MARTINGALE SCHRÖDINGER BRIDGES AND OPTIMAL SEMISTATIC PORTFOLIOS

MARCEL NUTZ, JOHANNES WIESEL, AND LONG ZHAO

ABSTRACT. In a two-period financial market where a stock is traded dynamically and European options at maturity are traded statically, we study the so-called martingale Schrödinger bridge Q_*; that is, the minimal-entropy martingale measure among all models calibrated to option prices. This minimization is shown to be in duality with an exponential utility maximization over semistatic portfolios. Under a technical condition on the physical measure P, we show that an optimal portfolio exists and provides an explicit solution for Q_*. This result overcomes the remarkable issue of non-closedness of semistatic strategies discovered by Acciaio, Larsson and Schachermayer. Specifically, we exhibit a dense subset of calibrated martingale measures with particular properties to show that the portfolio in question has a well-defined and integrable option position.

1. Introduction and Main Results

The martingale Schrödinger bridge was introduced by [24] as a pricing model achieving perfect calibration to all Vanilla options while retaining stylized facts of a reference model. Starting from a reference stochastic volatility model (SVM) which typically cannot be calibrated perfectly, the martingale Schrödinger bridge is constructed as the calibrated measure which is closest to the SVM in the sense of relative entropy. In contrast to the classical Schrödinger bridge [27] and [3, 4], this problem features an additional martingale constraint to generate an arbitrage-free model. A similar approach is used by [22, 23] in a two-period setting to solve the longstanding joint S&P 500/VIX smile calibration puzzle; here entropy minimization is utilized to construct a model that is jointly calibrated to the S&P 500, VIX futures and VIX options.

The aforementioned works rest on (sometimes implicit) mathematical assumptions of strong duality and attainment. These are plausible as natural extensions of standard results in markets without option trading (see [13, 18, 37, 40], among others). However, [1] exhibited a surprising obstacle to obtaining such extensions: the space of semistatic portfolios of stocks and options is not closed (both in a two-period model and in continuous time). In classical mathematical finance, closedness

Date: April 25, 2022.

2010 *Mathematics Subject Classification.* 91G10; 60G42; 60H05; 94A17; 26B40.

Key words and phrases. Martingale Schrödinger bridge; semistatic trading.

MN acknowledges support by an Alfred P. Sloan Fellowship and NSF Grants DMS-1812661, DMS-2106056.

results are at the very heart of the separation arguments underlying the Fundamental Theorem of Asset Pricing and the existence of optimal portfolios for utility maximization. As a consequence, it is not obvious how to formulate and prove the desired results.

The purpose of the present paper is to provide such results, at least in one setting. On the one hand, we prove strong duality between the martingale Schrödinger bridge problem and an exponential utility maximization problem over semistatic portfolios. This duality, as well as the existence of the martingale Schrödinger bridge itself (primal attainment), is obtained along the lines of classical entropy minimization and Schrödinger bridge theory. On the other hand, we prove (under a technical condition) that the dual problem is attained in a natural space of admissible portfolios, and that this dual solution yields the log-density of the martingale Schrödinger bridge. We thus derive from first principles the type of implicit condition assumed on the optimal log-density, e.g., in [23, Theorem 16], and overcome the non-closedness issue discovered in [1]. To wit, while in general a convergent sequence of semistatic portfolios may have an undesirable limit with unclear financial interpretation, the specific limit of a utility-maximizing sequence in our problem is shown to be an admissible portfolio.

We consider a two-period model where the price of a stock is modeled by the canonical process (X, Y) on $\mathbb{R}^2$ under a (physical) reference probability P. Here X is the stock price at date $t = 1$ and Y is the price at the terminal date $t = 2$. In addition, European options $g(Y)$ are liquidly traded at time zero. By the Breeden–Litzenberger formula [9], the risk-neutral distribution ν of Y can be derived from the prices of call options with arbitrary strikes, and then the arbitrage-free price of a general option $g(Y)$ is given by the integral $E^\nu[g]$. The martingale Schrödinger bridge problem can now be formalized as

$$\inf_{Q \in \mathcal{M}(\nu)} H(Q|P), \tag{1.1}$$

where H is the relative entropy (or Kullback–Leibler divergence)

$$H(Q|P) := \begin{cases} E^Q\left[\log \frac{dQ}{dP}\right], & Q \ll P \\ \infty, & Q \not\ll P \end{cases}$$

and $\mathcal{M}(\nu)$ is the set of calibrated equivalent martingale measures,

$$\mathcal{M}(\nu) := \left\{Q \in \mathcal{P}(\mathbb{R}^2) : \ Q \sim P, \ Q^2 = \nu, \ E^Q[Y|X] = X\right\}. \tag{1.2}$$

Here $\mathcal{P}(\mathbb{R}^2)$ is the set of probability measures on $\mathbb{R}^2$ and Q^2 denotes the second marginal of $Q \in \mathcal{P}(\mathbb{R}^2)$, or equivalently, the distribution of the price Y under Q.

We remark in passing that (1.1) relates to the classical (static) Schrödinger bridge problem $\inf_{Q\in\Pi(\mu,\nu)} H(Q|P)$ over the set $\Pi(\mu, \nu)$ of couplings of two measures μ, ν; see [15, 27, 29] for surveys. In this problem, there is no martingale constraint. On the other hand, (1.1) relates to the martingale optimal transport problem $\inf_{Q\in\mathsf{M}(\mu,\nu)} E^Q[c]$ which minimizes an integrated cost over the set $\mathsf{M}(\mu, \nu)$ of martingale couplings; see [5, 19, 25] and the literature thereafter. In that problem, there is no reference measure. The resulting value yields model-independent bounds for the price of the exotic option c and as a consequence of the linear structure, solutions tend to be degenerate. By contrast, solutions of (1.1) tend to preserve features of the reference model P, as emphasized in [24]. The classical Schrödinger bridge problem arises from the classical optimal transport problem by entropic regularization as used in the context of Sinkhorn's algorithm [11, 33]. Similarly, entropic regularization of martingale optimal transport leads to the martingale Schrödinger bridge, and this was used in [28] to develop a version of Sinkhorn's algorithm for martingale optimal transport. See also [21] for a related algorithm using a different relaxation.

Returning to our problem (1.1)—for it to be meaningful, we must assume that

$$\mathcal{M}_{\mathrm{fin}}(\nu) := \{Q \in \mathcal{M}(\nu) : H(Q|P) < \infty\} \neq \emptyset; \tag{1.3}$$

that is, there exists a calibrated martingale measure with finite entropy. This condition implies the absence of arbitrage in semistatic trading strategies. It implies the usual no-arbitrage condition on the stock alone, but also depends on but also depends on the interplay of P and ν. A precise characterization of (1.3), or even just $\mathcal{M}(\nu) \neq \emptyset$, in terms of trading strategies along the lines of a fundamental theorem of asset pricing [12], is an interesting open problem. (Like the question studied in the present paper, the answer is not obvious due to the failure of closedness [1].) We can now state the basic wellposedness result.

Proposition 1.1. *The problem* (1.1) *admits a unique minimizer* $Q_* \in \mathcal{M}(\nu)$*, called the martingale Schrödinger bridge.*

This will essentially follow from standard entropy minimization theory [10] and properties of $\mathcal{M}(\nu)$ which are variations of results found, e.g., in [5]. Proposition 1.1 lacks a more specific description: we expect by (formal) duality that the log-density of Q_* corresponds to a semistatic portfolio with certain admissibility criteria, and those criteria are crucial for any further analysis of the martingale Schrödinger bridge and its computation (as seen, e.g., in [23]). Specifically, trading in our market gives rise to a semistatic outcome of the form

$$V = h(X)(Y - X) + g(Y),$$

where $h(X)$ is the number of stocks held over the second period. Stock trading in the first period, starting from a deterministic initial stock price X_0, corresponds to a term $h_0(X_0)(X - X_0)$ which can be absorbed into the functions h, g above and hence will not be represented explicitly. We write

$$\mathcal{V} = \{V \text{ measurable: } V = h(X)(Y - X) + g(Y) \text{ for some } h, g : \mathbb{R} \to \mathbb{R}\}. \tag{1.4}$$

In order to have a well-defined option price, the function g needs to be (measurable and) integrable under the pricing measure ν. We thus set

$$\mathcal{V}_1 := \{V \in \mathcal{V} : \ h, g \text{ are measurable, } g \in L^1(\nu),\ E^\nu[g] = 0\} \tag{1.5}$$

for those outcomes whose option is available from zero initial capital. Finally, we want $h(X)(Y - X)$ to have suitable martingale properties. There is some flexibility here regarding the definition; one natural choice is to require the martingale property under all $Q \in \mathcal{M}_{\text{fin}}(\nu)$ (see also Remark 2.7 for another possible choice). For $V \in \mathcal{V}_1$, this is equivalent to $V \in L^1(Q)$ for all $Q \in \mathcal{M}_{\text{fin}}(\nu)$. In summary, our set of admissible portfolios (for zero initial capital) is

$$\mathcal{V}_{\text{adm}} = \left\{ V \in \mathcal{V} : \begin{array}{c} h, g : \mathbb{R} \to \mathbb{R} \text{ are measurable, } E^{\nu}[g] = 0, \\ E^{Q}[h(X)(Y-X)] = 0 \text{ for all } Q \in \mathcal{M}_{\text{fin}}(\nu) \end{array} \right\}.$$

We then have the following strong duality between the martingale Schrödinger bridge (primal) problem and the dual problem of exponential utility maximization over semistatic portfolios.

Proposition 1.2. *Let $u(x) = -e^{-\gamma x}/\gamma$ for some $\gamma > 0$. Then*

$$\frac{1}{\gamma} \inf_{Q \in \mathcal{M}(\nu)} H(Q|P) = \sup_{V \in \mathcal{V}_{\text{adm}}} u^{-1}\big(E^{P}[u(V)]\big). \tag{1.6}$$

The duality will be obtained by showing that the log-density of Q_* can be approximated by semistatic portfolios with good integrability properties; cf. Proposition 2.4. That proposition, in turn, is inspired by seminal results in the theory of (classical) Schrödinger bridges, especially Föllmer's construction of Schrödinger potentials [15]. Our argument does not require dual attainment and thus avoids discussing delicate properties of the portfolios: the supremum in (1.6) would be the same if taken, say, over portfolios $h(X)(Y-X) + g(Y)$ with bounded continuous functions h, g. But of course, this space would not allow for attainment in general.

Turning to the delicate part, we want to show that the dual problem is attained at an admissible portfolio V_* and that this maximizer yields the log-density of $\mathcal{Q}_*$. We denote by $P = P^1 \otimes P^{\bullet}$ the disintegration of P; that is, P^1 is the law of X under P and $P^{\bullet}(x, dy)$ is the conditional law of Y given $X = x$.

Theorem 1.3. *Suppose that $dP^{\bullet}/d\nu$ is P^1-a.s. uniformly bounded from above and below. Then the minimizer Q_* of* (1.1) *is given by the density*

$$Z_* := \frac{dQ_*}{dP} = e^{H(Q_*|P) + V_*}, \tag{1.7}$$

where $V_* \in \mathcal{V}_{\text{adm}}$ is the unique solution of the dual problem,

$$V_* = \operatorname*{arg\,max}_{V \in \mathcal{V}_{\text{adm}}} E^{P}[u(V)].$$

In particular, $V_* = h(X)(Y-X) + g(Y)$, where h, g are measurable functions with $g \in L^1(\nu)$ and $E^{\nu}[g] = 0$ as well as $h(X)(Y-X) \in L^1(Q)$ and $E^{Q}[h(X)(Y-X)] = 0$ for all $Q \in \mathcal{M}_{\text{fin}}(\nu)$.

The boundedness condition in Theorem 1.3 can be weakened to an integrability condition; see Remark 3.4. In contrast to the other results, this theorem does not seem to follow from classical arguments. If the space of admissible portfolios were closed, the theorem would follow from the approximation result in Proposition 2.4, broadly as in the classical framework of mathematical finance without options. To overcome the failure of closedness (specifically, of $\mathcal{V}_1$ and $\mathcal{V}_{\text{adm}}$, as shown in [1]), we first leverage a result from our companion paper [32], where it is shown that the functional form of semistatic portfolios is stable under pointwise limits. As a consequence, the approximation result still implies that V_* is of the general form $V_* = h(X)(Y - X) + g(Y)$ for some measurable functions h, g.

On the flip side, another insight from [32] is that the key failure in the counterexample of [1] is the integrability of the option g which is in turn crucial to associate a price. Hence, it is not surprising that establishing this integrability occupies the lion's share of the proof of Theorem 1.3; it uses novel arguments and seems to be the first result in this direction. Our line of attack is to construct a measure $\widetilde{Q}$ in (a relaxation of) $\mathcal{M}_{\text{fin}}(\nu)$ such that $h(X)(Y-X)$ is $\widetilde{Q}$-integrable; once that is achieved, soft arguments imply that $g \in L^1(\nu)$. In fact, we establish that such measures are dense: in Proposition 3.2 we show that any $Q \in \mathcal{M}_{\text{fin}}(\nu)$ is the limit of calibrated (absolutely continuous) martingale measures Q_n under which the dynamic trading strategy h is uniformly bounded a.s. The proof is intricate and develops, among other things, explicit stability properties of the convex order, building on ideas from martingale optimal transport [6]. See also Section 3.3 for further comments.

We do not know how far the technical condition on P in Theorem 1.3 can be relaxed. However, analogy with the classical Schrödinger bridge problem suggests that some condition may be necessary. Indeed, the corresponding question in that setting—without martingale constraint but with two marginal constraints—is to show that the log-density of the Schrödinger bridge is of the form $f(x) + g(y)$ and establish the measurability and integrability properties of those "Schrödinger potentials" (f, g). This problem has a long history (e.g., [7]). A series of results revealed that the additive form $f(x) + g(y)$ always holds, but also that the measurability of (f, g) fails without additional conditions; moreover, even when measurability holds, integrability fails without further conditions (see [8, 10, 16, 35, 36]). The study of Schrödinger potentials remains an area of active study (see for instance [2, 14, 20, 30, 31]) that we have benefited from, especially for our companion paper [32]. For the present work, we have not been able to transfer as many of those techniques.

Regarding potential future work, it seems likely that our line of argument can be extended to show the existence of optimal portfolios for more general utility functions. Generalizations in the structure of the market, for instance also adding options with maturity $t = 1$, are relatively straightforward in the general parts whereas replacing our argument for the integrability of the option is nontrivial. In a different direction, one may remember how [34] used existence for exponential utility to show the Fundamental Theorem of Asset Pricing. Of course, that is not immediately applicable here, as we have used (1.3) in our proof of existence.

The remainder of the paper has a simple structure: Section 2 derives the wellposedness and duality results (Propositions 1.1 and 1.2), and Section 3 provides the proof of dual attainment (Theorem 1.3).

2. Wellposedness and Duality

In this section, we first prove the wellposedness of the martingale Schrödinger bridge Q_* (Proposition 1.1). Then, we prove the duality with exponential utility maximization (Proposition 1.2) through an approximation of Q_* (Proposition 2.4).

We start by recalling a general result on entropy minimization.

Lemma 2.1. *Consider a measurable space $(\Omega, \mathcal{F})$ and denote by $\mathcal{P}(\Omega)$ its collection of probability measures. Fix $R \in \mathcal{P}(\Omega)$, let $\mathcal{Q} \subseteq \mathcal{P}(\Omega)$ be convex and closed in variation, and suppose that $\mathcal{Q}_{\mathrm{fin}} := \{Q \in \mathcal{Q} : H(Q|R) < \infty\} \neq \emptyset$. Then there exists a unique $Q_* \in \mathcal{Q}$ such that*

$$H(Q_*|R) = \inf_{Q \in \mathcal{Q}} H(Q|R) \in [0, \infty).$$

Moreover, $Q_* \gg Q$ for any $Q \in \mathcal{Q}_{\mathrm{fin}}$. In particular, if there exists $Q \in \mathcal{Q}_{\mathrm{fin}}$ with $Q \sim R$, then $Q_* \sim R$. Furthermore,

$$\log \frac{dQ_*}{dR} \in L^1(Q) \quad \textit{for all} \quad Q \in \mathcal{Q}_{\mathrm{fin}}. \tag{2.1}$$

Proof. In the stated form, the result can be found in [29, Theorem 1.10 and Corollary 1.13]. Its main part is very classical; cf. [10]. The integrability (2.1) is less known but can also be deduced from [10]. □

Lemma 2.1 is not directly applicable to the set $\mathcal{Q} = \mathcal{M}(\nu)$ of martingale measures defined in (1.2) as this set is not closed due to the equivalence constraint. Writing

$$\Pi(\nu) = \{Q \in \mathcal{P}(\mathbb{R}^2) : Q^2 = \nu\},$$

we consider instead the following relaxations defined with absolute continuity,

$$\begin{aligned} \widetilde{\mathcal{M}}(\nu) &:= \big\{Q \in \Pi(\nu) : \ Q \ll P, \ E^Q[Y|X] = X\big\} \supseteq \mathcal{M}(\nu), \\ \widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu) &:= \{Q \in \widetilde{\mathcal{M}}(\nu) : H(Q|P) < \infty\} \supseteq \mathcal{M}_{\mathrm{fin}}(\nu) \end{aligned} \tag{2.2}$$

and argue that $\widetilde{\mathcal{M}}(\nu)$ satisfies the hypotheses of Lemma 2.1. To this end, we first give an extension of [5, Lemma 2.2 and Theorem 2.4]. Recall that two measures $\mu, \nu \in \mathcal{P}(\mathbb{R})$ are in convex order [38], denoted $\mu \preceq_c \nu$, if they have finite first moments and $E^\mu[f] \leq E^\nu[f]$ holds for all convex functions $f : \mathbb{R} \to \mathbb{R}$. As an example, $E^Q[Y|X] = X$ implies $Q^1 \preceq_c Q^2$ by Jensen's inequality.

Lemma 2.2. *The set $\{Q \in \Pi(\nu) : \ E^Q[Y|X] = X\}$ is weakly closed.*

Proof. Let $(Q_n)_{n \geq 1}$ be a sequence of measures converging weakly to some limit Q, then $Q \in \Pi(\nu)$ by the continuity of the projection Y. To see that $E^{Q_n}[Y|X] = X$ implies $E^Q[Y|X] = X$, we show that $|X| + |Y|$ is (Q_n)-uniformly integrable. Indeed, as $\{\nu\}$ is uniformly integrable, the la Valleé–Poussin theorem yields a convex function $f : \mathbb{R} \to \mathbb{R}_+$ of superlinear growth with $\int f \, d\nu < \infty$. Thus

$$\sup_{\mu :\, \mu \preceq_c \nu} \int f \, d\mu \leq \int f \, d\nu < \infty$$

by the definition of the convex order, showing that $\{\mu : \ \mu \preceq_c \nu\}$ is uniformly integrable. As a result, $|X| + |Y|$ is $\{Q \in \Pi(\nu) : Q^1 \preceq_c \nu\}$-uniformly integrable and in particular (Q_n)-uniformly integrable. □

We can now show the wellposedness of the martingale Schrödinger bridge Q_*.

Proof of Proposition 1.1. Using Lemma 2.2, we readily verify that $\widetilde{\mathcal{M}}(\nu)$ is convex and closed in variation. Since $\widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu) \supseteq \mathcal{M}_{\mathrm{fin}}(\nu) \neq \emptyset$ by our assumption (1.3), applying Lemma 2.1 with $\mathcal{Q} = \widetilde{\mathcal{M}}(\nu)$ yields existence and uniqueness of

$$Q_* = \operatorname*{arg\,min}_{Q \in \widetilde{\mathcal{M}}(\nu)} H(Q|P) \tag{2.3}$$

as well as $Q_* \sim P$; that is, $Q_* \in \mathcal{M}_{\mathrm{fin}}(\nu)$. It now follows that Q_* is also the unique minimizer of $\inf_{Q \in \mathcal{M}(\nu)} H(Q|P)$. □

We record the following observation for use in Section 3.

Remark 2.3. For any $Q \in \mathcal{M}_{\mathrm{fin}}(\nu)$, a straightforward calculation shows that the density $Z := dQ/dP$ can be written as $Z = e^{H(Q|P)+V}$ for some $V \in L^1(Q)$ with $E^Q[V] = 0$. For the density

$$Z_* := \frac{dQ_*}{dP} = e^{H(Q_*|P)+V_*}, \tag{2.4}$$

of the minimizer, we have not only that $E^{Q_*}[V_*] = 0$ but also that $V_* \in L^1(\widetilde{Q})$ for all $\widetilde{Q} \in \widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu)$. This follows from (2.1) by way of (2.3).

The next result characterizes the minimizer Q_* through certain approximating sequences of semistatic portfolios and will serve as the basis to prove the duality (Proposition 1.2). We write $\mathcal{V}_{\mathrm{b}}$ for the set of portfolios $V = h(X)(Y - X) + g(Y)$ where $h, g : \mathbb{R} \to \mathbb{R}$ are bounded measurable and $E^\nu[g] = 0$; clearly $\mathcal{V}_{\mathrm{b}} \subset \mathcal{V}_{\mathrm{adm}}$.

Proposition 2.4. *Given $Q_* \in \mathcal{M}_{\mathrm{fin}}(\nu)$ with density* (2.4)*, the following statements are equivalent:*

(i) *Q_* is the minimizer of* (1.1).

(ii) *There exist probability measures $(Q_n)_{n \geq 1}$ with densities*

$$Z_n := dQ_n/dP = e^{H(Q_n|P)+V_n} \quad \text{with} \quad V_n \in \mathcal{V}_{\mathrm{b}}$$

such that

$$H(Q_n|P) \to H(Q_*|P) \quad \text{and} \quad V_n \to V_* \text{ in } L^1(Q_*) \quad \text{as } n \to \infty.$$

(iii) *There exist $(V_n)_{n \geq 1} \subseteq \mathcal{V}_{\mathrm{b}}$ such that*

$$E^P\left[e^{V_n}\right] \to E^P\left[e^{V_*}\right] \quad \text{as } n \to \infty.$$

(iv) *There exist $(V_n)_{n \geq 1} \subseteq \mathcal{V}_{\mathrm{b}}$ such that*

$$E^Q\left|e^{V_n - V_*} - 1\right| \to 0 \quad \text{as } n \to \infty.$$

This result is inspired by a characterization of (classical) Schrödinger bridges in [16, Proposition 3.6], see also [10], and Föllmer's construction of Schrödinger potentials [15]. The key feature is that the approximating random variables V_n are portfolios and have good integrability properties (whereas the properties of V_* are unclear at this stage).

Remark 2.5. The assertion of Proposition 2.4 remains valid if $\mathcal{V}_{\mathrm{b}}$ is replaced by $\mathcal{V}_{\mathrm{adm}}$ or by $\mathcal{V}_1$. This will be clear from the proof.

Proof of Proposition 2.4. $(i) \Rightarrow (ii)$: By separability of $L^1(\mathbb{R})$ we can write

$$\mathcal{M}(\nu) = \{Q \sim P : E^Q[h_i(X)(Y-X)] = 0,\ E^Q[g_i(Y)] = 0,\ i = 1, 2, \dots\}$$

for a countable collection of bounded measurable functions $h_i, g_i : \mathbb{R} \to \mathbb{R}$. Denote the set of measures for which only the first n constraints are enforced by

$$\mathcal{M}_n(\nu) = \{Q \sim P : E^Q[h_i(X)(Y-X)] = 0,\ E^Q[g_i(Y)] = 0,\ i = 1, 2, \dots, n\}.$$

Clearly $\mathcal{M}(\nu) \subseteq \mathcal{M}_n(\nu)$ and $\mathcal{M}_n(\nu)$ is convex and closed in variation. Consider the problem $\inf_{Q \in \mathcal{M}_n(\nu)} H(Q|P)$. This minimization problem over measures with finitely many linear constraints is well known to be in duality with exponential utility maximization over (static) trading in the finitely many assets $h_i(X)(Y-X), g_i(Y)$ defining the constraints. Specifically, by [17, Section 3, esp. Corollary 3.25], the minimizer Q_n of $\inf_{Q \in \mathcal{M}_n(\nu)} H(Q|P)$ is of the form

$$Z_n := \frac{dQ_n}{dP} = \exp\Big(c_n + \tilde{h}_n(X)(Y-X) + \tilde{g}_n(Y)\Big)$$

for some $c_n \in \mathbb{R}$, where $\tilde{h}_n(X) = \sum_{i=1}^n a_{i,n} h_i(X)$ and $\tilde{g}_n(Y) = \sum_{i=1}^n b_{i,n} g_i(Y)$ for some $a_{i,n}, b_{i,n}, \in \mathbb{R}$. As $Q_n \in \mathcal{M}_n(\nu)$ we have $c_n = H(Q_n|P)$. That is, $\log Z_n$ is of the form $H(Q_n|P) + V_n$ for some $V_n \in \mathcal{V}_{\mathrm{b}}$. Applying [29, Theorem 1.17] to the sets $\mathcal{Q}_n := \mathcal{M}_n(\nu)$ and $\mathcal{Q} := \mathcal{M}(\nu)$ satisfying $\cap_n \mathcal{Q}_n = \mathcal{Q}$, and recalling that $\mathcal{M}_{\mathrm{fin}}(\nu) \neq \emptyset$ by our assumption (1.3), we conclude that

$$H(Q_*|Q_n) \to 0, \quad H(Q_n|P) \to H(Q_*|P) \quad \text{and} \quad \log Z_n \to \log Z_* \text{ in } L^1(Q_*).$$

In particular, $V_n \to V_*$ in $L^1(Q_*)$ follows.
$(ii) \Rightarrow (iii)$: Since Z_n, Z are probability densities, we have $e^{-H(Q_n|P)} = E^P[e^{V_n}]$ and $e^{-H(Q_*|P)} = E^P[e^{V_*}]$. Thus $H(Q_n|P) \to H(Q_*|P)$ is equivalent to $E^P[e^{V_n}] \to E^P[e^{V_*}]$.
$(iii) \Leftrightarrow (iv)$: By a change of measure, (iii) is equivalent to

$$E^{Q_*}[e^{V_n - V_*}] = e^{H(Q_*|P)} E^P[e^{V_n}] \to e^{H(Q_*|P)} E^P[e^{V_*}] = 1,$$

and now Scheffé's lemma yields the equivalence with (iv).
$(iii) \Rightarrow (i)$: Without loss of generality, we assume $E^P[e^{V_n}] < \infty$ for all n. Define probability measures Q_n by

$$Z_n := \frac{dQ_n}{dP} = e^{H(Q_n|P) + V_n}$$

and recall that (iii) is equivalent to $H(Q_n|P) \to H(Q_*|P)$. Take any $Q \in \mathcal{M}_{\mathrm{fin}}(\nu)$. Using the definition of $H(\cdot|P)$ and Lemma 2.6 below,

$$H(Q|P) - H(Q|Q_n) = E^Q[\log Z_n] = H(Q_n|P) + E^Q[V_n] = H(Q_n|P).$$

As $H(Q|Q_n) \geq 0$, it follows that

$$H(Q|P) \geq \lim_{n \to \infty} H(Q_n|P) = H(Q_*|P).$$

Since $Q \in \mathcal{M}_{\mathrm{fin}}(\nu)$ was arbitrary, we conclude that Q_* is the minimizer of (1.1). $\square$

The following technical result was used in the preceding proof.

Lemma 2.6. *Let $V \in \mathcal{V}_1$ satisfy $E^P[e^V] < \infty$. Then $V \in L^1(Q)$ and $E^Q[V] = 0$ for all $Q \in \mathcal{M}_{\mathrm{fin}}(\nu)$.*

Proof. Define an auxiliary probability measure Q' via

$$Z' := \frac{dQ'}{dP} = e^{c+V},$$

where $c \in \mathbb{R}$ is the normalization constant. Moreover, let $Q \in \mathcal{M}_{\mathrm{fin}}(\nu)$ and denote by Z its density. Applying the inequality $\log x \le x - 1$ to $x = z'/z > 0$ yields $\log z' \le \log z + z'/z - 1$ and hence

$$\log Z' \le \log Z + Z'/Z - 1 \quad \text{on} \quad \{Z > 0\},$$

where $\log 0 := -\infty$. In view of $H(Q|P) < \infty$, we have $\log Z + Z'/Z - 1 \in L^1(Q)$ and conclude that $(\log Z')^+ \in L^1(Q)$. By the definition of $\mathcal{V}_1$,

$$\log Z' = c + V = h(X)(Y - X) + g(Y)$$

for some $g \in L^1(\nu)$ with $E^\nu[g] = 0$, and hence $(\log Z')^+ \in L^1(Q)$ translates to the positive part of the martingale transform $h(X)(Y - X)$ being Q-integrable. As Q is a martingale measure, this already implies (see [26, Theorem 2b]) that $h(X)(Y - X) \in L^1(Q)$ and $E^Q[h(X)(Y - X)] = 0$. The claim follows. □

We are now in a position to prove the duality result.

Proof of Proposition 1.2. Let Q_* be the minimizer from Proposition 1.1 and recall from (2.4) the notation $Z_* = dQ_*/dP = e^{H(Q_*|P)+V_*}$ where $E^{Q_*}[V_*] = 0$. Let $u(x) = -e^{-\gamma x}/\gamma$ for some $\gamma > 0$. A change of measure and Jensen's inequality yield that for any $V \in \mathcal{V}_{\mathrm{adm}}$,

$$\begin{aligned} E^P[u(V)] = E^{Q_*}\left[Z_*^{-1} u(V)\right] &= -\frac{1}{\gamma} E^{Q_*}\left[e^{-H(Q_*|P) - V_* - \gamma V}\right] \\ &\le -\frac{1}{\gamma} e^{-H(Q_*|P) - E^{Q_*}[V_*] - \gamma E^{Q_*}[V]} = -\frac{1}{\gamma} e^{-H(Q_*|P)}, \end{aligned}$$

where the equality used that $E^{Q_*}[V] = 0$ due to $Q_* \in \mathcal{M}_{\mathrm{fin}}(\nu)$ and $V \in \mathcal{V}_{\mathrm{adm}}$.

On the other hand, Proposition 2.4 (iv) shows that there exist $(V_n) \subseteq \mathcal{V}_{\mathrm{b}}$ such that $E^{Q_*}\left[e^{-V_* - \gamma V_n}\right] \to 1$ and consequently $-E^{Q_*}\left[e^{-H(Q_*|P) - V_* - \gamma V_n}\right] \to -e^{-H(Q_*|P)}$. In view of $\mathcal{V}_{\mathrm{b}} \subseteq \mathcal{V}_{\mathrm{adm}}$, this yields $\sup_{V \in \mathcal{V}_{\mathrm{adm}}} E^P[u(V)] \ge -\frac{1}{\gamma} e^{H(Q_*|P)}$. Lastly,

$$\inf_{Q \in \mathcal{M}(\nu)} u\left(\frac{1}{\gamma} H(Q|P)\right) = u\left(\frac{1}{\gamma} H(Q_*|P)\right) = -\frac{1}{\gamma} e^{-H(Q_*|P)},$$

so that combining the two inequalities yields

$$\sup_{V \in \mathcal{V}_{\mathrm{adm}}} E^P[u(V)] = \inf_{Q \in \mathcal{M}(\nu)} u\left(\frac{1}{\gamma} H(Q|P)\right)$$

as claimed. □

Remark 2.7. By the proof, the duality (1.6) still holds if the supremum is taken over the larger set $\mathcal{V}_1 \cap L^1(Q_*) \supset \mathcal{V}_{\mathrm{adm}}$, providing an alternative definition of admissibility.

3. Admissibility and Dual Attainment

3.1. **Preliminary Considerations.** Let Q_* be the minimizer from Proposition 1.1 and recall from (2.4) the notation $Z_* = dQ_*/dP = e^{H(Q_*|P)+V_*}$. With a view towards the duality relation, note that

$$E^P\left[u\left(-\frac{1}{\gamma}V_*\right)\right] = -E^P\left[\frac{1}{\gamma}e^{V_*}\right] = -E^P\left[\frac{1}{\gamma}e^{-H(Q_*|P)}Z_*\right] = -\frac{1}{\gamma}e^{-H(Q_*|P)}.$$

It is thus tempting to conclude that $-V_*/\gamma$ "attains" the supremum in (1.6). However, it far from obvious whether V_* belongs to the dual domain $\mathcal{V}_{\mathrm{adm}}$ (or is a portfolio in any sense). At this stage, we know that $E^{Q_*}[V_*] = 0$ and that V_* is the limit of certain portfolios $V_n \in \mathcal{V}_{\mathrm{b}} \subset \mathcal{V}_{\mathrm{adm}} \subset \mathcal{V}_1$; cf. Proposition 2.4. The missing conclusion would be obvious if any of these spaces had a good closure property. However, as mentioned in the Introduction, [1] has shown that this is not the case: specifically, the authors exhibit a two-period model and an L^p-convergent sequence $V_n \in \mathcal{V}_{\mathrm{b}}$ whose limit is outside $\mathcal{V}_1$. The proof uses a clever contradiction argument avoiding a detailed study of the limiting random variable, and so it may not be clear what exactly goes wrong in the limit.

The first possible issue is whether the limit still has the functional form $h(X)(Y-X)+g(Y)$ for some functions h, g. A second issue is whether (these functions are measurable and) g is integrable as required by the definition of $\mathcal{V}_1$. The first issue is analyzed in our companion paper [32] which shows that the functional form is stable even under pointwise limits. Under the mild condition that $P \sim P^1 \otimes P^2$ (which is implied by the condition in Theorem 1.3), we can also guarantee that h, g remain measurable.

Lemma 3.1. *We have $V_* \in \mathcal{V}$; that is, $V_* = h(X)(Y-X)+g(Y)$ for some functions $h, g : \mathbb{R} \to \mathbb{R}$. If $P \sim P^1 \otimes P^2$, the functions h, g are a.s. uniquely determined and measurable.*

Proof. By Proposition 2.4 we can find $V_n \in \mathcal{V}_{\mathrm{b}}$ with $V_n \to V_*$ P-a.s. The two claims then follow from [32, Theorem 2.2] and [32, Theorem 3.1], respectively. □

This stability of the functional form indicates that the key failure in the counterexample of [1] is the integrability of the option. It is then clear that some original arguments will be required to obtain that the option position in our specific limit V_* is nevertheless integrable—which motivates the rest of this section.

3.2. **Proof of Theorem 1.3.** Recall that the disintegration of a probability measure $R \in \mathcal{P}(\mathbb{R}^2)$ is denoted $R = R^1 \otimes R^\bullet$, where R^1 is the first marginal (distribution of X) and $R^\bullet : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a stochastic kernel (conditional distribution of Y given X). We interchangeably use $R^\bullet(x)$, $R^\bullet(x,\cdot)$ or $R^\bullet(x,dy)$ to denote the conditional distribution given $X = x$.

Our basic line of attack is simple (yet seems to be novel): recalling Remark 2.3,

$$V_* = h(X)(Y-X)+g(Y) \in L^1(Q) \quad \text{for all} \quad Q \in \widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu), \tag{3.1}$$

where $\widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu)$ was defined in (2.2). We shall construct $\widetilde{Q} \in \widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu)$ such that h is uniformly bounded $\widetilde{Q}^1$-a.s. Then clearly $h(X)(Y-X) \in L^1(\widetilde{Q})$ and now (3.1) yields $g(Y) \in L^1(\widetilde{Q})$, or equivalently $g \in L^1(\nu)$, as desired.

On the other hand, the construction of $\widetilde{Q}$ is somewhat intricate. It is based on an approximation of Q_* by a sequence of probability measures $(\widetilde{Q}_n)$, which simultaneously retain the martingale property and satisfy h is $\widetilde{Q}^1_n$-a.s. uniformly bounded for all $n \in \mathbb{N}$. Next, we state a general version of this approximation result, applicable to any measurable function $h : \mathbb{R} \to \mathbb{R}$ and any $Q \in \mathcal{M}_{\mathrm{fin}}(\nu)$ satisfying the technical condition (3.2) below. In the proof of Theorem 1.3, the result will be applied to the specific function h in $V_* = h(X)(Y-X) + g(Y)$ and $Q = Q_*$.

Proposition 3.2. *Let $h : \mathbb{R} \to \mathbb{R}$ be measurable and $Q = Q^1 \otimes Q^\bullet \in \mathcal{M}_{\mathrm{fin}}(\nu)$. Suppose that there exists a Q^1-integrable function $I : \mathbb{R} \to [0,\infty)$ such that*

$$H\big(Q^\bullet(x')|P^\bullet(x)\big) \le I(x') \qquad \textit{for } (Q^1 \otimes Q^1)\textit{-a.a. } (x,x'). \tag{3.2}$$

Then there exist measures $\widetilde{Q}_n = \widetilde{Q}^1_n \otimes \widetilde{Q}^\bullet_n \in \widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu)$ such that

(i) h is $\widetilde{Q}^1_n$-a.s. uniformly bounded for each $n \in \mathbb{N}$,
(ii) $\widetilde{Q}_n \to Q$ in variation,
(iii) $H(\widetilde{Q}_n|P) \to H(Q|P)$.

In particular there exists $\widetilde{Q} \in \widetilde{\mathcal{M}}_{\mathrm{fin}}(\nu)$ such that h is uniformly bounded $\widetilde{Q}^1$-a.s.

The proof is lengthy and deferred to Section 3.3. For ease of reference, we record some standard facts in the next lemma.

Lemma 3.3. *Given probability measures $Q = Q^1 \otimes Q^\bullet$ and $R = R^1 \otimes R^\bullet$ on $\mathbb{R}^2$,*

(i) $Q \ll R$ if and only if $Q^1 \ll R^1$ and $Q^\bullet \ll R^\bullet$ Q^1-a.s.,
(ii) if $Q \ll R$, then
$$\frac{dQ}{dR} = \frac{dQ^1}{dR^1}\frac{dQ^\bullet}{dR^\bullet} \quad R\textit{-a.s.},$$
(iii) $H(Q|R) = H(Q^1|R^1) + E^{Q^1}[H(Q^\bullet|R^\bullet)]$.

We are now ready to detail the proof of Theorem 1.3.

Proof of Theorem 1.3. Set $\mu := Q^1_*$. We first construct a function I satisfying (3.2) with $Q = Q_*$. By our assumption, there are constants $0 < l < L < \infty$ such that

$$l \le \frac{dP^\bullet(x)}{d\nu}(y) \le L \qquad (\mu \otimes \nu)\text{-a.a. } (x,y). \tag{3.3}$$

Using Lemma 3.3(i), $Q_* \sim P$ implies that $Q^\bullet_* \sim P^\bullet \sim \nu$ μ-a.s. Note also

$$\begin{aligned} \frac{dQ^\bullet_*(x')}{dP^\bullet(x)}(y) &= \frac{dQ^\bullet_*(x')}{dP^\bullet(x')}(y)\,\frac{dP^\bullet(x')}{dP^\bullet(x)}(y) \\ &= \frac{dQ^\bullet_*(x')}{dP^\bullet(x')}(y)\,\frac{dP^\bullet(x')}{d\nu}(y)\,\frac{d\nu}{dP^\bullet(x)}(y), \qquad (\mu \otimes \mu \otimes \nu)\text{-a.a. } (x,x',y). \end{aligned} \tag{3.4}$$

Combining (3.3) and (3.4) we obtain

$$\log \frac{dQ_*^\bullet(x')}{dP^\bullet(x)}(y) \le \log \frac{dQ_*^\bullet(x')}{dP^\bullet(x')}(y) + \log(L/l)$$

and now integrating against $Q_*^\bullet(x')$ yields

$$H\big(Q_*^\bullet(x')|P^\bullet(x)\big) \le H\big(Q_*^\bullet(x')|P^\bullet(x')\big) + \log(L/l) =: I(x'), \qquad (\mu\otimes\mu)\text{-a.a. } (x,x').$$

Lemma 3.3 (iii) together with $H(Q_*|P)<\infty$ then implies $I \in L^1(\mu)$.

Noting that $Q_* \sim \mu\otimes\nu$, Lemma 3.1 yields that

$$V_* = h(X)(Y-X) + g(Y)$$

for some measurable functions h, g. Next, we verify that g is ν-integrable with $E^\nu[g]=0$. Indeed, the function I satisfies (3.2) with $Q=Q_*$, hence Proposition 3.2 provides $\widetilde{Q} \in \widetilde{\mathcal{M}}_{\rm fin}(\nu)$ such that h is $\widetilde{Q}^1$-a.s. uniformly bounded. This clearly implies $h(X)(Y-X) \in L^1(\widetilde{Q})$. Recalling from Remark 2.3 that V_* is $\widetilde{Q}$-integrable, we can deduce that $g(Y) \in L^1(\widetilde{Q})$; that is, $g \in L^1(\widetilde{Q}^2) = L^1(\nu)$. We can now conclude from $E^{Q_*}[V_*]=0$ that $E^\nu[g] = E^{Q_*}[g(Y)] = 0$, completing the proof that $V_* \in \mathcal{V}_1$.

Recall from Remark 2.3 that $V_* \in L^1(Q)$ for all $Q \in \widetilde{\mathcal{M}}_{\rm fin}(\nu)$. Having established $g \in L^1(\nu)$, this implies $h(X)(Y-X) \in L^1(Q)$ and then $E^Q[h(X)(Y-X)]=0$ by the martingale property. As $\widetilde{\mathcal{M}}_{\rm fin}(\nu) \supset \mathcal{M}_{\rm fin}(\nu)$, this shows that $V_* \in \mathcal{V}_{\rm adm}$. □

Remark 3.4. As seen in the proof, the boundedness condition in Theorem 1.3 can be weakened to the following integrability condition:

(i) $P \sim P^1 \otimes \nu$,

(ii) there exists a Q_*^1-integrable function $I:\mathbb{R}\to[0,\infty)$ such that

$$E^{Q_*^\bullet(x')}\left[\left|\log \frac{dP^\bullet(x')}{dP^\bullet(x)}\right|\right] \le I(x') \qquad \text{for } (P^1\otimes P^1)\text{-a.a. } (x,x').$$

3.3. **Proof of Proposition 3.2.** The program for this proof can be sketched as follows. First, we shall identify a sequence (μ_n) of sub-probability measures $\mu_n \ll Q^1$ such that h is uniformly bounded μ_n-a.e. and, when renormalized, $\mu-\mu_n$ dominates μ_n in convex order. Strassen's theorem then guarantees the existence of martingale measures M_n with first marginal μ_n and second marginal $\mu-\mu_n$. The desired measures $\widetilde{Q}_n$ have marginals $\mu_n/\mu_n(\mathbb{R})$ and ν: they will be built by embedding mass $\mu_n(\mathbb{R})$ according to $\widetilde{Q}_n^\bullet$ and mass $1-\mu_n(\mathbb{R})$ according to the composition of $M_n^\bullet$ with $\widetilde{Q}_n^\bullet$.

Let us first recall that the convex order of two probability measures μ,ν can be characterized via their quantile functions F_μ^{-1}, F_ν^{-1}. Indeed $\mu \preceq_c \nu$ if and only if

$$\int_u^1 F_\mu^{-1}(p)\,dp \le \int_u^1 F_\nu^{-1}(p)\,dp \tag{3.5}$$

for all $u\in[0,1]$, with equality for $u=0$, see [38, Theorem 3.A.5]. If μ,ν are finite measures with the same total mass, then $\mu\preceq_c\nu$ if and only if $\mu/\mu(\mathbb{R}) \preceq_c \nu/\nu(\mathbb{R})$. In particular, we can apply the characterization (3.5) to these normalized measures.

To simplify notation, we omit the normalizing constant and write F_{μ}^{-1} instead of $F_{\mu/\mu(\mathbb{R})}^{-1}$ in this case.

As a preparation for the proof of Proposition 3.2, we first establish two lemmas. Lemma 3.5 (i) has the same assertion as [6, Example 2.4] but is obtained with a different, more quantitative argument which is then used in Lemma 3.5 (ii) to elaborate on finer properties. Those properties are instrumental for the proof of Lemma 3.6 which describes a stability property of the convex order that will be applied in the proof of Proposition 3.2.

Lemma 3.5. *Let $A = [a, b] \subseteq \mathbb{R}$. Suppose that μ_A and μ_B are finite measures with the same mass and zero barycenter such that μ_A is concentrated on A and μ_B is concentrated on $B := \mathbb{R} \setminus (a, b)$.*

(i) We have $\mu_A \preceq_c \mu_B$.

(ii) Define

$$E := \left\{ u \in (0,1) : \int_u^1 F_{\mu_A}^{-1}(p)\, dp = \int_u^1 F_{\mu_B}^{-1}(p)\, dp \right\}.$$

Then E is of the form $(0, \alpha] \cup [\beta, 1)$ for some $0 \leq \alpha \leq \beta \leq 1$.[1] Furthermore,

(a) $E = (0, 1)$ if and only if $\mu_A = \mu_B$, in which case both measures are concentrated on $\{a, b\}$,

(b) $\mu_B((-\infty, a)) = 0$ and $\mu_B(a) > \mu_A(a) = \alpha$, whenever $\alpha > 0$ and $E \neq (0, 1)$,

(c) $\mu_B((b, \infty)) = 0$ and $\mu_B(b) > \mu_A(b) = 1 - \beta$, whenever $\beta < 1$ and $E \neq (0, 1)$.

Proof. We may assume that μ_A and μ_B are probability measures. We first show (i) by verifying (3.5) for $u \in (0, 1)$. Indeed, define

$$u^* := \sup\{p \in (0,1) : \ F_{\mu_B}^{-1}(p) \leq a\}.$$

Note that $F_{\mu_B}^{-1}(p) \geq b$ for all $p \in (u^*, 1)$ and $F_{\mu_A}^{-1}(p) \in [a, b]$ for all $p \in (0, 1)$. Hence

$$\int_u^1 F_{\mu_A}^{-1}(p)\, dp \leq \int_u^1 F_{\mu_B}^{-1}(p)\, dp \tag{3.6}$$

for all $u \in [u^*, 1)$. Suppose, towards a contradiction, that there exists $\hat{u} \in (0, u^*)$ such that (3.6) holds with the reverse, strict inequality at $u = \hat{u}$. As $F_{\mu_A}^{-1}(p) \geq a$ for all $p \in (0, 1)$ and $F_{\mu_B}^{-1}(p) \leq a$ for all $p \in (0, u^*)$, we deduce that

$$\int_0^1 F_{\mu_A}^{-1}(p)\, dp > \int_0^1 F_{\mu_B}^{-1}(p)\, dp, \tag{3.7}$$

contradicting that μ_A and μ_B have the same barycenter. This shows (i).

Turning to (ii), the proof of (a) is immediate. We can thus assume that there exists $\tilde{u} \in (0, 1)$ such that (3.6) holds with strict inequality at $u = \tilde{u}$. If there exists

[1] The conventions $(0, 0] := \emptyset$ and $[1, 1) := \emptyset$ are used.

no $\hat{u} \in (0, \tilde{u})$ such that

$$\int_{\hat{u}}^{1} F_{\mu_A}^{-1}(p)\,dp = \int_{\hat{u}}^{1} F_{\mu_B}^{-1}(p)\,dp, \tag{3.8}$$

then $\alpha = 0$. Whereas if such $\hat{u}$ exists, then necessarily $F_{\mu_B}^{-1}(p) \leq a$ for all $p \in (0, \hat{u}]$, for otherwise $F_{\mu_B}^{-1}(\hat{u}) \geq b$ and the equality in (3.8) cannot hold. It follows that

$$\int_{u}^{1} F_{\mu_A}^{-1}(p)\,dp \geq \int_{u}^{1} F_{\mu_B}^{-1}(p)\,dp \tag{3.9}$$

for all $u \in (0, \hat{u}]$. Since we have shown the reverse inequality in (i), we conclude that (3.9) holds with equality for all $u \in (0, \hat{u}]$. That is, E contains an interval of the form $(0, \alpha]$ for some $\alpha \geq 0$. Changing the integral bounds from $(u, 1)$ to $(0, u)$ by subtracting the barycenter on both sides of the above equations, an analogous argument shows that E contains an interval of the form $[\beta, 1)$ for some $\beta \in [0, 1]$. In conclusion, E is of the form $(0, \alpha] \cup [\beta, 1)$ for $0 \leq \alpha \leq \beta \leq 1$.

To show (b), suppose that $E \neq (0, 1)$ and $\alpha > 0$. The assumption implies that

$$\int_{0}^{\alpha} F_{\mu_A}^{-1}(p)\,dp = \int_{0}^{\alpha} F_{\mu_B}^{-1}(p)\,dp.$$

Consequently, $F_{\mu_A}^{-1}(p) = a = F_{\mu_B}^{-1}(p)$ for all $p \in (0, \alpha]$. That is, $\mu_B((-\infty, a)) = 0$, along with $\mu_A(a) \geq \alpha$ and $\mu_B(a) \geq \alpha$. Since

$$\int_{0}^{u} F_{\mu_A}^{-1}(p)\,dp > \int_{0}^{u} F_{\mu_B}^{-1}(p)\,dp \quad \text{for} \quad u \in (\alpha, \beta),$$

it is necessary that $F_{\mu_A}^{-1}(p) \in (a, b]$ for $p \in (\alpha, 1)$ and that $F_{\mu_B}^{-1}(p) = a$ for all $p > \alpha$ that are sufficiently close to α. We conclude that $\mu_B(a) > \mu_A(a) = \alpha$, showing (b). Part (c) is proved analogously. □

Lemma 3.6. *In the setting of Lemma 3.5, suppose that $\mu_A \neq \mu_B$ are probability measures. Let (μ_A^n), (μ_B^n) be sequences of probability measures with barycenter zero such that $\mu_A^n \ll \mu_A$ for all n as well as $d_{\mathrm{TV}}(\mu_A^n, \mu_A) \to 0$, $d_{\mathrm{TV}}(\mu_B^n, \mu_B) \to 0$ and $W_1(\mu_B^n, \mu_B) \to 0$ for $n \to \infty$. Then $\mu_A^n \preceq_c \mu_B^n$ for all n sufficiently large.*

Here W_1 denotes 1-Wasserstein distance, and we emphasize that the lemma does not require $\mu_B^n \ll \mu_B$.

Proof. Consider the set E in Lemma 3.5 (ii). As $\mu_A \neq \mu_B$, we have $E \neq (0, 1)$ and $\alpha < \beta$. Let us consider the cases $\alpha > 0$ and $\alpha = 0$ separately. If $\alpha > 0$, Lemma 3.5 (ii) (b) states that $\bar{\alpha} := \mu_B(a) > \mu_A(a) = \alpha$. Let $\bar{\alpha}_n := \mu_B^n(a)$ and $\alpha_n := \mu_A^n(a)$. Since $d_{\mathrm{TV}}(\mu_A^n, \mu_A) \to 0$ and $d_{\mathrm{TV}}(\mu_B^n, \mu_B) \to 0$ as $n \to \infty$, it follows that $\bar{\alpha}_n \to \bar{\alpha}$ and $\alpha_n \to \alpha$. In view of $\mu_A^n \ll \mu_A$, we conclude that $F_{\mu_A^n}^{-1}(p) \in [a, b]$ for $p \in (0, 1)$. Fix $\varepsilon < \bar{\alpha} - \alpha$. Then $\bar{\alpha}_n > \alpha + \varepsilon$ when n is sufficiently large, and

$$\int_{0}^{u} F_{\mu_A^n}^{-1}(p)\,dp \geq \int_{0}^{u} F_{\mu_B^n}^{-1}(p)\,dp \quad \text{for} \quad u \in (0, \alpha + \varepsilon]. \tag{3.10}$$

Whereas in the case $\alpha = 0$, we define $\bar{\alpha} := \mu_B((-\infty, a]) > 0$ and $\bar{\alpha}_n := \mu_B^n((-\infty, a])$. Again, for a fixed $\varepsilon < \bar{\alpha} - \alpha = \bar{\alpha}$, we have $\bar{\alpha}_n > \alpha + \varepsilon$ when n is sufficiently large, and (3.10) holds.

Similarly, we consider the cases $\beta < 1$ and $\beta = 1$ and define $\bar{\beta}$ accordingly. In either case we can fix $\varepsilon < \beta - \bar{\beta}$ and find n sufficiently large so that $\bar{\beta}_n < \beta - \varepsilon$ and

$$\int_u^1 F_{\mu_A^n}^{-1}(p)\,dp \le \int_u^1 F_{\mu_B^n}^{-1}(p)\,dp \quad \text{for} \quad u \in [\beta - \varepsilon, 1). \tag{3.11}$$

To complete the proof, it remains to show that the inequality in (3.11) holds for $u \in O := (\alpha + \varepsilon, \beta - \varepsilon)$, where $\varepsilon < \min\{\bar{\alpha} - \alpha, \beta - \bar{\beta}\}$ is fixed. Note that $(0,1) \setminus E = (\alpha, \beta)$ and $O \subsetneq (\alpha, \beta)$. As the integrals below are continuous functions of u, there exists $\gamma > 0$ such that

$$\int_u^1 F_{\mu_A}^{-1}(p)\,dp + \gamma \le \int_u^1 F_{\mu_B}^{-1}(p)\,dp \quad \text{for all} \quad u \in O, \tag{3.12}$$

thanks to the definition of E. In view of $d_{\mathrm{TV}}(\mu_A^n, \mu_A) \to 0$ and $d_{\mathrm{TV}}(\mu_B^n, \mu_B) \to 0$, the quantile functions converge pointwise. Moreover, we recall that the 1-Wasserstein distance satisfies

$$W_1(\mu_B^n, \mu_B) = \int_0^1 \left| F_{\mu_B^n}^{-1}(p) - F_{\mu_B}^{-1}(p) \right| dp.$$

Dominated convergence and $W_1(\mu_B^n, \mu_B) \to 0$ thus imply that

$$\lim_{n\to\infty} \int_u^1 F_{\mu_A^n}^{-1}(p)\,dp = \int_u^1 F_{\mu_A}^{-1}(p)\,dp, \qquad \lim_{n\to\infty} \int_u^1 F_{\mu_B^n}^{-1}(p)\,dp = \int_u^1 F_{\mu_B}^{-1}(p)\,dp$$

uniformly in $u \in O$. It now follows from (3.12) that

$$\int_u^1 F_{\mu_A^n}^{-1}(p)\,dp \le \int_u^1 F_{\mu_A}^{-1}(p)\,dp + \frac{\gamma}{2} \le \int_u^1 F_{\mu_B}^{-1}(p)\,dp - \frac{\gamma}{2} \le \int_u^1 F_{\mu_B^n}^{-1}(p)\,dp$$

for all $u \in O$ and $n \in \mathbb{N}$ large enough. This completes the proof. $\square$

Given measures λ, μ on $\mathbb{R}$, we write $\lambda \le \mu$ if $\lambda(A) \le \mu(A)$ for all $A \in \mathcal{B}(\mathbb{R})$. The total variation distance between λ and μ is defined as

$$d_{\mathrm{TV}}(\lambda, \mu) = \sup\{|\lambda(A) - \mu(A)| : A \in \mathcal{B}(\mathbb{R})\}.$$

If $\lambda \le \mu$, it is clear that $d_{\mathrm{TV}}(\lambda, \mu) = (\mu - \lambda)(\mathbb{R})$. For ease of reference, we record the following consequence.

Lemma 3.7. *Let $0 \ne \lambda \le \mu$ be finite measures on $\mathbb{R}$. Then the probability measures $\bar{\lambda} := \lambda/\lambda(\mathbb{R})$ and $\bar{\mu} := \mu/\mu(\mathbb{R})$ satisfy $\bar{\lambda} \ll \bar{\mu}$ and*

$$d_{\mathrm{TV}}(\bar{\lambda}, \bar{\mu}) \le \frac{(\mu - \lambda)(\mathbb{R})}{\mu(\mathbb{R})}. \tag{3.13}$$

Proof. We have $\bar{\lambda} \ll \bar{\mu}$ and $\lambda(\mathbb{R}) \leq \mu(\mathbb{R})$, so that

$$d_{\mathrm{TV}}(\bar{\lambda}, \bar{\mu}) = E^{\bar{\mu}}\left[\left(1 - \frac{d\bar{\lambda}}{d\bar{\mu}}\right)^{+}\right] = \frac{1}{\mu(\mathbb{R})} E^{\mu}\left[\left(1 - \frac{\mu(\mathbb{R})}{\lambda(\mathbb{R})}\frac{d\lambda}{d\mu}\right)^{+}\right]$$
$$\leq \frac{1}{\mu(\mathbb{R})} E^{\mu}\left[\left(1 - \frac{d\lambda}{d\mu}\right)^{+}\right] = \frac{1}{\mu(\mathbb{R})}(\mu - \lambda)(\mathbb{R}). \quad \square$$

We are now ready to give the proof of Proposition 3.2.

Proof of Proposition 3.2. To simplify notation we write $\mu := Q^1$ and assume without loss of generality that μ, ν have zero barycenter. We may also assume that $\mu \neq \delta_0$; otherwise the claim is trivial. The proof proceeds in six steps.

Step 1. Given $\delta > 0$ sufficiently small we shall construct sub-probability measures μ_A, μ_B (depending on δ) that satisfy the hypotheses of Lemma 3.5, and hence $\mu_A \preceq_c \mu_B$.

As μ has zero barycenter and is not a Dirac measure, there exists $c > 0$ such that $\tilde{\delta} := \mu((-\infty, -c]) \wedge \mu([c, \infty)) > 0$. Then, for all $0 < \delta < \tilde{\delta}$, there exist an interval $A := [a, b]$ and a measure λ_A such that

$$\mu((-\infty, a] \cup [b, \infty)) \leq \delta, \qquad \mu((a, -c]) \geq \delta, \qquad \mu([c, b)) \geq \delta$$

and

- $\lambda_A \leq \mu|_A$,
- λ_A has zero barycenter,
- $\mu - \lambda_A$ is concentrated on $\mathbb{R} \setminus (a, b)$ and nonzero.

In particular $\mu|_{(a,b)} \leq \lambda_A \leq \mu|_A$. In fact, if μ has no atoms, we can set $\lambda_A = \mu|_A$. In the presence of atoms at a or b, we may have to remove part of that mass so that λ_A has zero barycenter and $\lambda_A \neq \mu$. Define

$$\mu_A := \left(\frac{1}{\lambda_A(\mathbb{R})} - 1\right)\lambda_A \quad \text{and} \quad \mu_B := \mu - \lambda_A.$$

Evidently, the hypotheses of Lemma 3.5 are satisfied, and hence $\mu_A \preceq_c \mu_B$.

Step 2. Fix $\varepsilon > 0$. As h takes values in $\mathbb{R}$ there exists a set $A^\varepsilon \subseteq A$ such that h is uniformly bounded on A^ε and

$$\mu(A \setminus A^\varepsilon) \leq \frac{\delta \wedge \varepsilon}{2}\left(1 \wedge \frac{c}{|a| \vee |b|}\right). \tag{3.14}$$

The choice of the upper bound in (3.14) ensures that

$$\mu((a, -c] \cap A^\varepsilon) \geq \frac{\delta}{2}, \quad \mu([c, b) \cap A^\varepsilon) \geq \frac{\delta}{2}$$

and, writing X for the identity function on $\mathbb{R}$ in an abuse of notation,

$$\begin{aligned}\left|E^{\lambda_A}[\mathbf{1}_{A^\varepsilon} X]\right| &\le \left|E^{\lambda_A}[\mathbf{1}_A X]\right| + \left|E^{\lambda_A}\left[\mathbf{1}_{A\setminus A^\varepsilon} X\right]\right| \\ &\le 0 + (|a| \vee |b|)\frac{(\delta\wedge\varepsilon)c}{2(|a|\vee|b|)} = \frac{(\delta\wedge\varepsilon)c}{2}.\end{aligned}$$

By restricting λ_A to A^ε and possibly removing some mass on $(a,-c]$ or $[c,b)$, we can construct a measure λ_A^ε that is concentrated on A^ε, satisfies $\lambda_A^\varepsilon \le \lambda_A \le \mu$, has zero barycenter and

$$d_{\mathrm{TV}}(\lambda_A^\varepsilon, \lambda_A) = (\lambda_A - \lambda_A^\varepsilon)(\mathbb{R}) \le \mu(A\setminus A^\varepsilon) + \frac{1}{c}\left|E^{\lambda_A}[\mathbf{1}_{A^\varepsilon} X]\right| \le \delta\wedge\varepsilon.$$

In consequence,

$$\lambda_A^\varepsilon(\mathbb{R}) = \lambda_A(\mathbb{R}) - (\lambda_A - \lambda_A^\varepsilon)(\mathbb{R}) \ge 1 - \delta - \delta\wedge\varepsilon. \tag{3.15}$$

Step 3. Let $\delta \le \tilde{\delta}$ be fixed and consider a sequence $(\varepsilon_n)_{n\in\mathbb{N}} \downarrow 0$. For each $n \in \mathbb{N}$ we apply Step 2 to obtain

$$\mu_A^n := \left(\frac{1}{\lambda_A^n(\mathbb{R})} - 1\right)\lambda_A^n \quad \text{and} \quad \mu_B^n := \mu - \lambda_A^n, \tag{3.16}$$

where $\lambda_A^n := \lambda_A^{\varepsilon_n}$. Both measures have the same total mass and zero barycenter for all $n\in\mathbb{N}$. Moreover we have $\lambda_A^n \le \lambda_A$ and $d_{\mathrm{TV}}(\lambda_A, \lambda_A^n) = (\lambda_A - \lambda_A^n)(\mathbb{R}) \le \varepsilon_n \downarrow 0$.

In order to apply Lemma 3.6, we first need to scale the measures μ_A, μ_B, μ_A^n and μ_B^n so that they are probability measures. Set

$$\bar{\mu}_A = \frac{\lambda_A}{\lambda_A(\mathbb{R})} \quad \text{and} \quad \bar{\mu}_B = \frac{\mu - \lambda_A}{1 - \lambda_A(\mathbb{R})} \tag{3.17}$$

and define $\bar{\mu}_A^n$ and $\bar{\mu}_B^n$ analogously. Observe that $\bar{\mu}_A^n \ll \bar{\mu}_A$ and

$$d_{\mathrm{TV}}(\bar{\mu}_A^n, \bar{\mu}_A) \le \frac{(\lambda_A - \lambda_A^n)(\mathbb{R})}{\lambda_A(\mathbb{R})} \le \frac{\varepsilon_n}{\lambda_A(\mathbb{R})} \downarrow 0$$

by (3.13). Similarly we have $d_{\mathrm{TV}}(\bar{\mu}_B^n, \bar{\mu}_B) \to 0$. In particular it suffices to show that $E^{\mu_B^n}[|X|] \to E^{\mu_B}[|X|]$ in order to verify $W_1(\bar{\mu}_B^n, \bar{\mu}_B) \to 0$. In light of the definition of $\bar{\mu}_B$ in (3.17) and $d_{\mathrm{TV}}(\lambda_A^n, \lambda_A) \to 0$ this readily follows from $E^{\lambda_A^n}[|X|] \to E^{\lambda_A}[|X|]$.

Now we are in a position to apply Lemma 3.6 to $\bar{\mu}_A, \bar{\mu}_B, \bar{\mu}_A^n$ and $\bar{\mu}_B^n$, which yields $n_0 \in \mathbb{N}$ such that $\bar{\mu}_A^{n_0} \preceq_c \bar{\mu}_B^{n_0}$. Since the convex order is invariant under scaling, it follows that

$$\mu_A^* := \mu_A^{n_0} \preceq_c \mu_B^{n_0} =: \mu_B^*. \tag{3.18}$$

We recall that h is uniformly bounded on $A^* := A^{\varepsilon_{n_0}}$ by construction and define

$$\lambda_A^* := \lambda_A^{n_0} \tag{3.19}$$

in preparation for Step 5 below.

Step 4. By [39, Theorem 8] the relation $\mu_A^* \preceq_c \mu_B^*$ implies the existence of a mean-preserving probability kernel $M_\delta^\bullet : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ sending μ_A^* to μ_B^*; that is, $M_\delta := \mu_A^* \otimes M_\delta^\bullet \in \Pi(\mu_A^*, \mu_B^*)$ and $M_\delta^\bullet(x)$ has barycenter x for all $x\in\mathbb{R}$.

Recall that $Q = \mu \otimes Q^\bullet \in \mathcal{M}_{\mathrm{fin}}(\nu)$ and denote by $Q^\bullet_\delta$ the composition of $M^\bullet_\delta$ with $Q^\bullet$,

$$Q^\bullet_\delta(x, C) := E^{M^\bullet_\delta(x)}[Q^\bullet(\cdot, C)] \quad \text{for} \quad C \in \mathcal{B}(\mathbb{R}). \tag{3.20}$$

Note that $Q^\bullet_\delta$ is again mean-preserving.

Step 5. For $\delta \le \tilde\delta$, we set

$$\widetilde{Q}_\delta := \lambda^*_A \otimes Q^\bullet + \mu^*_A \otimes Q^\bullet_\delta, \tag{3.21}$$

where μ^*_A, λ^*_A and $Q^\bullet_\delta$ were defined in (3.18), (3.19) and (3.20), respectively (the set A and the measures λ^*_A, μ^*_A depend on δ). In view of (3.16), the first marginal of $\widetilde{Q}_\delta$ is the probability measure $\bar\mu^*_A := \lambda^*_A/\lambda^*_A(\mathbb{R})$.

We claim that $\widetilde{Q}_\delta \in \widetilde{\mathcal{M}}(\nu)$. Indeed, recall $\mu \otimes Q^\bullet = Q \sim P$ and observe that $\mu^*_A \sim \lambda^*_A \ll \mu$, $Q^\bullet \sim \nu$ μ-a.s. and $Q^\bullet_\delta \ll \nu$ μ^*_A-a.s. Lemma 3.3 (i) then yields $\widetilde{Q}_\delta \ll P$. Moreover, the martingale property $E^{\widetilde{Q}_\delta}[Y|X] = X$ follows from the fact that $Q^\bullet$ and $Q^\bullet_\delta$ are mean-preserving. Finally, recall that $Q^\bullet_\delta$ is the composition of $M^\bullet_\delta$ with $Q^\bullet$, and $\mu^*_A \otimes M^\bullet_\delta \in \Pi(\mu^*_A, \mu^*_B)$. Thus $\mu = \lambda^*_A + \mu^*_B$ and $\mu \otimes Q^\bullet \in \Pi(\mu, \nu)$ imply $\widetilde{Q}_\delta \in \Pi(\bar\mu^*_A, \nu)$ and in particular $\widetilde{Q}^2_\delta = \nu$. In summary, $\widetilde{Q}_\delta \in \widetilde{\mathcal{M}}(\nu)$ as desired.

From Step 2 and (3.15) we see that $\lambda^*_A \le \mu$ and $\lambda^*_A(\mathbb{R}) \ge 1 - 2\delta$, hence $\lambda^*_A \to \mu$ and $\mu^*_A \to 0$ in variation as $\delta \to 0$. It is now clear from the definition (3.21) that $\widetilde{Q}_\delta \to \mu \otimes Q^\bullet = Q$ in variation. Moreover, as λ^*_A is concentrated on A^* (cf. Step 3) and h is uniformly bounded on A^*, we see that h is $\bar\mu^*_A$-a.s. uniformly bounded for every $\delta \le \tilde\delta$. This proves Proposition 3.2 (i),(ii) after choosing $\delta = \delta(n)$ small enough, modulo showing that $H(\widetilde{Q}_\delta|P) < \infty$ for small δ (which will follow from the next step).

Step 6. As $\liminf_{\delta\to 0} H(\widetilde{Q}_\delta|P) \ge H(Q|P)$ due to the lower semicontinuity of $H(\cdot|P)$ and the convergence $\widetilde{Q}_\delta \to Q$, it remains to show

$$\limsup_{\delta\to 0} H(\widetilde{Q}_\delta|P) \le H(Q|P). \tag{3.22}$$

Note that $\widetilde{Q}_\delta$ is a convex combination of two probability measures:

$$\widetilde{Q}_\delta = \lambda^*_A(\mathbb{R})\bar\mu^*_A \otimes Q^\bullet + [1 - \lambda^*_A(\mathbb{R})]\bar\mu^*_A \otimes Q^\bullet_\delta.$$

As $H(\cdot|P)$ is convex, it follows that

$$H(\widetilde{Q}_\delta|P) \le \lambda^*_A(\mathbb{R}) H(\bar\mu^*_A \otimes Q^\bullet|P) + [1 - \lambda^*_A(\mathbb{R})] H(\bar\mu^*_A \otimes Q^\bullet_\delta|P). \tag{3.23}$$

We show that the first term converges to $H(Q|P)$ and the second converges to zero. Indeed, Lemma 3.3 (iii) yields $H(\bar\mu^*_A \otimes Q^\bullet|P) = H(\bar\mu^*_A|P^1) + E^{\bar\mu^*_A}[H(Q^\bullet|P^\bullet)]$, where

$$\begin{aligned} H(\bar\mu^*_A|P^1) &= \frac{1}{\lambda^*_A(\mathbb{R})} E^{\lambda^*_A}\left[\log\left(\frac{d\lambda^*_A}{dP^1}\right)\right] - \log \lambda^*_A(\mathbb{R}) \\ &\to E^{\mu}\left[\log\left(\frac{d\mu}{dP^1}\right)\right] = H(\mu|P^1) \end{aligned} \tag{3.24}$$

by dominated convergence and $\lambda_A^*(\mathbb{R}) \geq 1-2\delta$. Similarly, $E^{\bar{\mu}_A^*}[H(Q^\bullet|P^\bullet)] \to E^\mu[H(Q^\bullet|P^\bullet)]$, so that the first term in (3.23) satisfies

$$\lambda_A^*(\mathbb{R})H(\bar{\mu}_A^* \otimes Q^\bullet|P) \to H(\mu|P^1) + E^\mu[H(Q^\bullet|P^\bullet)] = H(Q|R).$$

It remains to show that the second term in (3.23) converges to zero,

$$[1-\lambda_A^*(\mathbb{R})]H(\bar{\mu}_A^* \otimes Q_\delta^\bullet|P) \to 0.$$

Using again Lemma 3.3 (iii),

$$H(\bar{\mu}_A^* \otimes Q_\delta^\bullet|P) = H\big(\bar{\mu}_A^*|P^1\big) + E^{\bar{\mu}_A^*}[H(Q_\delta^\bullet|P^\bullet)]. \tag{3.25}$$

In view of (3.24) and $\lambda_A^*(\mathbb{R}) \to 1$, it follows that $[1-\lambda_A^*(\mathbb{R})]H\big(\bar{\mu}_A^*|P^1\big) \to 0$. For the second term in (3.25), we use the definitions of $\bar{\mu}_A^*$ and μ_A^* to see that

$$[1-\lambda_A^*(\mathbb{R})]E^{\bar{\mu}_A^*}[H(Q_\delta^\bullet|P^\bullet)] = \frac{1-\lambda_A^*(\mathbb{R})}{\lambda_A^*(\mathbb{R})}E^{\lambda_A^*}[H(Q_\delta^\bullet|P^\bullet)] = E^{\mu_A^*}[H(Q_\delta^\bullet|P^\bullet)].$$

In view of (3.20), Jensen's inequality and convexity of H imply

$$E^{\mu_A^*}[H(Q_\delta^\bullet|P^\bullet)] = E^{\mu_A^*}\Big[H\Big(E^{M_\delta^\bullet(X)}[Q^\bullet]|P^\bullet(X)\Big)\Big] \leq E^{\mu_A^*}\Big[E^{M_\delta^\bullet(X)}[H(Q^\bullet|P^\bullet(X))]\Big].$$

Lastly, the assumptions that $H(Q^\bullet(x')|P^\bullet(x)) \leq I(x')$ for $(\mu\otimes\mu)$-a.a (x,x') and $I \in L^1(\mu)$ together with the facts that $\mu_A^* \otimes M_\delta^\bullet \in \Pi(\mu_A^*, \mu_B^*)$ and $\mu_B^* \to 0$ yield

$$E^{\mu_A^*}\Big[E^{M_\delta^\bullet(X)}[H(Q^\bullet|P^\bullet(X))]\Big] \leq E^{\mu_A^*}\Big[E^{M_\delta^\bullet(X)}[I]\Big] = E^{\mu_B^*}[I] \to 0$$

by the dominated convergence theorem. This shows (3.22) and hence Proposition 3.2 (iii), completing the proof. □

References

[1] B. Acciaio, M. Larsson, and W. Schachermayer. The space of outcomes of semi-static trading strategies need not be closed. *Finance Stoch.*, 21(3):741–751, 2017.

[2] J. M. Altschuler, J. Niles-Weed, and A. J. Stromme. Asymptotics for semi-discrete entropic optimal transport. *Preprint arXiv:2106.11862v1*, 2021.

[3] M. Avellaneda. Minimum-relative-entropy calibration of asset-pricing models. *Int. J. Theor. Appl. Finance*, 1(4):447–472, 1998.

[4] M. Avellaneda, R. Buff, C. Friedman, N. Grandechamp, L. Kruk, and J. Newman. Weighted Monte Carlo: a new technique for calibrating asset-pricing models. *Int. J. Theor. Appl. Finance*, 4(1):91–119, 2001.

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

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

[7] A. Beurling. An automorphism of product measures. *Ann. of Math. (2)*, 72:189–200, 1960.

[8] J. M. Borwein and A. S. Lewis. Decomposition of multivariate functions. *Canad. J. Math.*, 44(3):463–482, 1992.

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

[10] I. Csiszár. *I*-divergence geometry of probability distributions and minimization problems. *Ann. Probability*, 3:146–158, 1975.

[11] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems 26*, pages 2292–2300. 2013.

[12] R. C. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics and Stochastic Rep.*, 29(2):185–201, 1990.

[13] F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker. Exponential hedging and entropic penalties. *Math. Finance.*, 12:99–123, 2002.

[14] G. Deligiannidis, V. De Bortoli, and A. Doucet. Quantitative uniform stability of the iterative proportional fitting procedure. *Preprint arXiv:2108.08129v1*, 2021.

[15] H. Föllmer. Random fields and diffusion processes. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 101–203. Springer, Berlin, 1988.

[16] H. Föllmer and N. Gantert. Entropy minimization and Schrödinger processes in infinite dimensions. *Ann. Probab.*, 25(2):901–926, 1997.

[17] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. W. de Gruyter, Berlin, 3rd edition, 2011.

[18] M. Frittelli. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Math. Finance*, 10(1):39–52, 2000.

[19] A. Galichon, P. Henry-Labordère, and N. Touzi. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.*, 24(1):312–336, 2014.

[20] N. Gigli and L. Tamanini. Second order differentiation formula on $\mathrm{RCD}^*(K, N)$ spaces. *J. Eur. Math. Soc. (JEMS)*, 23(5):1727–1795, 2021.

[21] G. Guo and J. Obłój. Computational methods for martingale optimal transport problems. *Ann. Appl. Probab.*, 29(6):3311–3347, 2019.

[22] J. Guyon. The joint S&P 500/VIX smile calibration puzzle solved. *Risk*, 2020.

[23] J. Guyon. Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *Preprint SSRN:3853237*, 2021.

[24] P. Henry-Labordère. From (martingale) Schrödinger bridges to a new class of stochastic volatility model. *Preprint SSRN:3353270*, 2019.

[25] D. Hobson. Robust hedging of the lookback option. *Finance Stoch.*, 2(4):329–347, 1998.

[26] J. Jacod and A. N. Shiryaev. Local martingales and the fundamental asset pricing theorems in the discrete-time case. *Finance Stoch.*, 2(3):259–273, 1998.

[27] C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.

[28] H. D. March and P. Henry-Labordère. Building arbitrage-free implied volatility: Sinkhorn's algorithm and variants. *Preprint SSRN:3326486*, 2019.

[29] M. Nutz. *Introduction to Entropic Optimal Transport*. Lecture notes, Columbia University, 2021. `https://www.math.columbia.edu/~mnutz/docs/EOT_lecture_notes.pdf`.

[30] M. Nutz and J. Wiesel. Entropic optimal transport: Convergence of potentials. *Probab. Theory Related Fields, to appear.* arXiv:2104.11720v2.

[31] M. Nutz and J. Wiesel. Stability of Schrödinger potentials and convergence of Sinkhorn's algorithm. *Preprint arXiv:2201.10059v1*, 2022.

[32] M. Nutz, J. Wiesel, and L. Zhao. Limits of semistatic trading strategies. *Preprint*, 2022.

[33] G. Peyré and M. Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

[34] L. C. G. Rogers. Equivalent martingale measures and no-arbitrage. *Stochastics Stochastics Rep.*, 51(1-2):41–49, 1994.

[35] L. Rüschendorf and W. Thomsen. Note on the Schrödinger equation and I-projections. *Statist. Probab. Lett.*, 17(5):369–375, 1993.

[36] L. Rüschendorf and W. Thomsen. Closedness of sum spaces and the generalized Schrödinger problem. *Teor. Veroyatnost. i Primenen.*, 42(3):576–590, 1997.

[37] W. Schachermayer. Optimal investment in incomplete markets when wealth may become negative. *Ann. Appl. Probab.*, 11(3):694–734, 2001.

[38] M. Shaked and J. G. Shanthikumar. *Stochastic orders*. Springer Series in Statistics. Springer, New York, 2007.

[39] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.

[40] T. Zariphopoulou. A solution approach to valuation with unhedgeable risks. *Finance Stoch.*, 5(1):61–82, 2001.

(MN) Departments of Statistics and Mathematics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA
E-mail address: mnutz@columbia.edu

(JW) Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA
E-mail address: johannes.wiesel@columbia.edu

(LZ) Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA
E-mail address: long.zhao@columbia.edu