REPLICATION OF FINANCIAL DERIVATIVES UNDER EXTREME MARKET MODELS GIVEN MARGINALS

TONGSEOK LIM

ABSTRACT. The Black-Scholes-Merton model is a mathematical model for the dynamics of a financial market that includes derivative investment instruments, and its formula provides a theoretical price estimate of European-style options. The model's fundamental idea is to eliminate risk by hedging the option by purchasing and selling the underlying asset in a specific way, that is, to replicate the payoff of the option with a portfolio (which continuously trades the underlying) whose value at each time can be verified. One of the most crucial, yet restrictive, assumptions for this task is that the market follows a geometric Brownian motion, which has been relaxed and generalized in various ways.

The concept of robust finance revolves around developing models that account for uncertainties and variations in financial markets. Martingale Optimal Transport, which is an adaptation of the Optimal Transport theory to the robust financial framework, is one of the most prominent directions. In this paper, we consider market models with arbitrarily many underlying assets whose values are observed over arbitrarily many time periods, and demonstrates the existence of a portfolio sub- or super-hedging a general path-dependent derivative security in terms of trading European options and underlyings, as well as the portfolio replicating the derivative payoff when the market model yields the extremal price of the derivative given marginal distributions of the underlyings. In mathematical terms, this paper resolves the question of dual attainment for the multi-period vectorial martingale optimal transport problem.

Keywords: Robust finance, Hedging, Martingale, Optimal transport, Duality, Dual attainment, Multi-period, Infinite-dimensional linear programming
MSC2010 Classification: 90Bxx, 90Cxx, 49Jxx, 49Kxx, 60Dxx, 60Gxx

1

1. INTRODUCTION

The Black-Scholes-Merton model, often referred to as the Black-Scholes model, is a mathematical framework for calculating the theoretical pricing of options and other derivatives [7]. It was developed by economists Fischer Black and Myron Scholes in collaboration with mathematician Robert C. Merton in the early 1970s.

The Black-Scholes-Merton model revolutionized the way options are priced in financial markets by providing a mathematical formula to determine the fair value of options based on various factors, and their pricing model has become a fundamental tool in options trading and valuation. The Black-Scholes-Merton model combines concepts from mathematics, statistics, and economics to provide a rigorous framework for understanding option pricing and risk management. As a result, it influenced subsequent research in financial economics and inspired the development of more sophisticated models and techniques [18, 23, 24, 29, 47].

The idea of replication, also known as the principle of no-arbitrage, is arguably the most fundamental concept underlying the Black-Scholes-Merton model. It states that the value of an option can be recreated by establishing a risk-neutral portfolio of the underlying asset in such a way that the portfolio replicates the option's cash flows and payoffs. This idea has influenced the development of the efficient market hypothesis and has had a profound impact on the study of financial economics.

Despite its importance, it is worth noting that the Black-Scholes-Merton model has certain limitations, such as the assumption of the market following a geometric Brownian Motion and the assumption of continuous trading. These assumptions may not hold in real-world markets, and there have been subsequent models and refinements that address some of these limitations [12, 25, 30, 39, 56].

The concept of robust finance revolves around the development of models and approaches that account for market uncertainties. It recognizes the limitations of traditional financial models, which presume exact and accurate information about the market and its underlying assets, and instead attempts to construct frameworks that can adapt to the inherent uncertainties and risks in real-world financial systems [10, 19, 40, 42, 49, 51].

One of the most renowned and prospering directions is the theory of Martingale Optimal Transport (MOT), which is an adaptation of the Optimal Transport

(OT) theory to the robust financial framework. The OT theory, also known as the theory of Monge-Kantorovich transportation, is a mathematical framework that deals with the problem of efficiently moving mass from one distribution to another. OT theory seeks to find the optimal way to transport one distribution of mass to another while minimizing the cost or distance associated with the transportation. The theory originated in the 18th century with the works of mathematicians Gaspard Monge and Leonid Kantorovich, and has since been developed and extended by various researchers. The importance of OT theory lies in its wide range of applications across various disciplines, including mathematics, statistics, economics, physics and computer science [11, 14–16, 28, 32, 54, 55, 58].

The theory of Martingale Optimal Transport is an extension of OT theory that combines it with the concept of martingales from probability theory. By combining OT theory with the concept of martingales, MOT extends the applicability of optimal transportation models to dynamic and stochastic settings. It provides a valuable tool for understanding and solving transportation problems under uncertainty, with applications in various fields including finance, risk management, stochastic control, and data analysis [1, 2, 4, 17, 26, 31, 33–38, 46].

One of the most prominent applications of MOT is in mathematical finance and option pricing. It provides a framework for modeling and valuing derivative securities in the presence of uncertainty and stochastic dynamics. MOT-based approaches can be used to analyze optimal hedging strategies and calculate prices and risk measures for options and other financial derivatives [3, 27, 41, 43, 44].

The OT and MOT are infinite-dimensional linear programming problems, hence the problem has a dual programming problem. In MOT, the dual problem has an important interpretation in terms of determining the best sub- or super-hedging portfolio against a derivative security payout liability. As a result, one of the most significant concerns in MOT theory has been the topic of *dual attainment*, which refers to whether the dual problem attains an appropriate solution [5, 6, 13, 45].

While the dual attainment has been well established to be affirmative in OT, researchers discovered that the added term reflecting trading strategy appearing in the dual of MOT makes establishing the dual attainment for MOT theory far more subtle. Specifically, unlike the OT theory, the dual attainment in MOT is highly sensitive to the spatial dimension, which represents the number of assets in the financial market. Because many derivative instruments are traded in the

market and their payoffs are dependent on the values of various underlying assets,
it is critical to understand the dual problem and its solution in a multidimensional
context, i.e., a market with multiple assets as well as derivatives depending on
them. As a result, efforts have been made to comprehend the dual attainment in
the higher-dimensional case [20–22, 33, 48, 52], but a thorough knowledge of this
question appears to be far from its completion. In particular, to the best of the
author's knowledge, all research articles on the dual attainment of multidimen-
sional MOT assume two future maturities. In OT, this means that there are two
distributions for which we attempt to move mass efficiently from one to the other.
While this is a natural and sufficiently general setting in OT, it represents a sig-
nificant restriction in MOT because it indicates that the reward of the derivative
security under consideration can only depend on the values of the underlyings
at two future maturities. It is desirable, both theoretically and practically, to
investigate the martingale transport theory of price path-dependent derivative
instruments whose reward can depend on arbitrarily many assets and maturities.

The goal of this paper is to establish the dual attainment of the martingale op-
timal transport problem over an arbitrary number of time periods and assets. We
show that there exists a portfolio sub- or super-hedging a path-dependent deriva-
tive security in terms of trading European options and underlyings, such that the
portfolio replicates the derivative payoff when the market model yields the deriva-
tive's extremal price given marginal distributions of the underlyings. We contend
this is fundamental and relevant given that the price path of many underlying
assets over time frequently affects financial instruments and their payouts.

This paper is organized as follows. In Section 2, we introduce the multi-period
vectorial martingale optimal transport problem. In Section 3, we discuss duality
and the dual attainment result, whose proof is then provided in Section 4.

2. Martingale transport problem with multi assets and periods

Consider the asset price processes $(X_{t,i})_{t\geq 0}$ in the market indexed by $i \in [d] :=$
$\{1, 2, ..., d\}$ representing the ith underlying asset. We will not assume in this paper
that the joint probability law of the underlyings, often called as the market model,
is known, because we cannot determine this joint law from market information.
On the other hand, using a standard reasoning by Breeden and Litzenberger [9],
we will suppose that the market can witness the distribution of each price at

each fixed maturity $t > 0$, denoted by $\mathrm{Law}(X_{t,i}) \in \mathcal{P}(\mathbb{R})$, where $\mathcal{P}(\mathcal{X})$ denotes the set of all probability distributions over $\mathcal{X}$. We consider an arbitrary number of finite maturity times $0 < T_1 < T_2 < ... < T_N$, denote $X_{t,i} := X_{T_t,i}$ and $X_t = (X_{t,1}, ..., X_{t,d})$, and assume that $(X_t)_{t \in [N]}$ is a $\mathbb{R}^d$-valued martingale, which is generally assumed in the financial literature through the concept of the risk-neutral probability. Now according to the above consideration, we do not assume that the market model $\mathrm{Law}(X_t)_t \in \mathcal{P}(\mathbb{R}^{Nd})$ is known, but only the Nd-number of marginal distributions $\mu_{t,i} := \mathrm{Law}(X_{t,i}) \in \mathcal{P}(\mathbb{R})$ are known and fixed. This leads us to consider the *Vectorial Martingale Optimal Transport* (VMOT) problem: Assume the marginals $(\mu_{t,i})_{t,i}$ have finite first moment, write $\mu_t = (\mu_{t,1}, ..., \mu_{t,d})$, $\mu = (\mu_1, ..., \mu_N)$, and $X = (X_1, ..., X_N)$. We consider the space of *Vectorial Martingale Transports* from μ_t to μ_{t+1}, $t = 1, ..., N-1$, defined as follows:

(2.1) $\mathrm{VMT}(\mu) := \{\pi \in \mathcal{P}(\mathbb{R}^{Nd}) \mid \pi = \mathrm{Law}(X),\ \mathbb{E}_\pi[X_{t+1}|X_t] = X_t,$

$$\mathrm{Law}(X_{t,i}) = \mu_{t,i} \text{ for all } t \in [N],\ i \in [d]\}.$$

Given a *cost function* $c : \mathbb{R}^{Nd} \to \mathbb{R}$, we define the VMOT problem as

(2.2) $\max / \mathrm{minimize}\ \ \mathbb{E}_\pi[c(X)]\ \ \mathrm{over}\ \ \pi \in \mathrm{VMT}(\mu).$

In finance, the function c is naturally interpreted as a derivative security whose payoff $c(X)$ is fully determined at the terminal maturity T_N by the price path $X = (X_t)_t$ of the d-number of underlyings. In this case, $\mathbb{E}_\pi[c(X)]$ can be regarded as a fair price for the derivative security c under the market model π. Because π cannot be observed in the market, we must take into account all feasible models $\mathrm{VMT}(\mu)$ which are consistent with the marginal information $\mu = (\mu_{t,i})_{t,i}$. With this knowledge, the maximum and minimum values in (2.2) can be interpreted as the upper and lower price bounds for the derivative security c, respectively.

The VMOT problem is distinguished from the (ordinary) optimal transport problems by the martingale constraint $\mathbb{E}_\pi[X_{t+1}|X_t] = X_t$ for all $t = 1, ..., N-1$, which necessitates that every pair of marginals $\mu_{t,i}, \mu_{t+1,i}$ must be in *convex order*:

$\mu_{t,i} \preceq_c \mu_{t+1,i}$ if $\mu_{t,i}(f) \le \mu_{t+1,i}(f)$ for every convex function f on $\mathbb{R}$,

where $\mu(f) := \int f(x)\mu(dx)$, in which case $\mathrm{VMT}(\mu) \ne \emptyset$ and vice versa, as shown by Strassen [57]. Thus we will assume $\mu_{t,i} \preceq_c \mu_{t+1,i}$ for all $t \le N-1$ and $i \in [d]$.

The VMOT problem belongs to the class of *infinite-dimensional linear programming*, hence the problem has its *dual programming* problem. When the primal problem (2.2) is a minimization problem, its dual problem is given by

$$(2.3) \qquad\qquad \sup_{(\varphi,h)\in\Psi} \mu(\varphi),$$

where $\varphi = (\varphi_1, ..., \varphi_N)$, $\varphi_t = (\varphi_{t,1}, ..., \varphi_{t,d})$, $\varphi_{t,i} : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$, and $\mu(\varphi) := \sum_{t,i} \mu_{t,i}(\varphi_{t,i})$. Meanwhile, $h = (h_1, ..., h_N)$ with the convention $h_N \equiv 0$, $h_t = (h_{t,1}, ..., h_{t,d})$ where $h_{t,i} : \mathbb{R}^{td} \to \mathbb{R}$ is a function of $(X_1, ..., X_t)$. Finally, $(\varphi, h) \in \Psi$ means that $\varphi_{t,i} \in L^1(\mu_{t,i})$, $h_{t,i}$ is bounded, and the following inequality holds:

$$(2.4) \qquad \sum_{t=1}^{N}\sum_{i=1}^{d} \varphi_{t,i}(x_{t,i}) + h_{t,i}(x_1, ..., x_t)(x_{t+1,i} - x_{t,i}) \le c(x),$$

where $x = (x_1, ..., x_N) \in \mathbb{R}^{Nd}$, $x_t = (x_{t,1}, ..., x_{t,d}) \in \mathbb{R}^d$ represents a price path. Note that

$$(2.5) \qquad \mu(\varphi) = \mathbb{E}_\pi\left[\sum_{t=1}^{N}\sum_{i=1}^{d} \varphi_{t,i}(X_{t,i}) + h_{t,i}(X_1, ..., X_t)(X_{t+1,i} - X_{t,i}) \right]$$

for any $\pi \in \mathrm{VMT}(\mu)$ and $(\varphi, h) \in \Psi$, since π has marginals $\{\mu_{t,i}\}_{t,i}$ and the martingale property of π implies $\mathbb{E}_\pi[h_{t,i}(X_1, ..., X_t)(X_{t+1,i} - X_{t,i})] = 0$. Now if (2.2) is a maximization problem, its dual problem reads $\inf_{(\varphi,h)\in\Psi} \mu(\varphi)$, with the inequality (2.4) reversed (and $\varphi_{t,i}$ taking on their values in $\mathbb{R} \cup \{+\infty\}$).

The dual problem also has a significant interpretation and implication in finance. Let us assume that a financial firm is obligated to pay $c(X)$ at the terminal maturity N for a derivative instrument c. To mitigate risk, the firm may consider purchasing or selling European options $\varphi_{t,i}$ available from the market, the payoff of which is based solely on the price $X_{t,i}$ at the maturity t. Furthermore, the company may consider holding $h_{t,i}$ shares of the ith asset between the t and $t+1$ maturities, so that its return at time $t+1$ is $h_{t,i}(X_1, ..., X_t) \cdot (X_{t+1} - X_t)$. It is worth noting that $h_{t,i}$ is a function of all underlying prices up to time t. Then the left hand side of (2.4) represents the payout of the hedging portfolio (φ, h), and the inequality (2.4) mandates that the position must subhedge the liability for all possible market realizations $x \in \mathbb{R}^{Nd}$. Having stated that, notice the dual of the maximization problem in (2.2) represents an optimal superhedging problem.

3. Duality and our contribution

The celebrated *duality* result asserts, under a mild assumption on c and the marginals, that the primal and dual optimal values coincide (see e.g. [27, 59]):

$$(3.1) \qquad P(c) := \inf_{\pi \in \mathrm{VMT}(\mu)} \mathbb{E}_\pi[c(X)]$$

$$= \sup_{(\varphi, h) \in \Psi} \mu(\varphi) =: D(c).$$

In addition, the primal problem is known to be attained; there exist an optimizer (VMOT) π that yields $\mathbb{E}_\pi[c(X)] = P(c)$. Observe the VMOTs describe extreme market movements in the sense of maximizing or minimizing the fair price of c.

Unlike the primal problem, whose attainment can be easily established by standard arguments under a mild condition on c,[1] establishing *dual attainment*, i.e., proving the existence of a suitable form of dual optimizers, turns out to be a very nontrivial problem, as demonstrated by Brenier's work [8] on optimal transport problems. The situation is worse for the martingale optimal transport problems, where the martingale constraint makes the dual attainment problem even more complex, as shown by [4–6] even for a single asset setup ($d = 1$). Interestingly, these works show that the dual attainment problem frequently boils down to establishing convergence of a certain sequence of convex potentials.

The difficulty of the dual attainment problem is problematic not only from a mathematical standpoint, but also from a financial standpoint, because dual optimizers describe how to (sub-/super-)hedge a particular derivative investment most effectively, and furthermore, they often provide the most critical information on the structure of primal optimizers that describe the extreme market models. Due to its significance, numerous literature addresses the dual attainment, the majority of which focus on a single asset $d = 1$ and two period setup $N = 2$ with a few exceptions, such as [17, 50, 53] which investigated duality in a multi-period with a single asset setup, while [20–22, 33, 48, 52] investigated the structure of vector-valued martingale transports in a two-period setup. The purpose of this paper is to establish the dual attainment of the martingale optimal transport problem over an arbitrary number of time periods and assets, which is fundamental and relevant given that financial instruments and their payouts frequently

---

[1]This does not imply that the primal optimizers are easy to understand, describe or characterize.

depend on the price path of many underlyings over time. Specifically, the goal of this paper is to prove the following result (see Section 4 for irreducibility.)

Theorem 3.1. *Let $(\mu_{t,i})_{t\in[N]}$ be an irreducible sequence of marginals on $\mathbb{R}$ for each $i \in [d]$. Let $c(x)$ be a lower semi-continuous cost satisfying $|c(x)| \leq \sum_{t,i} v_{t,i}(x_{t,i})$ for some continuous functions $v_{t,i} \in L^1(\mu_{t,i})$. Then there exists a dual optimizer, that is a pair of function sequences $(\varphi, h) = (\varphi_{t,i}, h_{t,i})_{t\in[N],i\in[d]}$ that satisfies (2.4) tightly in the following pathwise sense (but needs not be in Ψ):*

$$(3.2) \qquad \sum_{t=1}^{N}\sum_{i=1}^{d} \varphi_{t,i}(x_{t,i}) + h_{t,i}(x_1,...,x_t)(x_{t+1,i} - x_{t,i}) = c(x) \quad \pi - a.s.$$

for every VMOT π *which solves the minimization problem in* (2.2).

The (φ, h) presented in the theorem is known as a *dual maximizer* since it is a solution concept to the problem (2.3), which is dual to the primal minimization problem. The equality (3.2) indicates that the portfolio composed of a dual maximizer replicates the derivative security c by yielding the same payout for all feasible price paths x given by any VMOT π minimizing the price $\mathbb{E}_\pi[c(X)]$. And (2.4) shows the portfolio otherwise subhedges the derivative c for all price paths. Theorem 3.1 also yields the existence of a dual minimizer such that the portfolio built of a dual minimizer superhedges an upper semi-continuous c by reversing the inequality in (2.4), and it replicates c with respect to the primal maximizers.

We emphasize that a dual optimizer does not necessarily belong to Ψ. Studies showed the dual problem (2.3) is generally not attained within the class Ψ even when $(d, N) = (1, 2)$ (see [4,6]), unless c satisfies a specific regularity property [5]. Although Ψ can be seen as a natural domain for the dual problem, it is rather "narrow" as it lacks suitable compactness due to its infinite dimensionality. This implies that establishing dual attainment is substantially more involved than establishing duality (3.1), which can usually be derived via standard argument in functional analysis and variational calculus. In summary, a dual optimizer (φ, h) needs not be in Ψ, but it does hold that $\varphi_{t,i}$ is real-valued $\mu_{t,i}$-almost surely, and that $h_{t,i}$ is also real valued. Since the marginal distributions $\{\mu_{t,i}\}_{t,i}$ are assumed in the VMOT problem, all of the functions in Theorem 3.1 are essentially real-valued (and measurable), while no further regularity is imposed a priori.

4. Proof of Theorem 3.1

When $N = 2$, [48] proved Theorem 3.1, and one of the most essential ingredients for the proof is the stability of the *convex cores* $\{\chi_n\}_n$, which was earlier recognized and proved when $d = 1$ in [6]. To describe, we need to explain the concept of *irreducibility* of probability distributions $\mu \preceq_c \nu$ on $\mathbb{R}$ in convex order.

Two probabilities (with finite first moment) in convex order $\mu \preceq_c \nu$ is called irreducible if $I := \{x \in \mathbb{R} \mid u_\mu(x) < u_\nu(x)\}$ is a connected interval and $\mu(I) = \mu(\mathbb{R})$, where $u_\mu(x) := \int_{\mathbb{R}} |x - y| d\mu(y)$ is called the potential function of μ. Note that I is open since u_μ is continuous. In this case, (I, J) is called the domain of (μ, ν) where J is the smallest interval satisfying $\nu(J) = \nu(\mathbb{R})$, that is, J is the union of I and any endpoints of I that are atoms of ν. Thus in particular, it holds $I = \text{int}(J)$, where $\text{int}(A)$ is the interior of A and $\text{conv}(A)$ is the convex hull of A. Note that J can be of the form $(a, b]$, $[a, b)$, (a, b) or $[a, b]$; in the first case it holds $\nu(b) > 0$, and in the second $\nu(a) > 0$. And in all cases $I = \text{int}(J) = (a, b)$. Of course, I and J can be (half)-infinite intervals. Roughly speaking, the irreducibility of $\mu \preceq_c \nu$ means that ν is regularly dispersed from μ. We underline that irreducibility is a natural and generic property that practically any pair of probability distributions on $\mathbb{R}$ in convex order fulfills, and that even if a pair is not irreducible, it can be perturbed arbitrarily small to become irreducible.

Now for irreducible pairs $(\mu_i, \nu_i)_{i \in [d]}$ with domains (I_i, J_i), set $I = I_1 \times \ldots \times I_d$, $J = J_1 \times \ldots \times J_d$ and $\mu^\otimes = \mu_1 \otimes \ldots \otimes \mu_d$, $\nu^\otimes = \nu_1 \otimes \ldots \otimes \nu_d$. I is an open rectangle in $\mathbb{R}^d$, $I = \text{int}(J)$, and $\mu^\otimes, \nu^\otimes \in \mathcal{P}(\mathbb{R}^d)$ are the product measures of μ_i's and ν_i's respectively. Now the following was shown in [6] for $d = 1$ and in [48] for $d \geq 2$.

Proposition 4.1. *Let $(\mu_i, \nu_i)_{i \in [d]}$ be irreducible pairs of probability distributions on $\mathbb{R}$ with domains $(I_i, J_i)_{i \in [d]}$. Let $a \in I$, $C \in \mathbb{R}$. Consider the following class of functions $\Lambda = \Lambda(a, C, \vec{\mu}, \vec{\nu})$ where every $\chi \in \Lambda$ satisfies the following:*

(1) χ is a real-valued convex function on J,
(2) $\chi \geq 0$ and $\chi(a) = 0$,
(3) $\int \chi \, d(\nu^\otimes - \mu^\otimes) \leq C$.

Then Λ is locally bounded in the following sense: for each compact subset K of J, there exists $M = M(K)$ such that $\chi \leq M$ on K for every $\chi \in \Lambda$. Furthermore, for any sequence $\{\chi_n\}_n$ in Λ, there exists a subsequence $\{\chi_{n_j}\}_j$ of $\{\chi_n\}_n$ and a real-valued convex function χ on J such that $\lim_{j \to \infty} \chi_{n_j}(x) = \chi(x)$ for all $x \in J$.

In Theorem 3.1, there are Nd number of marginal distributions $\mu_{t,i}$, indexed by $t \in [N]$ and $i \in [d]$, where t represents the time period and i is the martingale (or financial asset) index. Let $(I_{t,i}, J_{t,i})$ denote the domain of $\mu_{t,i} \preceq_c \mu_{t+1,i}$, $t = 1, ..., N - 1$, with the convention $J_{0,i} := J_{1,i}$ and $I_{N,i} := J_{N-1,i}$. Throughout the proof, keep in mind that the bounding constant C does not depend on n.

Proof of Theorem 3.1. **Step 1.** The assumption $|c(x)| \leq \sum_{t,i} v_{t,i}(x_{t,i})$ for continuous $v_{t,i} \in L^1(\mu_{t,i})$ ensures $P(c) = D(c)$ in (3.1) (see e.g. [59]). Clearly, a dual optimizer exists for $c(x)$ iff so does for $\tilde{c}(x) := c(x) - \sum_{t,i} v_{t,i}$. Thus by replacing c with $\tilde{c}$, from now on we will assume that $c \leq 0$.

As $P(c) = D(c) \in \mathbb{R}$, we can find an *approximating dual maximizer* $(\varphi_n, h_n)_{n \in \mathbb{N}}$ which consists of real-valued continuous functions $\varphi_{t,i,n} \in L^1(\mu_{t,i})$ and continuous bounded $h_{t,i,n}$ for every $t \in [N]$, $i \in [d]$ and $n \in \mathbb{N}$ (we assume $h_{N,i,n} \equiv 0$), such that the following duality holds:

$$(4.1) \qquad \sum_{t=1}^{N} \sum_{i=1}^{d} \varphi_{t,i,n}(x_{t,i}) + h_{t,i}(\bar{x}_t)\Delta x_{t,i} \leq c(x) \leq 0,$$

$$(4.2) \qquad \mu(\varphi_n) := \sum_{t=1}^{N} \sum_{i=1}^{d} \int \varphi_{t,i,n}(x_{t,i}) \, d\mu_{t,i}(x_{t,i}) \nearrow P(c) \quad \text{as} \quad n \to \infty,$$

where $\bar{x}_t = (x_1, ..., x_t)$, $\Delta x_{t,i} = x_{t+1,i} - x_{t,i}$ and $\Delta x_t = x_{t+1} - x_t$. Denote $\varphi_{t,n}^{\oplus}(x_t) = \sum_{i=1}^{d} \varphi_{t,i,n}(x_{t,i})$, $h_{t,n}(\bar{x}_t) = \left(h_{t,1,n}(\bar{x}_t), ..., h_{t,d,n}(\bar{x}_t)\right)$. Define

$$(4.3) \qquad \chi_{t,n}(x_t) := \sup_{x_1,...,x_{t-1}} \sum_{s=1}^{t-1} \left(\varphi_{s,n}^{\oplus}(x_s) + h_{s,n}(\bar{x}_s) \cdot \Delta x_s\right)$$

with the convention $\chi_{1,n} = \chi_{N+1,n} \equiv 0$. Notice $\chi_{t,n}$ is a convex function on $\mathbb{R}^d$, since it is a supremum of affine functions of x_t. We now show

$$(4.4) \qquad \chi_{t,n} \leq \chi_{t+1,n} - \varphi_{t,n}^{\oplus} \quad \text{for all } t \in [N] \text{ and } n \in \mathbb{N}.$$

This inequality can be shown as follows:

$$\chi_{t+1,n}(x_{t+1}) = \sup_{x_1,\dots,x_t} \sum_{s=1}^{t} \left(\varphi_{s,n}^{\oplus}(x_s) + h_{s,n}(\bar{x}_s) \cdot \Delta x_s \right)$$

$$\geq \sup_{\substack{x_1,\dots,x_{t-1} \\ x_t=x_{t+1}}} \sum_{s=1}^{t-1} \left(\varphi_{s,n}^{\oplus}(x_s) + h_{s,n}(\bar{x}_s) \cdot \Delta x_s \right) + \varphi_{t,n}^{\oplus}(x_{t+1})$$

$$= \chi_{t,n}(x_{t+1}) + \varphi_{t,n}^{\oplus}(x_{t+1}).$$

We can now establish the following crucial bound

$$(4.5) \qquad \int \chi_{t,n}\, d(\mu_t^{\otimes} - \mu_{t-1}^{\otimes}) \leq C \quad \text{for all } t = 2,\dots,N \text{ and } n \in \mathbb{N}.$$

To see this, by repeated application of (4.4), we have

$$\mu_t^{\otimes}(\chi_{t,n}) \leq \mu_t^{\otimes}(\chi_{t+1,n}) - \mu_t^{\otimes}(\varphi_{t,n}^{\oplus})$$

$$\leq \mu_{t+1}^{\otimes}(\chi_{t+1,n}) - \mu_t^{\otimes}(\varphi_{t,n}^{\oplus})$$

$$\leq \mu_{t+1}^{\otimes}(\chi_{t+2,n}) - \mu_{t+1}^{\otimes}(\varphi_{t+1,n}^{\oplus}) - \mu_t^{\otimes}(\varphi_{t,n}^{\oplus})$$

$$\leq \cdots \leq -\sum_{s=t}^{N} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus}),$$

where the second inequality is due to $\mu_t^{\otimes} \preceq_c \mu_{t+1}^{\otimes}$ and convexity of χ. Similarly,

$$\mu_{t-1}^{\otimes}(\chi_{t,n}) \geq \mu_{t-1}^{\otimes}(\chi_{t-1,n}) + \mu_{t-1}^{\otimes}(\varphi_{t-1,n}^{\oplus})$$

$$\geq \mu_{t-2}^{\otimes}(\chi_{t-1,n}) + \mu_{t-1}^{\otimes}(\varphi_{t-1,n}^{\oplus})$$

$$\geq \mu_{t-2}^{\otimes}(\chi_{t-2,n}) + \mu_{t-2}^{\otimes}(\varphi_{t-2,n}^{\oplus}) + \mu_{t-1}^{\otimes}(\varphi_{t-1,n}^{\oplus})$$

$$\geq \cdots \geq \sum_{s=1}^{t-1} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus}).$$

The two inequalities combine to give

$$(4.6) \qquad \int \chi_{t,n}\, d(\mu_t^{\otimes} - \mu_{t-1}^{\otimes}) \leq -\sum_{s=1}^{N} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus}) = -\mu(\varphi_n)$$

which, in conjunction with (4.2), yields (4.5).

From (4.5), we can obtain local uniform boundedness of $\{\chi_{t,n}\}_n$ via Proposition 4.1. We need to meet the proposition's second condition. For this, fix any $a \in I_1$, and let $L_{2,n}$ be an affine functions satisfying $L_{2,n} \leq \chi_{2,n}$ and $L_{2,n}(a) = \chi_{2,n}(a)$. By

linearity, $L_{2,n}(x_2) = \nabla L_{2,n}(x_1)\cdot(x_2-x_1)+L_{2,n}(x_1)$, which allows us to modify (4.1) by replacing $\varphi_{1,n}^{\oplus}(x_1)$ with $\varphi_{1,n}^{\oplus}(x_1) - L_{2,n}(x_1)$, $\varphi_{2,n}^{\oplus}(x_2)$ with $\varphi_{2,n}^{\oplus}(x_2) + L_{2,n}(x_2)$, and $h_{1,n}(x_1)$ with $h_{1,n}(x_1) - \nabla L_{2,n}(x_1)$ ($\nabla L_{2,n}$ is constant and does not depend on x_1). Notice this yields $\chi_{2,n} \geq 0$ and $\chi_{2,n}(a) = 0$. We can continue subtracting appropriate linear functions $L_{t,n}$, $t = 2, ..., N$, and obtain

$$(4.7) \qquad \chi_{t,n} \geq 0 \ \text{ and } \ \chi_{t,n}(a) = 0 \ \text{ for all } n \in \mathbb{N} \text{ and } t \in [N].$$

Note that the modifications have no effect on the value $\mu(\varphi_n)$.

To quantitatively state the local bound of χ, let $\{\epsilon_k\}_k$ be a positive decreasing sequence tending to zero as $k \to \infty$, and write $I_{t,i} =]a_{t,i}, b_{t,i}[$ where $-\infty \leq a_{t,i} < b_{t,i} \leq +\infty$. Then we define the compact interval $J_{t,i,k} := [c_{t,i,k}, d_{t,i,k}]$ for $t \in [N-1]$, $i \in [d]$ and $k \in \mathbb{N}$ as follows:

  If $a_{t,i} > -\infty$, then define $c_{t,i,k}$ by:

  $\mu_{t+1,i}(a_{t,i}) = 0 \Rightarrow c_{t,i,k} := a_{t,i} + \epsilon_k$; $\mu_{t+1,i}(a_{t,i}) > 0 \Rightarrow c_{t,i,k} := a_{t,i}$;

  If $b_{t,i} < +\infty$, then define $d_{t,i,k}$ by:

  $\mu_{t+1,i}(b_{t,i}) = 0 \Rightarrow d_{t,i,k} := b_{t,i} - \epsilon_k$; $\mu_{t+1,i}(b_{t,i}) > 0 \Rightarrow d_{t,i,k} := b_{t,i}$;

  If $a_{t,i} = -\infty$, then $c_{t,i,k} := -1/\epsilon_k$;

  If $b_{t,i} = +\infty$, then $d_{t,i,k} := +1/\epsilon_k$.

Set $J_{0,i,k} := J_{1,i,k}$. For example, if $\mu_{t+1,i}(a_{t,i}) = 0$ and $\mu_{t+1,i}(b_{t,i}) > 0$, then $J_{t,i,k} = [a_{t,i}+\epsilon_k, b_{t,i}]$. Let ϵ_1 be so small so that $\mu_{t,i}(J_{t,i,1}) > 0$, $\mu_{t+1,i}(J_{t,i,1}) > 0$ for every t, i. Observe that $J_{t,i,k} \nearrow J_{t,i}$ as $k \to \infty$. Let $J_{t,k} := J_{t,1,k} \times J_{t,2,k} \times ... \times J_{t,d,k}$. Then by Proposition 4.1, we deduce there exists $M_k \geq 0$ for each $k \in \mathbb{N}$ such that

$$(4.8) \qquad\qquad 0 \leq \sup_n \chi_{t,n} \leq M_k \ \text{ in } \ J_{t-1,k}.$$

Step 2. Given an approximating dual maximizer $(\varphi_n, h_n)_{n \in \mathbb{N}}$ (where we recall that each of $\varphi_n = (\varphi_{t,i,n})_{t,i}$ and $h_n = (h_{t,i,n})_{t,i}$ is in $\mathbb{R}^{Nd}$), our goal is to show pointwise convergence of $\varphi_{t,i,n}$ to some function $\varphi_{t,i}$ $\mu_{t,i}$-a.s. as $n \to \infty$, where $\varphi_{t,i} \in \mathbb{R} \cup \{-\infty\}$ is $\mu_{t,i}$-a.s. finite. But as noted in [48], there is an obstacle for the convergence when $d \geq 2$, that is, in the duality relation (4.1) one can always replace $(\varphi_{t,i,n})_{t,i}$ by $(\varphi_{t,i,n}+C_{t,i,n})_{t,i}$ for any constants $C_{t,i,n}$ satisfying $\sum_{t,i} C_{t,i,n} = 0$. This implies that the convergence cannot hold for any approximating dual

maximizer. In view of this, our goal is to show that there exists a suitable approximating dual maximizer that yields the almost sure convergence.

To this end, take any approximating dual maximizer $(\varphi_n, h_n)_{n \in \mathbb{N}}$ satisfying (4.1), (4.2). By repeatedly applying (4.4), we deduce

$$C \geq -\sum_{s=1}^{N} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus})$$

$$\geq \mu_N^{\otimes}(\chi_{N,n}) - \sum_{s=1}^{N-1} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus})$$

$$\geq \mu_{N-1}^{\otimes}(\chi_{N,n}) - \sum_{s=1}^{N-1} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus})$$

$$= \mu_{N-1}^{\otimes}(\chi_{N,n}) - \mu_{N-1}^{\otimes}(\chi_{N-1,n}) + \mu_{N-1}^{\otimes}(\chi_{N-1,n}) - \sum_{s=1}^{N-1} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus})$$

$$= \| \chi_{N,n} - \chi_{N-1,n} - \varphi_{N-1,n}^{\oplus} \|_{L^1(\mu_{N-1}^{\otimes})} + \mu_{N-1}^{\otimes}(\chi_{N-1,n}) - \sum_{s=1}^{N-2} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus})$$

$$\geq \| \chi_{N,n} - \chi_{N-1,n} - \varphi_{N-1,n}^{\oplus} \|_{L^1(\mu_{N-1}^{\otimes})} + \mu_{N-2}^{\otimes}(\chi_{N-1,n}) - \sum_{s=1}^{N-2} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus})$$

$$\geq \cdots \geq \sum_{s=2}^{N} \| \chi_{s,n} - \chi_{s-1,n} - \varphi_{s-1,n}^{\oplus} \|_{L^1(\mu_{s-1}^{\otimes})}$$

where the third and sixth inequality is due to the convexity of χ with $\mu_t^{\otimes} \preceq_c \mu_{t+1}^{\otimes}$, and the fifth equality is by the nonnegativity $\chi_{t,n} - \chi_{t-1,n} - \varphi_{t-1,n}^{\oplus} \geq 0$ from (4.4). On the other hand, the nonpositivity (4.1) yields

$$\sum_{s=1}^{N} \left(\varphi_{s,n}^{\oplus}(x_s) + h_{s,n}(\bar{x}_s) \cdot \Delta x_s \right) \leq \chi_{N,n}(x_N) + \varphi_{N,n}^{\oplus}(x_N) \leq 0$$

where the first inequality follows by taking supremum over $x_1, ..., x_{N-1}$ (recall (4.3) and $h_{N,n} \equiv 0$). Integrating with any $\pi \in \text{VMT}(\mu)$ yields $\| \chi_{N,n}(x_N) + \varphi_{N,n}^{\oplus} \|_{L^1(\mu_N^{\otimes})} \leq -\sum_{s=1}^{N} \mu_s^{\otimes}(\varphi_{s,n}^{\oplus}) \leq C$. We thus conclude

(4.9) $\| \chi_{t+1,n} - \chi_{t,n} - \varphi_{t,n}^{\oplus} \|_{L^1(\mu_t^{\otimes})} \leq C$ for all $n \in \mathbb{N}, t \in [N]$.

This uniform L^1 bound, in conjunction with the local uniform bound (4.8) and Komlós compactness theorem, can imply the desired almost sure convergence that

we now present. For this, we will extend the argument given in [48] for two-period case $N = 2$ into arbitrary $N \geq 2$. For each $k \in \mathbb{N}$, let $\mu_{t,i,k}$ be the restriction of $\mu_{t,i}$ on $J_{t-1,i,k}$ (where $J_{0,i,k} := J_{1,i,k}$) then normalized to be a probability distribution. Let $\mu_{t,k}^{\otimes} = \otimes_i \mu_{t,i,k}$, so that $\mu_{t,k}^{\otimes}(J_{t,k}) = 1$ where $J_{t,k} := \otimes_i J_{t,i,k}$. Define

$$v_{t,i,k,n} := \int \varphi_{t,i,n} \, d\mu_{t,i,k}, \quad t \in [N], i \in [d], k \in \mathbb{N}, n \in \mathbb{N}.$$

For each $k \in \mathbb{N}$, we will show that there exists $C = C(k)$ such that

$$(4.10) \qquad \sup_n \| \varphi_{t,i,n} - v_{t,i,k,n} \|_{L^1(\mu_{t,i,k})} \leq C.$$

To see this, observe that (4.8), (4.9) and the fact $J_{t-1,k} \subseteq J_{t,k}$ imply

$$\| \varphi_{t,n}^{\oplus} \|_{L^1(\mu_{t,k}^{\otimes})}$$
$$\leq \| \chi_{t+1,n} - \chi_{t,n} - \varphi_{t,n}^{\oplus} \|_{L^1(\mu_{t,k}^{\otimes})} + \| \chi_{t,n} - \chi_{t+1,n} \|_{L^1(\mu_{t,k}^{\otimes})}$$
$$\leq C + M_k =: C$$

where we used $\chi_{t+1,n} - \chi_{t,n} - \varphi_{t,n}^{\oplus} \geq 0$ to get the bound $\| \chi_{t+1,n} - \chi_{t,n} - \varphi_{t,n}^{\oplus} \|_{L^1(\mu_{t,k}^{\otimes})} \leq C$ from (4.9). From this, we obtain the bound

$$(4.11) \qquad \left| \sum_{i=1}^{d} v_{t,i,k,n} \right| \leq \| \varphi_{t,n}^{\oplus} \|_{L^1(\mu_{t,k}^{\otimes})} \leq C \text{ for all } n,$$

where the first inequality is by Jensen's inequality. Next, because $\varphi_{t,n}^{\oplus} \leq M_k$ on $J_{t-1,k}$ by (4.4) and (4.8), by taking supremum, we have

$$\sum_{i=1}^{d} \sup_{x_{t,i} \in J_{t,i,k}} \varphi_{t,i,n}(x_{t,i}) \leq M_k \text{ for all } n,$$

and note that clearly $v_{t,i,k,n} \leq \sup_{x_{t,i} \in J_{t,i,k}} \varphi_{t,i,n}(x_{t,i})$, so in particular,

$$\sup_{x_{t,1} \in J_{t,1,k}} \varphi_{t,1,n}(x_{t,1}) + \sum_{i=2}^{d} v_{t,i,k,n} \leq M_k.$$

Define $\hat{v}_{t,1,k,n} := -\sum_{i=2}^{d} v_{t,i,k,n}$. Since $\varphi_{t,n}^{\oplus} \leq M_k$ on $J_{t-1,k}$, we have

$$C \geq \| M_k - \varphi_{t,n}^{\oplus} \|_{L^1(\mu_{t,k}^{\otimes})} = M_k - \int (\varphi_{t,1,n} + \sum_{i=2}^{d} v_{t,i,k,n}) d\mu_{t,1,k}.$$

This implies that $\sup_n \parallel \varphi_{t,1,n} - \hat{v}_{t,1,k,n} \parallel_{L^1(\mu_{t,1,k})}$ is bounded, and then by (4.11), $\sup_n \parallel \varphi_{t,1,n} - v_{t,1,k,n} \parallel_{L^1(\mu_{t,1,k})}$ is bounded. This yields (4.10).

We can now apply the Komlós lemma, which states that every L^1-bounded sequence of real functions contains a subsequence such that the arithmetic means of all its subsequences converge pointwise almost everywhere.

Let $\tilde{v}_{t,1,k,n} = \frac{1}{n} \sum_{m=1}^n \hat{v}_{t,1,k,m}$, and $\tilde{v}_{t,i,k,n} = \frac{1}{n} \sum_{m=1}^n v_{t,i,k,m}$ for $i \geq 2$. Now for each $k \in \mathbb{N}$, a repeated application of Komlós lemma yields that there exists a subsequence $\{\varphi_{t,i,k,n}\}_n$ of $\{\varphi_{t,i,n}\}_n$, such that

(i) $\{\varphi_{t,i,k+1,n}\}_n$ is a further subsequence of $\{\varphi_{t,i,k,n}\}_n$, and

(ii) $\tilde{\varphi}_{t,i,k,n}(x_{t,i}) - \tilde{v}_{t,i,k,n}$ converges $\mu_{t,i,k}$ - a.s.   as $n \to \infty$,

where $\tilde{\varphi}_{t,i,k,n} := \frac{1}{n} \sum_{m=1}^n \varphi_{t,i,k,m}$. Note that for each k, our choice of a subsequence index can be made identical for every t and i, since there are finitely many indices of t, i. Then we select the diagonal sequence

$$\Phi_{t,i,n} := \varphi_{t,i,n,n}$$

and again define

$$w_{t,i,k,n} = \int \Phi_{t,i,n} d\mu_{t,i,k}, \quad 2 \leq i \leq d,$$

$$\hat{w}_{t,1,k,n} = -\sum_{i=2}^d w_{t,i,k,n},$$

$$\tilde{\Phi}_{t,i,n}(x_{t,i}) = \frac{1}{n} \sum_{m=1}^n \Phi_{t,i,m}(x_{t,i}),$$

$$\tilde{w}_{t,1,k,n} = \frac{1}{n} \sum_{m=1}^n \hat{w}_{t,1,k,m},$$

$$\tilde{w}_{t,i,k,n} = \frac{1}{n} \sum_{m=1}^n w_{t,i,k,m}, \quad 2 \leq i \leq d.$$

We finally claim that

(4.12) $\tilde{\Phi}_{t,i,n}(x_{t,i}) - \tilde{w}_{t,i,1,n}$ converges $\mu_{t,i}$ − a.s. for all $t \in [N], i \in [d]$.

Note that the dependence on k has now been removed. To prove the claim, since $\{\Phi_{t,i,n}\}_n$ is a subsequence of $\{\varphi_{t,i,k,n}\}_n$ for every $k \in \mathbb{N}$, Komlós lemma implies

(4.13) $\tilde{\Phi}_{t,i,n}(x_{t,i}) - \tilde{w}_{t,i,k,n}$ converges $\mu_{t,i,k}$ − a.s. for all t and i.

In particular, both $\{\tilde{\Phi}_{t,i,n}(x_{t,i}) - \tilde{w}_{t,i,1,n}\}_n$ and $\{\tilde{\Phi}_{t,i,n}(x_{t,i}) - \tilde{w}_{t,i,k,n}\}_n$ converge $\mu_{t,i,1}$ - a.s. as $n \to \infty$, hence their difference $\{\tilde{w}_{t,i,1,n} - \tilde{w}_{t,i,k,n}\}_n$ also converges for any fixed k. With (4.13), this implies (4.12). Finally, the fact that $\sum_{i=1}^d \tilde{w}_{t,i,1,n} = 0$ allows us to replace the approximating dual maximizer $(\varphi_n, h_n)_n$ by $(\psi_n, \tilde{h}_n)_n$, where $\psi_{t,i,n} := \tilde{\Phi}_{t,i,n}(x_{t,i}) - \tilde{w}_{t,i,1,n}$ and $\tilde{h}_{t,i,n}$ is a suitable Cesàro mean of a subsequence of $(h_{t,i,n})_n$ which is chosen consistently with the selection of $\tilde{\Phi}_{t,i,n}$.

Step 3. We will prove the convergence of $\{\chi_{t,n}\}_n$ defined in (4.3). Let $(\varphi_n, h_n)_n$ be an approximating dual maximizer such that

(4.14) $\{\varphi_{t,i,n}\}_n$ converges to a function $\varphi_{t,i}$ $\mu_{t,i} -$ a.s. as $n \to \infty$.

Now it is unclear that those nice properties $\chi_{t,n}$ enjoyed in Step 2, e.g., (4.7) or (4.8), keep holding, because $(\varphi_n, h_n)_n$ that we now consider is a Cesàro mean of a subsequence of an approximating dual maximizer. But notice that this implies at least the upper bound in (4.8) continues to hold, that is,

(4.15) $\sup_n \chi_{t,n} \leq M_k$ on $J_{t-1,k}$.

We will show that the normalization of $\chi_{t,n}$ (4.7) can be restored. Fix any $a \in I_1$ as in Step 2, and define $\varphi_t^{\oplus} := \sum_i \varphi_{t,i}$. First, (4.14) implies that there exists $a_t \in I_t$ for every $t \in [N]$ such that

(4.16) $\lim_{n \to \infty} \varphi_{t,n}^{\oplus}(a_t) = \varphi_t^{\oplus}(a_t) \in \mathbb{R}$.

In view of (4.4) which gives $\varphi_{1,n}^{\oplus} \leq \chi_{2,n}$, this implies

(4.17) $\inf_n \chi_{2,n}(a_1) > -\infty$.

On the other hand, since $I_1 = \text{int}(J_1)$ and $J_{1,k} \nearrow J_1$, for large enough k we have $\{a, a_1\} \subseteq \text{int}(J_{1,k})$. Now (4.15), (4.17) imply that both $\chi_{2,n}(a)$ and $\nabla\chi_{2,n}(a)$ are uniformly bounded in n, where $\nabla\chi_{2,n}(a) \in \partial\chi_{2,n}(a)$ is a subgradient of the convex function $\chi_{2,n}$ at a. Hence by taking a subsequence, we can assume that $\{\chi_{2,n}(a)\}_n$ and $\{\nabla\chi_{2,n}(a)\}_n$ both converge. Then as in Step 1, define an affine function $L_{2,n}(y) = \chi_{2,n}(a) + \nabla\chi_{2,n}(a) \cdot (y - a)$, and replace $\varphi_{1,n}^{\oplus}(x_1)$ with $\varphi_{1,n}^{\oplus}(x_1) - L_{2,n}(x_1)$, $\varphi_{2,n}^{\oplus}(x_2)$ with $\varphi_{2,n}^{\oplus}(x_2) + L_{2,n}(x_2)$, and finally $h_{1,n}(x_1)$ with $h_{1,n}(x_1) - \nabla\chi_{2,n}(a)$. This yields $\chi_{2,n}(a) = \nabla\chi_{2,n}(a) = 0$ for all n, while the a.s. convergence of $\varphi_{t,i,n}$ and the bound (4.15) are retained. Next, the inequality

$\chi_{3,n} \geq \chi_{2,n} + \varphi_{2,n}^{\oplus}$ with $\chi_{2,n} \geq 0$ yields $\inf_n \chi_{3,n}(a_2) > -\infty$. Thus we can repeat the argument and achieve the normalization (4.7), while the a.s. convergence of $\varphi_{t,i,n}$ and the bound (4.15) are still retained. Now (4.5) and Proposition 4.1 yield $\lim_{n \to \infty} \chi_{t,n} = \chi_t$ pointwise on J_{t-1}.

Step 4. We have obtained the almost sure limit functions $(\varphi_{t,i})_{t,i}$. We may define $\varphi_{t,i} := -\infty$ on a $\mu_{t,i}$-null set which includes $\mathbb{R} \setminus I_t$, so that they are defined everywhere on $\mathbb{R}$. We will now show there exist functions $h_t = (h_{t,i})_i : \mathbb{R}^{td} \to \mathbb{R}^d$ for all $t \in [N]$ with $h_N \equiv 0$, such that

$$(4.18) \qquad \sum_{t=1}^{N} \left(\varphi_t^{\oplus}(x_t) + h_t(\bar{x}_t) \cdot \Delta x_t \right) \leq c(x) \ \text{ for all } x \in \mathbb{R}^{Nd}.$$

For any function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ which is bounded below by an affine function, let $\text{conv}[f] : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ denote the lower semi-continuous convex envelope of f, that is the supremum of all affine functions l satisfying $l \leq f$ (If there is no such l, let $\text{conv}[f] \equiv -\infty$.) We will inductively obtain $h_{N-1}, h_{N-2}, ..., h_1$. Let us rewrite (4.1) as

$$(4.19) \qquad \sum_{t=1}^{N-1} \left(\varphi_{t,n}^{\oplus}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) \leq c(x) - \varphi_{N,n}^{\oplus}(x_N).$$

Define $H_{N-1,n}(\bar{x}_{N-1}, x_N) = \text{conv}[c(\bar{x}_{N-1}, \cdot) - \varphi_{N,n}^{\oplus}(\cdot)](x_N)$. We have

$$\sum_{t=1}^{N-1} \left(\varphi_{t,n}^{\oplus}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) \leq H_{N-1,n}(\bar{x}_{N-1}, x_N) \leq c(x) - \varphi_{N,n}^{\oplus}(x_N)$$

because the left hand side is affine in x_N. If we let $x_N = x_{N-1}$, we get

$$(4.20) \quad \sum_{t=1}^{N-2} \left(\varphi_{t,n}^{\oplus}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) \leq H_{N-1,n}(\bar{x}_{N-1}, x_{N-1}) - \varphi_{N-1,n}^{\oplus}(x_{N-1}).$$

Notice (4.19) and (4.20) have the same structure. With the convention $H_{N,n}(x) := c(x)$, this allows us to inductively deduce, backward in t,

$$(4.21) \qquad \sum_{s=1}^{t} \left(\varphi_{s,n}^{\oplus}(x_s) + h_{s,n}(\bar{x}_s) \cdot \Delta x_s \right) \leq H_{t+1,n}(\bar{x}_{t+1}, x_{t+1}) - \varphi_{t+1,n}^{\oplus}(x_{t+1})$$

for $t = 1, ..., N - 1$, where

(4.22) $\qquad H_{t,n}(\bar{x}_t, x_{t+1}) := \text{conv}[H_{t+1,n}(\bar{x}_t, \cdot) - \varphi^{\oplus}_{t+1,n}(\cdot)](x_{t+1})$

with an abuse of notation $H_{t+1,n}(\bar{x}_t, x_{t+1}) := H_{t+1,n}(\bar{x}_{t+1}, x_{t+1})$.

Now by dropping n, we analogously define

(4.23) $\qquad H_t(\bar{x}_t, x_{t+1}) := \text{conv}[H_{t+1}(\bar{x}_t, \cdot) - \varphi^{\oplus}_{t+1}(\cdot)](x_{t+1})$

with $H_N(x) := c(x)$. Next, since the $\limsup$ of convex functions is convex, in conjunction with the almost sure convergence, we have

$$\begin{aligned}
\limsup_{n\to\infty} H_{N-1,n}(\bar{x}_{N-1}, x_N) &= \limsup_{n\to\infty} \text{conv}[c(\bar{x}_{N-1}, \cdot) - \varphi^{\oplus}_{N,n}(\cdot)](x_N) \\
&\leq \text{conv}[\limsup_{n\to\infty} \left(c(\bar{x}_{N-1}, \cdot) - \varphi^{\oplus}_{N,n}(\cdot)\right)](x_N) \\
&\leq \text{conv}[c(\bar{x}_{N-1}, \cdot) - \varphi^{\oplus}_{N}(\cdot)](x_N) \\
&= H_{N-1}(\bar{x}_{N-1}, x_N).
\end{aligned}$$

This allows us to inductively deduce, for $t = 1, ..., N - 1$,

$$\begin{aligned}
\limsup_{n\to\infty} H_{t,n}(\bar{x}_t, x_{t+1}) &= \limsup_{n\to\infty} \text{conv}[H_{t+1,n}(\bar{x}_t, \cdot) - \varphi^{\oplus}_{t+1,n}(\cdot)](x_{t+1}) \\
&\leq \text{conv}[\limsup_{n\to\infty} \left(H_{t+1,n}(\bar{x}_t, \cdot) - \varphi^{\oplus}_{t+1,n}(\cdot)\right)](x_{t+1}) \\
&\leq \text{conv}[H_{t+1}(\bar{x}_t, \cdot) - \varphi^{\oplus}_{t+1}(\cdot)](x_{t+1}) \\
&= H_t(\bar{x}_t, x_{t+1}).
\end{aligned}$$

We need to discuss continuity of the convex function $x_{t+1} \mapsto H_t(\bar{x}_t, x_{t+1})$. The following inequality from (4.23)

$$H_t(\bar{x}_t, x_{t+1}) \leq H_{t+1}(\bar{x}_{t+1}, x_{t+1}) - \varphi^{\oplus}_{t+1}(x_{t+1})$$

becomes $H_{N-1}(\bar{x}_{N-1}, x_N) \leq c(x) - \varphi^{\oplus}_{N}(x_N)$ when $t = N - 1$. Then the $\mu^{\otimes}_N$-a.s. finiteness of $\varphi^{\oplus}_N$ gives, by convexity, $H_{N-1}(\bar{x}_{N-1}, x_N) < \infty$ if $x_N \in J_{N-1}$. Backward induction in t then gives $H_t(\bar{x}_t, x_{t+1}) < \infty$ if $x_{t+1} \in J_t$. This implies that for any $\bar{x}_t \in \mathbb{R}^{td}$, if there exists $y_0 \in \mathbb{R}^d$ such that $H_t(\bar{x}_t, y_0) > -\infty$, then $y \mapsto H_t(\bar{x}_t, y)$ is real-valued thus continuous in J_t. Now (4.21), (4.22) gives

$$\varphi^{\oplus}_{1,n}(x_1) + h_{1,n}(x_1) \cdot \Delta x_1 \leq H_{1,n}(x_1, x_2) \leq H_{2,n}(x_1, x_2, x_2) - \varphi^{\oplus}_{2,n}(x_2).$$

Letting $x_2 = x_1$ gives $\varphi_{1,n}^\oplus(x_1) \leq H_{1,n}(x_1, x_1)$. From this and the almost sure convergence, taking $\limsup$ yields

$$\varphi_1^\oplus(x_1) \leq H_1(x_1, x_1) \quad \text{and} \quad H_1(x_1, x_2) \leq H_2(x_1, x_2, x_2) - \varphi_2^\oplus(x_2).$$

Set $A_t := \{x_t \in \mathbb{R}^d \mid \varphi_t^\oplus(x_t) \in \mathbb{R}\}$, $t \in [N]$, and note that $A_t \subseteq I_t$. Since $x_2 \mapsto H_1(x_1, x_2)$ is continuous in J_1 for every $x_1 \in A_1$, the subdifferential $\partial H_1(x_1, \cdot)(x_2)$ is nonempty, convex and compact for every $x_2 \in I_1 = \text{int}(J_1)$. This allows us to choose a measurable function $h_1 : A_1 \to \mathbb{R}^d$ satisfying $h_1(x_1) \in \partial H_1(x_1, \cdot)(x_1)$. Then for $x_1 \in A_1$, we have

$$\varphi_1^\oplus(x_1) + h_1(x_1) \cdot (x_2 - x_1) \leq H_1(x_1, x_1) + h_1(x_1) \cdot (x_2 - x_1)$$
$$\leq H_1(x_1, x_2)$$
$$\leq H_2(x_1, x_2, x_2) - \varphi_2^\oplus(x_2).$$

In particular, for $x_1 \in A_1$ and $x_2 \in A_2$, it holds $H_2(x_1, x_2, x_2) > -\infty$. Hence again we can choose $h_2 : A_1 \times A_2 \to \mathbb{R}^d$ that satisfies $h_2(x_1, x_2) \in \partial H_2(x_1, x_2, \cdot)(x_2)$. Then for every $x_1 \in A_1$ and $x_2 \in A_2$, we have

$$\varphi_1^\oplus(x_1) + \varphi_2^\oplus(x_2) + h_1(x_1) \cdot (x_2 - x_1) + h_2(x_1, x_2) \cdot (x_3 - x_2)$$
$$\leq H_2(x_1, x_2, x_2) + h_2(x_1, x_2) \cdot (x_3 - x_2)$$
$$\leq H_2(x_1, x_2, x_3)$$
$$\leq H_3(x_1, x_2, x_3, x_3) - \varphi_3^\oplus(x_3).$$

By induction in t, we obtain $h_t : A_1 \times \cdots \times A_t \to \mathbb{R}^d$ (with $h_N \equiv 0$), satisfying (4.18) as desired. We may define $h_t = 0$ in $\mathbb{R}^{td} \setminus A_1 \times \cdots \times A_t$, noting that the left hand side of (4.18) is $-\infty$ if $x_t \notin A_t$ for some t.

Step 5. We will show that for any functions $h_t : \mathbb{R}^{td} \to \mathbb{R}^d$, $t = 1, ..., N$ with $h_N \equiv 0$ satisfying (4.18) (whose existence was shown in Step 4), and for any minimizer π^* for the problem (2.2), it holds

(4.24) $$\sum_{t=1}^N \left(\varphi_t^\oplus(x_t) + h_t(\bar{x}_t) \cdot \Delta x_t \right) = c(x), \quad \pi^* - a.s..$$

In other words, every minimizer π^* is concentrated on the contact set

$$\Gamma := \left\{ x \in \mathbb{R}^{Nd} \;\middle|\; \sum_{t=1}^{N} \left(\varphi_t^{\oplus}(x_t) + h_t(\bar{x}_t) \cdot \Delta x_t \right) = c(x) \right\}$$

whenever $\{h_t\}_t$ is chosen to satisfy (4.18). This will complete the proof.

Recall $\varphi_{t,i,n} \to \varphi_{t,i}$ $\mu_{t,i}$-a.s. and $\chi_{t,n} \to \chi_t$ in J_{t-1} where $\chi_{t,n}$ is defined in (4.3) with χ_t being its limit. For any $\pi \in \mathrm{VMT}(\mu)$ (not necessarily an optimizer), we have $c \in L^1(\pi)$ by the assumption of Theorem 3.1. We claim:

$$(4.25) \qquad \limsup_{n \to \infty} \int \sum_{t=1}^{N} \left(\varphi_{t,n}^{\oplus}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) d\pi$$

$$\leq \int \sum_{t=1}^{N} \left(\varphi_t^{\oplus}(x_t) + h_t(\bar{x}_t) \cdot \Delta x_t \right) d\pi.$$

To see how the claim implies (4.24), let π^* be any minimizer for (2.2) (which exists by the assumption on c). Then $P(c) = \int c \, d\pi^*$, hence

$$P(c) = \lim_{n \to \infty} \int \sum_{t=1}^{N} \left(\varphi_{t,n}^{\oplus}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) d\pi^*$$

$$\leq \int \sum_{t=1}^{N} \left(\varphi_t^{\oplus}(x_t) + h_t(\bar{x}_t) \cdot \Delta x_t \right) d\pi^*$$

$$\leq \int c(x) \, d\pi^* = P(c)$$

hence equality holds throughout. Notice that this implies (4.24).

To prove (4.25), we will extend the argument given in [6], [48] into the current multi-period vector-valued setting. Fix any $\pi \in \mathrm{VMT}(\mu)$. Denoting $\pi = \mathrm{Law}(X)$ where $X = (X_1, ..., X_N)$ is an $\mathbb{R}^d$-valued martingale under π, we let $\pi_t := \mathrm{Law}(X_t)$. Then as in Step 2 (but using $\pi_t \preceq_c \pi_{t+1}$ instead of $\mu_t^{\otimes} \preceq_c \mu_{t+1}^{\otimes}$), by (4.2), (4.4), we have (cf. (4.9))

$$(4.26) \qquad \| \chi_{t+1,n} - \chi_{t,n} - \varphi_{t,n}^{\oplus} \|_{L^1(\pi_t)} \leq C \quad \text{for all } n \in \mathbb{N}, t \in [N].$$

From this, as $\varphi_{t,n}^{\oplus} \to \varphi_t^{\oplus}$ and $\chi_{t,n} \to \chi_t$, we deduce by Fatou's lemma,

$$\chi_{t+1} - \chi_t - \varphi_t^{\oplus} \in L^1(\pi_t), \quad \text{and}$$

$$\limsup_{n\to\infty} \int (\varphi_{t,n}^{\oplus} + \chi_{t,n} - \chi_{t+1,n})\, d\pi_t \le \int (\varphi_t^{\oplus} + \chi_t - \chi_{t+1})\, d\pi_t,$$

recalling $\chi_{1,n} = \chi_{N+1,n} \equiv 0$ and $\pi_t(J_{t-1}) = 1$. This allows us to deduce

$$\limsup_{n\to\infty} \int \sum_{t=1}^{N} \left(\varphi_{t,n}^{\oplus}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\right) d\pi$$

$$= \limsup_{n\to\infty} \int \sum_{t=1}^{N} \left(\varphi_{t,n}^{\oplus}(x_t) + \chi_{t,n}(x_t) - \chi_{t+1,n}(x_t)\right.$$

$$\left. - \chi_{t,n}(x_t) + \chi_{t+1,n}(x_t) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\right) d\pi$$

$$(4.27) \qquad \le \sum_{t=1}^{N} \int \left(\varphi_t^{\oplus} + \chi_t - \chi_{t+1}\right) d\pi_t$$

$$+ \limsup_{n\to\infty} \int \sum_{t=1}^{N-1} \left(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\right) d\pi,$$

since $\sum_{t=1}^{N} \chi_{t,n}(x_t) = \sum_{t=1}^{N} \chi_{t+1,n}(x_{t+1})$. Now denote $\bar{X}_t = (X_1, ..., X_t)$, $\pi^t :=$ Law$(\bar{X}_t) \in \mathcal{P}(\mathbb{R}^{td})$. Then we can write $\pi^{t+1} = \pi_{\bar{x}_t} \otimes \pi^t$, where $\pi_{\bar{x}_t} \in \mathcal{P}(\mathbb{R}^d)$ is the conditional distribution of X_{t+1} given $\bar{X}_t = \bar{x}_t$ under the martingale law π. Martingale property means that $\int y\, d\pi_{\bar{x}_t}(y) = x_t$. For each t, choose a sequence

of functions $\xi_{t,n} : I_t \to \mathbb{R}^d$ satisfying $\xi_{t,n}(x_t) \in \partial\chi_{t+1,n}(x_t)$. Then we can compute

$$\int \sum_{t=1}^{N-1} \big(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\big)\, d\pi$$

$$= \iint \sum_{t=1}^{N-1} \big(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\big) d\pi_{\bar{x}_{N-1}}(x_N)\, d\pi^{N-1}(\bar{x}_{N-1})$$

$$= \int \bigg(\int \big(\chi_{N,n}(x_{N-1}) - \chi_{N,n}(x_N) + h_{N-1,n}(\bar{x}_{N-1}) \cdot \Delta x_{N-1}\big) d\pi_{\bar{x}_{N-1}}(x_N)$$

$$+ \sum_{t=1}^{N-2} \big(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\big) \bigg) d\pi^{N-1}(\bar{x}_{N-1})$$

$$= \int \bigg(\int \big(\chi_{N,n}(x_{N-1}) - \chi_{N,n}(x_N) + \xi_{N-1,n}(x_{N-1}) \cdot \Delta x_{N-1}\big) d\pi_{\bar{x}_{N-1}}(x_N)$$

$$+ \sum_{t=1}^{N-2} \big(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\big) \bigg) d\pi^{N-1}(\bar{x}_{N-1}),$$

where the last equality is due to the martingale property

$$(4.28) \qquad \int h_{N-1,n}(\bar{x}_{N-1}) \cdot \Delta x_{N-1}\, d\pi_{\bar{x}_{N-1}}(x_N)$$

$$= \int \xi_{N-1,n}(x_{N-1}) \cdot \Delta x_{N-1}\, d\pi_{\bar{x}_{N-1}}(x_N)$$

$$= 0.$$

Furthermore, by definition of ξ, we have

$$\chi_{N,n}(x_{N-1}) - \chi_{N,n}(x_N) + \xi_{N-1,n}(x_{N-1}) \cdot \Delta x_{N-1} \leq 0.$$

This allows us to further disintegrate $\pi^{N-1} = \pi_{\bar{x}_{N-2}} \otimes \pi^{N-2}$ and apply the same argument. By repeating, we obtain

$$\int \sum_{t=1}^{N-1} \big(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t\big)\, d\pi$$

$$= \int \cdots \int \big(\chi_{N,n}(x_{N-1}) - \chi_{N,n}(x_N) + \xi_{N-1,n}(x_{N-1}) \cdot \Delta x_{N-1}\big) d\pi_{\bar{x}_{N-1}}(x_N)$$

$$+ \big(\chi_{N-1,n}(x_{N-2}) - \chi_{N-1,n}(x_{N-1}) + \xi_{N-2,n}(x_{N-2}) \cdot \Delta x_{N-2}\big) d\pi_{\bar{x}_{N-2}}(x_{N-1})$$

$$+ \cdots + \big(\chi_{2,n}(x_1) - \chi_{2,n}(x_2) + \xi_{1,n}(x_1) \cdot \Delta x_1\big) d\pi_{\bar{x}_1}(x_2)\, d\pi^1(x_1).$$

Since $\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + \xi_{t,n}(x_t) \cdot \Delta x_t \leq 0$ for all t, repeated application of Fatou's lemma allows $\limsup$ to continue to penetrate into the innermost integral. Along the way, we also use the inequality $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$. This eventually yield

$$\limsup_{n \to \infty} \int \sum_{t=1}^{N-1} \left(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) d\pi$$

$$\leq \int \cdots \int \left(\chi_N(x_{N-1}) - \chi_N(x_N) + \xi_{N-1}(x_{N-1}) \cdot \Delta x_{N-1} \right) d\pi_{\bar{x}_{N-1}}(x_N)$$

$$+ \left(\chi_{N-1}(x_{N-2}) - \chi_{N-1}(x_{N-1}) + \xi_{N-2}(x_{N-2}) \cdot \Delta x_{N-2} \right) d\pi_{\bar{x}_{N-2}}(x_{N-1})$$

$$+ \cdots + \left(\chi_2(x_1) - \chi_2(x_2) + \xi_1(x_1) \cdot \Delta x_1 \right) d\pi_{\bar{x}_1}(x_2) d\pi^1(x_1)$$

for some $\xi_t(x_t) \in \partial \chi_{t+1}(x_t)$ which is a limit point of the bounded sequence $\{\xi_{t,n}(x_t)\}_n$. Lastly, the martingale property (4.28) allows us to substitute $\xi_t(x_t)$ back to $h_t(\bar{x}_t)$ and get

$$\limsup_{n \to \infty} \int \sum_{t=1}^{N-1} \left(\chi_{t+1,n}(x_t) - \chi_{t+1,n}(x_{t+1}) + h_{t,n}(\bar{x}_t) \cdot \Delta x_t \right) d\pi$$

$$\leq \int \cdots \int \left(\chi_N(x_{N-1}) - \chi_N(x_N) + h_{N-1}(\bar{x}_{N-1}) \cdot \Delta x_{N-1} \right) d\pi_{\bar{x}_{N-1}}(x_N)$$

$$+ \left(\chi_{N-1}(x_{N-2}) - \chi_{N-1}(x_{N-1}) + h_{N-2}(\bar{x}_{N-2}) \cdot \Delta x_{N-2} \right) d\pi_{\bar{x}_{N-2}}(x_{N-1})$$

$$+ \cdots + \left(\chi_2(x_1) - \chi_2(x_2) + h_1(\bar{x}_1) \cdot \Delta x_1 \right) d\pi_{\bar{x}_1}(x_2) d\pi^1(x_1)$$

$$= \int \sum_{t=1}^{N-1} \left(\chi_{t+1}(x_t) - \chi_{t+1}(x_{t+1}) + h_t(\bar{x}_t) \cdot \Delta x_t \right) d\pi.$$

Finally, in (4.27) we can combine the integrals which then yields the claim (4.25), hence the theorem. $\square$

References

[1] E. Bayraktar, A.M.G. Cox and Y. Stoev. Martingale Optimal Transport with Stopping. *SIAM Journal on Control and Optimization*, 56.1 (2018): 417–433.

[2] M. Beiglböck, A.M.G. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *Invent. Math.*, 208 (2017) 327–400.

[3] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: a mass transport approach. *Finance and Stochastics*, 17 (2013) 477–501.

[4] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1) (2016) 42–106.

[5] M. Beiglböck, T. Lim and J. Obłój. Dual attainment for the martingale transport problem. *Bernoulli.* 25 (2019), no. 3, 1640–1658.

[6] M. Beiglböck, M. Nutz and N. Touzi. Complete duality for martingale optimal transport on the line. *Ann. Probab.*, 45(5) (2017) 3038–3074.

[7] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, Vol. 81, No. 3 (1973) 637–654.

[8] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, Volume 44, Issue 4 (1991) 375–417.

[9] D.T. Breeden and R.H. Litzenberger. Prices of state-contingent claims implicit in option prices. *J. Business*, 51(4):621–651 (1978).

[10] M. Burzoni, F. Riedel and H.M. Soner. Viability and arbitrage under Knightian uncertainty. *Econometrica*, 89(3) (2021) 1207–1234.

[11] G. Carlier, V. Chernozhukov and A. Galichon. Vector quantile regression: an optimal transport approach. *The Annals of Statistics*, 2016, Vol. 44, No. 3, 1165–1192.

[12] L. Carraro, N.E. Karoui and J. Obłój. On Azéma-Yor processes, their optimal properties and the Bachelier-drawdown equation. *Ann. Probab.*, Vol. 40, No. 1 (2012) 372–400.

[13] P. Cheridito, M. Kiiski, D.J. Prömel and H.M. Soner. Martingale Optimal Transport Duality. *Mathematische Annalen*, 379 (2021) 1685–1712.

[14] V. Chernozhukov, I. Fernandez-Val and A. Galichon. Quantile and Probability Curves without Crossing. *Econometrica* 78(3) (2010) 1093–1125.

[15] V. Chernozhukov, A. Galichon, M. Hallin and M. Henry. Monge-Kantorovich depth, quantiles, ranks and signs. *The Annals of Statistics*, 2017, Vol. 45, No. 1, 223–256.

[16] V. Chernozhukov, A. Galichon, M. Henry and B. Pass. Identification of hedonic equilibrium and nonseperable simultaneous equations. *Journal of Political Economy*, 129(3) (2021).

[17] A.M.G. Cox, J. Obłój and N. Touzi. The Root solution to the multi-marginal embedding problem: an optimal stopping and time-reversal approach. *Probab. Theory Related Fields.* 173 (2019), no. 1-2, 211–259.

[18] M. Dai, S. Kou, H.M. Soner and C. Yang. Leveraged Exchange-Traded Funds with Market Closure and Frictions. *Management Science*, Volume 69, Issue 4 (2023) 1935–2545.

[19] M.H.A. Davis. Model-Free Methods in Valuation and Hedging of Derivative Securities. *The Handbook of Post Crisis Financial Modeling*, 2016.

[20] H. De March. Local structure of multi-dimensional martingale optimal transport. arXiv preprint. https://arxiv.org/abs/1805.09469

[21] H. De March. Quasi-sure duality for multi-dimensional martingale optimal transport. arXiv preprint. https://arxiv.org/abs/1805.01757

[22] H. De March and N. Touzi. Irreducible convex paving for decomposition of multidimensional martingale transport plans. *Ann. Probab.* 47 (2019), No. 3, 1726–1774.

[23] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, volume 300, 463–520 (1994)

[24] F. Delbaen and W. Schachermayer. The mathematics of arbitrage. *Springer Finance*, 2006.

[25] D. Duffie, D. Filipović and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3): 984–1053 (2003)

[26] Y. Dolinsky and H.M. Soner. Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Relat. Fields.*, 160 (2014) 391–427.

[27] S. Eckstein, G. Guo, T. Lim, and J. Obłój. Robust Pricing and Hedging of Options on Multiple Assets and Its Numerics. *SIAM J. Financial Math.*, 12 (2021), no. 1, 158–188.

[28] A. Figalli, Y-H Kim and R.J. McCann. When is multidimensional screening a convex program? *J. Econom. Theory* 146 (2011) 454–478.

[29] J.P. Fouque, G. Papanicolaou and K.R. Sircar. Derivatives in financial markets with stochastic volatility. *Cambridge University Press*, 2000.

[30] J.P. Fouque, G. Papanicolaou, K.R. Sircar and K. Sølna. Multiscale stochastic volatility for equity, interest rate, and credit derivatives. *Cambridge University Press*, 2011.

[31] A. Galichon, P. Henry-Labordère, and N. Touzi. A Stochastic Control Approach to No-Arbitrage Bounds Given Marginals, with an Application to Lookback Options. *Annals of Applied Probability.*, Volume 24, Number 1 (2014) 312–336.

[32] W. Gangbo and R.J. McCann. The geometry of optimal transportation. *Acta Math.*, Volume 177, Issue 2 (1996) 113–161.

[33] N. Ghoussoub, Y-H. Kim, and T. Lim. Structure of optimal martingale transport plans in general dimensions. *Ann. Probab.* 47(1): 109–164 (2019).

[34] N. Ghoussoub, Y-H. Kim and T. Lim. Optimal Brownian stopping when the source and target are radially symmetric distributions. *SIAM J. Control Optim.* 58 (2020), no. 5, 2765–2789.

[35] N. Ghoussoub, Y-H. Kim and A.Z. Palmer. A solution to the Monge transport problem for Brownian martingales. *Ann. Probab.* 49 (2021), no. 2, 877–907.

[36] N. Ghoussoub, Y-H. Kim and A.Z. Palmer. Optimal stopping of stochastic transport minimizing submartingale costs. *Trans. Amer. Math. Soc.*, 374 (2021), no. 10, 6963–6989.

[37] G. Guo, X. Tan and N. Touzi. On the monotonicity principle of optimal Skorokhod embedding problem. *SIAM J. Control Optim.*, 54-5 (2016) 2478–2489.

[38] G. Guo, X. Tan and N. Touzi. Optimal Skorokhod embedding under finitely-many marginal constraints. *SIAM J. Control Optim.*, 54-4 (2016) 2174–2201.

[39] J. Guyon and P. Henry-Labordère. Nonlinear option pricing. *Chapman & Hall/CRC Financ. Math. Ser.* CRC Press, Boca Raton, FL, 2014.

[40] P. Henry-Labordère. Model-free Hedging: A Martingale Optimal Transport Viewpoint. *Chapman & Hall/CRC Financ. Math. Ser.* CRC Press, Boca Raton, FL, 2017.

[41] D. Hobson. Robust hedging of the lookback option. *Finance and Stochastics*, 2 (1998) 329–347.

[42] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, Springer, Berlin (2011) 267–318.

[43] D. Hobson and M. Klimmek. Robust price bounds for the forward starting straddle. *Finance and Stochastics*, Volume 19, Issue 1 (2014) 189–214.

[44] D. Hobson and A. Neuberger. Robust bounds for forward start options. *Mathematical Finance*, Volume 22, Issue 1 (2012) 31–56.

[45] H. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete.*, 67(4) (1984) 399–432.

[46] D. Kramkov and Y. Xu. An optimal transport problem with backward martingale constraints motivated by insider trading. *Ann. Appl. Probab.*, 32 (2022), no. 1, 294–326.

[47] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.*, 9(3): 904–950 (1999)

[48] T. Lim. Geometry of vectorial martingale optimal transportations and duality. *Mathematical Programming*, https://doi.org/10.1007/s10107-023-01954-4 (2023)

[49] A. Neufeld, A. Papapantoleon and Q. Xiang. Model-free bounds for multi-asset options using option-implied information and their exact computation. *Management Science*, 69(4) (2023) 2051–2068.

[50] M. Nutz, F. Stebegg and X. Tan. Multiperiod martingale transport. *Stochastic Process. Appl.*, 130 (2020), no. 3, 1568–1615.

[51] J. Obłój. The Skorokhod embedding problem and its offspring. *Probability Surveys*, 1 (2004) 321–392.

[52] J. Obłój and P. Siorpaes. Structure of martingale transports in finite dimensions. arXiv preprint. https://arxiv.org/abs/1702.08433

[53] J. Obłój and P. Spoida. An iterated Azéma-Yor type embedding for finitely many marginals. *Ann. Probab.*, 45(4): 2210–2247 (2017).

[54] B. Pass. Convexity and multi-dimensional screening for spaces with different dimensions. *J. Econom. Theory.* 147 (2012) 2399–2418.

[55] F. Santambrogio. Optimal transport for applied mathematicians. Calculus of variations, PDEs, and modeling. *Progress in Nonlinear Differential Equations and their Applications*, 87 Birkhuser/Springer, Cham, (2015)

[56] W. Schachermayer. Asymptotic Theory of Transaction Costs. *Zur. Lect. Adv. Math.*, European Mathematical Society (EMS), Zürich, 2017.

[57] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36 (1965) 423-439.

[58] C. Villani. *Optimal Transport. Old and New*, Vol. 338, *Grundlehren der mathematischen Wissenschaften.* Springer (2009).

[59] D. Zaev. On the Monge-Kantorovich problem with additional linear constraints. *Mathematical Notes.*, 98(5-6) (2015) 725–741.

Tongseok Lim: Mitchell E. Daniels, Jr. School of Business
Purdue University, West Lafayette, Indiana 47907, USA

Email address: lim336@purdue.edu