nu)$-almost surely.

In the following, we denote $\Gamma_x \coloneqq \{y \in \mathbb{R} \mid (x, y) \in \Gamma\}$ and $\Gamma_y \coloneqq \{x \in \mathbb{R} \mid (x, y) \in \Gamma\}$ as well as $X_\Gamma \coloneqq \{x \in \mathbb{R} \mid \exists y \in \mathbb{R} : (x, y) \in \Gamma\}$ and $Y_\Gamma \coloneqq \{y \in \mathbb{R} \mid \exists x \in \mathbb{R} : (x, y) \in \Gamma\}$.

Using these sets and the previous thoughts, we now turn to the super hedging strategy. Clearly, an ideal super hedging strategy can easily be found solving the linear program in (6.7) using classic methods. However, we assume that the payoff function $c : \mathbb{R}^2 \to \mathbb{R}$ satisfies (6.6), the discrete analogue of the martingale Spence Mirrlees condition, and use that we know the structure of $\mathbb{Q}_{lc}(\mu, \nu)$. This leads to a more explicit approach.

Indeed, we provide an algorithm that yields an optimal super hedging strategy in two steps. In a first step, many of the numbers of interest are determined explicitly before ultimately a lower dimensional linear inequality system has to be solved.

This is somehow similar to the approach suggested in Guo & Obłój [32, Sec. 4.2]. Using the so-called discrete concave envelope, the authors rewrite the super hedging problem in (6.7) as a convex optimization problem for which various solving methods apply.

Herrmann & Stebegg [39, Sec. 5.1] use a somehow opposite approach. They first determine a dual optimizer solving a semi-infinite linear program and then derive the primal optimizer from that dual optimizer.

Though both articles consider finitely supported marginals, their results are more general, as the discrete version of the martingale Spence Mirrlees condition does not apply. As a consequence, the approaches are more technical and harder to understand and implement.

Let us now turn to our two step algorithm, where in the first step we choose numbers such that the $\mathbb{Q}_{lc}(\mu, \nu)$-almost sure equality in (6.8) holds, while in the second step we address the general inequality condition. For this purpose, we partition the set X_Γ as

$$X_\Gamma = X_\Gamma^\varphi \cup X_\Gamma^h \cup X_\Gamma^\psi,$$

where we use the disjoint subsets

$$\begin{aligned} X_\Gamma^\varphi &\coloneqq \{x \in X_\Gamma \mid |\Gamma_x| = 1\}, \\ X_\Gamma^h &\coloneqq \{x \in X_\Gamma \mid |\Gamma_x| = 2\}, \\ X_\Gamma^\psi &\coloneqq \{x \in X_\Gamma \mid |\Gamma_x| \geq 3\}. \end{aligned}$$

That is, we partition by the number of atoms of ν, an atom of μ is coupled with. The notations have a natural interpretation. In order to achieve equality in (6.8), for $x \in X_\Gamma^\varphi$, it suffices to choose $\varphi(x)$ suitably. For $x \in X_\Gamma^h$, additionally $h(x)$ has to be usefully adapted. Finally, for $x \in X_\Gamma^\psi$, also $\psi(y)$ has to be properly defined for at least one $y \in \Gamma_x$.

As in Algorithm 6.40, we go through the atoms of μ in a certain order. However, we start with the greatest atom, which exists by Assumption 6.48. Recall that Γ is the monotonicity set of $\mathbb{Q}_{lc}(\mu,\nu)$ and that $c : \mathbb{R}^2 \to \mathbb{R}$ satisfies (6.6). We define

$$D(x,y) \coloneqq \varphi(x) + \psi(y) + h(x)(y-x) - c(x,y).$$

Finally, we may now provide an algorithm to define numbers such that $D(x,y) = 0$ for all $(x,y) \in \Gamma$.

Algorithm 6.49. Let $\mu,\nu \in \mathcal{P}(\mathbb{R})$ satisfy Assumption 6.48 and $\Gamma \subseteq \mathbb{R}^2$ be the monotonicity set of $\mathbb{Q}_{lc}(\mu,\nu)$.

1. We begin considering x_N and distinguish three cases.

   1.1. $x_N \in X_\Gamma^\varphi$, i.e. $|\Gamma_{x_N}| = 1$. That is, $x_N \in \operatorname{supp}(\nu)$ and $\Gamma_{x_N} = \{x_N\}$. Define

   $$\varphi(x_N) \coloneqq c(x_N, x_N) - \psi(x_N).$$

   1.2. $x_N \in X_\Gamma^h$, i.e. $|\Gamma_{x_N}| = 2$ say $\Gamma_{x_N} = \{y_1^N, y_2^N\}$, where $y_1^N < x_N < y_2^N$. Define

   $$\begin{aligned} h(x_N) &\coloneqq \frac{c\left(x_N, y_2^N\right) - c\left(x_N, y_1^N\right) - \left(\psi\left(y_2^N\right) - \psi\left(y_1^N\right)\right)}{y_2^N - y_1^N}, \\ \varphi(x_N) &\coloneqq \frac{x_N - y_1^N}{y_2^N - y_1^N}\left(c\left(x_N, y_2^N\right) - \psi\left(y_2^N\right)\right) + \frac{y_2^N - x_N}{y_2^N - y_1^N}\left(c\left(x_N, y_1^N\right) - \psi\left(y_1^N\right)\right). \end{aligned}$$

   1.3. $x_N \in X_\Gamma^\psi$, i.e. $|\Gamma_{x_N}| \geq 3$, say $\Gamma_{x_N} = \{y_1^N, \ldots, y_{n_N}^N\}$ where $y_1^N < \ldots < y_{n_N}^N$. Define

   $$h(x_N) \coloneqq \frac{c\left(x_N, y_{n_N}^N\right) - c\left(x_N, y_1^N\right) - \left(\psi\left(y_{n_N}^N\right) - \psi\left(y_1^N\right)\right)}{y_{n_N}^N - y_1^N}.$$

   Now let $y \in \{y_2^N, \ldots, y_{n_N-1}^N\}$. Define

   $$\begin{aligned} \psi(y) \coloneqq c(x_N, y) &- \frac{y - y_1^N}{y_{n_N}^N - y_1^N}\left(c\left(x_N, y_{n_N}^N\right) - \psi\left(y_{n_N}^N\right)\right) \\ &- \frac{y_{n_N}^N - y}{y_{n_N}^N - y_1^N}\left(c\left(x_N, y_1^N\right) - \psi\left(y_1^N\right)\right). \end{aligned}$$

   Finally, define

   $$\begin{aligned} \varphi(x_N) \coloneqq &\frac{x_N - y_1^N}{y_{n_N}^N - y_1^N}\left(c\left(x_N, y_{n_N}^N\right) - \psi\left(y_{n_N}^N\right)\right) \\ &+ \frac{y_{n_N}^N - x_N}{y_{n_N}^N - y_1^N}\left(c\left(x_N, y_1^N\right) - \psi\left(y_1^N\right)\right). \end{aligned}$$

k+1. Assume that we have applied the previous step on x_N and the procedure of this step on $x_{N-1}, \ldots, x_{N-k+1}$. Then we have defined numbers such that $D(x,y) = 0$ for all $(x,y) \in \bigcup_{j=0}^{k-1} \left(\{x_{N-j}\} \times \Gamma_{x_{N-j}} \right)$. In this step, we proceed to achieve $D(x_{N-k}, y) = 0$ for all $y \in \Gamma_{x_{N-k}}$. Clearly, we may not change previously defined numbers. This is only relevant if $|\Gamma_{x_{N-k}}| \geq 3$. In order to improve readability, we write $K \coloneqq N - k$.

k+1.1. $x_K \in X_\Gamma^\varphi$, i.e. $|\Gamma_{x_K}| = 1$. That is, $x_K \in \operatorname{supp}(\nu)$ and $\Gamma_{x_K} = \{x_K\}$. Define

$$\varphi(x_K) \coloneqq c(x_K, x_K) - \psi(x_K).$$

k+1.2. $x_K \in X_\Gamma^h$, i.e. $|\Gamma_{x_K}| = 2$, say $\Gamma_{x_K} = \{y_1^K, y_2^K\}$, where $y_1^K < x_K < y_2^K$. Define

$$h(x_K) \coloneqq \frac{c\left(x_K, y_2^K\right) - c\left(x_K, y_1^K\right) - \left(\psi\left(y_2^K\right) - \psi\left(y_1^K\right)\right)}{y_2^K - y_1^K},$$
$$\varphi(x_K) \coloneqq \frac{x_K - y_1^K}{y_2^K - y_1^K}\left(c\left(x_K, y_2^K\right) - \psi\left(y_2^K\right)\right) + \frac{y_2^K - x_K}{y_2^K - y_1^K}\left(c\left(x_K, y_1^K\right) - \left(y_1^K\right)\right).$$

k+1.3. $x_K \in X_\Gamma^\psi$, i.e. $|\Gamma_{x_K}| \geq 3$, say $\Gamma_{x_K} = \{y_1^K, \ldots, y_{n_K}^K\}$, where $y_1^K < \ldots < y_{n_K}^K$. Define

$$h(x_K) \coloneqq \frac{c\left(x_K, y_{n_K}^K\right) - c\left(x_K, y_1^K\right) - \left(\psi\left(y_{n_K}^K\right) - \psi\left(y_1^K\right)\right)}{y_{n_K}^K - y_1^K}.$$

Now let $y \in \{y_2^K, \ldots, y_{n_K-1}^K\}$. Define

$$\psi(y) \coloneqq c(x_K, y) - \frac{y - y_1^K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_{n_K}^K\right) - \psi\left(y_{n_K}^K\right)\right)$$
$$- \frac{y_{n_K}^K - y}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_1^K\right) - \psi\left(y_1^K\right)\right).$$

Finally, define

$$\varphi(x_K) \coloneqq \frac{x_K - y_1^K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_{n_K}^K\right) - \psi\left(y_{n_K}^K\right)\right)$$
$$+ \frac{y_{n_K}^K - x_K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_1^K\right) - \psi\left(y_1^K\right)\right).$$

In step $k+1$, we define $\psi\left(y_2^K\right), \ldots, \psi\left(y_{n_K-1}^K\right)$, as it is clear that these numbers have not been defined in a previous step, while for $\psi\left(y_1^K\right)$ and $\psi\left(y_{n_K}^K\right)$ this can not be guaranteed.

Indeed, to get a contradiction, assume that there is a $j \in \{2, \ldots, n_K - 1\}$ such that $\psi\left(y_j^K\right)$ has been defined prior to step $k+1$. Then there is an $L = N - \ell > N - k = K$ such that $y_j^K \in \Gamma_{x_L}^\circ = \Gamma_{x_L} \setminus \left\{y_1^L, y_{n_L}^L\right\}$. But then, as $x_K < x_L$ and $y_1^K < y_j^K = y_i^L < y_{n_K}^K$ for some $i \in \{1, \ldots, n_L\}$, we have a contradiction to the left monotonicity.

Now let us prove that Algorithm 6.49 indeed yields $D(x,y) = 0$ for all $(x,y) \in \Gamma$, where in the proof we follow the steps and case distinctions of the algorithm.

Lemma 6.50. *Algorithm 6.49 yields* $D(x,y)=0$ *for all* $(x,y)\in\Gamma$.

Proof. 1.1. Implying $D(x_N,x_N)=0$, by definition we have

$$\varphi(x_N)=c(x_N,x_N)-\psi(x_N)=c(x_N,x_N)-\psi(x_N)-h(x_N)(x_N-x_N).$$

1.2. Implying $D\left(x_N,y_2^N\right)=D\left(x_N,y_1^N\right)$, by definition we have

$$h(x_N)=\frac{c\left(x_N,y_2^N\right)-c\left(x_N,y_1^N\right)-\left(\psi\left(y_2^N\right)-\psi\left(y_1^N\right)\right)}{y_2^N-y_1^N},$$

Implying $D\left(x_N,y_2^N\right)=0$, we also have

$$\begin{aligned}\varphi(x_N)&=\frac{x_N-y_1^N}{y_2^N-y_1^N}\left(c\left(x_N,y_2^N\right)-\psi\left(y_2^N\right)\right)+\frac{y_2^N-x_N}{y_2^N-y_1^N}\left(c\left(x_N,y_1^N\right)-\psi\left(y_1^N\right)\right)\\&=c\left(x_N,y_2^N\right)-\psi\left(y_2^N\right)\\&\quad-\frac{y_2^N-x_N}{y_2^N-y_1^N}\left(c\left(x_N,y_2^N\right)-c\left(x_N,y_1^N\right)-\left(\psi\left(y_2^N\right)-\psi\left(y_1^N\right)\right)\right)\\&=c\left(x_N,y_2^N\right)-\psi\left(y_2^N\right)-h(x_N)\left(y_2^N-x_N\right).\end{aligned}$$

1.3. Implying $D\left(x_N,y_1^N\right)=D\left(x_N,y_{n_N}^N\right)$, by definition we have

$$h(x_N)=\frac{c\left(x_N,y_{n_N}^N\right)-c\left(x_N,y_1^N\right)-\left(\psi\left(y_{n_N}^N\right)-\psi\left(y_1^N\right)\right)}{y_{n_N}^N-y_1^N}.$$

For $y\in\left\{y_2^N,\dots,y_{n_N-1}^N\right\}$, implying $D(x_N,y)=D\left(x_N,y_{n_N}^N\right)$, we have

$$\begin{aligned}\psi(y)&=c(x_N,y)-\frac{y-y_1^N}{y_{n_N}^N-y_1^N}\left(c\left(x_N,y_{n_N}^N\right)-\psi\left(y_{n_N}^N\right)\right)\\&\quad-\frac{y_{n_N}^N-y}{y_{n_N}^N-y_1^N}\left(c\left(x_N,y_1^N\right)-\psi\left(y_1^N\right)\right)\\&=c(x_N,y)-c\left(x_N,y_1^N\right)+\psi\left(y_1^N\right)-h(x_N)\left(y-y_1^N\right).\end{aligned}$$

Implying $D\left(x_N,y_{n_N}^N\right)=0$, we finally have

$$\begin{aligned}\varphi(x_N)&=\frac{x_N-y_1^N}{y_{n_N}^N-y_1^N}\left(c\left(x_N,y_{n_N}^N\right)-\psi\left(y_{n_N}^N\right)\right)+\frac{y_{n_N}^N-x_N}{y_{n_N}^N-y_1^N}\left(c\left(x_N,y_1^N\right)-\psi\left(y_1^N\right)\right)\\&=c\left(x_N,y_{n_N}^N\right)-\psi\left(y_{n_N}^N\right)\\&\quad-\frac{y_{n_N}^N-x_N}{y_{n_N}^N-y_1^N}\left(c\left(x_N,y_{n_N}^N\right)-c\left(x_N,y_1^N\right)-\left(\psi\left(y_{n_N}^N\right)-\psi\left(y_1^N\right)\right)\right)\\&=c\left(x_N,y_{n_N}^N\right)-\psi\left(y_{n_N}^N\right)-h(x_N)\left(y_{n_N}^N-x_N\right).\end{aligned}$$

k+1.1. Implying $D\left(x_K, x_K\right) = 0$, by definition we have

$$\varphi\left(x_K\right) = c\left(x_K, x_K\right) - \psi\left(x_K\right) = c\left(x_K, x_K\right) - \psi\left(x_K\right) - h\left(x_K\right)\left(x_K - x_K\right).$$

k+1.2. Implying $D\left(x_K, y_2^K\right) = D\left(x_K, y_1^K\right)$, by definition we have

$$h\left(x_K\right) = \frac{c\left(x_K, y_2^K\right) - c\left(x_K, y_1^K\right) - \left(\psi\left(y_2^K\right) - \psi\left(y_1^K\right)\right)}{y_2^K - y_1^K}.$$

Implying $D\left(x_K, y_2^K\right) = 0$, we also have

$$\begin{aligned}\varphi\left(x_K\right) &= \frac{x_K - y_1^K}{y_2^K - y_1^K}\left(c\left(x_K, y_2^K\right) - \psi\left(y_2^K\right)\right) + \frac{y_2^K - x_K}{y_2^K - y_1^K}\left(c\left(x_K, y_1^K\right) - \left(y_1^K\right)\right)\\ &= c\left(x_K, y_2^K\right) - \psi\left(y_2^K\right)\\ &\quad - \frac{y_2^K - x_K}{y_2^K - y_1^K}\left(c\left(x_K, y_2^K\right) - c\left(x_K, y_1^K\right) - \left(\psi\left(y_2^K\right) - \psi\left(y_1^K\right)\right)\right)\\ &= c\left(x_K, y_2^K\right) - \psi\left(y_2^K\right) - h\left(x_K\right)\left(y_2^K - x_K\right).\end{aligned}$$

k+1.3. Implying $D\left(x_K, y_1^K\right) = D\left(x_K, y_{n_K}^K\right)$, by definition we have

$$h\left(x_K\right) = \frac{c\left(x_K, y_{n_K}^K\right) - c\left(x_K, y_1^K\right) - \left(\psi\left(y_{n_K}^K\right) - \psi\left(y_1^K\right)\right)}{y_{n_K}^K - y_1^K}.$$

For $y \in \left\{y_2^K, \ldots, y_{n_K-1}^K\right\}$, implying $D\left(x_K, y\right) = D\left(x_K, y_{n_K}^K\right)$, we have

$$\begin{aligned}\psi\left(y\right) &= c\left(x_K, y\right) - \frac{y - y_1^K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_{n_K}^K\right) - \psi\left(y_{n_K}^K\right)\right)\\ &\quad - \frac{y_{n_K}^K - y}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_1^K\right) - \psi\left(y_1^K\right)\right)\\ &= c\left(x_K, y\right) - c\left(x_K, y_1^K\right) + \psi\left(y_1^K\right) - h\left(x_K\right)\left(y - y_1^K\right).\end{aligned}$$

Implying $D\left(x_K, y_{n_K}^K\right) = 0$, we finally have

$$\begin{aligned}\varphi\left(x_K\right) &= \frac{x_K - y_1^K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_{n_K}^K\right) - \psi\left(y_{n_K}^K\right)\right) + \frac{y_{n_K}^K - x_K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_1^K\right) - \psi\left(y_1^K\right)\right)\\ &= c\left(x_K, y_{n_K}^K\right) - \psi\left(y_{n_K}^K\right)\\ &\quad - \frac{y_{n_K}^K - x_K}{y_{n_K}^K - y_1^K}\left(c\left(x_K, y_{n_K}^K\right) - c\left(x_K, y_1^K\right) - \left(\psi\left(y_{n_K}^K\right) - \psi\left(y_1^K\right)\right)\right)\\ &= c\left(x_K, y_{n_K}^K\right) - \psi\left(y_{n_K}^K\right) - h\left(x_K\right)\left(y_{n_K}^K - x_K\right).\end{aligned}$$

$\square$

Now we proceed with the second part of our algorithm in order to achieve the inequality $D(x, y) \geq 0$ for all $(x, y) \in \text{supp}(\mu) \times \text{supp}(\nu)$. For this purpose, recall that we have already defined $\varphi(x)$ for all $x \in \text{supp}(\mu)$, $h(x)$ for all $x \in X_\Gamma^h \cup X_\Gamma^\psi$ and $\psi(y)$ for all $y \in \text{supp}(\nu)$ such that there is an $x \in X_\Gamma^\psi$ with $y \in \Gamma_x^\circ$. Conversely, we have to define $h(x)$ for all $x \in X_\Gamma^\varphi$ and $\psi(y)$ for all $y \in Y_\Gamma^\psi$, where

$$Y_\Gamma^\psi := \{z \in \text{supp}(\nu) \mid \nexists x \in X_\Gamma^\psi \text{ such that } \exists z^-, z^+ \in \Gamma_x \text{ with } z^- < z < z^+\}.$$

Now we iteratively plugin all numbers $\varphi(x)$, $\psi(y)$ and $h(x)$ into (6.8) that are defined by Algorithm 6.49. Then the resulting inequality system is formulated depending on $x \in \text{supp}(\mu)$, $y \in \text{supp}(\nu)$ and $c(x, y)$ for $x \in \text{supp}(\mu)$ and $y \in \text{supp}(\nu)$, which are constants, and on $h(x)$ for $x \in X_\Gamma^\varphi$ and $\psi(y)$ for $y \in Y_\Gamma^\psi$, which are the variables of the lower dimensional inequality system. In particular, the number of variables is reduced from $2N + M$ to at most $N + M$, as at least $\varphi(x)$ is defined by Algorithm 6.49 for all $x \in \text{supp}(\mu) = \{x_1, \ldots, x_N\}$. Also, the number of (in-)equalities is reduced from NM to at most $N(M-1)$, as Γ contains at least N elements for which equality is already guaranteed. Solving this lower dimensional system then yields a dual minimizer satisfying (6.8) and thus the conditions of Definition 5.18.

The existence of a solution to the simplified linear inequality system is implied by the existence of a dual minimizer, i.e. a solution to the linear inequality system in (6.8).

Indeed, having any dual minimizer (φ^*, ψ^*, h^*), we may adapt it according to Algorithm 6.49 without causing any inequalities to be unsatisfied. This can be seen by a recursive case distinction as indicated in the following. Remark that we do not argue completely formal, as the necessary calculations, though mathematically straightforward, are rather lengthy.

Starting from the generic inequality

$$\varphi(x) + \psi(y) + h(x)(y - x) \geq c(x, y)$$

for some $(x, y) \in \text{supp}(\mu) \times \text{supp}(\nu)$, we distinguish whether or not $\varphi(x), h(x)$ and $\psi(y)$ are specified by Algorithm 6.49. Then we either plugin the algorithmically specified numbers or those from the dual minimizer. After a first step the inequality does only depend on $\psi(\cdot), \psi^*(\cdot)$ and $c(\cdot, \cdot)$. Then, recursively plugging in the definition from Algorithm 6.49 if $\psi(\cdot)$ is specified there and $\psi^*(\cdot)$ if it is not, we arrive at an inequality of the form

$$\sum_{i=1}^{M} \left[\lambda_i \psi^*(y_i) + \sum_{j=1}^{N} \kappa_{j,i} c(x_j, y_i) \right] \geq 0 \tag{6.9}$$

for some $\lambda_i, \kappa_{j,i} \in \mathbb{R}$. Then we show that the left hand side of (6.9) may be written as

$$\sum_{j=1}^{N} \sum_{i=1}^{M} \lambda_{j,i}^* \left[\varphi^*(x_j) + \psi^*(y_i) + h^*(x_j)(y_i - x_j) - c(x_j, y_i) \right],$$

where $\lambda_{j,j}^* \geq 0$ whenever $\varphi^*(x_j) + \psi^*(y_i) + h^*(x_j)(y_i - x_j) > c(x_j, y_i)$. This clearly implies that (6.9) is satisfied and thus shows the claim.

6.4.5. An exemplary construction

Example 6.51. Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$ be two discrete measures with three atoms each, say

$$\mu = \omega_d \delta_{x_d} + \omega_m \delta_{x_m} + \omega_u \delta_{x_u} \quad \text{and} \quad \nu = \vartheta_d \delta_{y_d} + \vartheta_m \delta_{y_m} + \vartheta_u \delta_{y_u}.$$

We consider a combination of binary type options on two time points of an asset price process $(S_t)_{t=0,\dots,2}$ with payoff function

$$c(x,y) = \mathbb{1}_{\{(x,y)>(\overline{S}_1,\overline{S}_2)\}} + \mathbb{1}_{\{x>\overline{S}_1\}} + \mathbb{1}_{\{y>\overline{S}_2\}} - \mathbb{1}_{\{x<\underline{S}_1\}} - \mathbb{1}_{\{y<\underline{S}_2\}} - \mathbb{1}_{\{(x,y)<(\underline{S}_1,\underline{S}_2)\}},$$

where we choose $\underline{S}_1 = 100 = \underline{S}_2$, $\overline{S}_1 = 110$, and $\overline{S}_2 = 120$.

Before we illustrate this function in Figure 6.4, let us first prove that it is indeed of the kind such that the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu,\nu)$ is optimal for the discrete upper price bound problem in (6.4). By Theorem 6.34, we have to show that the condition in (6.6) is satisfied, where we denote the left hand side by $D(x', x, y^+, y', y^-)$. By the structure of the measures we have to consider three cases.

1. Let $x_u = x' > x = x_m$ and $y^+ = y_u > y_m = y' > y^- = y_d$. Then we have

$$\begin{aligned} D(x', x, y^+, y', y^-) &= \frac{1}{y_u - y_m}\left(\frac{3-1}{x_u - x_m} - \frac{1-0}{x_u - x_m}\right) \\ &\quad - \frac{1}{y_m - y_d}\left(\frac{1-0}{x_u - x_m} - \frac{0-(-1)}{x_u - x_m}\right) \\ &= \frac{1}{(y_u - y_m)(x_u - x_m)} > 0. \end{aligned}$$

2. Let $x_u = x' > x = x_d$ and $y^+ = y_u > y_m = y' > y^- = y_d$. Then we have

$$\begin{aligned} D(x', x, y^+, y', y^-) &= \frac{1}{y_u - y_m}\left(\frac{3-0}{x_u - x_d} - \frac{1-(-1)}{x_u - x_d}\right) \\ &\quad - \frac{1}{y_m - y_d}\left(\frac{1-(-1)}{x_u - x_d} - \frac{0-(-3)}{x_u - x_d}\right) \\ &= \frac{1}{y_u - y_m}\frac{1}{x_u - x_d} - \frac{1}{y_m - y_d}\frac{1}{x_u - x_d} \\ &= \frac{1}{x_u - x_d}\frac{-y_u + 2y_m - y_d}{(y_u - y_m)(y_m - y_d)} \overset{!}{>} 0. \end{aligned}$$

   This inequality is satisfied if and only if $y_m > \frac{y_u + y_d}{2}$.

3. Let $x_m = x' > x = x_d$ and $y^+ = y_u > y_m = y' > y^- = y_d$. Then we have

$$\begin{aligned} D(x', x, y^+, y', y^-) &= \frac{1}{y_u - y_m}\left(\frac{1-0}{x_m - x_d} - \frac{0-(-1)}{x_m - x_d}\right) \\ &\quad - \frac{1}{y_m - y_d}\left(\frac{0-(-1)}{x_m - x_d} - \frac{(-1)-(-3)}{x_m - x_d}\right) \\ &= \frac{1}{(y_m - y_d)(x_m - x_d)} > 0. \end{aligned}$$

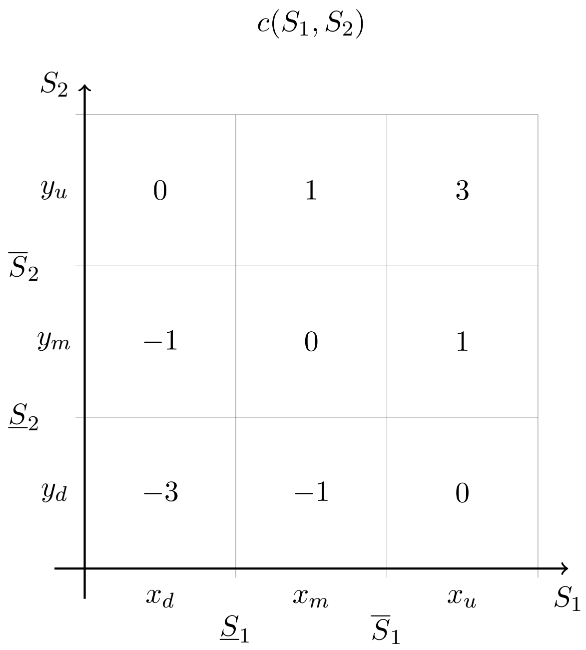


Figure 6.4.: The binary type payoff function c.

In total, we see that the atoms of ν should satisfy $y_m > \frac{y_u+y_d}{2}$. This is ensured by the explicit choice of the measures.

Let us think of a two-year horizon and choose $S_0 = 100$ for the expected value of the measures μ and ν. Also, choose $x_u = S_1^u = 120$, $x_m = S_1^m = 105$ and $x_d = S_1^d = 80$. Now the measure μ has to be chosen such that two conditions are satisfied. First, we need that the asset price is a martingale, which is satisfied if

$$120\omega_u + 105\omega_m + 80\omega_d = 100.$$

Also, we want μ to be a probability measure, i.e.

$$\omega_u + \omega_m + \omega_d = 1.$$

Furthermore, let us assume that the asset has an annualized volatility of 20 percent such that we get the condition

$$\left(\frac{120}{100} - 1\right)^2 \omega_u + \left(\frac{105}{100} - 1\right)^2 \omega_m + \left(\frac{80}{100} - 1\right)^2 \omega_d = 0.2^2. \tag{6.10}$$

The unique solution to this linear equality system is

$$(\omega_u, \omega_m, \omega_d) = \left(\frac{17}{49}, \frac{12}{49}, \frac{20}{49}\right).$$

In a second step let us choose ν properly. We proceed in the same way as for μ and assume $y_u = S_2^u = 135$, $y_m = S_2^m = 110$ and $y_d = S_2^d = 65$ such that indeed $y_m > \frac{y_u+y_d}{2}$. Let again the weights be such that the annualized volatility is 20 percent. This is, differently to before, that equation (6.10) has the right hand side $\left(\frac{1}{\sqrt{2}}0.2\right)^2$. We get the unique solution

$$(\vartheta_u, \vartheta_m, \vartheta_d) = \left(\frac{9}{35}, \frac{17}{45}, \frac{23}{63}\right).$$

Formally, we have to show that $\mu \leq_c \nu$. As the specifications of the measures do not contradict this property immediately, we prove the convex order by constructing the left monotone martingale transport plan.

Now let us determine the left monotone martingale transport plan, the upper price bound, the right monotone martingale transport plan and the lower price bound for the binary type option. Also, let us determine associated super and sub hedging strategies.

We begin with the calculation of $\mathbb{Q}_{lc}(\mu, \nu)$.

1. The measures to be considered in the first step are

$$\mu = \frac{20}{49}\delta_{80} + \frac{12}{49}\delta_{105} + \frac{17}{49}\delta_{120} \leq_c \frac{23}{63}\delta_{65} + \frac{17}{45}\delta_{110} + \frac{9}{35}\delta_{135} = \nu.$$

   We algorithmically couple the mass of δ_{80} with ν. We have $q_{x_d y_d} + q_{x_d y_m} = \frac{20}{49}$ and

$$65 q_{x_d y_d} + 110 q_{x_d y_m} = \frac{20}{49} \cdot 80.$$

   This implies

$$q_{x_d y_d} = \frac{40}{147} \quad \text{and} \quad q_{x_d y_m} = \frac{20}{147},$$

   which is admissible, as $q_{x_d y_d} = \frac{40}{147} < \frac{23}{63} = \vartheta_d$ and $q_{x_d y_m} = \frac{20}{147} < \frac{17}{45} = \vartheta_m$.

2. The residual measures are

$$\mu' = \frac{12}{49}\delta_{105} + \frac{17}{49}\delta_{120} \leq_c \frac{41}{441}\delta_{65} + \frac{533}{2205}\delta_{110} + \frac{9}{35}\delta_{135} = \nu'.$$

   We algorithmically couple the mass of δ_{105}. Here, we have $q_{x_m y_d} + q_{x_m y_m} = \frac{12}{49}$ and

$$65 q_{x_m y_d} + 110 q_{x_m y_m} = \frac{12}{49} \cdot 105.$$

   This implies

$$q_{x_m y_d} = \frac{4}{147} \quad \text{and} \quad q_{x_m y_m} = \frac{32}{147},$$

   which is again admissible, as $q_{x_m y_d} = \frac{4}{147} < \frac{41}{441} = \vartheta'_d$ and $q_{x_m y_m} = \frac{32}{147} < \frac{553}{2205} = \vartheta'_m$.

3. The residual measures are $\mu'' = \frac{17}{49}\delta_{120} \leq_c \frac{29}{441}\delta_{65} + \frac{53}{2205}\delta_{110} + \frac{9}{35}\delta_{135} = \nu''$. A simple test shows that an admissible coupling is possible. That is, $q_{x_u y_d} = \frac{29}{441}$, $q_{x_u y_m} = \frac{53}{2205}$ and $q_{x_u y_u} = \frac{9}{35}$.

An illustration of the above application of Algorithm 6.40 is given in Example A.2. We determined $\mathbb{Q}_{lc}(\mu, \nu)$ and thus we know that $\mu \leq_c \nu$, as otherwise no martingale transport plan would exist. Let us calculate the associated upper price bound. We have

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right] =& \frac{9}{35}\cdot 3+\frac{53}{2205}\cdot 1+\frac{29}{441}\cdot 0+\frac{32}{147}\cdot 0\\ &+\frac{4}{147}\cdot(-1)+\frac{20}{147}\cdot(-1)+\frac{40}{147}\cdot(-3)=-\frac{58}{315}.\end{aligned}$$

Similarly, we determine the right monotone martingale transport plan $\mathbb{Q}_{rc}(\mu,\nu)$ in order to compare the upper and lower price bounds.

1. The measures to be considered in the first step are

$$\mu=\frac{20}{49}\delta_{80}+\frac{12}{49}\delta_{105}+\frac{17}{49}\delta_{120}\leq_c\frac{23}{63}\delta_{65}+\frac{17}{45}\delta_{110}+\frac{9}{35}\delta_{135}=\nu.$$

Here, we have to first couple the mass of δ_{120} with ν. We have $q_{x_uy_m}+q_{x_uy_u}=\frac{17}{49}$ and

$$110q_{x_uy_m}+135q_{x_uy_u}=\frac{17}{49}\cdot 120.$$

This implies

$$q_{x_uy_m}=\frac{51}{245}\quad\text{and}\quad q_{x_uy_u}=\frac{34}{245},$$

which is admissible, as $q_{x_uy_m}=\frac{51}{245}<\frac{17}{45}=\vartheta_m$ and $q_{x_uy_u}=\frac{34}{245}<\frac{9}{35}=\vartheta_u$.

2. The residual measures are

$$\mu'=\frac{20}{49}\delta_{80}+\frac{12}{49}\delta_{105}\leq_c\frac{23}{63}\delta_{65}+\frac{374}{2205}\delta_{110}+\frac{29}{245}\delta_{135}=\nu'.$$

We couple the mass of δ_{105}. Here, we have $q_{x_my_d}+q_{x_my_m}=\frac{12}{49}$ and

$$65q_{x_my_d}+110q_{x_my_m}=\frac{12}{49}\cdot 105.$$

This implies

$$q_{x_my_d}=\frac{4}{147}\quad\text{and}\quad q_{x_my_m}=\frac{32}{147},$$

which is not admissible, as $q_{x_my_d}=\frac{4}{147}<\frac{23}{63}=\vartheta'_d$, but $q_{x_my_m}=\frac{32}{147}>\frac{374}{2205}=\vartheta'_m$. Thus, we choose

$$q_{x_my_m}=\frac{374}{2205}$$

and get two new equations for the desired weights, namely

$$q_{x_my_d}+q_{x_my_u}=\frac{12}{49}-\frac{374}{2205}=\frac{166}{2205}$$

and

$$65q_{x_my_d}+135q_{x_my_u}=\frac{12}{49}\cdot 105-\frac{374}{2205}\cdot 110=\frac{3112}{441}.$$

This implies

$$q_{x_my_d}=\frac{137}{3087}\quad\text{and}\quad q_{x_my_u}=\frac{53}{1715},$$

which then is admissible, as $q_{x_my_d}=\frac{137}{3087}<\frac{23}{63}=\vartheta'_d$ and $q_{x_my_u}=\frac{53}{1715}<\frac{29}{245}=\vartheta'_u$.

3. The residual measures are $\mu'' = \frac{20}{49}\delta_{80} \leq_c \frac{110}{343}\delta_{65} + \frac{30}{343}\delta_{135} = \nu''$. Again, an admissible coupling exists and it is clear how to find it.

Thus, we determined $\mathbb{Q}_{rc}(\mu, \nu)$. The lower price bound then is

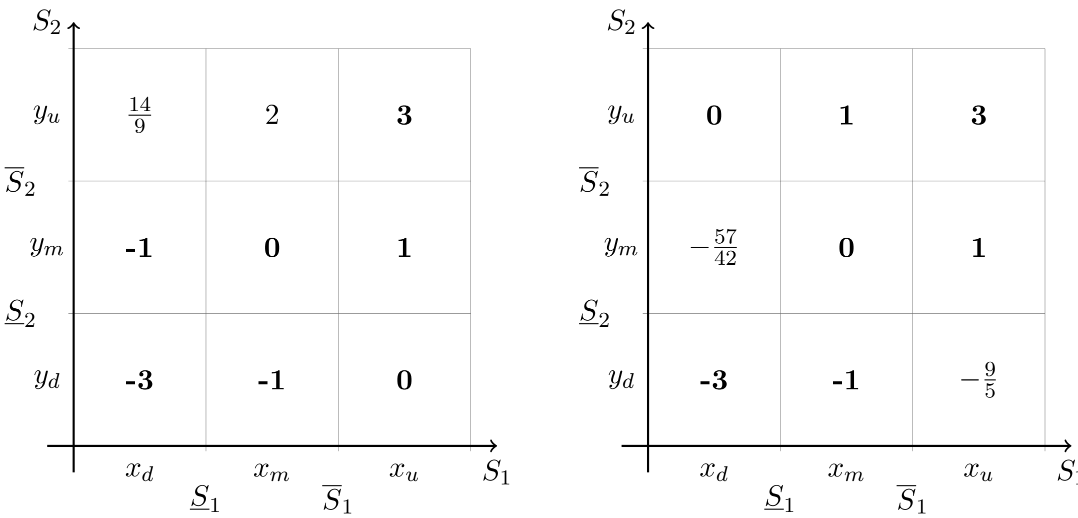
$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_{rc}(\mu,\nu)}\left[c(X,Y)\right] =& \frac{34}{245}\cdot 3 + \frac{51}{245}\cdot 1 + \frac{53}{1715}\cdot 1 \\ &+ \frac{374}{2205}\cdot 0 + \frac{30}{343}\cdot 0 + \frac{137}{3087}\cdot(-1) + \frac{110}{343}\cdot(-3) = -\frac{5419}{15435}.\end{aligned}$$

Let us now consider the hedging strategies. The super hedging strategy may be found using Algorithm 6.49 and solving the residual inequalities. The application of the two step algorithm is detailed in Example A.2. The resulting super hedging strategy is given by

$$\varphi(S_1) = \begin{cases} -\frac{85}{42}, & S_1 = 80 \\ \frac{5}{7}, & S_1 = 105 \\ \frac{33}{14}, & S_1 = 120, \end{cases} \quad \psi(S_2) = \begin{cases} 0, & S_2 = 65 \\ -\frac{13}{14}, & S_2 = 110 \\ 0, & S_2 = 135, \end{cases} \quad h(S_1) = \begin{cases} \frac{41}{630}, & S_1 = 80 \\ \frac{3}{70}, & S_1 = 105 \\ \frac{3}{70}, & S_1 = 120. \end{cases}$$

A sub hedging strategy is given by

$$\varphi(S_1) = \begin{cases} -\frac{103}{42}, & S_1 = 80 \\ -\frac{1}{9}, & S_1 = 105 \\ \frac{73}{45}, & S_1 = 120, \end{cases} \quad \psi(S_2) = \begin{cases} 0, & S_2 = 65 \\ 0, & S_2 = 110 \\ \frac{4}{9}, & S_2 = 135, \end{cases} \quad h(S_1) = \begin{cases} \frac{23}{630}, & S_1 = 80 \\ \frac{1}{45}, & S_1 = 105 \\ \frac{14}{225}, & S_1 = 120. \end{cases}$$



(a) $\mathbb{Q}_{lc}(\mu, \nu)$-almost sure equality (bold). (b) $\mathbb{Q}_{rc}(\mu, \nu)$-almost sure equality (bold).

Figure 6.5.: Super hedging strategy (left) and sub hedging strategy (right)

The two hedging strategies are displayed in Figures 6.5a and 6.5b. △

CHAPTER 7

PRICE BOUND APPROXIMATION

In Chapter 4, we discussed the connection between observable call option prices and the marginals of any potential pricing measure. In Chapters 5 and 6, we used this connection in the sense that we considered the price bound problems with respect to sets of martingale transport plans, for which we implicitly assume that the measures that serve as marginals are uniquely defined by observable call option prices.

In the standard market case restricted to $\mathbb{R}_+$ that we consider in this chapter, this formally means that we have two uniquely defined tuples (C_μ, μ) and (C_ν, ν), where $C_\mu, C_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ are price functions of call options on the asset prices X and Y at the two trading times t and T, and μ and ν are the associated marginals of the underlying price process $S = (X, Y)$ at the same trading times. The pairs satisfy

$$F_\mu(k) = 1 + C'_\mu(k+) \quad \text{and} \quad F_\nu(k) = 1 + C'_\nu(k+), \quad k \in \mathbb{R}_+.$$

In this setting, μ and ν are probability measures with common finite expected value. Hence, by Proposition 4.24 we have

$$C_\mu \leq C_\nu \iff \mu \leq_c \nu. \tag{7.1}$$

However, the tuples consisting of a call option price function and its associated marginal are uniquely defined if and only if call option prices are observable for strike prices that form a dense subset of $\mathbb{R}_+$. Otherwise, there are infinitely many consistent call option price functions and associated marginals. When it comes to application, this is the case as only finitely many different call option prices may be observed. Thus, it is natural to consider convergence issues, which shall be done in this chapter.

In detail, assume that there are sequences of measures $(\mu_n)_{n\in\mathbb{N}}$ and $(\nu_n)_{n\in\mathbb{N}}$ that approximate the measures μ and ν in some sense. Then we investigate whether the associated upper price bounds with respect to μ_n and ν_n approximate the upper price bound with respect to μ and ν as well, i.e. whether

$$\left| \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n,\nu_n)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)] - \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)] \right| \overset{n\to\infty}{\to} 0. \tag{7.2}$$

In order to further motivate the interest in convergence issues, recall the situation of a payoff function that satisfies the martingale Spence Mirrlees condition. In this case, the left monotone martingale transport plan is optimal for the upper price bound problem in (4.7) and there are techniques to determine it. However, for continuous marginals the determination is difficult, while for discrete marginals it is very simple. Thus, we prefer to approximate the actual upper price bound, given the convergence in (7.2) is valid.

In Section 7.1, we discuss the application-oriented point of view. We consider theoretical and empirical tuples of price functions and marginals, explain how we may use observable call option prices and define the notion of consistency rigorously. In Sections 7.2 and 7.3, we prove the convergence claimed in (7.2) and quantify the convergence speed for certain approximating sequences of marginals, assuming that the payoff function satisfies the martingale Spence Mirrlees condition and that the marginals are compactly supported. In Section 7.4, we generalize the results of Sections 7.2 and 7.3, allowing for more general payoff functions and arbitrary approximating sequences of marginals. Finally, we generalize the results allowing for marginals with unbounded support.

7.1. Basic considerations

7.1.1. Theoretical price functions & marginals

Let (C_μ, μ) and (C_ν, ν) be the tuples of the unknown theoretical call option price functions and the associated unknown theoretical marginals of the underlying S with respect to any potential pricing measure at two arbitrary trading times $0 < t < T$. In the following sections, we mostly use them as a reference. Let us additionally assume that the price of the underlying at time $t = 0$ is $s_0 = 1$.

We recall the no-arbitrage considerations of Section 3.2 and apply them to the tuples (C_μ, μ) and (C_ν, ν). As we assume the underlying to have prices in $\mathbb{R}_+$, the measures μ and ν are concentrated on $\mathbb{R}_+$. In particular, recall the notion of a candidate function.

Definition 7.1. A function $C : \mathbb{R}_+ \to \mathbb{R}_+$ is a *candidate function for call option prices*, if it satisfies the following conditions.

1. C is monotone non-increasing and convex.

2. $\lim_{k\to\infty} C(k) = 0$, $C'(0+) \geq -1$ and $C(0) = s_0 = 1$.

We denote the set of all candidate function by $\mathcal{K}^C$.

By the usual no-arbitrage considerations, we necessarily have that the theoretical call option price functions are candidate functions, i.e. $C_\mu, C_\nu \in \mathcal{K}^C$.

As the discounted underlying is a martingale with respect to every potential pricing measure, we require the theoretical marginals to satisfy $\mu \leq_c \nu$. Recalling equation (7.1), we know that then the theoretical call option price functions satisfy $C_\mu \leq C_\nu$.

These intrinsic properties may be used in plausibility checks for observed call option prices. Clearly, the observed prices must not contradict the existence of theoretical price functions with the above properties, as the observed prices are evaluations of theoretical call option price functions.

7.1.2. Empirical price functions & marginals

In the following, we additionally assume a compactness property for the theoretical call option price functions C_μ and C_ν that shall be satisfied in Sections 7.1.3, 7.2, 7.3 and 7.4.1.

Assumption 7.2. There are numbers $0 < K^* \leq L^* < \infty$ such that $C_\mu(k) = 0$ for all $k \geq K^*$ and $C_\nu(\ell) = 0$ for all $\ell \geq L^*$. We denote the set of all candidate functions that meet this assumption by $\mathcal{K}^C_{K^*}$ and $\mathcal{K}^C_{L^*}$ respectively.

Remark 7.3. Assumption 7.2 is not very strong when it comes to applications, as the call option prices decrease to be numerically negligible when the strike prices increase. The associated measures μ and ν have compact support under this assumption, i.e. we have $\operatorname{supp}(\mu) \subseteq [0, K^*]$ and $\operatorname{supp}(\nu) \subseteq [0, L^*]$. Indeed, for general $\rho \in \mathcal{P}(\mathbb{R}_+)$, we have

$$0 = C_\rho(x) = \int_{\mathbb{R}+} (s-x)^+ \rho(\mathrm{d}s) \iff \rho((x,\infty)) = 0. \qquad \Diamond$$

Now let us choose finitely many strike prices in the intervals $[0, K^*]$ and $[0, L^*]$ and use the associated observable call option prices for the approximation. In order to guarantee that the approximating marginals are in convex order and to reduce numerical complexity, we choose the strike prices suitably. Therefore, we define

$$m := \max\left\{ n \in \mathbb{N}_0 \;\middle|\; \frac{L^*}{K^*} \geq 2^n \right\},$$

and then $K := \frac{L^*}{2^m} \geq K^*$ and $L := L^*$. For every $n \in \mathbb{N}$, we choose the strike prices using equidistant partitions of the intervals $[0, K]$ and $[0, L]$ defined by

$$Z_n^\mu := \left\{ k_j^n := \frac{j}{2^n} K \;\middle|\; j = 0, \ldots, 2^n \right\} \quad \text{and} \quad Z_n^\nu := \left\{ \ell_i^n := \frac{i}{2^n} L \;\middle|\; i = 0, \ldots, 2^n \right\}.$$

Then the sets of associated call option prices are

$$C_n^\mu := \{C_\mu(k) \mid k \in Z_n^\mu\} = \left\{ C_\mu(0), C_\mu\left(\frac{K}{2^n}\right), \ldots, C_\mu\left(K - \frac{K}{2^n}\right), C_\mu(K) \right\},$$
$$C_n^\nu := \{C_\nu(\ell) \mid \ell \in Z_n^\nu\} = \left\{ C_\nu(0), C_\nu\left(\frac{L}{2^n}\right), \ldots, C_\nu\left(L - \frac{L}{2^n}\right), C_\nu(L) \right\}.$$

Now let us use the fact that the prices in C_n^μ and C_n^ν have to be evaluations of some candidate functions $C_\mu \in \mathcal{K}_K^C$ and $C_\nu \in \mathcal{K}_L^C$ respectively and the properties of such functions in order to state conditions that these prices necessarily have to satisfy. The conditions serve as a plausibility check when it comes to application.

Lemma 7.4. *The prices in C_n^μ and C_n^ν satisfy the following conditions for all $n \in \mathbb{N}$.*

1. *The sequences of price differences*

$$(C_\mu(k_j^n) - C_\mu(k_{j-1}^n))_{j=1,\dots,2^n} \quad \text{and} \quad \left(C_\nu\left(\ell_i^n\right) - C_\nu\left(\ell_{i-1}^n\right)\right)_{i=1,\dots,2^n}$$

are non-positive and monotone non-decreasing.

2. *We have $C_\mu(k_0^n) = C_\nu(\ell_0^n) = 1$.*
3. *We have $C_\mu(k_1^n) \geq 1 - \frac{K}{2^n}$ and $C_\nu(\ell_1^n) \geq 1 - \frac{L}{2^n}$.*
4. *We have $C_\mu(K) = C_\nu(L) = 0$.*

Proof. We only consider C_μ, as C_ν is treated analogously.

1. We use the first part of Definition 3.5. The candidate function C_μ is monotone non-increasing such that for all $k_{j-1}^n \leq k_j^n$, we have $C_\mu(k_{j-1}^n) \geq C_\mu(k_j^n)$, which yields the non-positivity. Further, C_μ is convex such that for all $k_{j-1}^n \leq k_j^n \leq k_{j+1}^n$, we have that the slopes of the secants are monotone non-decreasing, i.e.

$$\frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} - \frac{C_\mu(k_j^n) - C_\mu(k_{j-1}^n)}{k_j^n - k_{j+1}^n} \geq 0,$$

which is the claim by $k_{j+1}^n - k_j^n = k_j^n - k_{j-1}^n$.

2. We have $k_0^n = \ell_0^n = 0$ and $s_0 = 1$ such that the claim is implied by the third property from the second part of Definition 3.5.

3. By the convexity of the candidate function and the second property of the second part of Definition 3.5, we have

$$\frac{C_\mu(k_1^n) - C_\mu(k_0^n)}{k_1^n - k_0^n} \geq C'_\mu(k_0^n+) = C'_\mu(0+) \geq -1 \iff C_\mu(k_1^n) \geq 1 - \frac{K}{2^n}.$$

4. This is immediately implied by Assumption 7.2. It is sufficient for the first property of the second part of Definition 3.5. □

Now let us state conditions that the prices have to satisfy in order not to contradict the order conditions $C_\mu \leq C_\nu$ and thus $\mu \leq_c \nu$.

Lemma 7.5. *For all $n \in \mathbb{N}$, the prices in C_n^μ and C_n^ν satisfy*

$$C_\mu(2^m k_i^n) \leq C_\nu(\ell_i^n), \quad i = 0, \dots, 2^n.$$

Proof. As $K = \frac{L}{2^m}$, we have

$$2^m k_i^n = 2^m \frac{i}{2^n} K = \frac{i}{2^n} L = \ell_i^n.$$

If now $C_\mu(2^m k_i^n) > C_\nu(\ell_i^n)$, then this is a contradiction to $C_\mu \leq C_\nu$ and thus to $\mu \leq_c \nu$. □

If the conditions of the Lemmata 7.4 and 7.5 are satisfied, then we may define suitable price functions and measures for the approximation of (C_μ, μ) and (C_ν, ν) based on C_n^μ and C_n^ν. For the approximation of C_μ and C_ν, and of μ and ν, only such functions and measures should be considered that are consistent with C_n^μ and C_n^ν. Here, consistency is to be understood in the following sense.

Definition 7.6. A candidate function $C \in \mathcal{K}^C$ is called *consistent* with the prices of C_n^μ and C_n^ν, $n \in \mathbb{N}$, if respectively

$$C(k_j^n) = C_\mu(k_j^n), \quad j = 0, \ldots, 2^n,$$
$$C(\ell_i^n) = C_\nu(\ell_i^n), \quad i = 0, \ldots, 2^n.$$

The set of such candidate functions is denoted by $\mathcal{C}_n^\mu$ and $\mathcal{C}_n^\nu$, $n \in \mathbb{N}$, respectively.

Definition 7.7. A probability measure $\rho \in \mathcal{P}(\mathbb{R}_+)$ is called *consistent* with the prices of C_n^μ and C_n^ν, $n \in \mathbb{N}$, if respectively

$$\int_{\mathbb{R}_+} (x - k_j^n)^+ \rho(\mathrm{d}x) = C_\mu(k_j^n), \quad j = 0, \ldots, 2^n,$$
$$\int_{\mathbb{R}_+} (y - \ell_i^n)^+ \rho(\mathrm{d}y) = C_\nu(\ell_i^n), \quad i = 0, \ldots, 2^n.$$

The set of all such probability measures is denoted by $\mathcal{P}_n^\mu$ and $\mathcal{P}_n^\nu$, $n \in \mathbb{N}$, respectively.

Thus, we have defined the sets of all price functions and all measures that are possibly of interest. The consistency immediately implies the compactness of the support of any approximating measures $\mu_n \in \mathcal{P}_n^\mu$ and $\nu_n \in \mathcal{P}_n^\nu$.

Observe that the sets do in general not consist of only one element, as the observed call option prices do not determine the price function and the associated measures uniquely. The notation already suggests that the sets are in a one-to-one connection and this is indeed the case.

Lemma 4.14 guarantees that call option price functions and marginals are in a one-to-one connection. Thus, we only have to justify that the consistency is preserved. For this purpose, start with a consistent measure $\rho \in \mathcal{P}(\mathbb{R}_+)$ and observe that

$$C_\rho(x) = \int_{\mathbb{R}_+} (s - x)^+ \rho(\mathrm{d}s)$$

immediately yields the consistency of the price function $C_\rho \in \mathcal{K}^C$. If we otherwise start with a inconsistent measure ρ, then the same equation implies that C_ρ is inconsistent as well.

7.1.3. Towards price bound approximation

Recall that the actual aim of this chapter is to approximate the upper price bound $P_2^c(\mu,\nu)$ by a sequence of upper price bounds $(P_2^c(\mu_n,\nu_n))_{n\in\mathbb{N}}$. Thus, we may not consider arbitrary pairs of consistent call option price functions (C_{μ_n}, C_{ν_n}) and marginals (μ_n,ν_n). Indeed, in order to let $P_2^c(\mu_n,\nu_n)$ be meaningful, we have to consider pairs such that $C_{\mu_n} \leq C_{\nu_n}$ and $\mu_n \leq_c \nu_n$. Therefore, we define

$$\begin{aligned} \mathcal{C}_n^{\leq} &:= \left\{(C_{\mu_n}, C_{\nu_n}) \in \mathcal{C}_n^{\mu} \times \mathcal{C}_n^{\nu} \mid C_{\mu_n} \leq C_{\nu_n}\right\}, \\ \mathcal{P}_n^{\leq_c} &:= \left\{(\mu_n,\nu_n) \in \mathcal{P}_n^{\mu} \times \mathcal{P}_n^{\nu} \mid \mu_n \leq_c \nu_n\right\}. \end{aligned}$$

As $C_\mu \in \mathcal{C}_n^\mu$, $C_\nu \in \mathcal{C}_n^\nu$, and $C_\mu \leq C_\nu$, we have $(C_\mu, C_\nu) \in \mathcal{C}_n^{\leq}$ for all $n \in \mathbb{N}$. Analogously, as $\mu \in \mathcal{P}_n^\mu$, $\nu \in \mathcal{P}_n^\nu$, and $\mu \leq_c \nu$, we have $(\mu,\nu) \in \mathcal{P}_n^{\leq_c}$ for all $n \in \mathbb{N}$. That is, the theoretical price functions and marginals guarantee the non-emptiness of the sets of relevant pairs of approximating price functions and marginals respectively. Furthermore, by (7.1) and the same arguments as at the end of the previous section, $\mathcal{C}_n^{\leq}$ and $\mathcal{P}_n^{\leq_c}$ are isomorphic. In the following, we rather consider the measure theoretic point of view.

As $\mathcal{P}_n^{\leq_c} \neq \emptyset$, we may formulate the two well-defined price bound problems

$$\sup_{(\mu_n,\nu_n)\in\mathcal{P}_n^{\leq_c}} \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n,\nu_n)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)], \quad \text{and} \quad \inf_{(\mu_n,\nu_n)\in\mathcal{P}_n^{\leq_c}} \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n,\nu_n)} \mathbb{E}_{\mathbb{Q}}[c(X,Y)],$$

which yield the maximally and minimally possible values for the actual upper price bound $P_2^c(\mu,\nu)$ that are consistent with the observable call option prices in C_n^μ and C_n^ν. We observe that by the choice of $(Z_n^\mu)_{n\in\mathbb{N}}$ and $(Z_n^\nu)_{n\in\mathbb{N}}$, we have $\mathcal{P}_n^{\leq_c} \supseteq \mathcal{P}_{n+1}^{\leq_c}$, $n \in \mathbb{N}$. As

$$\inf_{(\mu_n,\nu_n)\in\mathcal{P}_n^{\leq_c}} P_2^c(\mu_n,\nu_n) \leq P_2^c(\mu,\nu) \leq \sup_{(\mu_n,\nu_n)\in\mathcal{P}_n^{\leq_c}} P_2^c(\mu_n,\nu_n),$$

this subset property suggests that for $n \to \infty$, we have

$$\inf_{(\mu_n,\nu_n)\in\mathcal{P}_n^{\leq_c}} P_2^c(\mu_n,\nu_n) \nearrow P_2^c(\mu,\nu) \swarrow \sup_{(\mu_n,\nu_n)\in\mathcal{P}_n^{\leq_c}} P_2^c(\mu_n,\nu_n).$$

That is, the largest and the smallest possible upper price bound considering marginals from $\mathcal{P}_n^{\leq_c}$ converge from above and from below to the actual upper price bound. This would in particular imply the convergence in (7.2) for any approximating sequence of price bounds.

7.2. Convergence in the martingale Spence Mirrlees case

In this section, we prove that the convergence in (7.2) holds true, at least in the case that the theoretical marginals are compactly supported and when the payoff function satisfies the martingale Spence Mirrlees condition. In the following section, for all $n \in \mathbb{N}$, we define a pair $(\mu_n^d, \nu_n^d) \in \mathcal{P}_n^{\leq_c}$ based on the observable call option prices C_n^μ and C_n^ν in order to investigate the convergence speed of the price bounds both empirically and theoretically.

In order to investigate the convergence itself, let in this section $0 < R_1 \leq R_2 < \infty$. Then we define the set of all pairs of probability measures with finite first moments and suitable compact support that are in convex order,

$$\mathcal{P}^{\leq_c}_{R_1,R_2} := \{(\rho_1, \rho_2) \in \mathcal{P}(\mathbb{R}_+) \times \mathcal{P}(\mathbb{R}_+) \mid \operatorname{supp}(\rho_1) \subseteq [0, R_1], \operatorname{supp}(\rho_2) \subseteq [0, R_2], \rho_1 \leq_c \rho_2\} .$$

By Assumption 7.2 and Remark 7.3, we have $(\mu, \nu) \in \mathcal{P}^{\leq_c}_n \subseteq \mathcal{P}^{\leq_c}_{K^*,L^*} \subseteq \mathcal{P}^{\leq_c}_{K,L}$ for the theoretical marginals μ and ν and all $n \in \mathbb{N}$.

Theorem 7.8. *Let the payoff function $c : \mathbb{R}^2_+ \to \mathbb{R}$ satisfy the martingale Spence Mirrlees condition and let $0 < R_1 \leq R_2 < \infty$. Then the mapping*

$$P_2^{c,MSM} : \begin{cases} \mathcal{P}^{\leq_c}_{R_1,R_2} \to \mathbb{R} \\ (\rho_1, \rho_2) \mapsto \mathbb{E}_{\mathbb{Q}_{lc}(\rho_1,\rho_2)}[c(X, Y)] \end{cases}$$

is continuous with respect to the topology $\mathcal{T}_{cb}(\mathbb{R}_+)^2$ as well as the topology $\mathcal{T}_1(\mathbb{R}_+)^2$.

In order to prove the theorem, we need a continuity result from the work of Juillet [51].

Theorem 7.9 ([51, Theorem 2.16])**.** *The mapping*

$$\text{Curt} : \begin{cases} \{(\rho_1, \rho_2) \in \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R}) \mid \rho_1 \leq_c \rho_2\} \to \mathcal{P}(\mathbb{R}^2) \\ (\rho_1, \rho_2) \mapsto \mathbb{Q}_{lc}(\rho_1, \rho_2) \end{cases}$$

is continuous with respect to the topologies $\mathcal{T}_{cb}(\mathbb{R})^2$ and $\mathcal{T}_{cb}(\mathbb{R}^2)$ as well as the topologies $\mathcal{T}_1(\mathbb{R})^2$ and $\mathcal{T}_1(\mathbb{R}^2)$.

Proof of Theorem 7.8. In order to use Theorem 7.9, we rewrite the mapping $P_2^{c,MSM}$ as a concatenation of a restriction of Curt to compactly supported measures and some additional mapping. Indeed, we may write

$$P_2^{c,MSM} = I \circ \text{Curt}|_{\mathcal{P}^{\leq_c}_{R_1,R_2}},$$

where

$$I : \begin{cases} \mathcal{P}([0, R_1] \times [0, R_2]) \to \mathbb{R} \\ \mathbb{Q} \mapsto \mathbb{E}_{\mathbb{Q}}[c(X, Y)] . \end{cases}$$

The continuity of $\text{Curt}|_{\mathcal{P}^{\leq_c}_{R_1,R_2}}$ is implied by Theorem 7.9. The continuity of I holds directly by the definition of weak convergence, since the payoff function c is in particular continuous and thus bounded on the compact set $[0, R_1] \times [0, R_2]$.

It remains to prove that $\text{Curt}(\mathcal{P}^{\leq_c}_{R_1,R_2}) \subseteq \mathcal{P}([0, R_1] \times [0, R_2])$. This is clear, as arbitrary couplings of measures are concentrated on some subset of the cartesian product of the supports of their respective marginals. □

7.3. Convergence speed in the martingale Spence Mirrlees case

In this section, we introduce a specially designed sequence $(\mu_n^d, \nu_n^d)_{n\in\mathbb{N}}$ of pairs of marginals such that $(\mu_n^d, \nu_n^d) \in \mathcal{P}_n^{\leq_c} \subseteq \mathcal{P}_{K,L}^{\leq_c}$ for all $n \in \mathbb{N}$, which converges to the theoretical marginals $(\mu, \nu) \in \mathcal{P}_{K,L}^{\leq_c}$ with respect to the Wasserstein distance. Then we investigate the convergence speed of the price bound difference

$$\left| \mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d,\nu_n^d)} \left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)} \left[c(X,Y)\right] \right|,$$

the convergence of which is implied by Theorem 7.8, as we assume the payoff function $c : \mathbb{R}_+^2 \to \mathbb{R}$ to satisfy the martingale Spence Mirrlees condition, both empirically and theoretically.

7.3.1. Specially designed marginals

For any $n \in \mathbb{N}$, the pair of marginals (μ_n^d, ν_n^d) is defined using a suitable pair of consistent call option price functions. We choose the consistent call option price functions $C_{\mu_n^d} \in \mathcal{C}_n^\mu$ and $C_{\nu_n^d} \in \mathcal{C}_n^\nu$ to be exactly the functions that result from interpolating between the observed call option prices in C_n^μ and C_n^ν. That is, for $k \in [k_j^n, k_{j+1}^n), j = 0, \ldots, 2^n - 1$, we define

$$C_{\mu_n^d}(k) := \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} C_\mu(k_j^n) + \frac{k - k_j^n}{k_{j+1}^n - k_j^n} C_\mu(k_{j+1}^n),$$

and for $\ell \in [\ell_i^n, \ell_{i+1}^n), i = 0, \ldots, 2^n - 1$, we define

$$C_{\nu_n^d}(\ell) := \frac{\ell_{i+1}^n - \ell}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_i^n) + \frac{\ell - \ell_i^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_{i+1}^n).$$

Then the associated measures μ_n^d and ν_n^d, $n \in \mathbb{N}$, are discrete measures of the form

$$\begin{aligned} \mu_n^d := \sum_{j=0}^{2^n} \omega_j^n \delta_{k_j^n} &:= \sum_{j=0}^{2^n} \left[\frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} - \frac{C_\mu(k_j^n) - C_\mu(k_{j-1}^n)}{k_j^n - k_{j-1}^n} \right] \delta_{k_j^n} \\ &= \frac{2^n}{K} \sum_{j=0}^{2^n} \left[C_\mu(k_{j+1}^n) - 2C_\mu(k_j^n) + C_\mu(k_{j-1}^n) \right] \delta_{k_j^n}, \end{aligned}$$

where we set $\frac{C_\mu(k_{2^n+1}^n) - C_\mu(k_{2^n}^n)}{k_{2^n+1}^n - k_{2^n}^n} = 0$ and $\frac{C_\mu(k_0^n) - C_\mu(k_{-1}^n)}{k_0^n - k_{-1}^n} = -1$, and

$$\begin{aligned} \nu_n^d := \sum_{i=0}^{2^n} \vartheta_i^n \delta_{\ell_i^n} &:= \sum_{i=0}^{2^n} \left[\frac{C_\nu(\ell_{i+1}^n) - C_\nu(\ell_i^n)}{\ell_{i+1}^n - \ell_i^n} - \frac{C_\nu(\ell_i^n) - C_\nu(\ell_{i-1}^n)}{\ell_i^n - \ell_{i-1}^n} \right] \delta_{\ell_i^n} \\ &= \frac{2^n}{L} \sum_{i=0}^{2^n} \left[C_\nu(\ell_{i+1}^n) - 2C_\nu(\ell_i^n) + C_\nu(\ell_{i-1}^n) \right] \delta_{\ell_i^n}, \end{aligned}$$

where we set $\frac{C_\nu(\ell_{2^n+1}^n) - C_\nu(\ell_{2^n}^n)}{\ell_{2^n+1}^n - \ell_{2^n}^n} = 0$ and $\frac{C_\nu(\ell_0^n) - C_\nu(\ell_{-1}^n)}{\ell_0^n - \ell_{-1}^n} = -1$.

The distribution functions of the measures satisfy

$$F_{\mu_n^d}(x) = \sum_{j=0}^{2^n} \left(1 + \frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n}\right) \mathbb{1}_{[k_j^n, k_{j+1}^n)}(x), \tag{7.3}$$

$$F_{\nu_n^d}(y) = \sum_{i=0}^{2^n} \left(1 + \frac{C_\nu(\ell_{i+1}^n) - C_\nu(\ell_i^n)}{\ell_{i+1}^n - \ell_i^n}\right) \mathbb{1}_{[\ell_i^n, \ell_{i+1}^n)}(y).$$

In Remark A.3, we show that $C_{\mu_n^d}$ and $C_{\nu_n^d}$ are recovered as the call option price functions corresponding to μ_n^d and ν_n^d in the sense of Definition 4.22. We also show that μ_n^d and ν_n^d are probability measures with expected value equal to $s_0 = 1$.

Thus, we may deduce the convex order $\mu_n^d \leq_c \nu_n^d$ showing that $C_{\mu_n^d} \leq C_{\mu_n^d}$. For this, let $x \in [k_j^n, k_{j+1}^n] \subset [\ell_i^n, \ell_{i+1}^n]$ and $\lambda_j^n, \gamma_i^n \in [0,1]$ be such that $x = \lambda_j^n k_{j+1}^n + (1 - \lambda_j^n) k_j^n$ and $x = \gamma_i^n \ell_{i+1}^n + (1 - \gamma_i^n)\ell_i^n$. Then we have

$$\begin{aligned} C_{\mu_n^d}(x) &= \lambda_j^n C_\mu(k_{j+1}^n) + (1 - \lambda_j^n) C_\mu(k_j^n) \\ &\leq \lambda_j^n C_\nu(k_{j+1}^n) + (1 - \lambda_j^n) C_\nu(k_j^n) \\ &\leq \lambda_j^n \left(\frac{k_{j+1}^n - \ell_i^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_{i+1}) + \frac{\ell_{i+1}^n - k_{j+1}^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_i) \right) \\ &\quad + (1 - \lambda_j^n) \left(\frac{k_j^n - \ell_i^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_{i+1}) + \frac{\ell_{i+1}^n - k_j^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_i) \right). \end{aligned}$$

In the first inequality, we use the fact that the theoretical price functions satisfy $C_\mu \leq C_\nu$. In the second inequality, we use the convexity of C_ν. If we now plugin the definition of λ_j^n, then we get

$$\begin{aligned} C_{\mu_n^d}(x) &\leq \frac{x - k_j^n}{k_{j+1}^n - k_j^n} \left(\frac{k_{j+1}^n - \ell_i^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_{i+1}) + \frac{\ell_{i+1}^n - k_{j+1}^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_i) \right) \\ &\quad + \frac{k_{j+1}^n - x}{k_{j+1}^n - k_j^n} \left(\frac{k_j^n - \ell_i^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_{i+1}) + \frac{\ell_{i+1}^n - k_j^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_i) \right) \\ &\leq \frac{x - \ell_i^n}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_{i+1}^n) + \frac{\ell_{i+1}^n - x}{\ell_{i+1}^n - \ell_i^n} C_\nu(\ell_i^n) \\ &= \gamma_i^n C_\nu(\ell_{i+1}^n) + (1 - \gamma_i^n) C_\nu(\ell_i^n) = C_{\nu_n^d}(x), \end{aligned}$$

where the second inequality follows from an easy calculation. Thus, $C_{\mu_n^d} \leq C_{\nu_n^d}$.

Recall that for payoff functions $c : \mathbb{R}_+^2 \to \mathbb{R}$ that satisfy the martingale Spence Mirrlees condition, by Theorem 7.8 we have

$$\mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d, \nu_n^d)}[c(X,Y)] \overset{n \to \infty}{\to} \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}[c(X,Y)].$$

By Algorithm 6.40, $\mathbb{Q}_{lc}(\mu_n^d, \nu_n^d)$ is easy to determine as μ_n^d and ν_n^d are discrete measures. Thus, we easily get the sequence of upper price bounds that approximates the actual upper price bound. In the following section, we illustrate this convergence for different theoretical marginals and payoff functions.

7.3.2. EMPIRICAL CONSIDERATIONS

Example 7.10. Let us now discuss the convergence speed of the price bound approximation for different compactly supported theoretical marginals and payoff functions that satisfy the martingale Spence Mirrlees condition empirically. Therefore, we calculate the approximating upper price bounds for several $n \in \mathbb{N}$ as well as the actual upper price bounds as far as possible. Additionally, we calculate the corresponding normalized price bound differences $d_n \coloneqq 2^{2n}(P_2^c(\mu,\nu) - P_2^c(\mu_n^d,\nu_n^d))$.

We calculate the approximating marginals slightly different from the theoretical definition in the previous section. In fact, we always choose $K = L$, i.e. we potentially lose some information available for the time t marginal μ. Also, we are not always able to calculate the actual upper price bound correctly. Indeed, we can only do so in the case of uniform marginals as discussed in Example 6.16. For all other distributions considered here, we estimate the actual upper price bound based on the convergence results so far.

We begin with uniform marginal distributions and use the results about the left monotone martingale transport plans from Example 6.16.

1. Let $\mu \sim \mathcal{U}[1,3], \nu \sim \mathcal{U}[0,4]$. Here, we partition the support of ν in maximally 2048 intervals, i.e. we have a difference of $\frac{1}{512}$ between two partition points.
   a) $c(x,y) = xy^2$. The exact upper price bound is

$$P_2^c(\mu,\nu) = \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] = \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[XY^2\right]$$
$$= \mathbb{E}_\mu\left[X\left(\frac{3}{4}\left(\frac{3}{2}X+\frac{1}{2}\right)^2+\frac{1}{4}\left(\frac{3}{2}-\frac{1}{2}X\right)^2\right)\right] = 12.5.$$

   For all $n = 3,\ldots,11$, we calculate the upper price bound $P_2^c(\mu_n^d,\nu_n^d)$ and d_n. This yields the results of Table 7.1.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d,\nu_n^d)$	12.808	12.57	12.517	12.504	12.501	12.5002	12.50006	12.50002	12.500004
d_n	19.7	17.9	17.1	16.6	16.4	16.35	16.3	16.274	16.274

Table 7.1.: Approximation results in the case 1.a)

   b) $c(x,y) = -\frac{1}{3}(y-x)^3$. The exact upper price bound is $P_2^c(\mu,\nu) = 0.5$. The upper price bounds $P_2^c(\mu_n^d,\nu_n^d)$ are given in Table 7.2.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d,\nu_n^d)$	0.493	0.499	0.5001	0.50004	0.50001	0.500004	0.500001	0.5000002	0.50000006
d_n	-0.44	-0.003	0.141	0.200	0.226	0.238	0.244	0.247	0.249

Table 7.2.: Approximation results in the case 1.b)

   c) $c(x,y) = \exp(x)\cdot y^2$. The exact upper price bound is $P_2^c(\mu,\nu) = 61.8801$. The upper price bounds $P_2^c(\mu_n^d,\nu_n^d)$ are given in Table 7.3.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d, \nu_n^d)$	65.8620	62.7911	62.0990	61.9338	61.8934	61.8834	61.8810	61.8803	61.8802
d_n	254.8	233.2	224.1	219.8	217.7	216.7	216.1	215.8	215.5

Table 7.3.: Approximation results in the case 1.c)

2. Let $\mu \sim \mathcal{U}[9,11], \nu \sim \mathcal{U}[0,20]$. Here, we partition the support of ν in maximally 2048 intervals, i.e. we have a difference of $\frac{5}{512}$ between two partition points.

   a) $c(x,y) = xy^2$. The exact upper price bound is

$$P_2^c(\mu,\nu) = \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] = \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[XY^2\right]$$
$$= \mathbb{E}_{\mu}\left[X\left(\frac{11}{20}\left(\frac{11}{2}X - \frac{81}{2}\right)^2 + \frac{9}{20}\left(\frac{99}{2} - \frac{9}{2}X\right)^2\right)\right] = 1356.5.$$

   For all $n = 3,\dots,11$, we calculate the upper price bound $P_2^c(\mu_n^d, \nu_n^d)$ and d_n. This yields the results of Table 7.4.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d, \nu_n^d)$	1421	1367.2	1359.35	1357.206	1356.676	1356.543	1356.511	1356.503	1356.501
d_n	4133	2739	2922	2894	2882	2805	2826	2808	2819

Table 7.4.: Approximation results in the case 2.a)

   b) $c(x,y) = -\frac{1}{3}(y-x)^3$. The exact upper price bound is $P_2^c(\mu,\nu) = 16.5$. The upper price bounds $P_2^c(\mu_n^d, \nu_n^d)$ are given in Table 7.5.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d, \nu_n^d)$	33.319	19.052	17.311	16.695	16.551	16.512	16.503	16.501	16.500
d_n	1076	653	830	800	831	799	821	802	813

Table 7.5.: Approximation results in the case 2.b)

   c) $c(x,y) = \exp(x)\cdot y^2$. The exact upper price bound is $P_2^c(\mu,\nu) = 4041627.609$. The upper price bounds $P_2^c(\mu_n^d, \nu_n^d)$ are given in Table 7.6.

n	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d, \nu_n^d)$	4826637	4236165	4093466	4054268	4044652	4042391	4041818	4041675
d_n	200962386	199206022	212331405	207103361	198219006	200101331	199976026	200114438

Table 7.6.: Approximation results in the case 2.c)

Now let us consider more complicated marginals. We choose triangular distributions, which still have a simple structure but already leave us no possibility to calculate the actual upper price bound explicitly. Triangular distributions are slightly more natural to serve as marginals for an asset price process than uniform distributions.

3. Let $\mu \sim \Delta[1,2,3], \nu \sim \Delta[0,2,4]$, i.e. the density functions of μ and ν are given by

$$f_\mu(x) = \begin{cases} 0, & x \leq 1 \\ x-1, & 1 < x \leq 2 \\ 3-x, & 2 < x \leq 3 \\ 0, & 3 < x \end{cases} \quad \text{and} \quad f_\nu(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq 2 \\ 4-x, & 2 < x \leq 4 \\ 0, & 4 < x. \end{cases}$$

Here, we partition the support of ν in maximally 2048 intervals, i.e. we have a difference of $\frac{1}{512}$ between two partition points.

a) $c(x,y) = xy^2$. For all $n = 3,\dots,11$, we calculate the upper price bound $P_2^c(\mu_n^d, \nu_n^d)$. Assuming the standardization is correct, solving the equation

$$2^{2n}\left(P_2^c\left(\mu_n^d,\nu_n^d\right) - P_2^c(\mu,\nu)\right) = 2^{2(n+1)}\left(P_2^c\left(\mu_{n+1}^d,\nu_{n+1}^d\right) - P_2^c(\mu,\nu)\right)$$

for $n = 10$, we receive a very good approximation to the exact price bound by $P_2^c(\mu,\nu) \approx 10.21832381$, which we use in the calculation of d_n. This yields the results of Table 7.7.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d,\nu_n^d)$	10.546	10.289	10.235	10.222	10.219	10.2186	10.21834	10.21834	10.21833
d_n	20.98	18.13	17.03	16.54	16.24	16.10	16.01	15.99	15.98

Table 7.7.: Approximation results in the case 3.a)

b) $c(x,y) = -\frac{1}{3}(y-x)^3$. The upper price bounds $P_2^c(\mu_n^d,\nu_n^d)$ are given in Table 7.8, where we use $P_2^c(\mu,\nu) \approx 0.218323819$ in the calculation of d_n.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d,\nu_n^d)$	0.22	0.218	0.2183	0.21833	0.21832	0.218323	0.2183236	0.2183238	0.21832381
d_n	-0.08	-0.1	-0.028	-0.028	-0.018	-0.0247	-0.0548	-0.0554	-0.0554

Table 7.8.: Approximation results in the case 3.b)

c) $c(x,y) = \exp(x) \cdot y^2$. The upper price bounds $P_2^c(\mu_n^d,\nu_n^d)$ are given in Table 7.9, where we use $P_2^c(\mu,\nu) \approx 44.79714628$ in the calculation of d_n.

n	3	4	5	6	7	8	9	10	11
$P_2^c(\mu_n^d,\nu_n^d)$	48.423	45.573	44.979	44.841	44.808	44.800	44.798	44.798	44.797
d_n	232.03	198.45	186.00	180.60	177.34	175.92	174.89	174.61	174.57

Table 7.9.: Approximation results in the case 3.c)

The above considerations allude a convergence speed of 2^{2n}. However, we can only investigate the case of uniform distributions exactly, as this is the only continuous distribution such that the mappings T_d and T_u can be calculated explicitly using the methods of Henry-Labordère & Touzi [38]. We shall later see that the general convergence speed is indeed worse than this suspicion. △

7.3.3. Theoretical considerations

Theorem 7.11. *Let* $(\mu, \nu) \in \mathcal{P}^{\leq_c}_{K,L}$. *Let* $c : \mathbb{R}^2_+ \to \mathbb{R}$ *be a Lipschitz continuous payoff function such that* c_{yy} *exists and* c *satisfies the martingale Spence Mirrlees condition. We denote by* $\hat{\Lambda}$ *the Lipschitz constant of* c *and assume* $\max\{\hat{\Lambda}, \sup_{(x,y)\in\mathbb{R}^2_+} |c_{yy}(x,y)|\} \leq \Lambda$. *Then, for any* $n \in \mathbb{N}$, *we have*

$$\left|\mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d,\nu_n^d)}\left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right]\right| \leq \frac{M_c}{2^n},$$

where $M_c := (7K + 5L) \cdot \tilde{\Lambda}$ *with* $\tilde{\Lambda} := \Lambda \cdot \max\{L, 1\}$. *If we additionally suppose that* $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$, *then, for any* $n \in \mathbb{N}$, *we have*

$$\left|\mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d,\nu_n^d)}\left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right]\right| \leq \frac{M_d}{2^{n+1}},$$

where $M_d := (7T_\mu K^2 + 5T_\nu L^2) \cdot \tilde{\Lambda}$ *with* $T_\mu := \sup_{\kappa\in[0,K]} |C''_\mu(\kappa)|$ *and* $T_\nu := \sup_{\lambda\in[0,L]} |C''_\nu(\lambda)|$.

Remark 7.12. Without assuming that $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$, we also have

$$\begin{aligned}&\left|\mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d,\nu_n^d)}\left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right]\right| \\ &\quad \leq 2\tilde{\Lambda}\left[7\sum_{j=0}^{2^n-1}\left\|\left(C_{\mu_n^d} - C_\mu\right)\Big|_{[k_j^n,k_{j+1}^n)}\right\|_\infty + 5\sum_{i=0}^{2^n-1}\left\|\left(C_{\nu_n^d} - C_\nu\right)\Big|_{[\ell_i^n,\ell_{i+1}^n)}\right\|_\infty\right], \quad n \in \mathbb{N}.\end{aligned}$$

This estimate is sharper than the ones in Theorem 7.11, but it directly depends on the theoretical marginals μ and ν. ◊

In the proof of Theorem 7.11, we need some of the results presented in Section 5.2.1. Thus, recall Definition 5.18, Theorem 5.21 and Remark 5.22. We also need to estimate the Wasserstein distances of μ_n^d and μ, and of ν_n^d and ν. Therefore, recall the definition of the Wasserstein distance in (2.1) and Lemma 2.3.

Theorem 7.13. *Let* $(\mu, \nu) \in \mathcal{P}^{\leq_c}_{K,L}$. *Then, for any* $n \in \mathbb{N}$, *we have*

$$W\left(\mu, \mu_n^d\right) = 2 \cdot \sum_{j=0}^{2^n-1}\left\|\left(C_{\mu_d^n} - C_\mu\right)\Big|_{[k_j^n,k_{j+1}^n)}\right\|_\infty \leq \frac{K}{2^n}, \tag{7.4}$$

$$W\left(\nu, \nu_n^d\right) = 2 \cdot \sum_{i=0}^{2^n-1}\left\|\left(C_{\nu_d^n} - C_\nu\right)\Big|_{[\ell_i^n,\ell_{i+1}^n)}\right\|_\infty \leq \frac{L}{2^n}. \tag{7.5}$$

If we additionally suppose that $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$, *then, for any* $n \in \mathbb{N}$, *we have*

$$W\left(\mu, \mu_n^d\right) \leq \frac{T_\mu \cdot K^2}{2^{n+1}}, \tag{7.6}$$

$$W\left(\nu, \nu_n^d\right) \leq \frac{T_\nu \cdot L^2}{2^{n+1}}. \tag{7.7}$$

Proof. We only consider $W\left(\mu, \mu_n^d\right)$, as the calculation of $W\left(\nu, \nu_n^d\right)$ is exactly the same. By Lemma 2.3, we have

$$W\left(\mu, \mu_n^d\right) = \int_{-\infty}^{\infty} \left| F_\mu(t) - F_{\mu_n^d}(t) \right| \mathrm{d}t.$$

In order to calculate the integral, we plugin the distribution function representations using the call option price function C_μ. In particular, we use Lemma 4.14 for F_μ and equation (7.3) for $F_{\mu_n^d}$. Then we have

$$W\left(\mu, \mu_n^d\right) = \int_0^K \left| 1 + C'_\mu(t+) - F_{\mu_n^d}(t) \right| \mathrm{d}t = \sum_{j=0}^{2^n-1} \int_{k_j^n}^{k_{j+1}^n} \left| C'_\mu(t+) - \frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} \right| \mathrm{d}t.$$

In the following, let us write $m_j^n := \frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n}$. For all $j = 0, \ldots, 2^n - 1$ there is a $k(j,n) \in [k_j^n, k_{j+1}^n)$ such that for all $t \in [k_j^n, k(j,n))$, we have $F_\mu(t) \le F_{\mu_n^d}(t)$, or equivalently $C'_\mu(t+) \le m_j^n$, and for all $t \in [k(j,n), k_{j+1}^n)$, we have $F_\mu(t) \ge F_{\mu_n^d}(t)$, or equivalently $C'_\mu(t+) \ge m_j^n$. Thus, we have

$$W\left(\mu, \mu_n^d\right) = \sum_{j=0}^{2^n-1} \left[\int_{k_j^n}^{k(j,n)} \left(m_j^n - C'_\mu(t+) \right) \mathrm{d}t + \int_{k(j,n)}^{k_{j+1}^n} \left(C'_\mu(t+) - m_j^n \right) \mathrm{d}t \right]. \tag{7.8}$$

Calculating the integrals leads to the exact representation in (7.4). We stress that the set of points $t \in \mathbb{R}_+$ such that $C'_\mu(t-) \neq C'_\mu(t+)$ is a Lebesgue null set. Hence, integrating over the right derivative $C'_\mu(\cdot+)$, we receive $C_\mu(\cdot)$. Based on (7.8), we thus obtain

$$\begin{aligned}
W\left(\mu, \mu_n^d\right) &= \sum_{j=0}^{2^n-1} \Big[m_j^n \left(k(j,n) - k_j^n \right) - \left(C_\mu(k(j,n)) - C_\mu(k_j^n) \right) \\
&\quad + \left(C_\mu(k_{j+1}^n) - C_\mu(k(j,n)) \right) - m_j^n \left(k_{j+1}^n - k(j,n) \right) \Big] \\
&= \sum_{j=0}^{2^n-1} \Big[\frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} \left(k(j,n) - k_j^n \right) - \left(C_\mu(k(j,n)) - C_\mu(k_j^n) \right) \\
&\quad + \left(C_\mu(k_{j+1}^n) - C_\mu(k(j,n)) \right) - \frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} \left(k_{j+1}^n - k(j,n) \right) \Big] \\
&= \sum_{j=0}^{2^n-1} \frac{1}{k_{j+1}^n - k_j^n} \Big[\left(C_\mu(k_{j+1}^n) - C_\mu(k_j^n) \right) \left(k(j,n) - k_j^n \right) \\
&\quad - \left(C_\mu(k(j,n)) - C_\mu(k_j^n) \right) \left(k_{j+1}^n - k_j^n \right) \\
&\quad + \left(C_\mu(k_{j+1}^n) - C_\mu(k(j,n)) \right) \left(k_{j+1}^n - k_j^n \right) \\
&\quad - \left(C_\mu(k_{j+1}^n) - C_\mu(k_j^n) \right) \left(k_{j+1}^n - k(j,n) \right) \Big],
\end{aligned}$$

where in the second and the third equality, we plugin the definition of m_j^n and put its denominator $k_{j+1}^n - k_j^n$ outside the brackets. If we now add a suitable zero and rearrange the terms, then we receive

$$\begin{aligned}W\left(\mu,\mu_n^d\right) &= \sum_{j=0}^{2^n-1} \frac{2}{k_{j+1}^n - k_j^n}\Big[\left(C_\mu(k_{j+1}^n) - C_\mu(k(j,n))\right)\left(k(j,n) - k_j^n\right) \\ &\quad - \left(C_\mu(k(j,n)) - C_\mu(k_j^n)\right)\left(k_{j+1}^n - k(j,n)\right)\Big] \\ &= 2\sum_{j=0}^{2^n-1} \lambda_j^n C_\mu(k_{j+1}^n) + (1-\lambda_j^n)C_\mu(k_j^n) - C_\mu\left(\lambda_j^n k_{j+1}^n + (1-\lambda_j^n)k_j^n\right) \\ &= 2\sum_{j=0}^{2^n-1} C_{\mu_n^d}(k(j,n)) - C_\mu(k(j,n)),\end{aligned}$$

where we use $\lambda_j^n := \frac{k(j,n)-k_j^n}{k_{j+1}^n-k_j^n}$ and the linearly interpolating definition of $C_{\mu_n^d}$. By the choice of $k(j,n)$, we have that the slope of the secant through $C_\mu(k_j^n)$ and $C_\mu(k_{j+1}^n)$ is contained in $[C_\mu'(k(j,n)-), C_\mu'(k(j,n)+)]$, i.e. it equals $C_\mu'(k(j,n))$ whenever the derivative exists. In particular, the distance of $C_{\mu_n^d}$ and C_μ on $[k_j^n, k_{j+1}^n)$ is maximal in $k(j,n)$. That is,

$$k(j,n) = \underset{k\in[k_j^n,k_{j+1}^n)}{\operatorname{argmax}} \left|C_{\mu_d^n}(k) - C_\mu(k)\right|.$$

Thus, we have the desired representation

$$W\left(\mu,\mu_n^d\right) = 2\cdot\sum_{j=0}^{2^n-1} \left\|\left(C_{\mu_d^n} - C_\mu\right)\Big|_{[k_j^n,k_{j+1}^n)}\right\|_\infty.$$

Now we turn to the estimate in (7.4). For this purpose, we estimate the slope $C_\mu'(t+)$ for $t\in[k_j^n,k_{j+1}^n)$. In particular, we have

$$C_\mu'(t+)\begin{cases}\geq C_\mu'(k_j^n+), & t\in[k_j^n,k(j,n)) \\ \leq C_\mu'(k_{j+1}^n+), & t\in[k(j,n),k_{j+1}^n).\end{cases}$$

Using the above estimate in (7.8), we get

$$\begin{aligned}W\left(\mu,\mu_n^d\right) &\leq \sum_{j=0}^{2^n-1}\left[\int_{k_j^n}^{k(j,n)}\left(m_j^n - C_\mu'(k_j^n+)\right)\mathrm{d}t + \int_{k(j,n)}^{k_{j+1}^n}\left(C_\mu'(k_{j+1}^n+) - m_j^n\right)\mathrm{d}t\right] \\ &= \sum_{j=0}^{2^n-1}\Bigg[\int_{k_j^n}^{k(j,n)}\left(\frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} - C_\mu'(k_j^n+)\right)\mathrm{d}t \\ &\quad + \int_{k(j,n)}^{k_{j+1}^n}\left(C_\mu'(k_{j+1}^n+) - \frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n}\right)\mathrm{d}t\Bigg] \\ &= \sum_{j=0}^{2^n-1}\Bigg[\left(\frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} - C_\mu'(k_j^n+)\right)\left(k(j,n) - k_j^n\right) \\ &\quad + \left(C_\mu'(k_{j+1}^n+) - \frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n}\right)\left(k_{j+1}^n - k(j,n)\right)\Bigg],\end{aligned} \tag{7.9}$$

where we first plugin the definition of m_j^n and then calculate the integrals. Then we immediately get $W(\mu, \mu_n^d) \le \frac{K}{2^n}$. Indeed, if we apply the inequalities $k(j,n) \le k_{j+1}^n$ and $-k(j,n) \le -k_j^n$ in the estimate in (7.9), then we have

$$\begin{aligned} W\left(\mu, \mu_n^d\right) &\le \sum_{j=0}^{2^n-1} \left(C'_\mu(k_{j+1}^n+) - C'_\mu(k_j^n+)\right)\left(k_{j+1}^n - k_j^n\right) \\ &= \frac{K}{2^n} \sum_{j=0}^{2^n-1} \left(C'_\mu(k_{j+1}^n+) - C'_\mu(k_j^n+)\right) \le \frac{K}{2^n}. \end{aligned} \tag{7.10}$$

In order to obtain the estimate in (7.6), we use the fact that the slopes get closer and closer when n increases. We assume that $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$ and rewrite the right hand side of (7.9). Then we have

$$\begin{aligned} W\left(\mu, \mu_n^d\right) \le \sum_{j=0}^{2^n-1} \Bigg[&\left(C_\mu(k_{j+1}^n) - C_\mu(k_j^n) - C'_\mu(k_j^n)(k_{j+1}^n - k_j^n)\right)\left(\frac{k(j,n) - k_j^n}{k_{j+1}^n - k_j^n}\right) \\ &+ \left(C'_\mu(k_{j+1}^n)(k_{j+1}^n - k_j^n) - C_\mu(k_{j+1}^n) + C_\mu(k_j^n)\right)\left(\frac{k_{j+1}^n - k(j,n)}{k_{j+1}^n - k_j^n}\right)\Bigg]. \end{aligned}$$

Now let us use the Theorem of Taylor. In particular, for two times continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$, we may use the formula of Taylor

$$f(x) = T_n f(x,a) + R_n f(x,a)$$

for $n = 1$, where $T_1 f(x,a) = f(a) + f'(a)(x-a)$ and $R_1 f(x,a) = \int_a^x (x-t) f''(t) \mathrm{d}t$. If we now apply this formula in the form

$$f(x) - f(a) - f'(a)(x-a) = f(x) - T_1 f(x,a) = R_1 f(x,a)$$

for $f \equiv C_\mu$ with $x = k_{j+1}^n$ and $a = k_j^n$, and with $x = k_j^n$ and $a = k_{j+1}^n$, then we obtain

$$W\left(\mu, \mu_n^d\right) \le \sum_{j=0}^{2^n-1} \left(\left(\frac{k(j,n) - k_j^n}{k_{j+1}^n - k_j^n}\right) R_1 C_\mu(k_{j+1}^n, k_j^n) + \left(\frac{k_{j+1}^n - k(j,n)}{k_{j+1}^n - k_j^n}\right) R_1 C_\mu(k_j^n, k_{j+1}^n)\right).$$

The well-known general Taylor residual estimate states that we have

$$|R_1 f(x,a)| \le \sup_{\xi \in (a-r, a+r)} \left| \frac{f''(\xi)}{2}(x-a)^2 \right|$$

for all $x \in (a-r, a+r)$. Choosing $r = \frac{K}{2^n} + \varepsilon$, $\varepsilon > 0$, we achieve

$$\begin{aligned} \left| R_1 C_\mu\left(k_{j+1}^n, k_j^n\right)\right| &\le \sup_{\kappa \in \left(k_{j-1}^n - \varepsilon, k_{j+1}^n + \varepsilon\right)} \left| \frac{C''_\mu(\kappa)}{2}\left(k_{j+1}^n - k_j^n\right)^2 \right| \\ &= \sup_{\kappa \in \left(k_{j-1}^n - \varepsilon, k_{j+1}^n + \varepsilon\right)} \left| \frac{C''_\mu(\kappa)}{2}\left(\frac{K}{2^n}\right)^2 \right| \le T_\mu \cdot K^2 \cdot 2^{-(2n+1)} \end{aligned}$$

and analogously $|R_1 C_\mu(k_j^n, k_{j+1}^n)| \leq T_\mu \cdot K^2 \cdot 2^{-(2n+1)}$. Thus, we get

$$W\left(\mu, \mu_n^d\right) \leq \sum_{j=0}^{2^n-1} \left(\frac{k(j,n) - k_j^n}{k_{j+1}^n - k_j^n} + \frac{k_{j+1}^n - k(j,n)}{k_{j+1}^n - k_j^n} \right) T_\mu \cdot K^2 \cdot 2^{-(2n+1)}$$
$$= 2^n \cdot T_\mu \cdot K^2 \cdot 2^{-(2n+1)} = \frac{T_\mu \cdot K^2}{2^{n+1}},$$

which is the desired estimate and thus ends the proof. □

Remark 7.14. Theorem 7.13 yields two different estimates for the Wasserstein distance $W(\mu, \mu_n^d)$ whenever we assume $C_\mu \in \mathcal{C}^2(\mathbb{R}_+)$. We analyze which one is the better estimate. The second estimate is better if and only if

$$\frac{T_\mu \cdot K^2}{2^{n+1}} \leq \frac{K}{2^n} \iff T_\mu \leq \frac{2}{K}.$$

Using the definition of T_μ, we observe that it is better whenever $C_\mu''(k) \leq \frac{2}{K}, k \in [0, K]$. Let us now consider two exemplary call option price functions in order to show that both estimates are relevant.

If we assume the structure $C_\mu(k) = (c_0 k^2 + c_1 k + c_2)\mathbb{1}_{\{0 \leq k \leq K\}}$ and as usual $C_\mu(0) = 1$, $C_\mu(K) = 0$ and $C_\mu'(K) = 0$, then we get

$$C_\mu(k) = \left(\left(\frac{k}{K}\right)^2 - 2 \cdot \frac{k}{K} + 1 \right) \mathbb{1}_{\{0 \leq k \leq K\}}.$$

This price function satisfies $C_\mu'' \equiv \frac{2}{K^2} \leq \frac{2}{K}$ on $[0, K]$, whenever $K \geq 1$. However, strictly speaking we have $C_\mu \notin \mathcal{C}^2(\mathbb{R}_+)$, as the second derivative $C_\mu''(k)$ is not continuous in $k = K$. Observe that the associated marginal μ has an atom of mass $\frac{K-2}{K}$ in 0, as $C_\mu'(0) = -\frac{2}{K}$.

As any function C_μ such that $C_\mu(0) = 1$, $C_\mu'(0) = -1$ and $C_\mu''(k) \leq \frac{2}{K}$ for all $k \in [0, K]$, takes negative values somewhere in $\left[0, \frac{K - \sqrt{K^2 - 4K}}{2}\right]$ assuming $K \geq 4$, in this case there is no continuous distribution μ on $[0, K]$ such that $C_\mu'' \leq \frac{2}{K}$.

Alternatively, choosing

$$C_\mu(k) = \left(\left(1 - \frac{k}{K}\right) \exp\left(-\frac{K-1}{K} k\right) \right) \mathbb{1}_{\{0 \leq k \leq K\}},$$

we have a price function such that the first estimate yields the better results. ◊

Proof of Theorem 7.11. We may rewrite $\left|\mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d, \nu_n^d)}\left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right]\right|$ as

$$\left| \sup_{\mathbb{Q} \in \mathcal{M}_2(\mu_n^d, \nu_n^d)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] - \sup_{\mathbb{Q} \in \mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] \right|,$$

since the payoff function c satisfies the martingale Spence Mirrlees condition.

Now we may reformulate this difference as the difference of the values of the dual problems, as the payoff function satisfies the strong duality properties of Corollary 5.2 and

Theorem 6.15 respectively. That is, we have

$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n^d,\nu_n^d)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] = \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\int_{\mathbb{R}_+} \varphi(x)\mu_n^d(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi(y)\nu_n^d(\mathrm{d}y)\right\},$$

$$\sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] = \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\int_{\mathbb{R}_+} \varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi(y)\nu(\mathrm{d}y)\right\}.$$

Now let us apply Theorem 5.21 and Remark 5.22. For this purpose, we have to prove that the conditions are satisfied. By assumption and by construction respectively, we have that $\mu \leq_c \nu$ and $\mu_n^d \leq_c \nu_n^d$ are compactly supported. The payoff function c is Lipschitz continuous on $[0,L]\times[0,L] = \mathrm{conv}(\mathrm{supp}(\nu)) \times \mathrm{conv}(\mathrm{supp}(\nu))$ with constant $\hat{\Lambda} \leq \Lambda$. It remains to show that there is a Lipschitz continuous function $u : [0,L] = \mathrm{conv}(\mathrm{supp}(\nu)) \to \mathbb{R}$ such that $y \mapsto c(x,y) + u(y)$ is concave on $[0,L]$ for μ-almost every $x \in \mathbb{R}_+$. As $c_{yy} \leq \Lambda$, it is clear that $u(y) \coloneqq -\frac{\Lambda}{2}y^2$ is such a function with Lipschitz constant ΛL. We define $\tilde{\Lambda} \coloneqq \Lambda \cdot \max\{L,1\}$.

Thus, by Theorem 5.21 there are solutions (φ^*,ψ^*,h^*) and $(\varphi_n^*,\psi_n^*,h_n^*)$ for the dual problems with respect to (μ,ν) and (μ_n^d,ν_n^d) respectively. By Remark 5.22, φ^* and φ_n^* are Lipschitz continuous with constant $7\tilde{\Lambda}$, and ψ^* and ψ_n^* are Lipschitz continuous with constant $5\tilde{\Lambda}$. Hence, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d,\nu_n^d)}\left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right]\\
&\quad= \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\int_{\mathbb{R}_+} \varphi(x)\mu_n^d(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi(y)\nu_n^d(\mathrm{d}y)\right\}\\
&\qquad - \inf_{(\varphi,\psi,h)\in\mathcal{D}_2^{\geq c}} \left\{\int_{\mathbb{R}_+} \varphi(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi(y)\nu(\mathrm{d}y)\right\}\\
&\quad\leq \int_{\mathbb{R}_+} \varphi^*(x)\mu_n^d(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi^*(y)\nu_n^d(\mathrm{d}y)\\
&\qquad - \left(\int_{\mathbb{R}_+} \varphi^*(x)\mu(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi^*(y)\nu(\mathrm{d}y)\right)\\
&\quad= \int_{\mathbb{R}_+} \varphi^*(x)\left(\mu_n^d - \mu\right)(\mathrm{d}x) + \int_{\mathbb{R}_+} \psi^*(y)\left(\nu_n^d - \nu\right)(\mathrm{d}y)\\
&\quad\leq 7\tilde{\Lambda}W\left(\mu,\mu_n^d\right) + 5\tilde{\Lambda}W\left(\nu,\nu_n^d\right),
\end{aligned}$$

where in the last inequality we scale the integrands by their Lipschitz constants and then use the dual representation of the Wasserstein distance in (2.2). Completely analogous, but using φ_n^* and ψ_n^* in the first inequality instead of φ^* and ψ^*, we obtain

$$\mathbb{E}_{\mathbb{Q}_{lc}(\mu,\nu)}\left[c(X,Y)\right] - \mathbb{E}_{\mathbb{Q}_{lc}(\mu_n^d,\nu_n^d)}\left[c(X,Y)\right] \leq 7\tilde{\Lambda}W\left(\mu,\mu_n^d\right) + 5\tilde{\Lambda}W\left(\nu,\nu_n^d\right).$$

Using the estimates in (7.4) - (7.7), we have the claimed convergence speed estimates. □

In Example A.4, we show that the convergence speed of Theorem 7.11 is best possible.

7.4. Generalized results

In this section, we generalize the results of Sections 7.2 and 7.3. In particular, we aim at generalizing Theorems 7.8 and 7.11. In Section 7.4.1, we get rid of the assumption that the payoff function satisfies the martingale Spence Mirrlees condition and also of the specially designed sequence of marginals $(\mu_n^d, \nu_n^d)_{n\in\mathbb{N}}$. We further present a result relying on a more general set of observable call option prices. All results of that section are derived for theoretical marginals with bounded support. In Section 7.4.2, we adapt the definition of the marginals μ_n^d and ν_n^d such that we get rid of Assumption 7.2, i.e. such that we may consider theoretical marginals with unbounded support.

7.4.1. The case of general payoff functions & call option prices

In order to generalize the results of the previous sections, we closely analyze the proof of Theorem 7.11. First we observe that the martingale Spence Mirrlees condition is not necessary. Thus, we get the following result.

Theorem 7.15. *Let $(\mu,\nu) \in \mathcal{P}_{K,L}^{\leq_c}$. Let $c : \mathbb{R}_+^2 \to \mathbb{R}$ be a Lipschitz continuous payoff function such that c_{yy} exists. We denote by $\hat{\Lambda}$ the Lipschitz constant of c and assume $\max\{\hat{\Lambda}, \sup_{(x,y)\in\mathbb{R}^2} |c_{yy}(x,y)|\} \leq \Lambda$. Then, for any $n \in \mathbb{N}$, we have*

$$\left| \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n^d,\nu_n^d)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] - \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] \right| \leq \frac{M_c}{2^n},$$

where $M_c = (7K + 5L) \cdot \tilde{\Lambda}$ with $\tilde{\Lambda} = \Lambda \cdot \max\{L, 1\}$. If we additionally suppose that $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$, then, for any $n \in \mathbb{N}$, we have

$$\left| \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n^d,\nu_n^d)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] - \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] \right| \leq \frac{M_d}{2^{n+1}},$$

where $M_d = (7T_\mu K^2 + 5T_\nu L^2) \cdot \tilde{\Lambda}$ with $T_\mu = \sup_{\kappa\in[0,K]} |C_\mu''(\kappa)|$ and $T_\nu = \sup_{\lambda\in[0,L]} |C_\nu''(\lambda)|$.

Furthermore, we observe that the proof is parted in the estimation of the price bound difference against the Wasserstein distances of the approximating marginals and the estimation of those Wasserstein distances. Removing the second part and replacing μ_n^d and ν_n^d by general approximating sequences μ_n and ν_n, we obtain the following result.

Theorem 7.16. *Let $(\mu,\nu) \in \mathcal{P}_{K,L}^{\leq_c}$. Let $c : \mathbb{R}_+^2 \to \mathbb{R}$ be a payoff function as in Theorem 7.15. Finally, let $(\mu_n,\nu_n) \in \mathcal{P}_n^{\leq_c}$, $n \in \mathbb{N}$. Then, for any $n \in \mathbb{N}$, we have*

$$\left| \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n,\nu_n)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] - \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X,Y)\right] \right| \leq \tilde{\Lambda}\left[7W(\mu,\mu_n) + 5W(\nu,\nu_n)\right].$$

We may quantify the Wasserstein distances of the previous theorem. For this purpose, we need a lemma provided by Guo & Obłój [32].

Lemma 7.17 ([32, Lemma 3.13])**.** *Let $\rho_1, \rho_2 \in \mathcal{P}([0, R])$, $R > 0$. If there is an $\varepsilon > 0$ such that*

$$|C_{\rho_1}(k) - C_{\rho_2}(k)| \leq \varepsilon, \quad k \in [0, R],$$

then $W(\rho_1, \rho_2) \leq R\sqrt{2\varepsilon}$.

Lemma 7.18. *Let $(\mu, \nu) \in \mathcal{P}_{K,L}^{\leq_c}$ and $(\mu_n, \nu_n) \in \mathcal{P}_n^{\leq_c}$. Then, for any $n \in \mathbb{N}$, we have*

$$W(\mu, \mu_n) \leq K\sqrt{2^{1-n}},$$
$$W(\nu, \nu_n) \leq L\sqrt{2^{1-n}}.$$

Proof. We prove the claim for $W(\mu, \mu_n)$ and use Lemma 7.17. Therefore, we have to estimate $|C_{\mu_n}(k) - C_\mu(k)|$ for all $k \in [0, K]$. There is a $0 \leq j \leq 2^n - 1$ such that $k \in [k_j^n, k_{j+1}^n]$. Then we have

$$C_{\mu_n}(k) - C_\mu(k) \leq C_{\mu_n}(k_j^n) - C_\mu(k_{j+1}^n) = C_\mu(k_j^n) - C_\mu(k_{j+1}^n) \leq k_{j+1}^n - k_j^n = \frac{1}{2^n},$$

where the first inequality is the monotonicity of C_μ and C_{μ_n}, the equality is implied by the consistency, and the second inequality holds by $C'_\mu(0+) \geq -1$ and the convexity of C_μ. Analogously, we have

$$C_{\mu_n}(k) - C_\mu(k) \geq C_{\mu_n}(k_{j+1}^n) - C_\mu(k_j^n) = C_\mu(k_{j+1}^n) - C_\mu(k_j^n) \geq k_j^n - k_{j+1}^n = -\frac{1}{2^n}.$$

Thus, $|C_{\mu_n}(k) - C_\mu(k)| \leq \frac{1}{2^n}$ for all $k \in [0, K]$. Using Lemma 7.17, the assertion holds. □

This quantification yields the following explicit version of Theorem 7.16.

Corollary 7.19. *Let the conditions of Theorem 7.16 be satisfied. Then, for any $n \in \mathbb{N}$, we have*

$$\left| \sup_{\mathbb{Q} \in \mathcal{M}_2(\mu_n, \nu_n)} \mathbb{E}_\mathbb{Q}[c(X, Y)] - \sup_{\mathbb{Q} \in \mathcal{M}_2(\mu, \nu)} \mathbb{E}_\mathbb{Q}[c(X, Y)] \right| \leq \tilde{\Lambda}\sqrt{2^{1-n}}(7K + 5L).$$

Analyzing the proof of Lemma 7.18, we observe that a result similar to Corollary 7.19 may be deduced for general partitions and thus general sequences of approximating marginals. For $R \in \mathbb{R}_+$, we denote by $Z^R = (Z_n^R)_{n \in \mathbb{N}}$ a sequence of partitions of $[0, R]$. If $Z_n^R = \{z_0^n, \ldots, z_{j_n}^n\}$, $j_n \in \mathbb{N}$ with $0 = z_0^n < \ldots < z_{j_n}^n = R$, then we denote $\Delta Z_n^R := \max_{j=1,\ldots,j_n} |z_j^n - z_{j-1}^n|$.

For $0 < R_1 \leq R_2 < \infty$, let $(\mu, \nu) \in \mathcal{P}_{R_1,R_2}^{\leq_c}$ and Z^{R_1} and Z^{R_2} be sequences of partitions of $[0, R_1]$ and $[0, R_2]$ respectively. Then, for all $n \in \mathbb{N}$, we denote by

$$C_{Z_n^{R_1}}^\mu := \left\{ C_\mu(z) \,\middle|\, z \in Z_n^{R_1} \right\} \quad \text{and} \quad C_{Z_n^{R_2}}^\nu := \left\{ C_\nu(z) \,\middle|\, z \in Z_n^{R_2} \right\}$$

the sets of observable call option prices associated to the strike prices in $Z_n^{R_1}$ and $Z_n^{R_2}$.

Finally, for all $n \in \mathbb{N}$, we define

$$\mathcal{P}^{\leq_c}_{Z_n^{R_1}, Z_n^{R_2}} := \left\{ (\rho_1, \rho_2) \in \mathcal{P}(\mathbb{R}_+) \times \mathcal{P}(\mathbb{R}_+) \;\middle|\; \begin{array}{l} \rho_1 \text{ is consistent with } C^{\mu}_{Z_n^{R_1}}, \\ \rho_2 \text{ is consistent with } C^{\nu}_{Z_n^{R_2}} \text{ and } \rho_1 \leq_c \rho_2 \end{array} \right\}.$$

Theorem 7.20. *Let $c : \mathbb{R}_+^2 \to \mathbb{R}$ be a payoff function as in Theorem 7.16. Let $K, L \in \mathbb{R}_+$ be arbitrary and $(\mu, \nu) \in \mathcal{P}^{\leq_c}_{K,L}$. Let Z^K and Z^L be sequences of partitions of $[0, K]$ and $[0, L]$ with $\Delta Z_n^K \overset{n\to\infty}{\to} 0$ and $\Delta Z_n^L \overset{n\to\infty}{\to} 0$ and let $(\mu_n, \nu_n)_{n\in\mathbb{N}}$ be a sequence of approximating marginals such that $(\mu_n, \nu_n) \in \mathcal{P}^{\leq_c}_{Z_n^K, Z_n^L}$ for all $n \in \mathbb{N}$. Then, for any $n \in \mathbb{N}$, we have*

$$\left| \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n,\nu_n)} \mathbb{E}_{\mathbb{Q}}\left[c(X, Y)\right] - \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X, Y)\right] \right| \leq \tilde{\Lambda} \left(7K\sqrt{2\Delta Z_n^K} + 5L\sqrt{2\Delta Z_n^L} \right).$$

In particular, the mapping

$$P_2^c : \begin{cases} \mathcal{P}^{\leq_c}_{K,L} \to \mathbb{R} \\ (\rho_1, \rho_2) \mapsto \sup_{\mathbb{Q}\in\mathcal{M}_2(\rho_1,\rho_2)} \mathbb{E}_{\mathbb{Q}}[c(X, Y)] \end{cases}$$

is continuous with respect to the topology $\mathcal{T}_{cb}(\mathbb{R}_+)^2$ as well as the topology $\mathcal{T}_1(\mathbb{R}_+)^2$.

Proof. As in the proof of Theorem 7.11, we get

$$\left| \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu_n,\nu_n)} \mathbb{E}_{\mathbb{Q}}\left[c(X, Y)\right] - \sup_{\mathbb{Q}\in\mathcal{M}_2(\mu,\nu)} \mathbb{E}_{\mathbb{Q}}\left[c(X, Y)\right] \right| \leq \tilde{\Lambda} \left[7W(\mu, \mu_n) + 5W(\nu, \nu_n)\right]. \quad (7.11)$$

Indeed, the techniques used to obtain the above estimate are independent of the approximating measures as well as the underlying partitions.

As in the proof of Lemma 7.18, for a general partition Z^K of $[0, K]$ and an arbitrary measure μ_n consistent with $C^{\mu}_{Z_n^K}$, we obtain $|C_{\mu_n}(k) - C_{\mu}(k)| \leq \Delta Z_n^K$ for all $n \in \mathbb{N}$.

By Lemma 7.17, we then have $W(\mu, \mu_n) \leq K\sqrt{2\Delta Z_n^K}$. Proceeding analogously, we obtain $W(\nu, \nu_n) \leq L\sqrt{2\Delta Z_n^L}$. The claim concerning the convergence speed holds by (7.11).

The continuity of the mapping P_2^c follows from the convergence speed assertion in equation (7.11), when we observe that for every sequence $(\mu_n, \nu_n)_{n\in\mathbb{N}}$ in $\mathcal{P}^{\leq_c}_{K,L}$ there are sequences of partitions Z^K and Z^L such that $(\mu_n, \nu_n) \in \mathcal{P}^{\leq_c}_{Z_n^K, Z_n^L}$ for all $n \in \mathbb{N}$. □

7.4.2. The case of general theoretical marginals

In this section, we adapt the definition of μ_n^d and ν_n^d to the case that Assumption 7.2 is not satisfied, i.e. that the theoretical marginals $\mu, \nu \in \mathcal{P}(\mathbb{R}_+)$ have unbounded support. We denote the resulting measures by μ_n^∞ and ν_n^∞.

In this case, Guo & Obłój [32] derive various results on closely related problems using similar techniques such as Theorem 5.21. However, we believe that their proofs are incorrect, as the conditions of the above mentioned theorem are not satisfied. We stress that the results of this chapter, though very similar, are independent from those of Guo & Obłój [32] in the sense that neither are special cases of the other. Still, the results are so similar that techniques which overcome the errors in the work of Guo & Obłój [32] are very likely to improve and generalize our results as well.

Differently to the compact situation, we do not allow the strike prices to be different for the observed option prices. Thus, for every $n \in \mathbb{N}$, we denote the set of strike prices by

$$Z_n^\infty := \left\{ k_j^n := \frac{j}{2^n} \;\middle|\; j = 0, \dots, 4^n \right\}.$$

Then we have $k_0^n = 0$ and $k_{4^n}^n = 2^n$. We denote the options prices associated to Z_n^∞ by

$$C_n^{\mu,\infty} := \{C_\mu(k) \mid k \in Z_n^\infty\} \quad \text{and} \quad C_n^{\nu,\infty} := \{C_\nu(k) \mid k \in Z_n^\infty\}.$$

Now let us define candidate functions $C_{\mu_n^\infty}, C_{\nu_n^\infty} \in \mathcal{K}^C$ consistent with the prices in $C_n^{\mu,\infty}$ and $C_n^{\nu,\infty}$ such that $C_{\mu_n^\infty} \leq C_{\nu_n^\infty}$, from which we may then derive consistent measures $\mu_n^\infty, \mu_n^\infty \in \mathcal{P}(\mathbb{R}_+)$ such that $\mu_n^\infty \leq_c \nu_n^\infty$. As in the compactly supported case, let us use piecewise linear functions. Then the main difference to the former case is the fact that $C_\mu(k_{4^n}^n) \neq 0$ and $C_\nu(k_{4^n}^n) \neq 0$. Hence, for $k \geq k_{4^n}^n$, it is unclear how to choose $C_{\mu_n^\infty}(k)$ and $C_{\nu_n^\infty}(k)$. However, for $k \in [0, k_{4^n}^n)$, we choose the functions $C_{\mu_n^\infty}(k)$ and $C_{\nu_n^\infty}(k)$ similar to Section 7.3.1, i.e. for $k \in [k_j^n, k_{j+1}^n), j = 0, \dots, 4^n - 1$, we define

$$C_{\mu_n^\infty}(k) := \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} C_\mu(k_j^n) + \frac{k - k_j^n}{k_{j+1}^n - k_j^n} C_\mu(k_{j+1}^n), \tag{7.12}$$

$$C_{\nu_n^\infty}(k) := \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} C_\nu(k_j^n) + \frac{k - k_j^n}{k_{j+1}^n - k_j^n} C_\nu(k_{j+1}^n). \tag{7.13}$$

It remains to decide how to choose $C_{\mu_n^\infty}$ and $C_{\nu_n^\infty}$ on $(k_{4^n}^n, \infty)$. For this, we introduce two methods that seem natural, explain why both do not work as desired, and adapt one of them in order to be purposeful.

1. Replace the observed option prices $C_{\mu_n^\infty}(k_{4^n}^n)$ and $C_{\nu_n^\infty}(k_{4^n}^n)$ by zero. Then we have $C_{\mu_n^\infty} \equiv C_{\nu_n^\infty} \equiv 0$ on $(k_{4^n}^n, \infty)$, and linearly interpolating on $[k_{4^n-1}^n, k_{4^n}^n)$, we have completely defined the price functions. This method yields two problems. While the first problem - the consistency is lost - is of minor interest when it comes to approximation and convergence, the second problem is more severe. In Figure 7.1, we see that the convexity may be violated by this choice. This is impractical, as it gives rise to signed measures.

2. Choose $C_{\mu_n^\infty}$ and $C_{\nu_n^\infty}$ on $(k_{4^n}^n, \infty)$ as the pointwise maximum of the linear continuation of $C_{\mu_n^\infty}$ and $C_{\nu_n^\infty}$ on $[k_{4^n-1}^n, k_{4^n}^n)$ respectively and the zero function. Then we remain to have consistency and convexity of the two price functions. However, in general we do not have $C_{\mu_n^\infty} \leq C_{\nu_n^\infty}$ such that the convex order of the resulting measures is lost. This is illustrated in Figure 7.2.

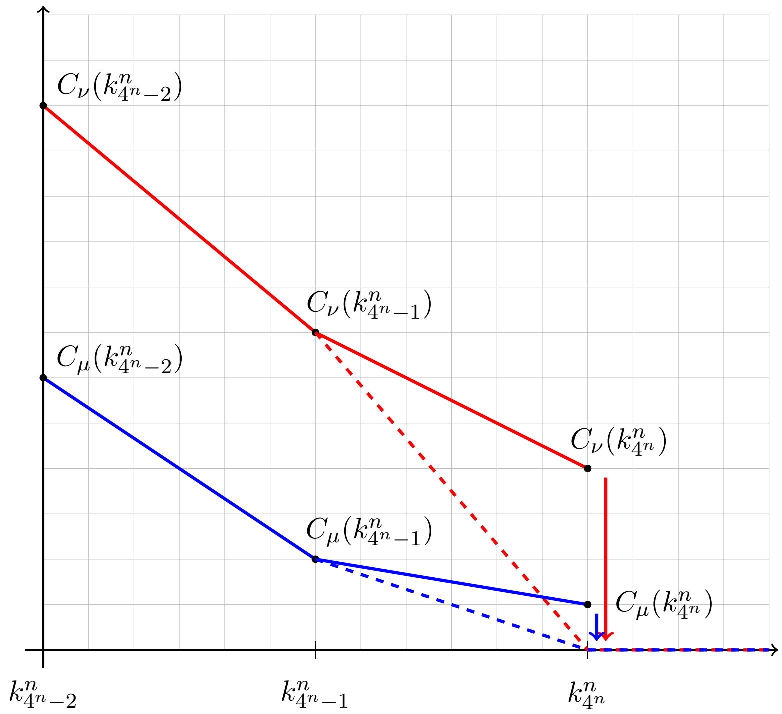


Figure 7.1.: The convexity of $C_{\nu_n^\infty}$ (red) is lost.

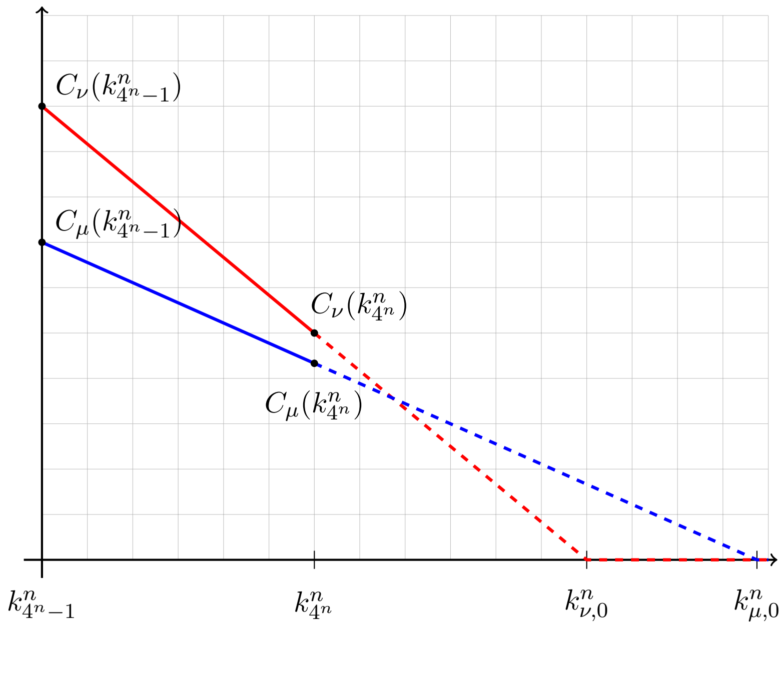


Figure 7.2.: The order $C_{\mu_n^\infty} \leq C_{\nu_n^\infty}$ is lost.

In order to obtain a useful choice for the price function on the critical interval, we adapt the second method slightly. For this purpose, we denote by $k^n_{\mu,0}$ and $k^n_{\nu,0}$ the smallest zeros of the continuations of $C_{\mu_n^\infty}$ and $C_{\nu_n^\infty}$ on $[k^n_{4^n-1}, k^n_{4^n})$ to $(k^n_{4^n}, \infty)$ as described in the second method, i.e. exactly the point at which the maximum switches. Formally, we have

$$k^n_{\mu,0} = \inf\left\{ k \in (k^n_{4^n}, \infty) \,\middle|\, \frac{k^n_{4^n} - k}{k^n_{4^n} - k^n_{4^n-1}} C_\mu(k^n_{4^n-1}) + \frac{k - k^n_{4^n-1}}{k^n_{4^n} - k^n_{4^n-1}} C_\mu(k^n_{4^n}) = 0 \right\},$$

$$k^n_{\nu,0} = \inf\left\{ k \in (k^n_{4^n}, \infty) \,\middle|\, \frac{k^n_{4^n} - k}{k^n_{4^n} - k^n_{4^n-1}} C_\nu(k^n_{4^n-1}) + \frac{k - k^n_{4^n-1}}{k^n_{4^n} - k^n_{4^n-1}} C_\nu(k^n_{4^n}) = 0 \right\}.$$

We have to distinguish two cases.

1. Let $k^n_{\mu,0} \leq k^n_{\nu,0}$. Then the procedure described in the second method already remains the order of the price functions. Hence, we choose $C_{\mu_n^\infty}$ and $C_{\nu_n^\infty}$ exactly as the described linear continuations truncated in zero. This is illustrated in Figure 7.3. For the resulting measures, we observe that μ_n^∞ and ν_n^∞ have an atom in $k^\infty_{\mu,0}$ and $k^n_{\nu,0}$ respectively, but neither have an atom in $k^n_{4^n}$.

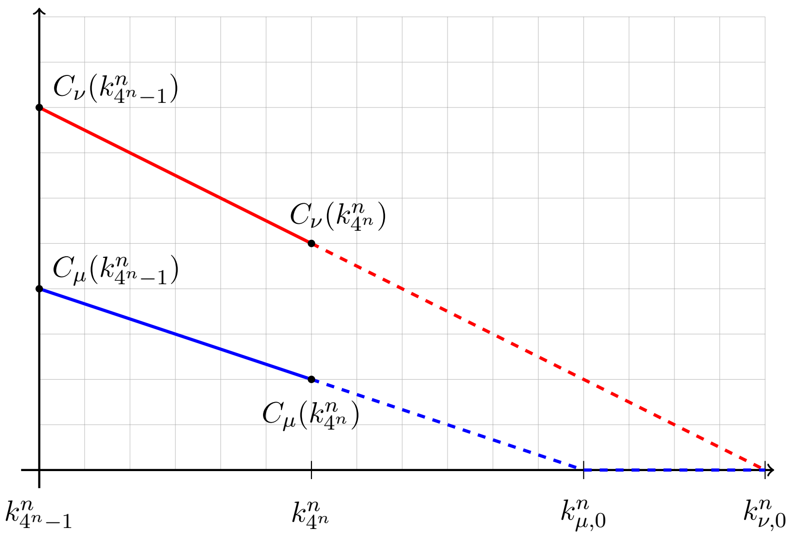


Figure 7.3.: If $k^n_{\mu,0} \leq k^n_{\nu,0}$, then the convexity and the order $C_{\mu_n^\infty} \leq C_{\nu_n^\infty}$ are preserved.

2. Let $k^n_{\nu,0} < k^n_{\mu,0}$. Then the procedure described in the second method fails to remain the order of the price functions. Hence, we choose $C_{\mu_n^\infty}$ as the described linear continuation and $C_{\nu_n^\infty}$ as linear interpolation between the points $(k^n_{4^n}, C_\nu(k^n_{4^n}))$ and $(k^n_{\mu,0}, 0)$ on $(k^n_{4^n}, k^n_{\mu,0})$ and constantly equal to zero on $[k^n_{\mu,0}, \infty)$. This procedure is illustrated in Figure 7.4. Doing so, we reduce the slope for none of the functions in order to keep the convexity and we guarantee the order of the functions by reducing both to zero in the same point. For the resulting measures, we observe that ν_n^∞ has atoms in $k^n_{4^n}$ and $k^n_{\mu,0}$, while μ_n^∞ has only one atom in $k^n_{\mu,0}$.

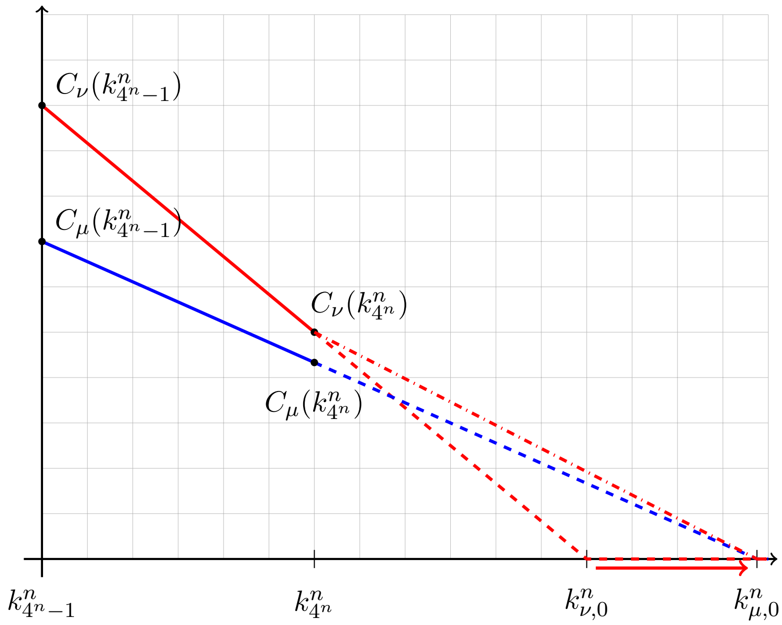


Figure 7.4.: If $k^n_{\nu,0} < k^n_{\mu,0}$, then the convexity is preserved but the order $C_{\mu_n^\infty} \leq C_{\nu_n^\infty}$ is lost.

After this heuristic description, we formalize the choices of $C_{\mu_n^\infty}$ and $C_{\nu_n^\infty}$ on $(k^n_{4^n}, \infty)$. For $k \in (k^n_{4^n}, k^n_{\mu,0})$, we define

$$C_{\mu_n^\infty}(k) := \frac{k^n_{\mu,0} - k}{k^n_{\mu,0} - k^n_{4^n}} C_\mu(k^n_{4^n}) \tag{7.14}$$

and $C_{\mu_n^\infty} \equiv 0$ on $(k^n_{\mu,0}, \infty)$. For $C_{\nu_n^\infty}$ a case distinction is necessary.

1. Let $k^n_{\mu,0} \leq k^n_{\nu,0}$. Then we define

$$C_{\nu_n^\infty}(k) := \frac{k^n_{\nu,0} - k}{k^n_{\nu,0} - k^n_{4^n}} C_\nu(k^n_{4^n}) \tag{7.15}$$

   for $k \in (k^n_{4^n}, k^n_{\nu,0})$ and $C_{\nu_n^\infty} \equiv 0$ on $(k^n_{\nu,0}, \infty)$.

2. $k^n_{\mu,0} > k^n_{\nu,0}$. Then we define

$$C_{\nu_n^\infty}(k) := \frac{k^n_{\mu,0} - k}{k^n_{\mu,0} - k^n_{4^n}} C_\nu(k^n_{4^n}) \tag{7.16}$$

   for $k \in (k^n_{4^n}, k^n_{\mu,0})$ and $C_{\nu_n^\infty} \equiv 0$ on $(k^n_{\mu,0}, \infty)$.

By (7.12) - (7.16), we define call option price functions $C_{\mu_n^\infty}, C_{\nu_n^\infty} \in \mathcal{K}^C$.

From those functions we may derive probability measures $\mu_n^\infty, \nu_n^\infty \in \mathcal{P}(\mathbb{R}_+)$ with expected value equal to one. The resulting measures are in convex order by construction, i.e. $\mu_n^\infty \leq_c \nu_n^\infty$, and they are discrete measures of the form

$$\mu_n^\infty := \sum_{j=0}^{4^n-1} \omega_j^n \delta_{k_j^n} + \mu_{n,r}^\infty,$$

where $\omega_j^n = \frac{C_\mu(k_{j+1}^n)-C_\mu(k_j^n)}{k_{j+1}^n-k_j^n} - \frac{C_\mu(k_j^n)-C_\mu(k_{j-1}^n)}{k_j^n-k_{j-1}^n}$ and

$$\mu_{n,r}^\infty := \omega_\mu^n \delta_{k_{\mu,0}^n} := \frac{C_\mu(k_{4^n-1}^n) - C_\mu(k_{4^n}^n)}{k_{4^n}^n - k_{4^n-1}^n}\delta_{k_{\mu,0}^n},$$

and slightly more complicated

$$\nu_n^\infty := \sum_{i=0}^{4^n-1} \vartheta_i^n \delta_{k_i^n} + \nu_{n,r}^\infty,$$

where $\vartheta_i^n = \frac{C_\nu(k_{i+1}^n)-C_\nu(k_i^n)}{k_{i+1}^n-k_i^n} - \frac{C_\nu(k_i^n)-C_\nu(k_{i-1}^n)}{k_i^n-k_{i-1}^n}$ and

$$\begin{aligned}\nu_{n,r}^\infty &:= \left(\vartheta_{4^n}^n \delta_{k_{4^n}^n} + \vartheta_\mu^n \delta_{k_{\mu,0}^n}\right) \mathbb{1}_{\{k_{\mu,0}^n > k_{\nu,0}^n\}} + \vartheta_\nu^n \delta_{k_{\nu,0}^n} \mathbb{1}_{\{k_{\mu,0}^n \le k_{\nu,0}^n\}} \\ &:= \left(\left[\frac{-C_\nu(k_{4^n}^n)}{k_{\mu,0}^n - k_{4^n}^n} - \frac{C_\nu(k_{4^n}^n) - C_\nu(k_{4^n-1}^n)}{k_{4^n}^n - k_{4^n-1}^n}\right]\delta_{k_{4^n}^n} + \frac{C_\nu(k_{4^n}^n)}{k_{\mu,0}^n - k_{4^n}^n}\delta_{k_{\mu,0}^n}\right)\mathbb{1}_{\{k_{\mu,0}^n > k_{\nu,0}^n\}} \\ &\quad + \frac{C_\nu(k_{4^n-1}^n) - C_\nu(k_{4^n}^n)}{k_{4^n}^n - k_{4^n-1}^n}\delta_{k_{\nu,0}^n}\mathbb{1}_{\{k_{\mu,0}^n \le k_{\nu,0}^n\}}.\end{aligned}$$

Now we investigate the convergence and convergence speed of $|P_2^c(\mu_n^\infty, \nu_n^\infty) - P_2^c(\mu, \nu)|$, where we assume $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$ and that $c : \mathbb{R}_+^2 \to \mathbb{R}$ satisfies the assumptions of Theorem 5.21, where we assume that the Lipschitz constants of c and the y-concavifier u are less than $\Lambda > 0$. Thus, using Theorem 5.21, Remark 5.22 and the same techniques as in the proof of Theorem 7.11, we obtain the estimate

$$|P_2^c(\mu_n^\infty, \nu_n^\infty) - P_2^c(\mu, \nu)| \le \Lambda \left[19W(\mu, \mu_n^\infty) + 17W(\nu, \nu_n^\infty)\right]. \tag{7.17}$$

Let us first calculate $W(\mu, \mu_n^\infty)$. In great parts, we may proceed exactly as we do to obtain (7.9) and (7.10) in the proof of Theorem 7.13. Thus, we get

$$\begin{aligned} W(\mu, \mu_n^\infty) &= \sum_{j=0}^{4^n-1}\left[\int_{k_j^n}^{k(j,n)} (m_j^n - C'_\mu(t))\mathrm{d}t + \int_{k(j,n)}^{k_{j+1}^n} (C'_\mu(t) - m_j^n)\mathrm{d}t\right] \\ &\quad + \int_{k_{4^n}^n}^\infty \left|F_\mu(t) - F_{\mu_n^\infty}(t)\right| \mathrm{d}t \\ &\le \frac{1 + C'_\mu(k_{4^n}^n)}{2^n} + \int_{k_{4^n}^n}^\infty \left|F_\mu(t) - F_{\mu_n^\infty}(t)\right| \mathrm{d}t.\end{aligned}$$

It remains to estimate the second summand. For this purpose, recall that μ_n^∞ has no atom in $k_{4^n}^n$. Thus, we have $F_\mu \ge F_{\mu_n^\infty}$ on $[k_{4^n}^n, k_{\mu,0}^n)$ and we have $F_\mu \le F_{\mu_n^\infty} \equiv 1$ on $[k_{\mu,0}^n, \infty)$. Hence, we have

$$\int_{k_{4^n}^n}^\infty \left|F_\mu(t) - F_{\mu_n^\infty}(t)\right| \mathrm{d}t = \int_{k_{4^n}^n}^{k_{\mu,0}^n} F_\mu(t) - F_{\mu_n^\infty}(t)\mathrm{d}t + \int_{k_{\mu,0}^n}^\infty 1 - F_\mu(t)\mathrm{d}t.$$

If we now plugin the distribution function representations using the associated call option price functions, then we get

$$\begin{aligned}
&\int_{k_{4^n}^n}^{\infty} \left|F_\mu(t) - F_{\mu_n^\infty}(t)\right| \mathrm{d}t \\
&= \int_{k_{4^n}^n}^{k_{\mu,0}^n} 1 + C_\mu'(t) - \left(1 + \frac{C_\mu(k_{4^n}^n) - C_\mu(k_{4^n-1}^n)}{k_{4^n}^n - k_{4^n-1}^n}\right) \mathrm{d}t + \int_{k_{\mu,0}^n}^{\infty} 1 - (1 + C_\mu'(t))\mathrm{d}t \\
&= \int_{k_{4^n}^n}^{k_{\mu,0}^n} C_\mu'(t) - \left(\frac{C_\mu(k_{4^n}^n) - C_\mu(k_{4^n-1}^n)}{k_{4^n}^n - k_{4^n-1}^n}\right) \mathrm{d}t + C_\mu(k_{\mu,0}^n) \\
&= C_\mu(k_{\mu,0}^n) - \left(C_\mu(k_{4^n}^n) + \left(\frac{C_\mu(k_{4^n}^n) - C_\mu(k_{4^n-1}^n)}{k_{4^n}^n - k_{4^n-1}^n}\right)(k_{\mu,0}^n - k_{4^n}^n)\right) + C_\mu(k_{\mu,0}^n) \\
&= 2C_\mu(k_{\mu,0}^n),
\end{aligned}$$

where in the last step we use that $\left(C_\mu(k_{4^n}^n) + \left(\frac{C_\mu(k_{4^n}^n) - C_\mu(k_{4^n-1}^n)}{k_{4^n}^n - k_{4^n-1}^n}\right)(k_{\mu,0}^n - k_{4^n}^n)\right) = 0$, as the fraction is exactly the slope of the linear function $C_{\mu_n^\infty}$ on $(k_{4^n}^n, k_{\mu,0}^n)$, which has, starting in $C_\mu(k_{4^n}^n)$, its zero in $k_{\mu,0}^n$.

Now let us calculate $W(\nu, \nu_n^\infty)$. Analogously to the situation with μ and μ_n^∞, we obtain

$$W(\nu, \nu_n^\infty) \leq \frac{1 + C_\nu'(k_{4^n}^n)}{2^n} + \int_{k_{4^n}^n}^{\infty} \left|F_\nu(t) - F_{\nu_n^\infty}(t)\right| \mathrm{d}t.$$

It remains to estimate the second summand. Therefore we have to distinguish two cases.

1. Let $k_{\mu,0}^n \leq k_{\nu,0}^n$. Then ν_n^∞ has no atom in $k_{4^n}^n$, but only in $k_{\nu,0}^n$. Thus, we are in the exact same situation as for $W(\mu, \mu_n^\infty)$, i.e. we have

$$\int_{k_{4^n}^n}^{\infty} \left|F_\nu(t) - F_{\nu_n^\infty}(t)\right| \mathrm{d}t = 2C_\nu(k_{\nu,0}^n).$$

2. Let $k_{\mu,0}^n > k_{\nu,0}^n$. Then ν_n^∞ has atoms in $k_{4^n}^n$ and $k_{\mu,0}^n$, but we do not know the mass of $\delta_{k_{4^n}^n}$ in particular. Thus, depending on this mass, we have to distinguished three further cases. We illustrate these cases in Figure 7.5.

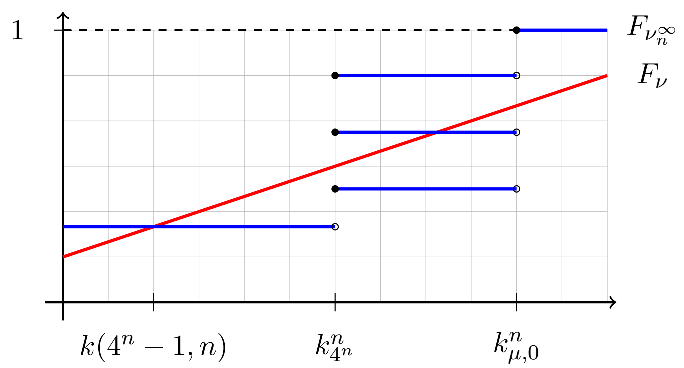


Figure 7.5.: Possible distribution functions of ν_n^∞ (blue).

2.1. Let $F_\nu(k^n_{4^n}) \geq F_{\nu_n^\infty}(k^n_{4^n})$. This is very similar to the first case as $F_\nu \geq F_{\nu_n^\infty}$ on $[k^n_{4^n}, k^n_{\mu,0})$ and $F_\nu \leq F_{\nu_n^\infty}$ on $[k^n_{\mu,0}, \infty)$. Though the value of $F_{\nu_n^\infty}$ is different and we have to replace $k^n_{\nu,0}$ by $k^n_{\mu,0}$ in the calculation, we analogously obtain

$$\int_{k^n_{4^n}}^\infty \left|F_\nu(t) - F_{\nu_n^\infty}(t)\right| \mathrm{d}t = 2C_\nu(k^n_{\mu,0}).$$

2.2. Let $F_\nu(k^n_{4^n}) < F_{\nu_n^\infty}(k^n_{4^n})$ and $F_\nu(k^n_{\mu,0}-) > F_{\nu_n^\infty}(k^n_{\mu,0}-)$. Then there exists some $k(4^n, n) \in (k^n_{4^n}, k^n_{\mu,0})$ such that $F_\nu \leq F_{\nu_n^\infty}$ on $(k^n_{4^n}, k(4^n, n))$ and $F_\nu \geq F_{\nu_n^\infty}$ on $(k(4^n, n), k^n_{\mu,0})$. Thus, for the integral of interest we have

$$\begin{aligned}\int_{k^n_{4^n}}^\infty \left|F_\nu(t) - F_{\nu_n^\infty}(t)\right| \mathrm{d}t &= \int_{k^n_{4^n}}^{k(4^n,n)} F_{\nu_n^\infty}(t) - F_\nu(t)\mathrm{d}t \\ &\quad + \int_{k(4^n,n)}^{k^n_{\mu,0}} F_\nu(t) - F_{\nu_n^\infty}(t)\mathrm{d}t + \int_{k^n_{\mu,0}}^\infty 1 - F_\nu(t)\mathrm{d}t \\ &= \int_{k^n_{4^n}}^{k(4^n,n)} \left(\frac{0 - C_\nu(k^n_{4^n})}{k^n_{\mu,0} - k^n_{4^n}}\right) - C'_\nu(t)\mathrm{d}t \\ &\quad + \int_{k(4^n,n)}^{k^n_{\mu,0}} C'_\nu(t) - \left(\frac{0 - C_\nu(k^n_{4^n})}{k^n_{\mu,0} - k^n_{4^n}}\right)\mathrm{d}t + C_\nu(k^n_{\mu,0})\end{aligned}$$

Now the two integrals may be estimated analogously to the estimates in (7.9) and (7.10) in the proof of Theorem 7.13. Thus, we obtain

$$\int_{k^n_{4^n}}^\infty \left|F_\nu(t) - F_{\nu_n^\infty}(t)\right| \mathrm{d}t \leq \left(k^n_{\mu,0} - k^n_{4^n}\right)\left(C'_\nu(k^n_{\mu,0}) - C'_\nu(k^n_{4^n})\right) + C_\nu(k^n_{\mu,0}).$$

2.3. Let $F_\nu(k^n_{4^n}) < F_{\nu_n^\infty}(k^n_{4^n})$ and $F_\nu(k^n_{\mu,0}-) \leq F_{\nu_n^\infty}(k^n_{\mu,0}-)$. Then we have $F_\nu \leq F_{\nu_n^\infty}$ on $(k^n_{4^n}, \infty)$. However, this is the case only if $C'_\nu \leq \frac{0 - C_\nu(k^n_{4^n})}{k^n_{\mu,0} - k^n_{4^n}}$ on $(k^n_{4^n}, k^n_{\mu,0})$. From $C_\nu(k^n_{4^n}) = C_{\nu_n^\infty}(k^n_{4^n})$ we may then derive $C_\nu \leq C_{\nu_n^\infty}$ on $(k^n_{4^n}, k^n_{\mu,0})$. This implies $C_\nu(k^n_{\mu,0}) \leq C_{\nu_n^\infty}(k^n_{\mu,0}) = 0$, which is a contradiction to $C_\mu \leq C_\nu$, as by construction and $C_\mu \in \mathcal{C}^2(\mathbb{R}_+)$, we have $C_\mu(k^n_{\mu,0}) > 0$. Thus, this case can not appear.

Let us conclude the results on the Wasserstein distances. We have

$$W(\mu, \mu_n^\infty) \leq \frac{1 + C'_\mu(k^n_{4^n})}{2^n} + 2C_\mu(k^n_{\mu,0}) = \frac{F_\mu(k^n_{4^n})}{2^n} + 2C_\mu(k^n_{\mu,0})$$

and

$$\begin{aligned}W(\nu, \nu_n^\infty) &\leq \frac{1 + C'_\nu(k^n_{4^n})}{2^n} + 2C_\nu(k^n_{\nu,0})\mathbb{1}_{\{k^n_{\mu,0} \leq k^n_{\nu,0}\}} \\ &\quad + \mathbb{1}_{\{k^n_{\mu,0} > k^n_{\nu,0}\}} \times \Big(2C_\nu(k^n_{\mu,0})\mathbb{1}_{\{F_\nu(k^n_{4^n}) \geq F_{\nu_n^\infty}(k^n_{4^n})\}} \\ &\qquad + \left(\left(k^n_{\mu,0} - k^n_{4^n}\right)\left(C'_\nu(k^n_{\mu,0}) - C'_\nu(k^n_{4^n})\right) + C_\nu(k^n_{\mu,0})\right)\mathbb{1}_{\{F_\nu(k^n_{4^n}) < F_{\nu_n^\infty}(k^n_{4^n})\}}\Big).\end{aligned}$$

Defining $h^\nu(k^n_{4^n}) \coloneqq \mathbb{1}_{\{F_\nu(k^n_{4^n}) \geq F_{\nu^\infty_n}(k^n_{4^n})\}}$ and using the structural properties, we achieve the simplified estimate

$$\begin{aligned} W(\nu,\nu^\infty_n) \leq \frac{F_\nu(k^n_{4^n})}{2^n} &+ C_\nu\left(\max\{k^n_{\nu,0},k^n_{\mu,0}\}\right) \\ &+ C_\nu\left(\max\{k^n_{\nu,0},k^n_{\mu,0}\}\right)\left(\mathbb{1}_{\{k^n_{\mu,0}\leq k^n_{\nu,0}\}} + \mathbb{1}_{\{k^n_{\mu,0}> k^n_{\nu,0}\}}h^\nu(k^n_{4^n})\right) \\ &+ \left(k^n_{\mu,0} - k^n_{4^n}\right)\left(C'_\nu(k^n_{\mu,0}) - C'_\nu(k^n_{4^n})\right)\mathbb{1}_{\{k^n_{\mu,0}> k^n_{\nu,0}\}}(1-h^\nu(k^n_{4^n})). \end{aligned}$$

Now plugging in the estimates for the Wasserstein distances in (7.17), we obtain

$$\begin{aligned} |P^c_2(\mu^\infty_n,\nu^\infty_n) - P^c_2(\mu,\nu)| \leq 19\Lambda\Bigg(&\frac{F_\mu(k^n_{4^n})}{2^n} + 2C_\mu(k^n_{\mu,0})\Bigg) \\ + 17\Lambda\Bigg(&\frac{F_\nu(k^n_{4^n})}{2^n} + C_\nu\left(\max\{k^n_{\nu,0},k^n_{\mu,0}\}\right) \\ &+ C_\nu\left(\max\{k^n_{\nu,0},k^n_{\mu,0}\}\right)\left(\mathbb{1}_{\{k^n_{\mu,0}\leq k^n_{\nu,0}\}} + \mathbb{1}_{\{k^n_{\mu,0}> k^n_{\nu,0}\}}h^\nu(k^n_{4^n})\right) \\ &+ \left(k^n_{\mu,0} - k^n_{4^n}\right)\left(C'_\nu(k^n_{\mu,0}) - C'_\nu(k^n_{4^n})\right)\mathbb{1}_{\{k^n_{\mu,0}> k^n_{\nu,0}\}}(1-h^\nu(k^n_{4^n}))\Bigg) \end{aligned}$$

The previous calculations prove the following theorem, which concludes the results derived in this section.

Theorem 7.21. *Let $\mu,\nu \in \mathcal{P}(\mathbb{R}_+)$. Let $c : \mathbb{R}^2_+ \to \mathbb{R}$ satisfy the assumptions of Theorem 5.21 such that the Lipschitz constants of c and its y-concavifier are less than $\Lambda > 0$. Let $C_\mu, C_\nu \in \mathcal{C}^2(\mathbb{R}_+)$. Then, for any $n \in \mathbb{N}$, we have*

$$\begin{aligned} |P^c_2(\mu^\infty_n,\nu^\infty_n) - P^c_2(\mu,\nu)| \leq 19\Lambda\Bigg(&\frac{F_\mu(k^n_{4^n})}{2^n} + 2C_\mu(k^n_{\mu,0})\Bigg) \\ + 17\Lambda\Bigg(&\frac{F_\nu(k^n_{4^n})}{2^n} + C_\nu\left(\max\{k^n_{\nu,0},k^n_{\mu,0}\}\right) \\ &+ C_\nu\left(\max\{k^n_{\nu,0},k^n_{\mu,0}\}\right)\left(\mathbb{1}_{\{k^n_{\mu,0}\leq k^n_{\nu,0}\}} + \mathbb{1}_{\{k^n_{\mu,0}> k^n_{\nu,0}\}}h^\nu(k^n_{4^n})\right) \\ &+ \left(k^n_{\mu,0} - k^n_{4^n}\right)\left(C'_\nu(k^n_{\mu,0}) - C'_\nu(k^n_{4^n})\right)\mathbb{1}_{\{k^n_{\mu,0}> k^n_{\nu,0}\}}(1-h^\nu(k^n_{4^n}))\Bigg). \end{aligned}$$

Remark 7.22. Assuming in Theorem 7.21 that the measures μ and ν have compact support, say contained in $[0,K]$, we recover an assertion similar to Theorem 7.15 for all $n \in \mathbb{N}$ such that $2^n \geq K$. Indeed if n is great enough, then under the compactness assumption we have by definition $k^n_{\mu,0} = k^n_{\nu,0} = k^n_{4^n}$ as well as $C_\mu(k^n_{\mu,0}) = C_\nu(\max\{k^n_{\mu,0},k^n_{\nu,0}\}) = 0$ and $F_\mu(k^n_{\mu,0}) = F_\mu(k^n_{4^n}) = F_\nu(k^n_{\nu,0}) = F_\nu(k^n_{4^n}) = 1$. Thus, we then have

$$|P^c(\mu^\infty_n,\nu^\infty_n) - P^c(\mu,\nu)| \leq \frac{36\cdot\Lambda}{2^n}. \qquad \diamondsuit$$

Finally, as in the previous section, we may generalize the above theorem concerning the sequence of approximating measures. In return, we lose the explicit numerical statement.

Theorem 7.23. *Let the assumptions of Theorem 7.21 be satisfied and let* $\mu_n, \nu_n \in \mathcal{P}(\mathbb{R}_+)$ *be such that* $\mu_n \leq_c \nu_n$ *for all* $n \in \mathbb{N}$. *Then, for any* $n \in \mathbb{N}$, *we have*

$$|P_2^c(\mu_n, \nu_n) - P_2^c(\mu, \nu)| \leq 19\Lambda W(\mu, \mu_n) + 17\Lambda W(\nu, \nu_n).$$

In particular, the mapping

$$P_2^c : \begin{cases} \mathcal{P}(\mathbb{R}_+) \times \mathcal{P}(\mathbb{R}_+) \to \mathbb{R} \\ (\rho_1, \rho_2) \mapsto \sup_{\mathbb{Q} \in \mathcal{M}_2(\rho_1, \rho_2))} \mathbb{E}_{\mathbb{Q}}[c(X, Y)] \end{cases}$$

is continuous with respect to the topology $\mathcal{T}_{cb}(\mathbb{R}_+)^2$ *as well as the topology* $\mathcal{T}_1(\mathbb{R}_+)^2$.

APPENDIX A

USEFUL SUPPLEMENTS

A.1. SUPPLEMENTS TO CHAPTER 4

Proof of Lemma 4.3. Replacing ϕ by $\phi - h$, we may without loss of generality assume that ϕ is non-negative and lower semi-continuous. By Lemma 2.1, we may write $\phi = \lim_{n\to\infty} \phi_n$, where $(\phi_n)_{n\in\mathbb{N}}$ is a non-decreasing sequence of functions $\phi_n \in \mathcal{C}_b(\mathbb{R})$. Then we have

$$\begin{aligned}\int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi(\mathrm{d}(x,y)) &= \lim_{n\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi_n(x,y)\pi(\mathrm{d}(x,y)) \\ &= \lim_{n\to\infty} \lim_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi_n(x,y)\pi_k(\mathrm{d}(x,y)),\end{aligned}$$

where the first equality holds by the monotone convergence theorem and the second equality holds by the definition of weak convergence. As the ϕ_n are continuous and such that we have $\phi_n \le \phi$, $n \in \mathbb{N}$, we obtain

$$\begin{aligned}\int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi(\mathrm{d}(x,y)) &= \lim_{n\to\infty} \liminf_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi_n(x,y)\pi_k(\mathrm{d}(x,y)) \\ &\le \lim_{n\to\infty} \liminf_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi_k(\mathrm{d}(x,y)) \\ &= \liminf_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} \phi(x,y)\pi_k(\mathrm{d}(x,y)).\end{aligned}$$

□

Proof of Lemma 4.4. Let $\mu \in \mathcal{P}$, $\nu \in \mathcal{Q}$ and $\pi \in \Pi_2(\mu,\nu)$. By assumption, for any $\varepsilon > 0$ there are a $K_\varepsilon \subseteq \mathcal{X}$ independent of the choice of $\mu \in \mathcal{P}$ such that $\mu(\mathcal{X} \setminus K_\varepsilon) \le \varepsilon$ and an $L_\varepsilon \subseteq \mathcal{Y}$ independent of the choice of $\nu \in \mathcal{Q}$ such that $\nu(\mathcal{Y} \setminus L_\varepsilon) \le \varepsilon$. Then, for any coupling

(X, Y) of (μ, ν), we have

$$\mathbb{P}((X,Y) \notin K_\varepsilon \times L_\varepsilon) \leq \mathbb{P}(X \notin K_\varepsilon) + \mathbb{P}(Y \notin L_\varepsilon) = \mu(\mathcal{X} \setminus K_\varepsilon) + \nu(\mathcal{Y} \setminus L_\varepsilon) \leq 2\varepsilon. \quad \text{(A.1)}$$

As $K_\varepsilon \times L_\varepsilon$ is a compact set and (A.1) is independent of the coupling (X, Y), we obtain that $\Pi_2(\mathcal{P}, \mathcal{Q})$ is indeed tight in $\mathcal{P}(\mathcal{X} \times \mathcal{Y})$. □

A.2. Supplements to Chapter 5

Proof of Lemma 5.4. We construct a suitable function $u \in \mathcal{S}$ starting with an approximating step-function $t \leq f$. Therefore let $\varepsilon > 0$. We proceed with the construction such that it is uniform for all μ_i, $i = 1, \dots, n$. Therefore we notice that the set $\{\mu_1, \dots, \mu_n\}$ of probability measures is tight as it is finite. Hence, there is a compact set $K = [a, b] \subset \mathbb{R}$ such that for all $i = 1, \dots, n$, we have

$$\mu_i(K) \geq 1 - \frac{\varepsilon}{12B},$$

where $B \geq 0$ is the smallest bound for the bounded function f. Then we choose

$$\mu := \operatorname{argmax} \{\nu(K) \mid \nu \in \{\mu_1, \dots, \mu_n\}\}.$$

Notice that the function f is uniformly continuous on K, as it is bounded and continuous and K is compact. Hence, there is a $k \in \mathbb{N}$ such that for all $|x - x'| < \frac{b-a}{k}$, we have

$$|f(x) - f(x')| < \frac{\varepsilon}{6\mu(K)}.$$

We partition $K = [a, b]$ in k intervals of length $\frac{b-a}{k}$. That is, we have

$$[a,b] = \bigcup_{i=1}^{k} [x_{i-1}, x_i) := \bigcup_{i=1}^{k} \left[a + (i-1) \cdot \frac{b-a}{k}, a + i \cdot \frac{b-a}{k}\right).$$

Let us further define $x^{(i)} := \operatorname{argmin}\{f(x) | x \in [x_{i-1}, x_i)\}$, $i = 1, \dots, k$, and the step function

$$t(x) := \sum_{i=1}^{k} \mathbb{1}_{[x_{i-1}, x_i)}(x) f\left(x^{(i)}\right) + 1_{\{b\}}(x) f\left(x^{(k)}\right).$$

Then we have $t \leq f$ and $|x - x^{(i)}| < \frac{b-a}{k}$ for $x \in [x_{i-1}, x_i)$, $i = 1, \dots, k$. Thus, we have

$$f(x) - t(x) = f(x) - f\left(x^{(i)}\right) < \frac{\varepsilon}{6\mu(K)}. \quad \text{(A.2)}$$

As the construction aims for a function in $\mathcal{S}$, we have to change this step function such that it is continuous. This has to be done with neither violating the order relation with respect to f nor letting the distance to f grow too big. Therefore we connect the „inner, lower corners“ of the step function such that the resulting function is continuous and slightly smaller than the step function. Below, we formally define the function $u(x)$ for $x \in K$ and give a visual aid in Figure A.1.

$$\begin{aligned} u(x) &:= \mathbb{1}_{[x_0,x_1)}(x)\frac{(x_1-x)t(x_0)+(x-x_0)(t(x_0)\wedge t(x_1))}{x_1-x_0} \\ &+ \sum_{i=2}^{k-1}\mathbb{1}_{[x_{i-1},x_i)}(x)\frac{(x_i-x)(t(x_{i-2})\wedge t(x_{i-1}))+(x-x_{i-1})(t(x_{i-1})\wedge t(x_i))}{x_i-x_{i-1}} \\ &+ \mathbb{1}_{[x_{k-1},x_k]}(x)\frac{(x_k-x)(t(x_{k-1})\wedge t(x_k))+(x-x_{k-1})t(x_k)}{x_k-x_{k-1}}. \end{aligned}$$

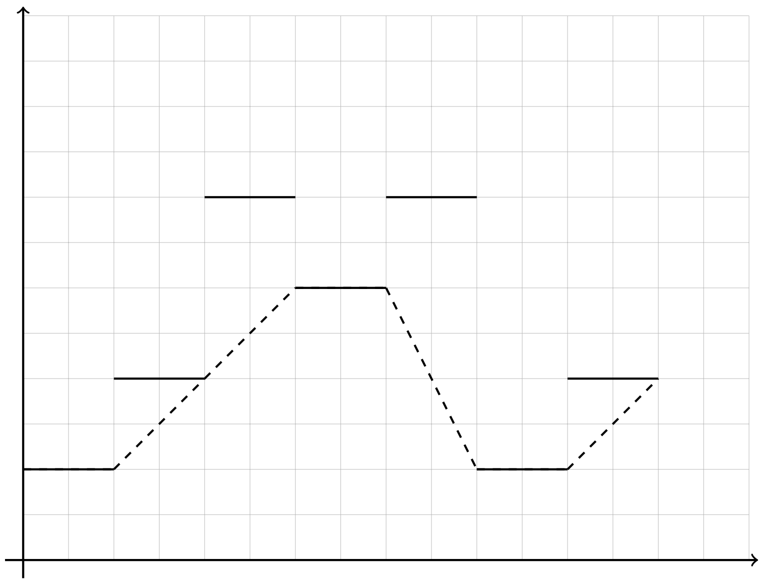


Figure A.1.: Some step function t (solid) and its associated continuous adaption u (dashed).

By definition, we have $u \leq t \leq f$ on K. Hence, the order relation is satisfied. Let $x \in [x_{i-1}, x_i)$, $i = 1, \ldots, k$. Then, for the distance between f and u, we have

$$f(x) - u(x) = |f(x) - u(x)| \leq |f(x) - t(x)| + |t(x) - u(x)| \leq \frac{\varepsilon}{6\mu(K)} + |t(x) - u(x)|,$$

where we use (A.2). Let us now analyze the second summand. We have

$$\begin{aligned} |t(x) - u(x)| &= \left| f\left(x^{(i)}\right) - \frac{(x_i - x)\left(t\left(x_{i-2}\right) \wedge t\left(x_{i-1}\right)\right) + \left(x - x_{i-1}\right)\left(t\left(x_{i-1}\right) \wedge t\left(x_i\right)\right)}{x_i - x_{i-1}} \right| \\ &= \left| \frac{(x_i - x)\left(f\left(x^{(i)}\right) - \left(f\left(x^{(i-1)}\right) \wedge f\left(x^{(i)}\right)\right)\right)}{x_i - x_{i-1}} \right. \\ &\qquad \left. + \frac{(x - x_{i-1})\left(f\left(x^{(i)}\right) - \left(f\left(x^{(i)}\right) \wedge f\left(x^{(i+1)}\right)\right)\right)}{x_i - x_{i-1}} \right|. \end{aligned}$$

We observe that the greatest value results when both minima are not realized by $f\left(x^{(i)}\right)$, as in any other case we get rid of at least one summand. Thus, we have

$$|t(x) - u(x)| \leq \left| \frac{(x_i - x)\left(f\left(x^{(i)}\right) - f\left(x^{(i-1)}\right)\right) + (x - x_{i-1})\left(f\left(x^{(i)}\right) - f\left(^{(i+1)}\right)\right)}{x_i - x_{i-1}} \right|.$$

As $\frac{x_i - x}{x_i - x_{i-1}}, \frac{x - x_{i-1}}{x_i - x_{i-1}} \leq 1$ and $|x^{(i)} - x_{i-1}|, |x_{i-1} - x^{(i-1)}|, |x^{(i)} - x_i|, |x_i - x^{(i+1)}| < \frac{b-a}{k}$, we finally obtain

$$\begin{aligned} |t(x) - u(x)| \leq& \left|f\left(x^{(i)}\right) - f\left(x^{(i-1)}\right)\right| + \left|f\left(x^{(i)}\right) - f\left(x^{(i+1)}\right)\right| \\ \leq& \left|f\left(x^{(i)}\right) - f\left(x_{i-1}\right)\right| + \left|f\left(x_{i-1}\right) - f\left(x^{(i-1)}\right)\right| \\ &+ \left|f\left(x^{(i)}\right) - f\left(x_i\right)\right| + \left|f\left(x_i\right) - f\left(x^{(i+1)}\right)\right| < \frac{4\varepsilon}{6\mu\left(K\right)}. \end{aligned}$$

We stress that a separate investigation of the cases $i = 1$ and $i = k$ is not necessary, as they are covered by the above in the sense that one of the summands in the expression before (⋄) does not appear in these cases. In total, we have

$$f(x) - u(x) < \frac{5\varepsilon}{6\mu(K)}$$

for all $x \in K$ such that in particular

$$\int_K (f(x) - u(x))\ \mu_i(\mathrm{d}x) < \int_K \frac{5\varepsilon}{6\mu(K)}\ \mu_i(\mathrm{d}x) = \frac{5\varepsilon}{6}\frac{\mu_i(K)}{\mu(K)} \leq \frac{5}{6}\varepsilon$$

for all $i = 1, \ldots n$. By construction u is piecewise linear and continuous on K.

It remains to show that we may expand u suitably on $\mathbb{R} \setminus K$ such that the integral condition is satisfied. Therefore let us define u outside of K. For $x \in (-\infty, a)$, we choose $u(x) \coloneqq \max\{-B, l(x)\}$, where $l(x)$ is the linear function with the smallest slope $m_l \geq 0$ such that $l(a) = u(a)$ and $l(x) \leq f(x)$ for all $x \leq a$. Analogously, for $x \in (b, \infty)$, we choose $u(x) \coloneqq \max\{-B, r(x)\}$, where $r(x)$ is the linear function with the greatest slope $m_r \leq 0$ such that $r(b) = u(b)$ and $r(x) \leq f(x)$ for all $x \geq b$.

Then u is piecewise linear and continuous on $\mathbb{R}$ and since u is not differentiable in only finitely many points, we have $u \in \mathcal{S}$. We also have $-u \leq B$ by definition and $f \leq B$ by assumption such that finally, for all $i = 1, \ldots, n$, we have

$$\int_{\mathbb{R}\setminus K} (f(x) - u(x))\mu_i(\mathrm{d}x) \leq \int_{\mathbb{R}\setminus K} 2B\mu_i(\mathrm{d}x) = 2B\mu_i(\mathbb{R} \setminus K) < \frac{2B\varepsilon}{12B} = \frac{\varepsilon}{6}.$$

Thus, $u \in \mathcal{S}$ satisfies $u \leq f$ on $\mathbb{R}$ and $\int_{\mathbb{R}} (f(x) - u(x))\mu_i(\mathrm{d}x) < \varepsilon$ for all $i = 1, \ldots, n$. □

Proof of Lemma 5.5. Let $(\pi_k)_{k\in\mathbb{N}}$ be a sequence in $\Pi_n(\mu_1, \ldots, \mu_n)$ weakly converging to some $\pi \in \Pi_n(\mu_1, \ldots, \mu_n)$, i.e $\pi_k \overset{w}{\to} \pi$. We have to show that

$$A_k \coloneqq \left| \int_{\mathbb{R}^n} f(x_1, \ldots, x_n)(\pi - \pi_k)(\mathrm{d}(x_1, \ldots, x_n)) \right| \overset{k\to\infty}{\to} 0.$$

Let $a > 0$. Then we have

$$\begin{aligned} A_k &\leq \left| \int_{[-a,a]^n} f(x_1, \ldots, x_n)(\pi - \pi_k)(\mathrm{d}(x_1, \ldots, x_n)) \right| \\ &\quad + \left| \int_{\mathbb{R}^n \setminus [-a,a]^n} f(x_1, \ldots, x_n)(\pi - \pi_k)(\mathrm{d}(x_1, \ldots, x_n)) \right| \\ &=: A_k(a) + \varepsilon_a(k). \end{aligned}$$

As $[-a, a]^n$ is compact and f is continuous and thus bounded on $[-a, a]^n$, we have $A_k(a) \to 0$ for all $a > 0$ as $k \to \infty$. This is an immediate consequence of the weak convergence of the sequence $(\pi_k)_{k\in\mathbb{N}}$.

We further observe

$$\begin{aligned} \varepsilon_a(k) &\leq \int_{\mathbb{R}^n \setminus [-a,a]^n} |f(x_1, \ldots, x_n)|(\pi + \pi_k)(\mathrm{d}(x_1, \ldots, x_n)) \\ &\leq \int_{\mathbb{R}^n \setminus [-a,a]^n} K \left(1 + \sum_{i=1}^n f_i(x_i)\right)(\pi + \pi_k)(\mathrm{d}(x_1, \ldots, x_n)), \end{aligned}$$

where we used the triangle inequality and the condition in (5.6). Then, partly calculating the integral, we get

$$\begin{aligned} \varepsilon_a(k) \leq& K \left((\pi + \pi_k)(\mathbb{R}^n \setminus [-a, a]^n) + \sum_{i=1}^n \left(\int_{\mathbb{R}^n \setminus [-a,a]^n} f_i(x_i)(\pi + \pi_k)(\mathrm{d}(x_1, \ldots, x_n)) \right) \right) \\ =& K \left((\pi + \pi_k)(\mathbb{R}^n \setminus [-a, a]^n) \right) + 2K \sum_{i=1}^n \int_{\mathbb{R} \setminus [-a,a]} f_i(x_i)\mu_i(\mathrm{d}x_i). \end{aligned} \tag{A.3}$$

With $a \to \infty$ we have $\mathbb{R}^n \setminus [-a, a]^n \to \emptyset$. Thus, the first summand of (A.3) tends to 0 as $a \to \infty$ for all $k \in \mathbb{N}$. For the second summand we get the same by the integrability of the f_i, $i = 1, \ldots, n$, i.e. we have $\varepsilon_a(k) \to 0$ uniformly for all $k \in \mathbb{N}$ as $a \to \infty$. Hence, for all $\varepsilon > 0$ there is an $a > 0$ such that $\varepsilon_a(k) < \varepsilon$ for all $k \in \mathbb{N}$.

Thus, for all $\varepsilon > 0$ there is an $a > 0$ such that

$$\lim_{k\to\infty} A_k \leq \lim_{k\to\infty} A_k(a) + \lim_{k\to\infty} \varepsilon_a(k) < \varepsilon.$$

Finally, we get $\lim_{k\to\infty} A_k = 0$ as $\varepsilon \to 0$. □

Theorem A.1. *[74, Theorem 45.8] Let $\mathcal{X}, \mathcal{Y}$ be vector spaces such that $\mathcal{X}$ is locally convex, and let $K \subseteq \mathcal{X}$ and $T \subseteq \mathcal{Y}$ be convex sets. Let $f : K \times T \to \mathbb{R}$. If*

1. *K is compact,*
2. *$x \mapsto f(x, y)$ it continuous and convex on K for all $y \in T$,*
3. *$y \mapsto f(x, y)$ is concave on T for all $x \in K$,*

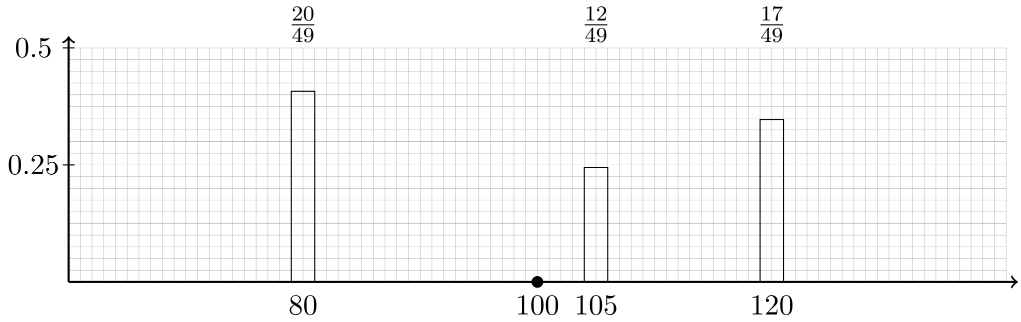
then

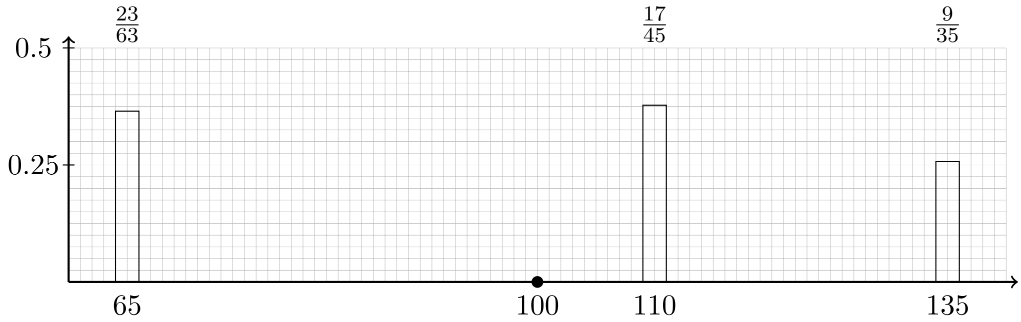
$$\sup_{y\in T} \inf_{x\in K} f(x, y) = \inf_{x\in K} \sup_{y\in T} f(x, y).$$

A.3. Supplements to Chapter 6

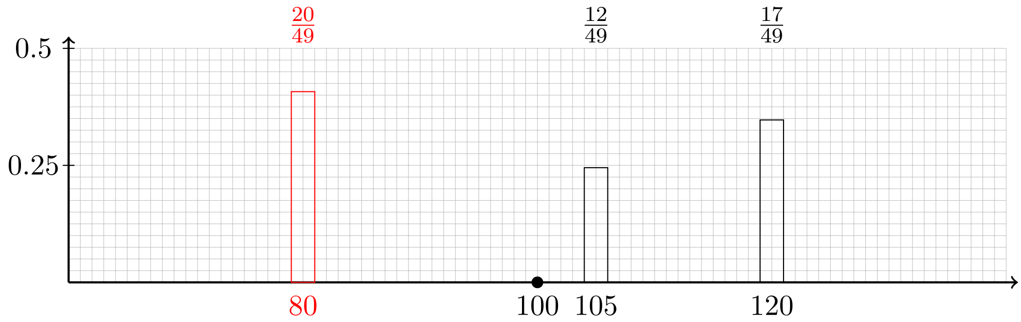
Example A.2. In this example, we complement Example 6.51 by presenting a step by step illustration of the algorithmic determination of the left monotone martingale transport plan and the super hedging strategy. We start with the illustration of Algorithm 6.40.

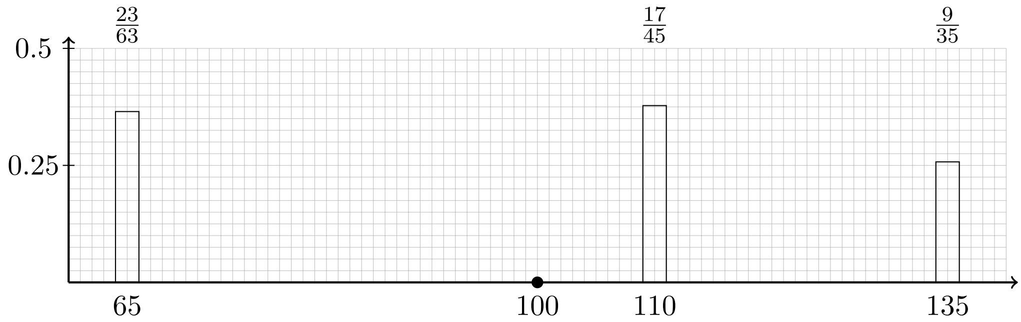
1. The measures μ and ν illustrated.



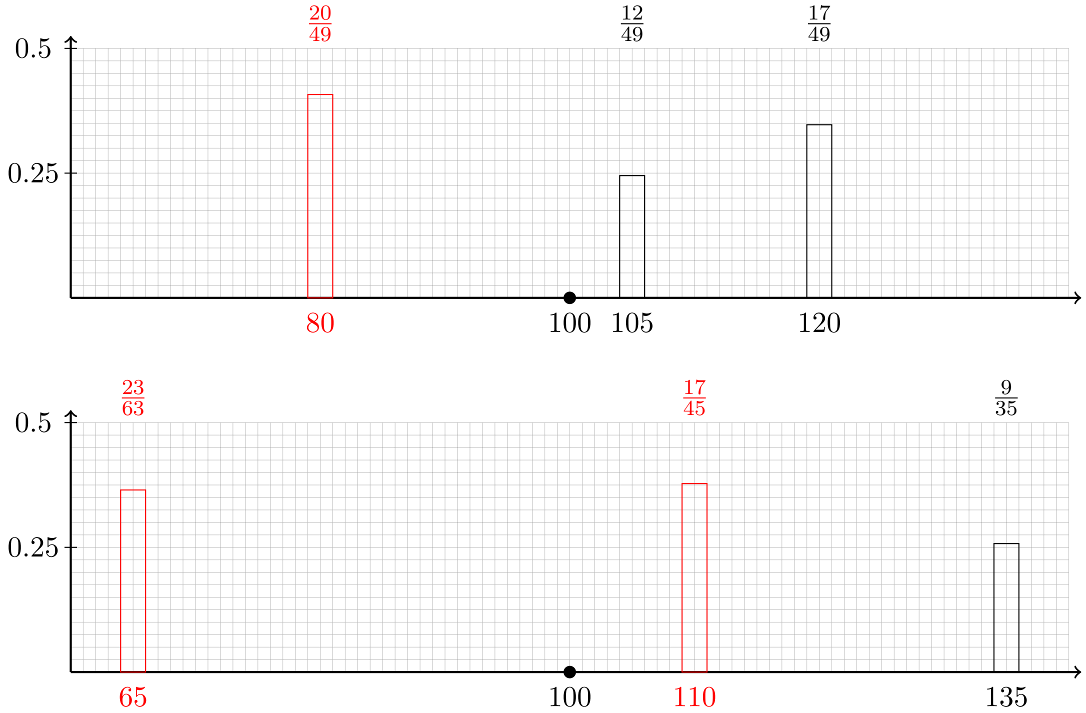


2. We start coupling the mass of the smallest atom of μ, δ_{80}.

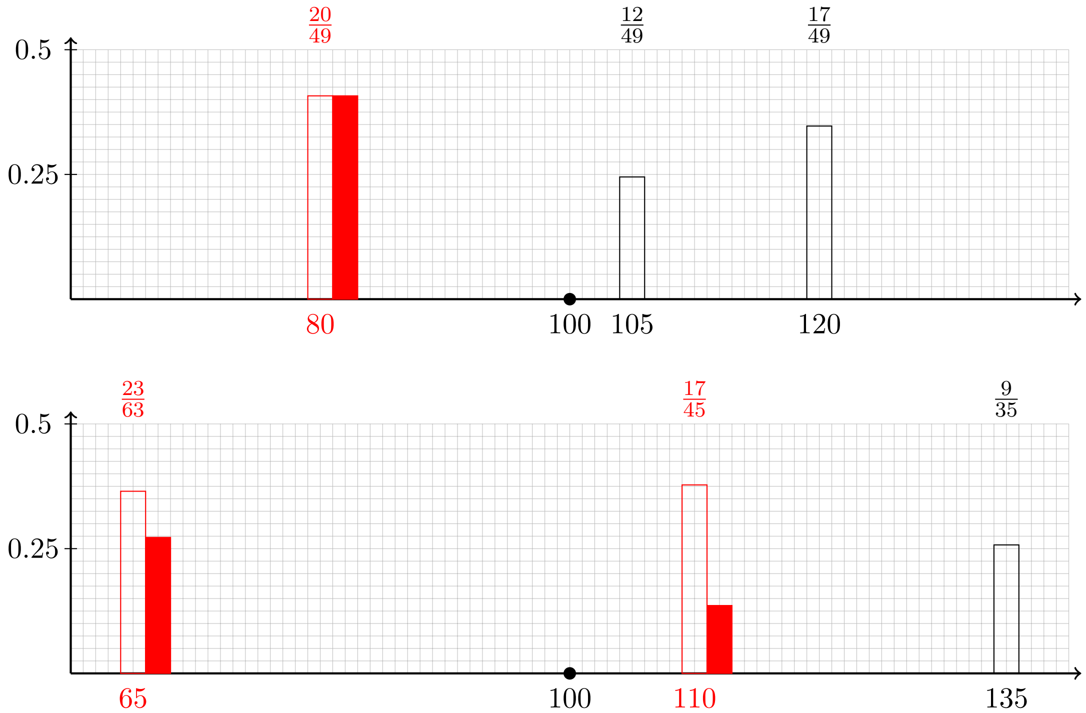




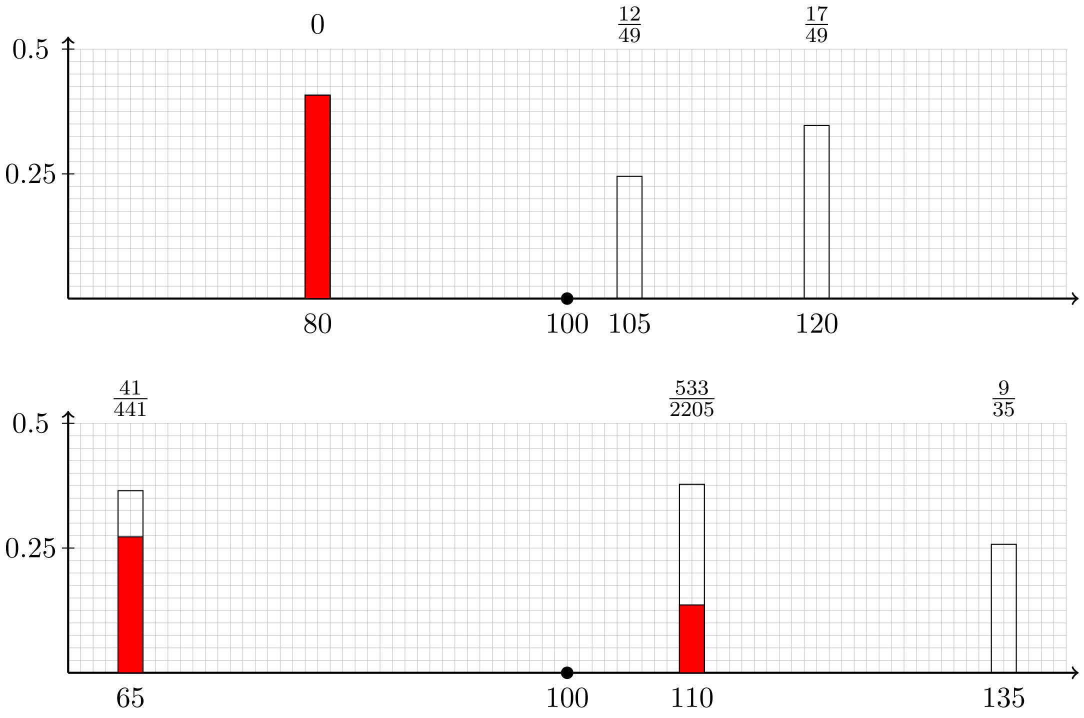
3. As δ_{80} is no atom of ν, we couple with δ_{65} and δ_{110}.



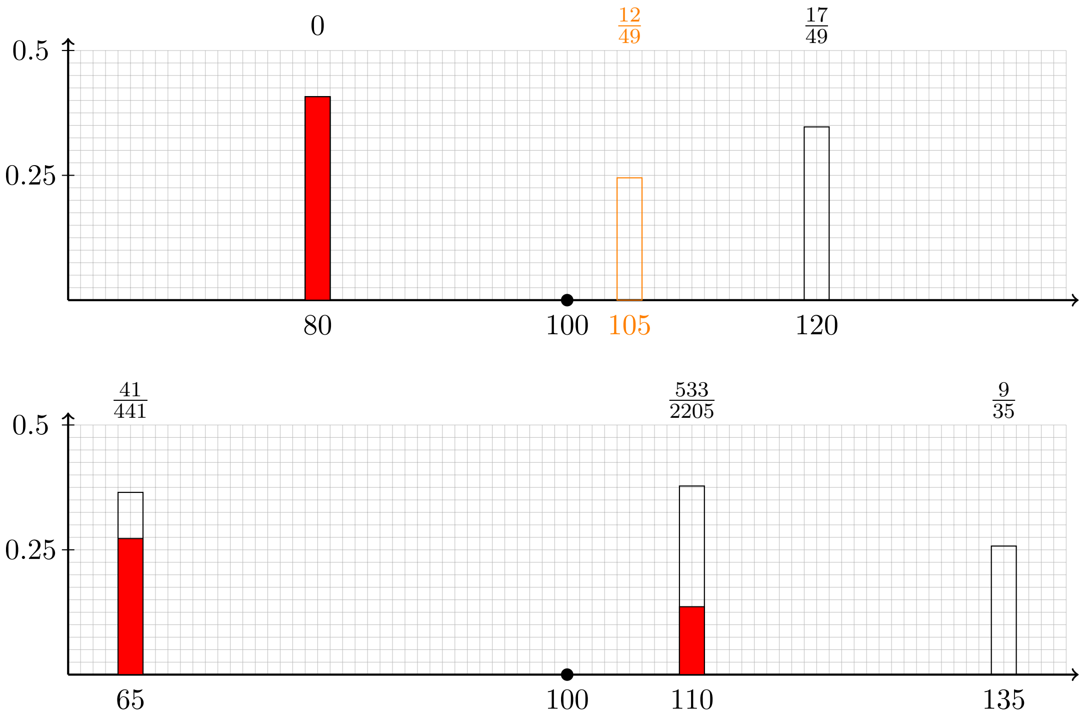
4. We determine the weights $q_{x_d y_d} = \frac{40}{147}$ and $q_{x_d y_m} = \frac{20}{147}$ and check the suitability.



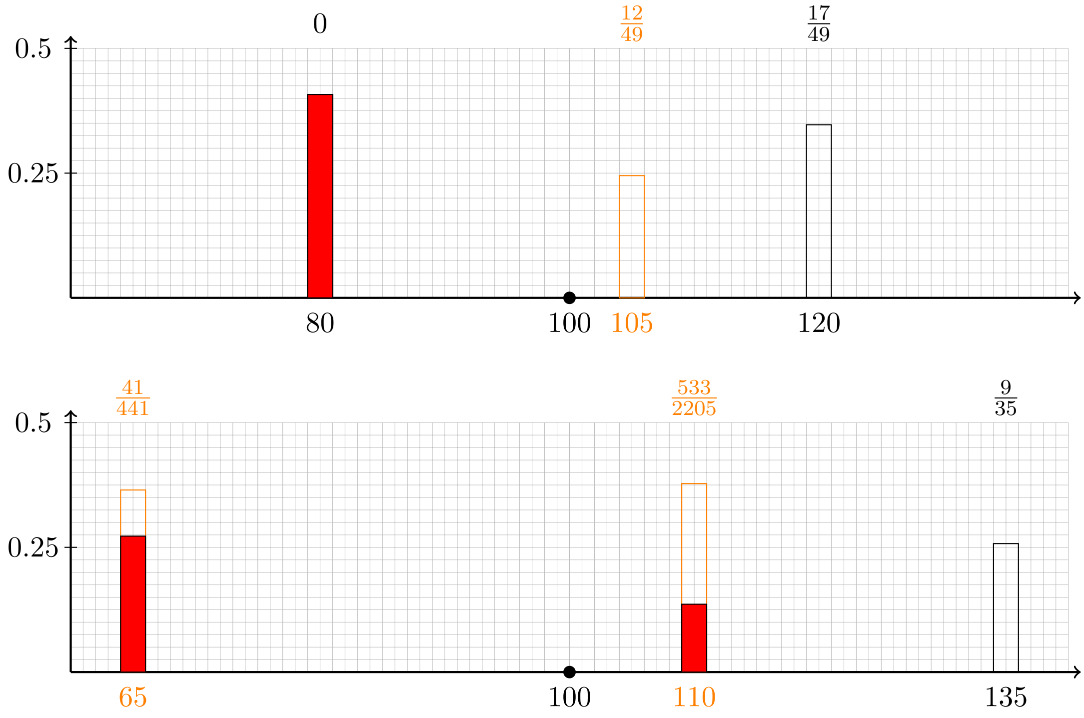
5. We proceed with the measures μ' and ν'.



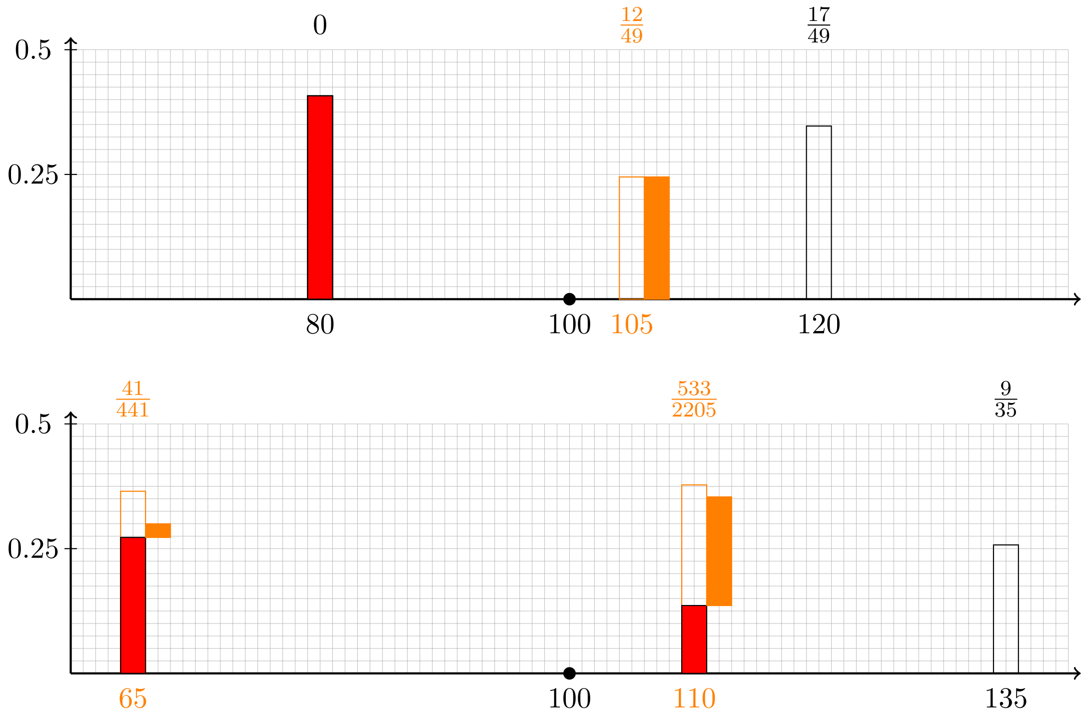
6. We couple the smallest atom of μ', δ_{105}.



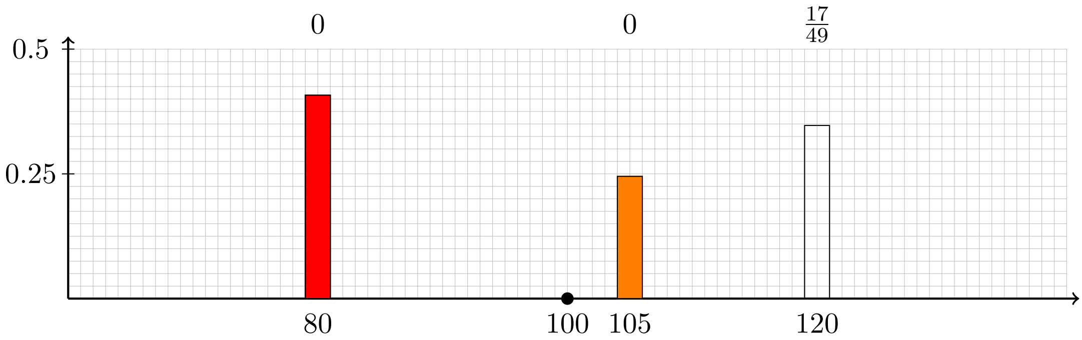
7. As δ_{105} is no atom of ν', we couple with δ_{65} and δ_{110}.

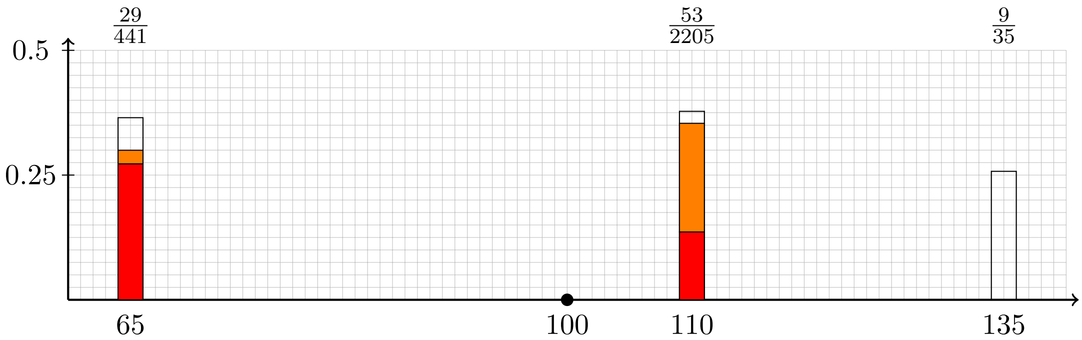


8. We determine the weights $q_{x_m y_d} = \frac{4}{147}$ and $q_{x_m y_m} = \frac{32}{147}$ and check for suitability.

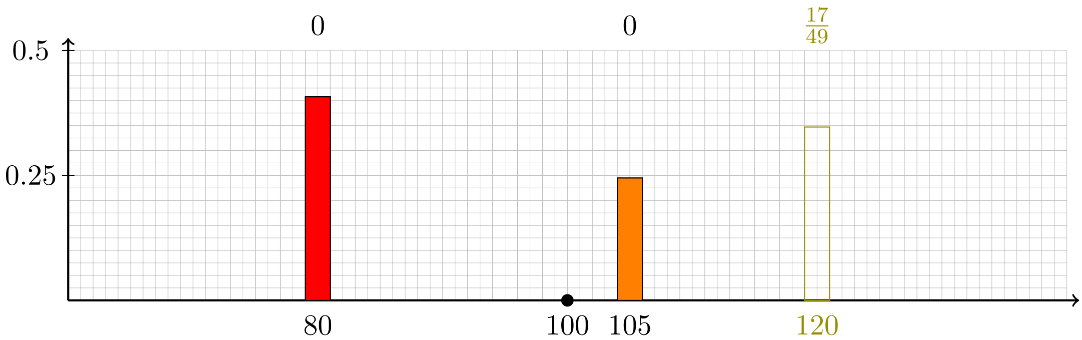


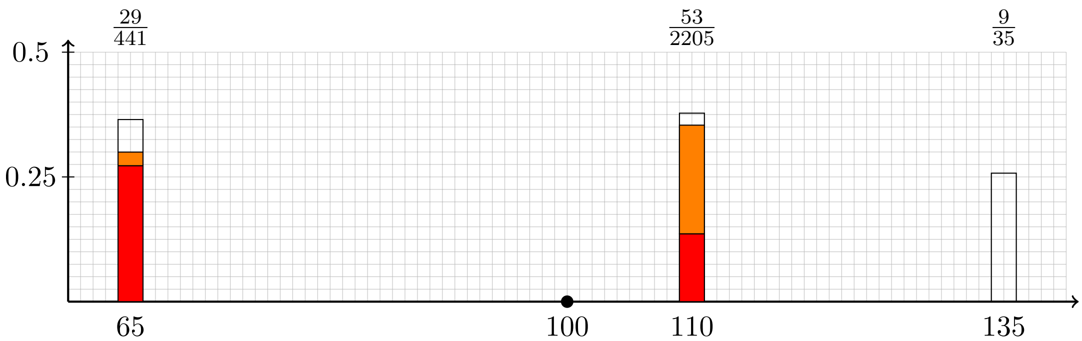
9. We proceed with the measures μ'' and ν".



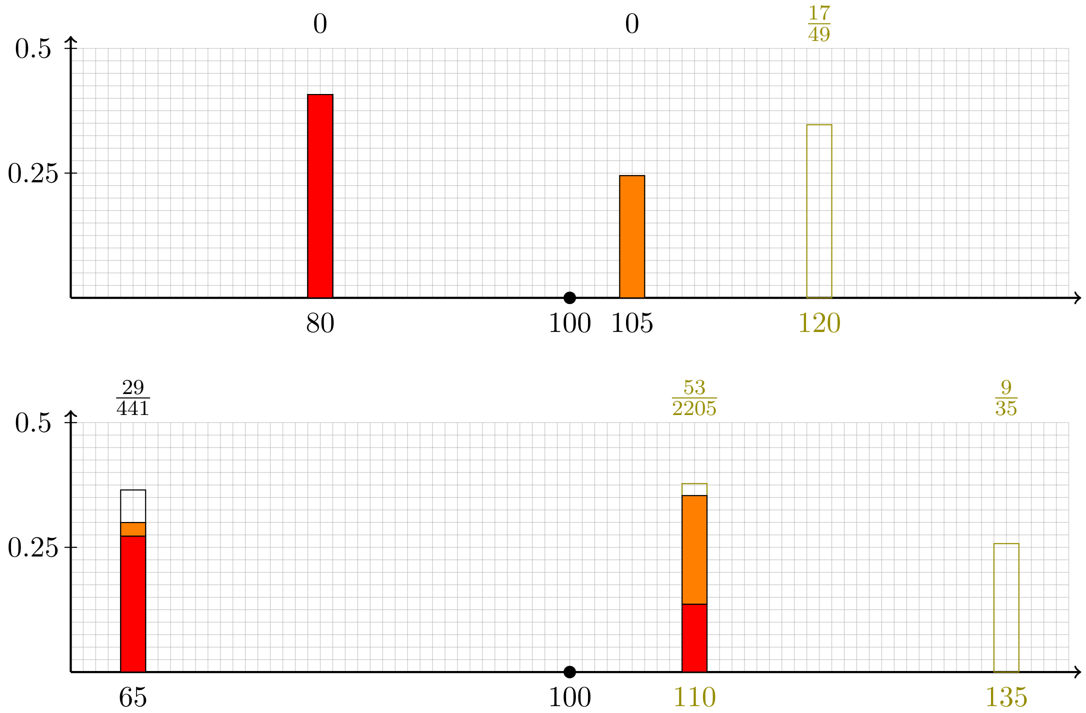


10. We finish the algorithm coupling the smallest atom of μ'', δ_{120}.

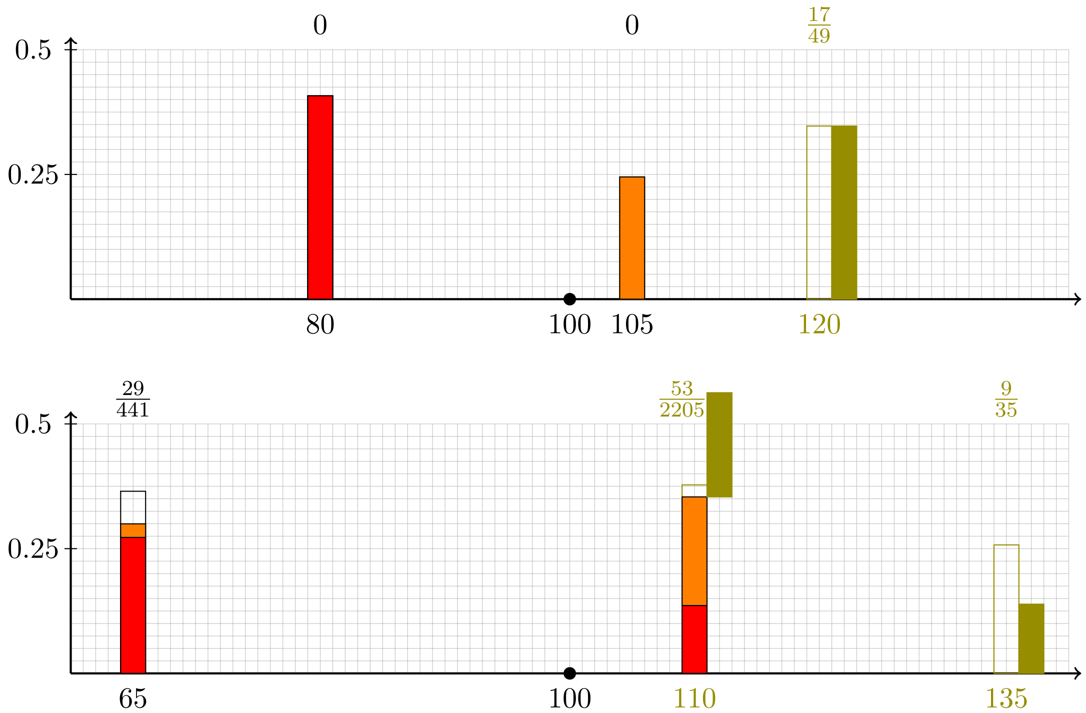




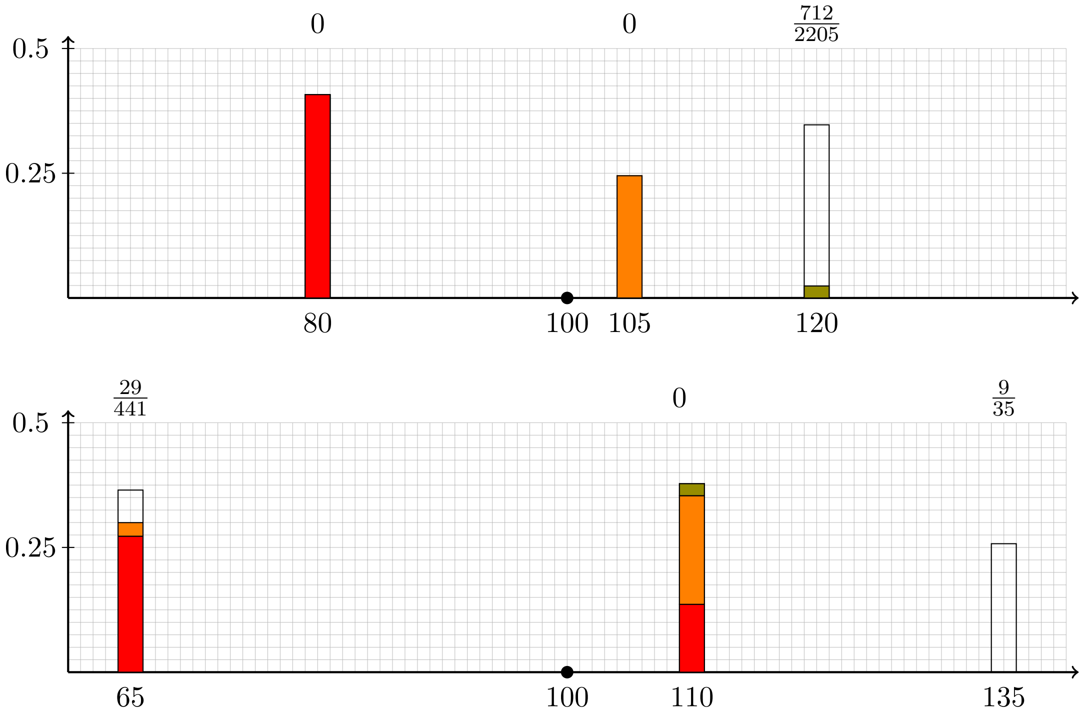
11. As δ_{120} is no atom of ν'', we couple with δ_{110} and δ_{135}.



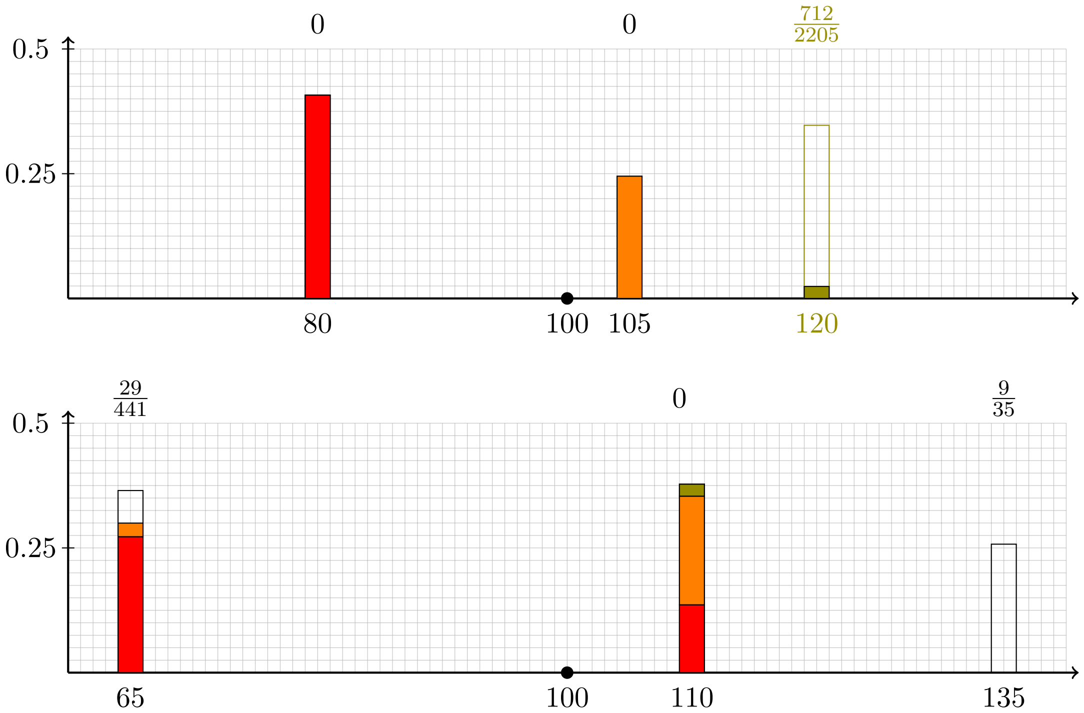
12. We determine the weights $q_{x_u y_m} = \frac{51}{245}$ and $q_{x_u y_u} = \frac{34}{245}$ and check for suitability.



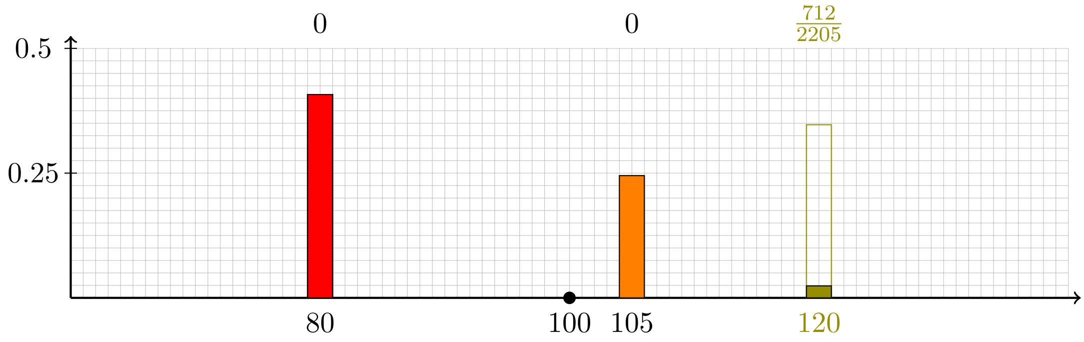
13. We choose $q_{x_u y_m} = \vartheta''_m = \frac{53}{2205}$.

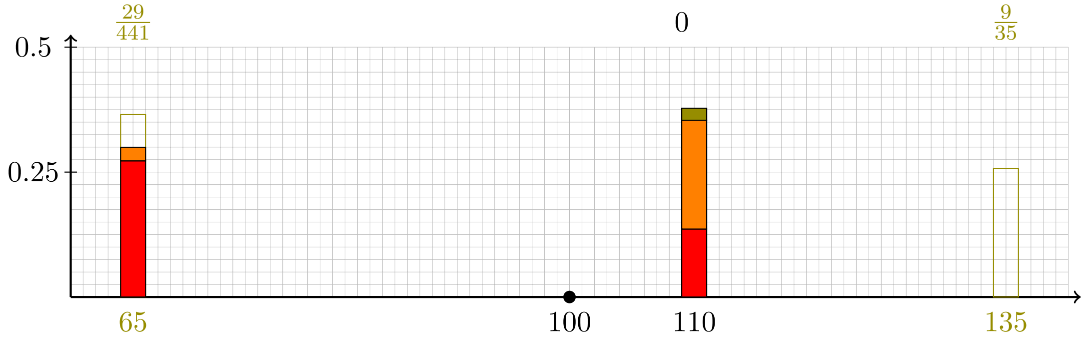


14. We couple the remaining mass of δ_{120}.

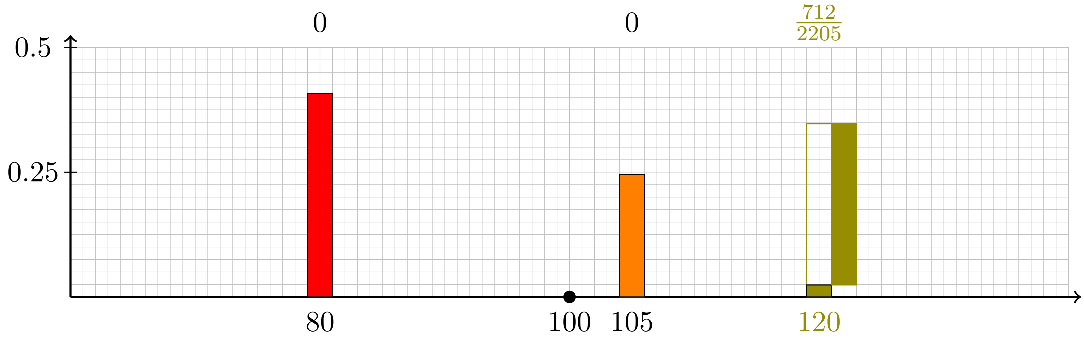


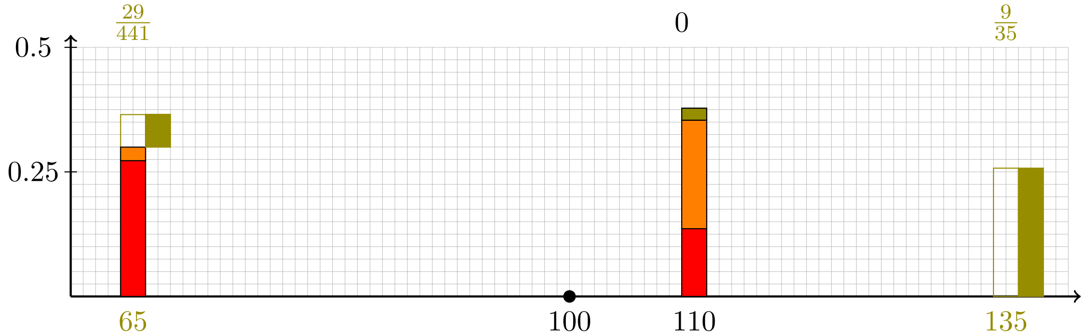
15. We couple the mass of δ_{120} with δ_{65} and δ_{135}.



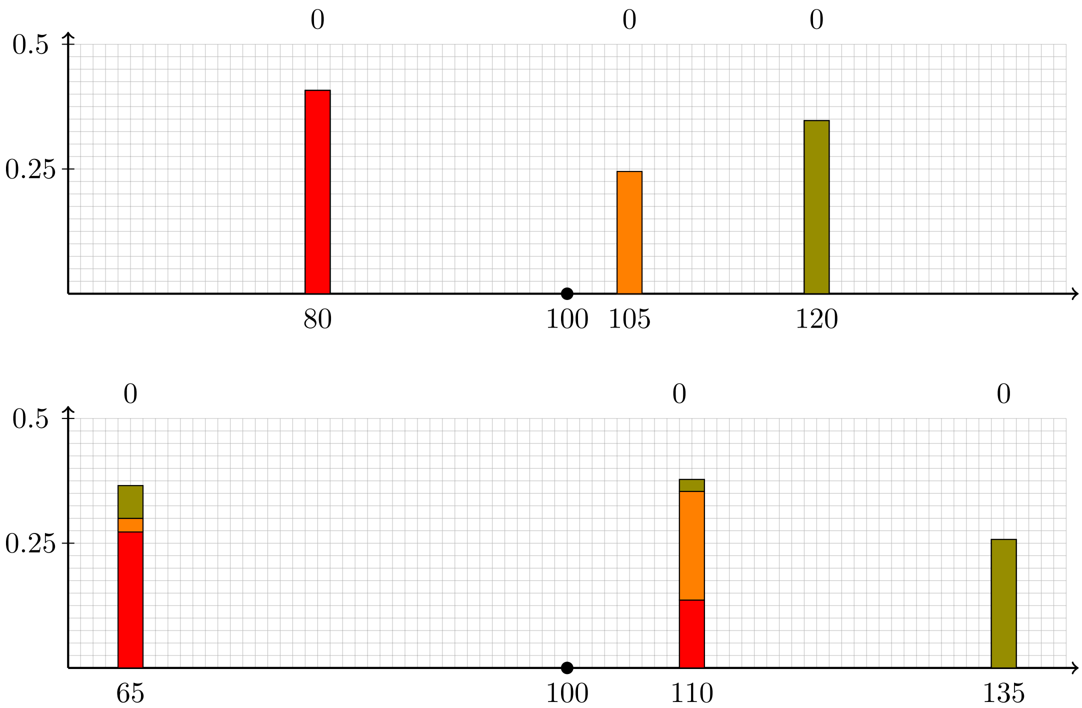


16. We determine the weights $q_{x_u y_d} = \frac{29}{441}$ and $q_{x_u y_u} = \frac{9}{35}$ and check for suitability.

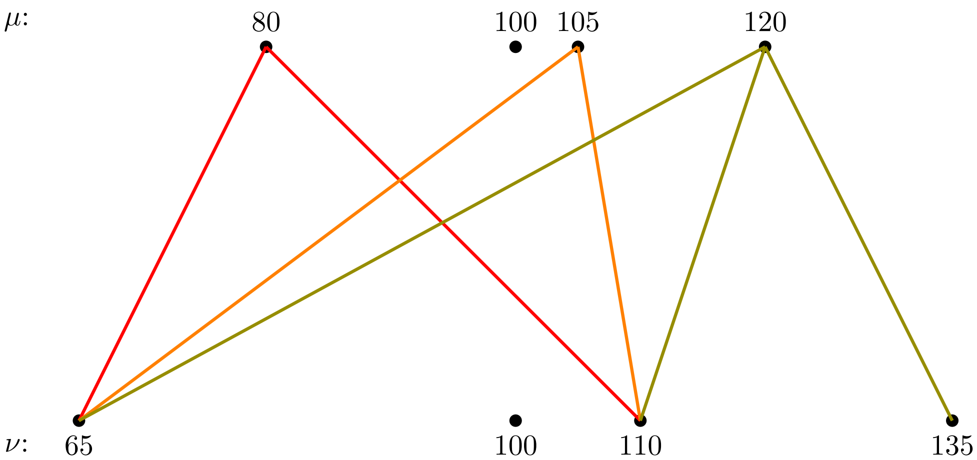




17. We determined the left monotone martingale transport plan $\mathbb{Q}_{lc}(\mu, \nu)$.



18. The left monotonicity is obvious.



Theoretically, the algorithm does not apply when coupling the mass of δ_{120}, as it is the last atom of μ to couple. Instead, we should just transport its mass to the residual of ν. However, in order to illustrate a situation in which a case distinction is necessary, we decided to determine the transport weights using the algorithm.

Let us now construct a super hedging strategy. For this purpose, observe that (x_d, y_u) and (x_m, y_u) are the only points in $\mathrm{supp}(\mu) \times \mathrm{supp}(\nu)$ not contained in the support of $\mathbb{Q}_{lc}(\mu, \nu)$. Thus, ultimately we have to solve the inequalities

$$\varphi(x_d) + \psi(y_u) + h(x_d)(y_u - x_d) \geq c(x_d, y_u), \tag{A.4}$$

$$\varphi(x_m) + \psi(y_u) + h(x_m)(y_u - x_m) \geq c(x_m, y_u). \tag{A.5}$$

Before we solve this linear inequality system, let us apply Algorithm 6.49 in order to guarantee equality on the support of $\mathbb{Q}_{lc}(\mu, \nu)$. We skip the detailed calculations and immediately state the resulting numbers.

We start with $x_u = 120$. As $x_u \in X_\Gamma^\psi = \{y_d, y_m, y_u\}$, we have

$$\begin{aligned} h(x_u) &= \frac{3 - (\psi(y_u) - \psi(y_d))}{70}, \\ \psi(y_m) &= -\frac{13}{14} + \frac{9}{14}\psi(y_u) + \frac{5}{14}\psi(y_d), \\ \varphi(x_u) &= \frac{33}{14} - \frac{11}{14}\psi(y_u) - \frac{3}{14}\psi(y_d). \end{aligned}$$

For $x_m = 105$, we have $x_m \in X_\Gamma^h = \{y_d, y_m\}$ and hence

$$\begin{aligned} h(x_m) &= \frac{3 - (\psi(y_u) - \psi(y_d))}{70}, \\ \varphi(x_m) &= \frac{5}{7} - \frac{4}{7}\psi(y_u) - \frac{3}{7}\psi(y_d). \end{aligned}$$

Finally, for $x_d = 80$, we have $x_d \in X_\Gamma^h = \{y_d, y_m\}$ and hence

$$\begin{aligned} h(x_d) &= \frac{41}{630} - \frac{1}{70}\psi(y_u) + \frac{1}{70}\psi(y_d), \\ \varphi(x_d) &= -\frac{85}{42} - \frac{3}{14}\psi(y_u) - \frac{11}{14}\psi(y_d). \end{aligned}$$

If we now plugin these numbers as well as (x_d, y_u) and (x_m, y_u) into the inequalities (A.4) and (A.5), then we obtain the inequalities $\frac{14}{9} \geq 0$ and $2 \geq 1$, which are satisfied independently of the choices of $\psi(y_d)$ and $\psi(y_u)$. Thus, we may choose $\psi(y_d) = \psi(y_u) = 0$. If we now plugin those values into the above hedging functions, then we obtain the super hedging strategy presented in Example 6.51. △

A.4. Supplements to Chapter 7

Remark A.3. In this remark, we prove that $C_{\mu_n^d}$ is recovered when considering the definition of a call option price functions corresponding to the measure μ_n^d. Let $k \in [k_j^n, k_{j+1}^n)$. Then we have

$$C_{\mu_n^d}(k) = \int_{\mathbb{R}_+} (x-k)^+ \mu_n^d(\mathrm{d}x) = \int_k^K (x-k)\mu_n^d(\mathrm{d}x) = \sum_{i=j+1}^{2^n} \omega_i^n (k_i^n - k).$$

If we now plugin the definition of ω_i^n, then we have

$$\begin{aligned}
C_{\mu_n^d}(k) &= \sum_{i=j+1}^{2^n} \left(\frac{C_\mu(k_{i+1}^n) - C_\mu(k_i^n)}{k_{i+1}^n - k_i^n} - \frac{C_\mu(k_i^n) - C_\mu(k_{i-1}^n)}{k_i^n - k_{i-1}^n} \right) (k_i^n - k) \\
&= \sum_{i=j+1}^{2^n} \left(\frac{C_\mu(k_{i+1}^n) - C_\mu(k_i^n)}{k_{i+1}^n - k_i^n} \right) (k_i^n - k) - \sum_{i=j+1}^{2^n} \left(\frac{C_\mu(k_i^n) - C_\mu(k_{i-1}^n)}{k_i^n - k_{i-1}^n} \right) (k_i^n - k) \\
&= \sum_{i=j+1}^{2^n-1} \left(\frac{C_\mu(k_{i+1}^n) - C_\mu(k_i^n)}{k_{i+1}^n - k_i^n} \right) (k_i^n - k) + 0 \\
&\quad - \left(\frac{C_\mu(k_{j+1}^n) - C_\mu(k_j^n)}{k_{j+1}^n - k_j^n} \right) (k_{j+1}^n - k) - \sum_{i=j+2}^{2^n} \left(\frac{C_\mu(k_i^n) - C_\mu(k_{i-1}^n)}{k_i^n - k_{i-1}^n} \right) (k_i^n - k),
\end{aligned}$$

where we just split off the last and the first summand of the first and the second sum respectively. Then shifting the index and writing the two remaining sums as one, we get

$$\begin{aligned}
C_{\mu_n^d}(k) &= \sum_{i=j+2}^{2^n} \left(\frac{C_\mu(k_i^n) - C_\mu(k_{i-1}^n)}{k_i^n - k_{i-1}^n} \right) (k_{i-1}^n - k_i^n) - \left(C_\mu(k_{j+1}^n) - C_\mu(k_j^n) \right) \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} \\
&= (-1) \cdot \left(\sum_{i=j+2}^{2^n} (C_\mu(k_i^n) - C_\mu(k_{i-1}^n)) \right) - \left(C_\mu(k_{j+1}^n) - C_\mu(k_j^n) \right) \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} \\
&= C_\mu(k_{j+1}^n) - \left(C_\mu(k_{j+1}^n) - C_\mu(k_j^n) \right) \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} \\
&= \frac{k - k_j^n}{k_{j+1}^n - k_j^n} C_\mu(k_{j+1}^n) + \frac{k_{j+1}^n - k}{k_{j+1}^n - k_j^n} C_\mu(k_j^n),
\end{aligned}$$

where in the third equality we simplify the telescope sum and use that $C_\mu(k_{2^n}^n) = 0$.

We also show that μ_n^d and ν_n^d both have mass and expected value equal to one. Having mass equal to 1 is immediately clear for both measures by definition of the weights. Concerning the expected value, let $X \sim \mu_n^d$ and $Y \sim \nu_n^d$. Then we have

$$\begin{aligned}
\mathbb{E}_{\mu_n^d}[X] &= \sum_{j=0}^{2^n} \omega_j^n k_j^n = \sum_{j=0}^{2^n} \frac{2^n}{K} \left(C_\mu(k_{j+1}^n) - 2C_\mu(k_j^n) + C_\mu(k_{j-1}^n) \right) k_j^n \\
&= \sum_{j=0}^{2^n} j \left(C_\mu(k_{j+1}^n) - 2C_\mu(k_j^n) + C_\mu(k_{j-1}^n) \right) \\
&= \sum_{j=0}^{2^n} jC_\mu(k_{j+1}^n) - 2\sum_{j=0}^{2^n} jC_\mu(k_j^n) + \sum_{j=0}^{2^n} jC_\mu(k_{j-1}^n) \\
&= \sum_{j=1}^{2^n+1} (j-1)C_\mu(k_j^n) - 2\sum_{j=0}^{2^n} jC_\mu(k_j^n) + \sum_{j=-1}^{2^n-1} (j+1)C_\mu(k_j^n) \\
&= 2^n \cdot C_\mu(k_{2^n+1}^n) + (2^n - 1) \cdot C_\mu(k_{2^n}^n) \\
&\quad - 2 \cdot 2^n \cdot C_\mu(k_{2^n}^n) - 2 \cdot 0 \cdot C_\mu(k_0^n) \\
&\quad + 0 \cdot C_\mu(k_{-1}^n) + C_\mu(k_0^n),
\end{aligned}$$

as only those summands of the three sums remain present that are not contained in all sums. Rearranging and using some of the earlier definitions, we have

$$\begin{aligned}\mathbb{E}_{\mu_n^d}[X] &= 2^n \left(C_\mu(k_{2^n+1}^n) - C_\mu(k_{2^n}^n)\right) - C_\mu(k_{2^n}^n) + C_\mu(k_0^n) \\ &= \frac{C_\mu(k_{2^n+1}^n) - C_\mu(k_{2^n}^n)}{k_{2^n+1}^n - k_{2^n}^n} - C_\mu(K) + C_\mu(0) = C_\mu(0) = s_0 = 1.\end{aligned}$$

Analogously, we obtain $\mathbb{E}_{\nu_n^d}[Y] = 1$. ◊

Example A.4. In this example, we show that the convergence speed proven in Theorem 7.11 is ideal in the sense that it can not be improved. This means that the speed in our empirical investigations is better than the guaranteed convergence speed. For this purpose, we consider the two discrete measures

$$\mu = \frac{1}{4}\delta_1 + \frac{1}{2}\delta_{\frac{7}{3}} + \frac{1}{4}\delta_3 \quad \text{and} \quad \nu = \frac{1}{4}\delta_0 + \frac{1}{2}\delta_{\frac{7}{3}} + \frac{1}{4}\delta_4.$$

These measures have mass 1, mean $\frac{7}{3}$ and the call option price functions

$$\begin{aligned}C_\mu(k) &= \left(\frac{13}{6} - k\right)\mathbb{1}_{\{0\le k\le 1\}} + \left(\frac{23}{12} - \frac{3}{4}k\right)\mathbb{1}_{\{1\le k\le \frac{7}{3}\}} + \left(\frac{3}{4} - \frac{1}{4}k\right)\mathbb{1}_{\{\frac{7}{3}\le k\le 3\}}, \\ C_\nu(\ell) &= \left(\frac{13}{6} - \frac{3}{4}\ell\right)\mathbb{1}_{\{0\le \ell\le \frac{7}{3}\}} + \left(1 - \frac{1}{4}\ell\right)\mathbb{1}_{\{\frac{7}{3}\le \ell\le 4\}}.\end{aligned}$$

We easily see that $C_\mu \le C_\nu$ and thus $\mu \le_c \nu$.

By the requirements of Theorem 7.11, we have $c_{xyy} > 0$. Hence, the upper price bounds are realized by the left monotone martingale transport $\mathbb{Q}_{lc}(\mu, \nu)$. Using Algorithm 6.40, we easily verify that

$$\mathbb{Q}_{lc}(\mu,\nu) = \frac{1}{7}\delta_{1,0} + \frac{3}{28}\delta_{1,\frac{7}{3}} + \frac{5}{112}\delta_{\frac{7}{3},0} + \frac{11}{28}\delta_{\frac{7}{3},\frac{7}{3}} + \frac{1}{16}\delta_{\frac{7}{3},4} + \frac{1}{16}\delta_{3,0} + \frac{3}{16}\delta_{3,4}.$$

The most simple payoff function satisfying the martingale Spence Mirrlees condition is $c(x, y) = xy^2$. For this function, we get

$$P_2^c(\mu,\nu) = \frac{3}{28}c\left(1,\frac{7}{3}\right) + \frac{11}{28}c\left(\frac{7}{3},\frac{7}{3}\right) + \frac{1}{16}c\left(\frac{7}{3},4\right) + \frac{3}{16}c(3,4) = \frac{913}{54}.$$

In order to prove the optimality of the convergence speed, we calculate the approximating measures μ_n^d and ν_n^d and the associated price bounds for general $n \ge 3$. The measures have the structure

$$\mu_n^d = \frac{1}{4}(\delta_1 + \delta_3) + \mu_n^r \text{ and } \nu_n^d = \frac{1}{4}(\delta_0 + \delta_4) + \nu_n^r,$$

for all $n \in \mathbb{N}$, where μ_n^r and ν_n^r are also measures with two atoms close to $\frac{7}{3}$ each. This follows from the determination technique of the approximating measures based on the associated piecewise linearly interpolated call option price functions $C_{\mu_n^d}$ and $C_{\nu_n^d}$. Indeed, these deviate from the functions C_μ and C_ν only on the interval $\left(k_{j(n)}^n, k_{j(n)+1}^n\right)$, where $j(n)$ is such that $k_{j(n)}^n < \frac{7}{3} < k_{j(n)+1}^n$. Consequently, $k_{j(n)}^n$ and $k_{j(n)+1}^n$ are the atoms of the residual measures.

We determine the general structure of these values and denote $j = j(n)$ and $k_j = k^n_{j(n)}$ for the rest of the example. We have

$$k_j < \frac{7}{3} < k_{j+1} \iff 4 \cdot \frac{j}{2^n} < \frac{7}{3} < 4 \cdot \frac{j+1}{2^n} \iff j < \frac{7}{3} \cdot \frac{2^n}{4} < j + 1.$$

As $j \in \mathbb{N}$, we clearly have

$$j = \left\lfloor \frac{7}{3} \cdot \frac{2^n}{4} \right\rfloor = \begin{cases} \frac{7}{3} \cdot \frac{2^n}{4} - \frac{1}{3}, & n \text{ even}, \\ \frac{7}{3} \cdot \frac{2^n}{4} - \frac{2}{3}, & n \text{ odd}, \end{cases}$$

from which we immediately get

$$k_j = 4 \cdot \frac{j}{2^n} = \begin{cases} \frac{7}{3} - \frac{4}{3 \cdot 2^n}, & n \text{ even}, \\ \frac{7}{3} - \frac{8}{3 \cdot 2^n}, & n \text{ odd}, \end{cases} \quad \text{and} \quad k_{j+1} = 4 \cdot \frac{j+1}{2^n} = \begin{cases} \frac{7}{3} + \frac{8}{3 \cdot 2^n}, & n \text{ even}, \\ \frac{7}{3} + \frac{4}{3 \cdot 2^n}, & n \text{ odd}. \end{cases}$$

The masses of μ^d_n and ν^d_n in the atoms δ_{k_j} and $\delta_{k_{j+1}}$ are the differences of the slopes of $C_{\mu^d_n}$ and $C_{\nu^d_n}$ on the intervals (k_j, k_{j+1}) and (k_{j-1}, k_j), and (k_{j+1}, k_{j+2}) and (k_j, k_{j+1}) respectively. As C_μ and C_ν have the same slopes in these areas, we know that the masses of the atoms are equal for μ^d_n and ν^d_n. Hence, we have

$$\omega^n_j = \vartheta^n_j = m^n_j - \left(-\frac{3}{4}\right) \text{ and } \omega^n_{j+1} = \vartheta^n_{j+1} = -\frac{1}{4} - m^n_j,$$

where $m^n_j = \frac{C_\mu(k_{j+1}) - C_\mu(k_j)}{k_{j+1} - k_j} = \frac{2^n}{4}(C_\mu(k_{j+1}) - C_\mu(k_j)) = \frac{2^n}{4}(C_\nu(k_{j+1}) - C_\nu(k_j))$. Using the representations of C_μ and C_ν, we deduce

$$\begin{aligned} m^n_j &= \frac{2^n}{4}\left(1 - \frac{1}{4}k_{j+1} - \left(\frac{13}{6} - \frac{3}{4}k_j\right)\right) \\ &= \frac{2^n}{4}\left(-\frac{7}{6} - \frac{1}{4}(k_{j+1} - k_j) + \frac{1}{2}k_j\right) \\ &= \frac{2^n}{4}\left(\frac{1}{2}k_j - \frac{7}{6}\right) - \frac{1}{4} \\ &= \begin{cases} \frac{2^n}{4}\left(\frac{1}{2}\left(\frac{7}{3} - \frac{4}{3 \cdot 2^n} - \frac{7}{6}\right)\right) - \frac{1}{4} = -\frac{5}{12}, & n \text{ even}, \\ \frac{2^n}{4}\left(\frac{1}{2}\left(\frac{7}{3} - \frac{8}{3 \cdot 2^n} - \frac{7}{6}\right)\right) - \frac{1}{4} = -\frac{7}{12}, & n \text{ odd}. \end{cases} \end{aligned}$$

This finally implies

$$\omega^n_j = \vartheta^n_j = \begin{cases} \frac{1}{3}, & n \text{ even}, \\ \frac{1}{6}, & n \text{ odd}, \end{cases} \quad \text{and} \quad \omega^n_{j+1} = \vartheta^n_{j+1} = \begin{cases} \frac{1}{6}, & n \text{ even}, \\ \frac{1}{3}, & n \text{ odd}. \end{cases}$$

In total, we have the general structure

$$\begin{aligned} \mu^d_n &= \frac{1}{4}(\delta_1 + \delta_3) + \frac{1}{6}\left(\delta_{k_j} + \delta_{k_{j+1}}\right) + \frac{1}{6}\left(\delta_{k_j}\mathbb{1}_{\{n \text{ even}\}} + \delta_{k_{j+1}}\mathbb{1}_{\{n \text{ odd}\}}\right), \\ \nu^d_n &= \frac{1}{4}(\delta_0 + \delta_4) + \frac{1}{6}\left(\delta_{k_j} + \delta_{k_{j+1}}\right) + \frac{1}{6}\left(\delta_{k_j}\mathbb{1}_{\{n \text{ even}\}} + \delta_{k_{j+1}}\mathbb{1}_{\{n \text{ odd}\}}\right). \end{aligned}$$

By construction we have $\mu_n^d \leq_c \nu_n^d$. Thus we may calculate $\mathbb{Q}_{lc}(\mu_n^d, \nu_n^d)$. In the following, let $n \geq 3$ be even. Then we have to transport the mass $\frac{1}{4}$ of δ_1 to δ_0 and δ_{k_j} such that

$$\begin{aligned} q_{1,0} + q_{1,k_j} &= \frac{1}{4}, \\ 0 \cdot q_{1,0} + k_j \cdot q_{1,k_j} &= \frac{1}{4} \cdot 1. \end{aligned}$$

These equalities imply $q_{1,0} = \frac{k_j - 1}{4k_j}$ and $q_{1,k_j} = \frac{1}{4k_j}$. As n is even, we have $\omega_{k_j}^n = \vartheta_{k_j}^n = \frac{1}{3}$. Since $k_j \geq 1$, we obviously have $\frac{1}{4k_j} < \frac{1}{3}$.

Now we transport as much mass as possible from δ_{k_j} to δ_{k_j}. Thus, we have $q_{k_j,k_j} = \frac{1}{3} - \frac{1}{4k_j}$. The leftover mass is transported to δ_0 and $\delta_{k_{j+1}}$ according to

$$\begin{aligned} q_{k_j,0} + q_{k_j,k_{j+1}} &= \frac{1}{4k_j}, \\ 0 \cdot q_{k_j,0} + k_{j+1} \cdot q_{k_j,k_{j+1}} &= \frac{1}{3} k_j - \left(\frac{1}{3} - \frac{1}{4k_j} \right) k_j = \frac{1}{4}. \end{aligned}$$

Thus, we obtain $q_{k_j,0} = \frac{1}{4k_j} - \frac{1}{4k_{j+1}}$ and $q_{k_j,k_{j+1}} = \frac{1}{4k_{j+1}}$.

Now we transport as much mass as possible from $\delta_{k_{j+1}}$ to $\delta_{k_{j+1}}$. Therefore we have $q_{k_{j+1},k_{j+1}} = \frac{1}{6} - \frac{1}{4k_{j+1}}$. The leftover mass is transported to δ_0 and δ_4 according to

$$\begin{aligned} q_{k_{j+1},0} + q_{k_{j+1},4} &= \frac{1}{6} - \frac{1}{4k_{j+1}}, \\ 0 \cdot q_{k_{j+1},0} + 4 \cdot q_{k_{j+1},4} &= \frac{1}{6} k_{j+1} - \left(\frac{1}{6} - \frac{1}{4k_{j+1}} \right) k_{j+1} = \frac{1}{4}. \end{aligned}$$

This yields $q_{k_{j+1},0} = \frac{1}{4k_{j+1}} - \frac{1}{16}$ and $q_{k_{j+1},4} = \frac{1}{16}$.

Finally, two equations have to be satisfied, namely

$$\begin{aligned} q_{3,0} + q_{3,4} &= \frac{1}{4}, \\ 0 \cdot q_{3,0} + 4 \cdot q_{3,4} &= \frac{1}{4} \cdot 3, \end{aligned}$$

which is satisfied by $q_{3,0} = \frac{1}{16}$ and $q_{3,4} = \frac{3}{16}$.

In total, we get

$$\begin{aligned} \mathbb{Q}_{lc}\left(\mu_n^d, \nu_n^d\right) =& \delta_{1,0} \cdot \frac{k_j - 1}{4k_j} + \delta_{1,k_j} \cdot \frac{1}{4k_j} \\ &+ \delta_{k_j,k_j} \cdot \left(\frac{1}{3} - \frac{1}{4k_j} \right) + \delta_{k_j,0} \cdot \left(\frac{1}{4k_j} - \frac{1}{4k_{j+1}} \right) + \delta_{k_j,k_{j+1}} \cdot \frac{1}{4k_{j+1}} \\ &+ \delta_{k_{j+1},k_{j+1}} \cdot \left(\frac{1}{6} - \frac{1}{4k_{j+1}} \right) + \delta_{k_{j+1},0} \cdot \left(\frac{1}{4k_{j+1}} - \frac{1}{16} \right) + \delta_{k_{j+1},4} \cdot \frac{1}{16} \\ &+ \delta_{3,0} \cdot \frac{1}{16} + \delta_{3,4} \cdot \frac{3}{16}. \end{aligned}$$

Analogously, we proceed for $n \geq 3$ odd. Then we get

$$\begin{aligned}\mathbb{Q}_{lc}\left(\mu_n^d, \nu_n^d\right) =& \delta_{1,0} \cdot \frac{k_j - 1}{4k_j} + \delta_{1,k_j} \cdot \frac{1}{4k_j} \\ &+ \delta_{k_j,k_j} \cdot \left(\frac{1}{6} - \frac{1}{4k_j}\right) + \delta_{k_j,0} \cdot \left(\frac{1}{4k_j} - \frac{1}{4k_{j+1}}\right) + \delta_{k_j,k_{j+1}} \cdot \frac{1}{4k_{j+1}} \\ &+ \delta_{k_{j+1},k_{j+1}} \cdot \left(\frac{1}{3} - \frac{1}{4k_{j+1}}\right) + \delta_{k_{j+1},0} \cdot \left(\frac{1}{4k_{j+1}} - \frac{1}{16}\right) + \delta_{k_{j+1},4} \cdot \frac{1}{16} \\ &+ \delta_{3,0} \cdot \frac{1}{16} + \delta_{3,4} \cdot \frac{3}{16}.\end{aligned}$$

With these transport plans, we may now calculate the approximating upper price bound sequences. For $n \geq 3$ even, we find

$$\begin{aligned}P_2^c\left(\mu_n^d, \nu_n^d\right) =& c\left(1, k_j\right) \cdot \frac{1}{4k_j} + c\left(k_j, k_j\right) \cdot \left(\frac{1}{3} - \frac{1}{4k_j}\right) + c\left(k_j, k_{j+1}\right) \cdot \frac{1}{4k_{j+1}} \\ &+ c\left(k_{j+1}, k_{j+1}\right) \cdot \left(\frac{1}{6} - \frac{1}{4k_{j+1}}\right) + c\left(k_{j+1}, 4\right) \cdot \frac{1}{16} + c\left(3, 4\right) \cdot \frac{3}{16} \\ =& \frac{k_j}{4} + \frac{k_j^3}{3} - \frac{k_j^2}{4} + \frac{k_j \cdot k_{j+1}}{4} + \frac{k_{j+1}^3}{6} - \frac{k_{j+1}^2}{4} + k_{j+1} + 9.\end{aligned}$$

Plugging in the derived representations of k_j and k_{j+1}, we get

$$P_2^c\left(\mu_n^d, \nu_n^d\right) = 9 + \frac{7}{12} - \frac{1}{3 \cdot 2^n} + \frac{7}{3} + \frac{8}{3 \cdot 2^n} + \frac{k_j^3}{3} + \frac{k_{j+1}^3}{6} + \frac{k_j \cdot k_{j+1} - k_j^2 - k_{j+1}^2}{4}.$$

If we now also plugin the representations for the higher degree terms and rearrange the former, then we achieve

$$\begin{aligned}P_2^c\left(\mu_n^d, \nu_n^d\right) =& \frac{143}{12} + \frac{7}{3 \cdot 2^n} + \frac{1}{3}\left(\left(\frac{7}{3}\right)^3 - \left(\frac{7}{3}\right)^2 \frac{4}{3 \cdot 2^n} + \frac{7}{3} \cdot \frac{16}{9 \cdot 2^{2n}} - \frac{64}{27 \cdot 2^{3n}}\right) \\ &+ \frac{1}{6}\left(\left(\frac{7}{3}\right)^3 + \left(\frac{7}{3}\right)^2 \frac{8}{3 \cdot 2^n} + \frac{7}{3} \cdot \frac{64}{9 \cdot 2^{2n}} + \frac{512}{27 \cdot 2^{3n}}\right) \\ &+ \frac{1}{4}\Bigg(\frac{49}{9} + \frac{28}{9 \cdot 2^n} - \frac{32}{9 \cdot 2^{2n}} - \left(\frac{49}{9} - \frac{56}{9 \cdot 2^n} + \frac{16}{9 \cdot 2^{2n}}\right) \\ &\quad - \left(\frac{49}{9} + \frac{112}{9 \cdot 2^n} + \frac{64}{9 \cdot 2^{2n}}\right)\Bigg) \\ =& \frac{913}{54} + \frac{84}{54 \cdot 2^n} + \mathcal{O}\left(\frac{1}{2^{2n}}\right).\end{aligned}$$

Analogously, for $n \geq 3$ odd, we have

$$P_2^c(\mu_n^d, \nu_n^d) = \frac{913}{54} + \frac{78}{54 \cdot 2^n} + \mathcal{O}\left(\frac{1}{2^{2n}}\right).$$

That is indeed the convergence speed from Theorem 7.11. △

Bibliography

[1] Luigi Ambrosio and Nicola Gigli, *A user's guide to optimal transport*, Modelling and optimisation of flows on networks, Springer, 2013, pp. 1–155.

[2] Ole E. Barndorff-Nielsen and Neil Shephard, *Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63** (2001), no. 2, 167–241.

[3] Mathias Beiglböck, Alexander M. G. Cox, and Martin Huesmann, *The geometry of multi-marginal Skorokhod embedding*, arXiv preprint arXiv:1705.09505 (2017).

[4] ———, *Optimal transport and Skorokhod embedding*, Inventiones mathematicae **208** (2017), no. 2, 327–400.

[5] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner, *Model-independent bounds for option prices—a mass transport approach*, Finance and Stochastics **17** (2013), no. 3, 477–501.

[6] Mathias Beiglböck, Pierre Henry-Labordere, and Nizar Touzi, *Monotone martingale transport plans and skorokhod embedding*, Stochastic Processes and their Applications **127** (2017), no. 9, 3005–3013.

[7] Mathias Beiglböck and Nicolas Juillet, *On a problem of optimal transport under marginal martingale constraints*, The Annals of Probability **44** (2016), no. 1, 42–106.

[8] Mathias Beiglböck, Tongseok Lim, and Jan Obłój, *Dual attainment for the martingale transport problem*, Bernoulli (2018), to appear.

[9] Mathias Beiglböck, Marcel Nutz, and Nizar Touzi, *Complete duality for martingale optimal transport on the line*, The Annals of Probability **45** (2017), no. 5, 3038–3074.

[10] Dimitir P. Bertsekas and Steven Shreve, *Stochastic optimal control: the discrete-time case*, 2004.

[11] Avi Bick, *Comments on the valuation of derivative assets*, Journal of Financial Economics **10** (1982), no. 3, 331–345.

[12] Fischer Black and Myron Scholes, *The pricing of options and corporate liabilities*, Journal of political economy **81** (1973), no. 3, 637–654.

[13] Douglas T. Breeden and Robert H. Litzenberger, *Prices of state-contingent claims implicit in option prices*, Journal of business (1978), 621–651.

[14] Haydyn Brown, David G. Hobson, and Leonard C. G. Rogers, *Robust hedging of barrier options*, Mathematical Finance **11** (2001), no. 3, 285–314.

[15] Peter Carr and Roger Lee, *Hedging variance options on continuous semimartingales*, Finance and Stochastics **14** (2010), no. 2, 179–207.

[16] Peter K. Clark, *A subordinated stochastic process model with finite variance for speculative prices*, Econometrica: journal of the Econometric Society (1973), 135–155.

[17] Alexander M. G. Cox and Jan Obłój, *Robust hedging of double touch barrier options*, SIAM Journal on Financial Mathematics **2** (2011), no. 1, 141–182.

[18] ———, *Robust pricing and hedging of double no-touch options*, Finance and Stochastics **15** (2011), no. 3, 573–605.

[19] Alexander M. G. Cox and Jiajie Wang, *Root's barrier: Construction, optimality and applications to variance options*, The Annals of Applied Probability **23** (2013), no. 3, 859–894.

[20] John C. Cox, Stephen A. Ross, and Mark Rubinstein, *Option pricing: A simplified approach*, Journal of financial Economics **7** (1979), no. 3, 229–263.

[21] Giorgio Dall'Aglio, *Sugli estremi dei momenti delle funzioni di ripartizione doppia*, Annali della Scuola Normale Superiore di Pisa-Classe di Scienze **10** (1956), no. 1-2, 35–74.

[22] Hadrien De March and Nizar Touzi, *Irreducible convex paving for decomposition of multi-dimensional martingale transport plans*, arXiv preprint arXiv:1702.08298 (2017).

[23] Emanuel Derman and Iraj Kani, *Riding on a smile*, Risk **7** (1994), no. 2, 32–39.

[24] Yan Dolinsky and H. Mete Soner, *Martingale optimal transport and robust hedging in continuous time*, Probability Theory and Related Fields **160** (2014), no. 1-2, 391–427.

[25] ———, *Martingale optimal transport in the Skorokhod space*, Stochastic Processes and their Applications **125** (2015), no. 10, 3893–3931.

[26] Bruno Dupire, *Model art*, Risk **6** (1993), no. 9, 118–124.

[27] ———, *Pricing with a smile*, Risk **7** (1994), no. 1, 18–20.

[28] Herbert Federer, *Geometric measure theory*, Springer, 2014.

[29] Alfred Galichon, Pierre Henry-Labordere, and Nizar Touzi, *A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options*, The Annals of Applied Probability **24** (2014), no. 1, 312–336.

[30] Nassif Ghoussoub, Young-Heon Kim, and Tongseok Lim, *Structure of optimal martingale transport plans in general dimensions*, arXiv preprint arXiv:1508.01806 (2015).

[31] Gaoyue Guo, *A stability result on optimal Skorokhod embedding*, arXiv preprint arXiv:1701.08204 (2017).

[32] Gaoyue Guo and Jan Obłój, *Computational methods for martingale optimal transport problems*, arXiv preprint arXiv:1710.07911 (2017).

[33] Gaoyue Guo, Xiaolu Tan, and Nizar Touzi, *On the monotonicity principle of optimal Skorokhod embedding problem*, SIAM Journal on Control and Optimization **54** (2016), no. 5, 2478–2489.

[34] ———, *Optimal Skorokhod embedding under finitely many marginal constraints*, SIAM Journal on Control and Optimization **54** (2016), no. 4, 2174–2201.

[35] ———, *Tightness and duality of martingale transport on the Skorokhod space*, Stochastic Processes and their Applications **127** (2017), no. 3, 927–956.

[36] Patrick S. Hagan, Deep Kumar, Andrew S. Lesniewski, and Diana E. Woodward, *Managing smile risk*, The Best of Wilmott **1** (2002), 249–296.

[37] Pierre Henry-Labordere, *Model-free hedging: A martingale optimal transport viewpoint*, Chapman and Hall/CRC, 2017.

[38] Pierre Henry-Labordère and Nizar Touzi, *An explicit martingale version of the one-dimensional Brenier theorem*, Finance and Stochastics **20** (2016), no. 3, 635–668.

[39] Sebastian Herrmann and Florian Stebegg, *Robust pricing and hedging around the globe*, arXiv preprint arXiv:1707.08545 (2017).

[40] Steven L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The review of financial studies **6** (1993), no. 2, 327–343.

[41] David G. Hobson, *Robust hedging of the lookback option*, Finance and Stochastics **2** (1998), no. 4, 329–347.

[42] ———, *The Skorokhod embedding problem and model-independent bounds for option prices*, Paris-Princeton Lectures on Mathematical Finance 2010, Springer, 2011, pp. 267–318.

[43] David G. Hobson and Martin Klimmek, *Model-independent hedging strategies for variance swaps*, Finance and Stochastics **16** (2012), no. 4, 611–649.

[44] ———, *Robust price bounds for the forward starting straddle*, Finance and Stochastics **19** (2015), no. 1, 189–214.

[45] David G. Hobson, Peter Laurence, and Tai-Ho Wang, *Static-arbitrage optimal subreplicating strategies for basket options*, Insurance: Mathematics and Economics **37** (2005), no. 3, 553–572.

[46] ———, *Static-arbitrage upper bounds for the prices of basket options*, Quantitative finance **5** (2005), no. 4, 329–342.

[47] David G. Hobson and Anthony Neuberger, *Robust bounds for forward start options*, Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics **22** (2012), no. 1, 31–56.

[48] David G. Hobson and Dominykas Norgilas, *The left-curtain martingale coupling in the presence of atoms*, arXiv preprint arXiv:1802.08337 (2018).

[49] Zhaoxu Hou and Jan Obłój, *Robust pricing–hedging dualities in continuous time*, Finance and Stochastics (2018), 1–57.

[50] Martin Huesmann and Florian Stebegg, *Monotonicity preserving transformations of MOT and SEP*, Stochastic Processes and their Applications **128** (2018), no. 4, 1114–1134.

[51] Nicolas Juillet, *Stability of the shadow projection and the left-curtain coupling*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 52, Institut Henri Poincaré, 2016, pp. 1823–1843.

[52] Sigrid Källblad, Xiaolu Tan, and Nizar Touzi, *Optimal Skorokhod embedding given full marginals and Azéma–Yor peacocks*, The Annals of Applied Probability **27** (2017), no. 2, 686–719.

[53] Leonid V. Kantorovich, *On a problem of Monge*, CR (Doklady) Acad. Sci. URSS (NS), vol. 3, 1948, pp. 225–226.

[54] ———, *On a problem of Monge*, Journal of Mathematical Sciences **133** (2006), no. 4, 1383–1383.

[55] Hans G. Kellerer, *Duality theorems for marginal problems*, Probability Theory and Related Fields **67** (1984), no. 4, 399–432.

[56] Steven G. Kou, *A jump-diffusion model for option pricing*, Management science **48** (2002), no. 8, 1086–1101.

[57] Steven G. Kou and Hui Wang, *Option pricing under a double exponential jump diffusion model*, Management science **50** (2004), no. 9, 1178–1192.

[58] Tongseok Lim, *Multi-martingale optimal transport*, arXiv preprint arXiv:1611.01496 (2016).

[59] Eva Lütkebohmert and Julian Sester, *Tightening robust price bounds for exotic derivatives*, Working Paper, available on `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3290503`.

[60] Dilip B. Madan, Peter P. Carr, and Eric C. Chang, *The variance gamma process and option pricing*, Review of Finance **2** (1998), no. 1, 79–105.

[61] Robert C. Merton, *Theory of rational option pricing*, The Bell Journal of economics and management science (1973), 141–183.

[62] ______, *Option pricing when underlying stock returns are discontinuous*, Journal of financial economics **3** (1976), no. 1-2, 125–144.

[63] Gaspard Monge, *Mémoire sur la théorie des déblais et des remblais*, Histoire de l'Académie Royale des Sciences de Paris (1781).

[64] Roger B. Nelsen, *An introduction to copulas*, Springer Science & Business Media, 2007.

[65] Marcel Nutz, Florian Stebegg, and Xiaowei Tan, *Multiperiod martingale transport*, arXiv preprint arXiv:1703.10588 (2017).

[66] J Obłój, *On some aspects of Skorokhod embedding problem and its applications in mathematical finance*, Notes for the students of the 5th European summer school in financial mathematics (2012).

[67] Svetlozar T. Rachev and Ludger Rüschendorf, *Mass transportation problems: Volume i: Theory*, vol. 1, Springer Science & Business Media, 1998.

[68] ______, *Mass transportation problems: Volume ii: Applications*, vol. 1, Springer Science & Business Media, 1998.

[69] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, vol. 293, Springer Science & Business Media, 2013.

[70] Wim Schoutens, *Lévy processes in finance*, Wiley, 2003.

[71] Wim Schoutens, Erwin Simons, and Jurgen Tistaert, *A perfect calibration! Now what?*, The best of Wilmott (2003), 281.

[72] Moshe Shaked and J. George Shanthikumar, *Stochastic orders*, Springer Science & Business Media, 2007.

[73] Volker Strassen, *The existence of probability measures with given marginals*, The Annals of Mathematical Statistics (1965), 423–439.

[74] Helmut Strasser, *Mathematical theory of statistics: statistical experiments and asymptotic decision theory*, vol. 7, Walter de Gruyter, 1985.

[75] Nizar Touzi, *Martingale inequalities, optimal martingale transport, and robust superhedging*, ESAIM: Proceedings and Surveys **45** (2014), 32–47.

[76] Robert J. Vanderbei, *Linear programming*, Springer, 2015.

[77] Cédric Villani, *Optimal transport: old and new*, vol. 338, Springer Science & Business Media, 2008.

[78] Dirk Werner, *Funktionalanalysis*, Springer, 2006.

[79] Danila A. Zaev, *On the Monge–Kantorovich problem with additional linear constraints*, Mathematical Notes **98** (2015), no. 5-6, 725–741.