Bid–Ask Martingale Optimal Transport

Bryan Liang* Marcel Nutz† Shunan Sheng‡ Valentin Tissot-Daguette§

March 26, 2026

Abstract

Martingale Optimal Transport (MOT) provides a framework for robust pricing and hedging of illiquid derivatives. Classical MOT enforces exact calibration of model marginals to the mid-prices of vanilla options. Motivated by the industry practice of fitting bid and ask marginals to vanilla prices, we introduce a relaxation of MOT in which model-implied volatilities are only required to lie within observed bid–ask spreads; equivalently, model marginals lie between the bid and ask marginals in convex order. The resulting Bid–Ask MOT (BAMOT) yields realistic price bounds for illiquid derivatives and, via strong duality, can be interpreted as the superhedging price when short and long positions in vanilla options are priced at the bid and ask, respectively. We further establish convergence of BAMOT to classical MOT as bid–ask spreads vanish, and quantify the convergence rate using a novel distance intrinsically linked to bid–ask spreads. Finally, we support our findings with several synthetic and real-data examples.

Keywords Martingale Optimal Transport, Financial Derivatives, Robust Hedging, Bid–Ask Friction

AMS 2020 Subject Classification 91G20, 62P05, 60G42

JEL Classification G11, G13, D52

Acknowledgments We would like to thank Bruno Dupire, Martin Forde, Sergey Nadtochiy, Walter Schachermayer, Mete Soner, and Josef Teichmann for helpful discussions on this topic. Shunan Sheng is supported by a Bloomberg Quantitative Finance Ph.D. Fellowship.

1 Introduction

Martingale Optimal Transport (MOT) [3, 6, 7, 19, 22] offers a powerful framework to price and hedge illiquid derivatives. Consider a claim H contingent on the evolution of a liquidly traded asset $(X_t)_{0\le t\le T}$ up to some fixed maturity $T>0$. Suppose also that the static hedging instruments consist of all T-vanilla options (i.e., all put and call options with maturity T and any strike $K\ge 0$), in addition to dynamic trading in $(X_t)_{0\le t\le T}$. Absent market frictions, one classically extracts a

*Quantitative Research, Office of the CTO, Bloomberg. Email: bliang17@bloomberg.net.

†Columbia University, Departments of Statistics and Mathematics. Research supported by NSF Grants DMS-2106056, DMS-2407074. Email: mnutz@columbia.edu.

‡Columbia University, Department of Statistics. Email: ss6574@columbia.edu. Research initiated during an internship at Bloomberg.

§Quantitative Research, Office of the CTO, Bloomberg. Email: vtissotdague@bloomberg.net

unique risk-neutral distribution μ consistent with the Black–Scholes implied volatility (IV) skew $\sigma(K,T)$. Indeed, assuming zero carry, we have by [11] that

$$\frac{\mu(\mathrm{d}K)}{\mathrm{d}K} = \partial_K^2 c(K,T), \quad c(K,T) := c_{\mathrm{BS}}(K,T,\sigma(K,T)), \tag{1}$$

where $c_{\mathrm{BS}}(K,T,\sigma)$ is the price of the (K,T) call option in the Black–Scholes model with volatility σ. MOT then considers the set of all arbitrage-free models (martingale measures) for $(X_t)_{0\leq t\leq T}$ such that the marginal law $\mu_T^{\mathbb{Q}}$ of X_T under $\mathbb{Q}$ coincides with μ, leading to the primal (or measure) problem

$$\sup_{\mathbb{Q}\in\mathcal{Q}(\mu)} \mathbb{E}_{\mathbb{Q}}[H], \qquad \mathcal{Q}(\mu) = \big\{\mathbb{Q} \mid \text{martingale measure, } \mu_T^{\mathbb{Q}} = \mu\big\}. \tag{2}$$

A similar logic applies when the static hedging instruments are available at multiple maturities $T_1,\dots,T_N$, leading to marginal constraints $\mu_{T_i}^{\mathbb{Q}} = \mu_i$ for $i = 1,\dots,N$. To wit, MOT requires admissible models to exactly match the market's implied volatility (IV) skews, thus imposing hard constraints on the marginal distributions. The dual of MOT, discussed in more detail below, is a semi-static superhedging problem, which assumes that vanilla options can be bought and sold at the prices implied by those IV skews, with no bid–ask friction.

1.1 Bid–Ask Friction

The present work develops an extension of MOT that accounts for the presence of bid–ask spreads for vanilla prices in real markets. Indeed, in practice, vanilla options are often represented in terms of bid and ask IV skews $\sigma^{\mathrm{b}}(K,T) \leq \sigma^{\mathrm{a}}(K,T)$. Then a martingale measure $\mathbb{Q}$ is consistent with the options market if and only if the implied volatilities $\sigma^{\mathbb{Q}}(K,T)$ that it generates (see Fig. 1) satisfy

$$\sigma^{\mathrm{b}}(K,T) \leq \sigma^{\mathbb{Q}}(K,T) \leq \sigma^{\mathrm{a}}(K,T), \quad K \geq 0. \tag{3}$$

Of course, order books only contain quotes for finitely many strikes at any given moment. Data vendors provide bid/ask IV skews by interpolating/extrapolating the available quotes. As detailed in Section 2.1, a common approach for this task is to fit *bid and ask marginal distributions* $\mu^{\mathrm{b}},\mu^{\mathrm{a}}$ within a parametric family to the available quotes. This motivates the primal formulation of our problem, which replaces the single (mid-price) marginal μ of the MOT problem (2) by a set of possible marginals determined via μ^{b} and μ^{a} as follows.

1.2 Primal Formulation

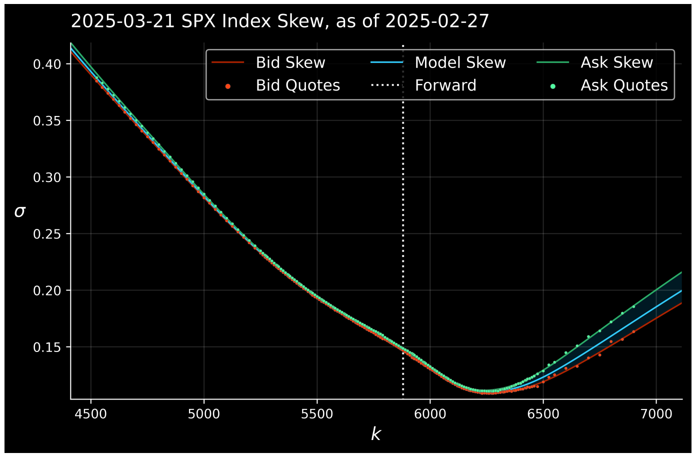
To be consistent with the bid and ask IVs, the possible marginals $\mu_T^{\mathbb{Q}} = \mathrm{Law}^{\mathbb{Q}}(X_T)$ for the asset price X_T must satisfy certain convex order constraints. Indeed, (3) implies

$$\mathbb{E}_{\mu^{\mathrm{b}}}[(X-K)^+] = c_{\mathrm{BS}}(K,T,\sigma^{\mathrm{b}}(K,T)) \leq c_{\mathrm{BS}}(K,T,\sigma^{\mathbb{Q}}(K,T)) = \mathbb{E}_{\mu_T^{\mathbb{Q}}}[(X-K)^+]$$

for all $K \geq 0$ since $c_{\mathrm{BS}}(K,T,\sigma)$ is nondecreasing in σ. This means that μ^{b} and $\mu_T^{\mathbb{Q}}$ are in *convex order,* denoted $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_T^{\mathbb{Q}}$, by the property of convex order recalled in (7) below. Similarly, $\mu_T^{\mathbb{Q}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. This leads us to formulate the following extension of the MOT problem (2), which we call the *Bid–Ask Martingale Optimal Transport* (BAMOT) problem:

$$\sup_{\mathbb{Q}\in\mathcal{Q}(\mu^{\mathrm{b}},\mu^{\mathrm{a}})} \mathbb{E}_{\mathbb{Q}}[H], \quad \mathcal{Q}(\mu^{\mathrm{b}},\mu^{\mathrm{a}}) = \big\{\mathbb{Q} \mid \text{martingale measure, } \mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_T^{\mathbb{Q}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}\big\}. \tag{4}$$

Figure 1: Bid, model, and ask implied volatility skews for S&P 500 Index (SPX) options expiring on 03-21-2025, as of 02-27-2025.



More generally, we will consider a version of this problem with multiple maturities $T_1, \ldots, T_N$. In contrast to its classical counterpart, BAMOT captures model uncertainty about the underlying asset $(X_t)_{0 \le t \le T}$ arising not only from its price dynamics, but also from its marginal distributions: while MOT prescribes the marginal μ exactly, the bid and ask marginals merely imply inequality constraints for the marginal distribution of the asset. In particular, BAMOT is nontrivial even for claims that depend only on X_T, such as a digital option, whereas MOT simply prices them by taking expectation under μ. In general, the value of (4) is no less than the value of (2) (for any $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$), reflecting that the seller of the claim faces additional model uncertainty.

1.3 Dual Formulation

Next, we consider the effect of bid–ask spreads from the hedging perspective. In the classical MOT setting with a single marginal μ, the dual (or portfolio) problem aims at minimizing the superhedging cost of the claim H, given by

$$\inf\{\mu(\psi) \mid \psi \in \Psi(H)\}, \quad \Psi(H) = \Big\{\psi \in L^1(\mu) \mid \exists\, \Delta \text{ such that } \mathrm{P\&L}^H_{\psi,\Delta} \ge 0\Big\}.$$

Here Δ is a dynamic hedging strategy trading in the asset $(X_t)_{0 \le t \le T}$ and

$$\mathrm{P\&L}^H_{\psi,\Delta}(X) = \psi(X_T) + (\Delta \bullet X)_T - H$$

is the profit and loss resulting from three terms: the static hedge ψ, which is constructed from vanilla options at cost $\mu(\psi) := \int \psi \, \mathrm{d}\mu$, dynamic (self-financing) trading according to Δ leading to the stochastic integral $(\Delta \bullet X)_T$, and selling the claim H. For the static hedge ψ, it is sufficient (cf. [22]) to look for profiles of the form

$$\psi(x) = \gamma^{\$} + \int_{\mathbb{R}_+} (x - K)^+ \lambda(\mathrm{d}K) \tag{5}$$

where $\gamma^{\$}$ is a cash amount and λ is a signed measure describing the weights of a call options portfolio (as eventual positions in the forward contract can be absorbed by the delta hedge).

Granted that Fubini's theorem applies, the cost of ψ reads $\mu(\psi) = \gamma^{\$} + \int_{\mathbb{R}_+} c(K,T)\lambda(\mathrm{d}K)$, where $c(K,T) = \mathbb{E}_\mu[(X-K)^+]$.

In our problem formulation, we keep the assumption that trading in $(X_t)_{0\le t\le T}$ is frictionless but highlight the static profile ψ, which is now exposed to the bid–ask frictions in the options market. We naturally assume that static hedging is limited to put and call options (and the forward). In fact, puts would be redundant given calls and the forward, hence we may suppose again that ψ consists of a cash amount and a portfolio of call options as in (5). In the presence of bid–ask spreads, we need to separate the contracts that were bought from those that were sold. To this end, consider the Jordan decomposition $\lambda = \lambda^+ - \lambda^-$ of the portfolio weights into nonnegative measures $\lambda^\pm$. We can then decompose the static profile as

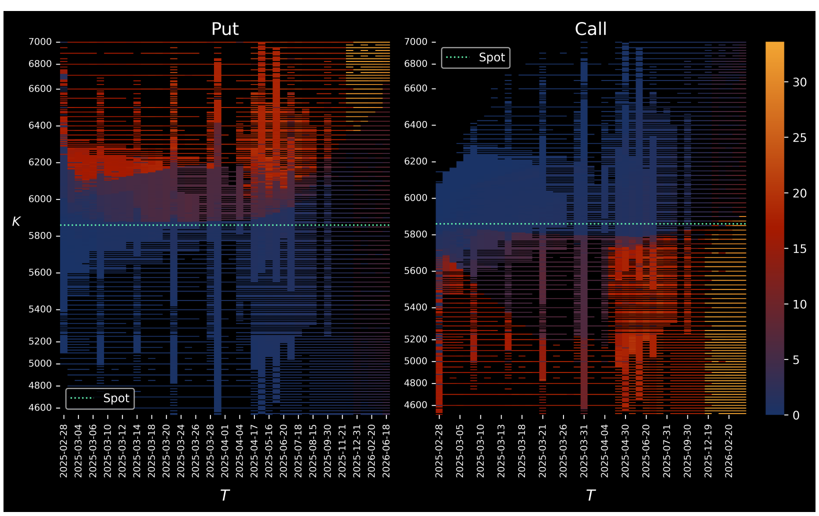
$$\psi = \psi^{\mathrm{a}} - \psi^{\mathrm{b}}, \quad \psi^{\mathrm{a}}(x) = \gamma^{\$} + \int_{\mathbb{R}_+} (x-K)^+ \lambda^+(\mathrm{d}K), \quad \psi^{\mathrm{b}}(x) = \int_{\mathbb{R}_+} (x-K)^+ \lambda^-(\mathrm{d}K).$$

Since $\lambda^\pm$ are nonnegative, both $\psi^{\mathrm{a}}, \psi^{\mathrm{b}}$ are *convex* profiles, the first being bought at the ask and the second sold at the bid. Writing $c^{\mathrm{b}}(K,T)$ and $c^{\mathrm{a}}(K,T)$ for the bid and ask call option prices, the cost of ψ becomes

$$\gamma^{\$} + \int_{\mathbb{R}_+} c^{\mathrm{a}}(K,T)\lambda^+(\mathrm{d}K) - \int_{\mathbb{R}_+} c^{\mathrm{b}}(K,T)\lambda^-(\mathrm{d}K).$$

Figure 2 displays the bid–ask spread for put and call options on the S&P 500 Index (SPX) across strikes and maturities as of the close of 02-07-2025. As can be seen, the spreads are non-negligible, even for liquid underlyings such as SPX. With bid and ask marginals $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$ as in Section 1.2, we can write the cost of the static profile succinctly as $\mu^{\mathrm{a}}(\psi^{\mathrm{a}}) - \mu^{\mathrm{b}}(\psi^{\mathrm{b}})$.

Figure 2: Bid–ask spread ($) of S&P 500 Index (SPX) options as of 02-07-2025.



More generally, we formulate our dual problem over *pairs* $\psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}})$ of convex functions, with ψ^{a} corresponding to call options bought (plus a cash position) and ψ^{b} corresponding to call options sold.[1] Writing

$$\mathrm{cost}(\psi^{\mathrm{b,a}}) = \mu^{\mathrm{a}}(\psi^{\mathrm{a}}) - \mu^{\mathrm{b}}(\psi^{\mathrm{b}}) \quad \text{for} \quad \psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}),$$

[1] We do not insist that $\psi^{\mathrm{a}} - \psi^{\mathrm{b}}$ be a Jordan decomposition. However, a Jordan decomposition has the minimal cost $\mu^{\mathrm{a}}(\psi^{\mathrm{a}}) - \mu^{\mathrm{b}}(\psi^{\mathrm{b}})$ among all decompositions into convex functions.

we arrive at the dual formulation

$$\inf_{\psi^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(H)} \mathrm{cost}(\psi^{\mathrm{b,a}}), \tag{6}$$

where

$$\Psi^{\mathrm{b,a}}(H) = \big\{\psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) \text{ convex} \mid \exists\, \Delta \text{ such that } \mathrm{P\&L}^{H}_{\psi^{\mathrm{a}}-\psi^{\mathrm{b}},\Delta} \geq 0\big\}.$$

Our strong duality result (Theorem 3.4) will prove that the optimal superhedging cost (6) coincides with the value of the primal problem (4). The presence of bid–ask spreads increases the value of the dual problem as $\mathrm{cost}(\psi^{\mathrm{b,a}}) \geq \mu(\psi^{\mathrm{a}} - \psi^{\mathrm{b}})$ for any $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. We will see in Section 5 that the value may differ substantially from the price under the mid-marginal $\mu = (\mu^{\mathrm{a}} + \mu^{\mathrm{b}})/2$ given real market data.

In contrast to the dual in classical MOT, (6) is nontrivial even for vanilla claims $H = h(X_T)$. Even if $h = \psi^0 - \psi^1$ is itself a difference of convex functions, e.g., via Jordan decomposition, there may be convex pairs $(\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) =: \psi^{\mathrm{b,a}}$ such that $\psi^{\mathrm{a}} - \psi^{\mathrm{b}} \geq h$ and $\mathrm{cost}(\psi^{\mathrm{b,a}}) < \mu^{\mathrm{a}}(\psi^0) - \mu^{\mathrm{b}}(\psi^1)$. That is, the *Jordan decomposition can be suboptimal* for the dual problem. A concrete example is shown in Fig. 8, where the risk-reversal strategy $h(x) = (x - 1.05)^+ - (0.95 - x)^+$ is optimally superhedged by a linear combination of *three* call options.

1.4 Contributions

On the conceptual side, we introduce a novel generalization of martingale optimal transport that is consistent with bid–ask spreads for vanilla options. While directly using order book data of bid and ask quotes would lead to a general linear programming problem, a key observation is that the existing market practice of fitting bid and ask marginal distributions opens the door to a tractable formulation in the realm of optimal transport.

On the theoretical side, in Section 2, we first formalize the corresponding primal (transport) problem and the dual (superhedging) problem. We also describe more precisely the relation between discrete bid and ask quotes in order books and the corresponding idealized marginal distributions. In Section 3, we then establish strong duality: the BAMOT primal value (4) coincides with the dual superhedging cost (6) for any upper semicontinuous payoff $h(X_{T_1}, \ldots, X_{T_N})$ with, at most, linear growth. Methodologically, we first show strong duality for the single-maturity case by a Hahn–Banach argument (Theorem 3.2). This case corresponds to superhedging a payoff $h(X_T)$ by a difference $\psi^{\mathrm{a}} - \psi^{\mathrm{b}}$ of convex functions of X_T and is genuinely novel compared to MOT. We then show strong duality for the general, multi-maturity case (Theorem 3.4) by combining the single-period result with MOT duality via a minimax argument.

Section 4 continues the theoretical analysis by studying the relationship between BAMOT and the classical MOT problem without bid–ask spreads. Proposition 4.1 establishes consistency, in the sense that the BAMOT value converges to the MOT limit as the bid–ask spread vanishes. By duality, the optimal superhedging costs then also converge. In the important special case of a single maturity, we further provide a quantitative analysis. To that end, we introduce a novel metric, the *bid–ask distance*, which aligns with the spreads of quoted vanilla options and captures the convex ordering between marginals. We then show a linear rate of convergence of the BAMOT price for payoffs, which are differences of convex functions (Proposition 4.5), and a square-root rate for upper semicontinuous payoffs (Proposition 4.6). Both rates are shown to be numerically sharp in examples, namely, for a risk-reversal strategy and an at-the-money digital option (Section 5.3).

On the practical side, we provide analytical and computational examples illustrating the formulation and its implications. In Section 5.1, we derive a closed-form solution for digital call options in a one-sided market and validate the impact of bid–ask frictions using real market data on the S&P 500 Index (SPX). Our findings reveal that the common practice of pricing digital options using the mid-price marginal can substantially underestimate (superhedging) prices relative to BAMOT in the presence of bid–ask frictions. In Section 5.2, we compare BAMOT super- and subhedging bounds for a forward-start payoff with those obtained under classical MOT. Finally, in Section 5.3, we document the convergence behavior as bid–ask spreads shrink for a risk-reversal strategy and an at-the-money digital call option.

1.5 Related Literature

Addressing market friction arising from bid–ask spreads has a long history in robust finance (see, e.g., [2, 10, 13, 17, 26] and the references therein). Transaction cost analysis was pioneered by [26], which introduced a sublinear pricing operator to capture the bid–ask spread of trading a contingent claim. It is shown that the superhedging cost of the claim agrees with the maximum contract value among all martingale measures consistent with the bid and ask prices of the hedging instruments. Closer to our robust finance framework, [15] utilized sublinear operators to describe the cost of static hedging and established general duality results that also include proportional transaction costs for dynamic hedging. In the present work, we assume that costs for trading in the underlying are negligible relative to the bid–ask spreads of the vanilla options, as is typically satisfied in practice for moderate trading frequency. Moreover, we focus on a particular model close to industry practices, rather than aiming for general results. A related duality was recently established by [17] in the context of generalized Limit Order Books (LOB) involving bundles of securities.

Over the past several years, researchers have increasingly leveraged MOT to address market frictions. However, many works focus on capturing bid–ask spreads in less liquid exotic options [18, 21]. In particular, these studies assume that vanilla options can still be bought and sold at a unique price, leading to exact calibration of the marginal distributions. Other approaches within the MOT framework specifically accommodate the uncertainty in inferring marginal distributions from market spreads. For instance, [24] introduced *η-market models*, which utilize a uniform tolerance parameter η for all vanilla options, effectively resulting in a constant bid–ask spread across all strikes. Alternatively, [40] directly optimized over a set of martingale measures whose marginals reside within a pre-defined ε-tolerance set. Nonetheless, selecting appropriate hyperparameters for η and ε is often challenging in practice. By contrast, the bid and ask marginals imposed in the present work are motivated by industry practice and can be directly calibrated to the bid and ask prices observed in market data. We mention that [32] took steps toward addressing these concerns, albeit from a more empirical perspective: the authors proposed a neural network-based approach to solve the dual problem, incorporating an objective function that explicitly accounts for the discrepancy between bid and ask prices.

In a different direction, not directly related to optimal transport, the modeling of bid–ask spreads using separate marginals has been explored through "lower" and "upper" measures by [28, 29]. In these studies, positions are grouped by directional sentiment: bullish views (long calls, short puts) are valued under the upper measure, while bearish views (long puts, short calls) utilize the lower measure. Consequently, the ask prices for calls and puts originate from different distributions.

While empirically robust, this framework does not provide the unified marginal constraints required for MOT, where bid and ask prices are recovered from a single consistent pair of distributions.

Finally, our results in Section 4 about convergence of BAMOT for shrinking bid–ask spreads are related to the literature on the stability of MOT, that is, its behavior as a function of the given marginals. Here [27] derived the stability of the left-curtain coupling, thereby establishing stability results for MOT. Independently, [5, 38] obtained continuity of the value functions and optimal couplings by exploiting the adapted Wasserstein topology. Subsequent works [4, 31] further confirmed this topology as the natural framework for studying stability in MOT. On the computational side, [20] investigated stability in order to provide guarantees for numerical methods when the marginals are approximated discretely.

1.6 Notation

Let $\mathcal{C}_L(\mathbb{R}_+^N)$ be the Banach space of continuous functions $f : \mathbb{R}_+^N \to \mathbb{R}$ with linear growth, endowed with the norm

$$\|f\|_L := \sup_{x \in \mathbb{R}_+^N} \frac{|f(x)|}{1 + \|x\|}.$$

Introduce also the larger set $\mathrm{USC}_L(\mathbb{R}_+^N)$ of upper semicontinuous (u.s.c.) functions with linear growth, and the smaller set $\mathrm{CVX}_L(\mathbb{R}_+^N)$ of convex functions with linear growth. When $N = 1$, we drop the dependence on the ambient space and simply write $\mathcal{C}_L, \mathrm{CVX}_L, \mathrm{USC}_L$.

Let $\mathcal{P}_1(\mathbb{R}_+^N)$ be the set of probability measures on $\mathbb{R}_+^N$ with finite first moment, and abbreviate again $\mathcal{P}_1 = \mathcal{P}_1(\mathbb{R}_+)$. The 1-Wasserstein distance $\mathcal{W}_1(\mu, \nu)$ is defined via $\mathcal{W}_1(\mu, \nu) = \inf \int \|x - y\| \mathbb{P}(\mathrm{d}x, \mathrm{d}y)$, where the infimum is taken over all couplings $\mathbb{P}$ of (μ, ν), i.e., all $\mathbb{P} \in \mathcal{P}(\mathbb{R}_+^N \times \mathbb{R}_+^N)$ with marginals (μ, ν). Integrals are denoted, interchangeably, by

$$\mu(f) = \mathbb{E}_\mu[f] = \int f \, \mathrm{d}\mu.$$

We say that $\mu, \nu \in \mathcal{P}_1$ are in *convex order* ($\mu \preceq_{\mathrm{c}} \nu$) if $\mu(\psi) \leq \nu(\psi)$ for all $\psi \in \mathrm{CVX}_L$. In fact, it is sufficient to verify the latter inequality for call payoffs, provided that the measures have identical barycenters

$$\mu \preceq_{\mathrm{c}} \nu \qquad \Longleftrightarrow \qquad \mathbb{E}_\mu[X] = \mathbb{E}_\nu[X] \quad \text{and} \quad \mathbb{E}_\mu[(X - K)^+] \leq \mathbb{E}_\nu[(X - K)^+] \quad \forall\, K \geq 0. \tag{7}$$

See, e.g., [34] for further background on convex order, and in particular [34, Theorem 3.A.1] for the above equivalence.

Organization The remainder of the paper is organized as follows. Section 2 presents a practitioners' construction of bid–ask marginals and formulates the primal and dual BAMOT problems. Section 3 establishes strong duality. Section 4 proves consistency and derives the convergence rate in the single-maturity case $N = 1$. Section 5 presents illustrative examples and numerical results. We conclude by discussing possible directions for future work in Section 6. Appendix A discusses the existence of bid and ask measures. Technical proofs and other complementary materials are presented in Appendices B and C.

2 Bid–Ask MOT

This section introduces the primal (measure) problem and the dual (portfolio) problem of BAMOT in detail. Both problems depend crucially on the assumption of bid and ask marginals that induce the prices of liquidly traded call and put options. Working with such marginals is market practice, offering a flexible method to calibrate bid–ask implied volatility skews, while ruling out static arbitrage. Fig. 1 replicates a typical market interface for implied volatility skews, such as the OVDV function on the Bloomberg Terminal. On the other hand, a snapshot of real-world order book data may contain quotes that cannot be exactly reproduced by two distributions. The next subsections explain how practitioners typically calibrate bid and ask marginals to quotes; see [12] and the OVDV documentation available on the Bloomberg Terminal.

2.1 Bid and Ask Marginals

Next, we describe a common procedure used by practitioners to construct bid and ask marginals by calibrating to bid and ask quotes of out-of-the-money (OTM) put and call options.[2] Fix a maturity T, let F_T, x_0 denote, respectively, the forward and spot price of the underlying, let $\{K_m\}_{m\in[M]}$ be the set of available strikes, and let $\{\mathrm{OTM}^{\mathrm{a}}(K_m,T)\}_{m\in[M]}$ and $\{\vartheta^{\mathrm{a}}(K_m,T)\}_{m\in[M]}$ denote, respectively, the ask prices and Vegas of the corresponding OTM options (put if $x_0 < K_m$ and call otherwise).

Practitioners then select a parametric family of densities $\{\phi^\theta \mid \theta\in\Theta\}$, where Θ is chosen so that the forward price is matched (i.e., $\int x\phi^\theta(x)\,\mathrm{d}x = F_T$ for all $\theta\in\Theta$), and solve the optimization

$$\theta^{\mathrm{a}} \in \arg\min_{\theta\in\Theta} \sum_{m\in[M]} \left| \frac{\mathrm{OTM}^{\mathrm{a}}(K_m,T) - \mathrm{OTM}^{\theta}(K_m,T)}{\vartheta^{\mathrm{a}}(K_m,T)} \right|^2,$$

where $\mathrm{OTM}^\theta(K,T)$ denotes the OTM option price under ϕ^θ. Here pricing errors are weighted by Vegas to approximately convert price discrepancies into errors in implied volatilities (IVs). The resulting distribution $\mu^{\mathrm{a}}(\mathrm{d}x) = \phi^{\theta^{\mathrm{a}}}(x)\,\mathrm{d}x$ is referred to as the calibrated *ask marginal*. One then perturbs the fitted parameters θ^{a} to obtain bid parameters θ^{b} that match bid quotes, and sets $\mu^{\mathrm{b}}(\mathrm{d}x) = \phi^{\theta^{\mathrm{b}}}(x)\,\mathrm{d}x$ as the corresponding *bid marginal*. In particular, one uses a perturbation such that $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$ to rule out *static arbitrage*.

A common choice of parametric family is mixtures of log-normal densities [12]. Suppose the ask marginal is calibrated as a log-normal mixture with J components, i.e.,

$$\phi^{\mathrm{a}}(x) = \sum_{j=1}^{J} w_j\,\phi_{\mathrm{BS}}(x; z_j, \sigma_j^{\mathrm{a}}), \qquad \sum_{j=1}^{J} w_j = 1,$$

where $\phi_{\mathrm{BS}}(\cdot; z,\sigma)$ denotes the density of a log-normal distribution with mean and volatility parameters $(z,\sigma)\in\mathbb{R}_+^2$. The bid marginal is then constructed by scaling down the volatility parameters, while keeping the remaining parameters unchanged:

$$\phi^{\mathrm{b}}(x) = \sum_{j=1}^{J} w_j\,\phi_{\mathrm{BS}}(x; z_j, \sigma_j^{\mathrm{b}}),$$

[2] OTM options are chosen as they are usually more liquid and therefore lead to more reliable and consistent quotes.

for some calibrated $\sigma_j^{\mathrm{b}} \le \sigma_j^{\mathrm{a}}$, $j \in [J]$.

Given multiple maturities $T_1 \le T_2 \le \cdots \le T_N$, we apply the above procedure at each maturity and obtain bid and ask marginals $\mu_i^{\mathrm{b}}, \mu_i^{\mathrm{a}}$ such that $\mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_i^{\mathrm{a}}$ for each $i \in [N] := \{1, \ldots, N\}$.

For simplicity, assume zero interest rate and no carry cost.[3] In this case, $F_{T_i} = x_0$ for all $i \in [N]$. The absence of *calendar arbitrage* across maturities suggests the cross-maturity condition $\mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_j^{\mathrm{a}}$ for any $1 \le i \le j \le N$. Indeed, if there exists $K \ge 0$ such that $\mathbb{E}_{\mu_i^{\mathrm{b}}}[(X-K)^+] > \mathbb{E}_{\mu_j^{\mathrm{a}}}[(X-K)^+]$, then one could sell the (K, T_i) call at a price at least $\mathbb{E}_{\mu_i^{\mathrm{b}}}[(X-K)^+]$ and buy the (K, T_j) call at a price at most $\mathbb{E}_{\mu_j^{\mathrm{a}}}[(X-K)^+]$; together with dynamic trading in the underlying, this yields an arbitrage.

The above discussion motivates the following assumption, which we adopt throughout the paper.

Assumption 2.1. There exist bid and ask marginals $\mu^{\mathrm{b}} = (\mu_1^{\mathrm{b}}, \ldots, \mu_N^{\mathrm{b}})$ and $\mu^{\mathrm{a}} = (\mu_1^{\mathrm{a}}, \ldots, \mu_N^{\mathrm{a}})$ with finite first moment, which determine the bid and ask prices of all call and put options with the respective maturities and satisfy

$$\mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_j^{\mathrm{a}}, \qquad \forall\, 1 \le i \le j \le N. \tag{8}$$

Remark 2.2. (a) Condition (8) rules out butterfly arbitrage; for any $i \in [N]$, $K \in \mathbb{R}_+$,

$$\mathbb{E}_{\mu_i^{\mathrm{a}}}[(X-K_-)^+] - 2\,\mathbb{E}_{\mu_i^{\mathrm{b}}}[(X-K)^+] + \mathbb{E}_{\mu_i^{\mathrm{a}}}[(X-K_+)^+] \ge 0, \quad K_\pm = K \pm \delta,\ \delta > 0. \tag{9}$$

Indeed, $\mathbb{E}_{\mu_i^{\mathrm{b}}}[(X-K)^+] \le \mathbb{E}_{\mu_i^{\mathrm{a}}}[(X-K)^+]$ by (8), so the left-hand side of (9) is bounded from below by $\mathbb{E}_{\mu_i^{\mathrm{a}}}[B(X,K)]$, with the butterfly spread $B(x,K) = (x-K_-)^+ - 2(x-K)^+ + (x-K_+)^+$. Since $B(\cdot, K) \ge 0$, its price under μ_i^{a} (and hence under any admissible bid–ask calibration) must be nonnegative.

(b) By a suitable choice of the bid and ask marginals, our framework covers situations in which the payoff depends on maturities for which vanilla options are not liquidly traded. For example, if $N = 2$ and the payoff is $H = h(X_{T_1}, X_{T_2})$, while vanillas are available at T_2 but not at T_1, one may take $\mu_1^{\mathrm{b}} := \delta_{x_0}$ and $\mu_1^{\mathrm{a}} := \mu_2^{\mathrm{a}}$, thereby leaving the T_1-marginal effectively unconstrained, while still enforcing the cross-maturity condition (8). Such a situation naturally arises when the product references an intermediate fixing date, e.g., forward-start, whereas market quotes are only liquid at coarser maturities.

2.2 Primal (Measure) Problem

Fix maturities $T_1 \le T_2 \le \cdots \le T_N$ with corresponding bid marginals $\mu^{\mathrm{b}} = (\mu_1^{\mathrm{b}}, \ldots, \mu_N^{\mathrm{b}})$ and ask marginals $\mu^{\mathrm{a}} = (\mu_1^{\mathrm{a}}, \ldots, \mu_N^{\mathrm{a}})$. We write $X_1, \ldots, X_N$ for the canonical process representing the spot price, where we identify $X_i = X_{T_i}$ for notational simplicity. By convention, the initial price $X_0 := x_0 \in \mathbb{R}_+$ is deterministic. Thus, a *martingale measure* is a probability measure $\mathbb{Q}$ on $\mathbb{R}_+^N$ such that $(X_i)_{0 \le i \le N}$ is a $\mathbb{Q}$-martingale. The set of *calibrated* martingale measures is

$$\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) := \left\{ \mathbb{Q} \mid \text{martingale measure},\ \mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_i^{\mathbb{Q}} \preceq_{\mathrm{c}} \mu_i^{\mathrm{a}},\ i \in [N] \right\}, \tag{10}$$

[3] With positive interest rates and/or dividends, the marginals $\mu_i^{\mathrm{b}}, \mu_i^{\mathrm{a}}$ should be calibrated from *discounted, dividend-adjusted quotes.*

where $\mu_i^{\mathbb{Q}}$ denotes the i-th marginal distribution of $\mathbb{Q}$. Fix a contingent claim $H = h(X_1, \ldots, X_N)$ for some payoff function $h \in \mathrm{USC}_L(\mathbb{R}_+^N)$. Then the *primal* (or *measure*) problem is defined as

$$P(h) := \sup_{\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})} \mathbb{E}_{\mathbb{Q}}[h(X_1, \ldots, X_N)]. \tag{11}$$

We observe that

$$\text{Assumption (8)} \quad \Longleftrightarrow \quad \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) \neq \emptyset.$$

Indeed, define $\mu_i = \min_{j \geq i} \mu_j^{\mathrm{a}}$, where the infimum is with respect to the convex order. Then $\mu_i \preceq_{\mathrm{c}} \mu_j$ for all $i \leq j$ and $\mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_i \preceq_{\mathrm{c}} \mu_i^{\mathrm{a}}$ for all i. By Strassen's theorem, the former implies the existence of a martingale measure $\mathbb{Q}$ with marginals $\mu = (\mu_1, \ldots, \mu_N)$, and $\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ by the latter. Conversely, if $\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$, let $(\mu_1, \ldots, \mu_N)$ be its vector of marginals. The martingale property implies $\mu_i \preceq_{\mathrm{c}} \mu_j$ for all $i \leq j$ and then $\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ implies (8).

Remark 2.3. Consider again the lower envelope $\underline{\mu}_i^{\mathrm{a}} := \min_{j \geq i} \mu_j^{\mathrm{a}}$ used in the above argument, and similarly the upper envelope $\overline{\mu}_i^{\mathrm{b}} := \max_{j \leq i} \mu_j^{\mathrm{b}}$. We see that (8), and in turn $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) \neq \emptyset$, are equivalent to $\overline{\mu}_i^{\mathrm{b}} \preceq_{\mathrm{c}} \underline{\mu}_i^{\mathrm{a}}$ for all i. That is, the prices of convex payoffs with respect to the lower and upper envelopes do not cross, as illustrated in Fig. 3 for an illiquid, in-the-money call on the S&P 500 Index (SPX).

Figure 3: Term structure of S&P 500 Index (SPX) call options struck at $K = 5550$ as of 2025-02-27, and corresponding price envelopes.

2.3 Dual (Portfolio) Problem

Consider a payoff function $h \in \mathrm{USC}_L(\mathbb{R}_+^N)$, meaning that the claim is $h(X_1, \ldots, X_N)$, and static hedging instruments $\psi = (\psi_1, \ldots, \psi_N)$ with payoffs $\psi_i(X_i)$. Moreover, let $\Delta = (\Delta_1, \ldots, \Delta_{N-1})$ be a progressively measurable trading strategy (delta hedge), that is, Δ_i is a measurable function of $X_1, \ldots, X_i$. The profit and loss associated with ψ, Δ, h is defined as

$$\mathrm{P\&L}_{\psi,\Delta}^h(x_1, \ldots, x_N) = \sum_{i=1}^N \psi_i(x_i) + \sum_{i=1}^{N-1} \Delta_i(x_1, \ldots, x_i)(x_{i+1} - x_i) - h(x_1, \ldots, x_N).$$

As customary in MOT [3, 22], it is sufficient to look for static profiles of the form

$$\psi_i(x) = \gamma_i^{\$} + \int_{\mathbb{R}_+} (x-K)^+ \lambda_i(\mathrm{d}K),$$

noting that eventual positions in the forward contract can be absorbed by the delta hedge. In our dual problem, the delta hedges are arbitrary, whereas the static positions are expressed as the difference $\psi_i = \psi_i^{\mathrm{a}} - \psi_i^{\mathrm{b}}$ of convex functions. Here ψ_i^{a} corresponds to a combination of call options with maturity T_i that are bought, while ψ_i^{b} corresponds to call options which are sold. (Put options are redundant given calls and the forward contract, hence can be ignored.) For brevity, we combine these functions into a vector

$$\psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) = (\psi_1^{\mathrm{a}}, \dots, \psi_N^{\mathrm{a}}, \psi_1^{\mathrm{b}}, \dots, \psi_N^{\mathrm{b}}) \in \mathrm{CVX}_L^{2N}$$

whose total cost is

$$\mathrm{cost}(\psi^{\mathrm{b,a}}) = \sum_{i=1}^N \mu_i^{\mathrm{a}}(\psi_i^{\mathrm{a}}) - \mu_i^{\mathrm{b}}(\psi_i^{\mathrm{b}}).$$

We denote the set of all admissible superhedging strategies by

$$\Psi^{\mathrm{b,a}}(h) = \Big\{ (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) \in \mathrm{CVX}_L^{2N} \;\mid\; \exists\, \Delta \text{ s.t. } \mathrm{P\&L}^h_{\psi^{\mathrm{a}} - \psi^{\mathrm{b}}, \Delta} \geq 0 \Big\}; \tag{12}$$

here and below, it is tacitly understood that Δ is progressively measurable. The *dual* (or *portfolio*) problem is then defined as

$$D(h) := \inf_{\psi^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(h)} \mathrm{cost}(\psi^{\mathrm{b,a}}), \tag{13}$$

The interpretation is straightforward: sell one unit of h; at each maturity T_i, buy a convex profile ψ_i^{a} priced at the ask, sell another profile ψ_i^{b} at the bid, and hold Δ_i shares of the underlying asset to hedge against potential losses and minimize cost.

3 Duality

The goal of this section is to show the strong duality $P(h) = D(h)$, where $P(h)$ and $D(h)$ denote the values of the primal (11) and dual (13) problems, respectively. We first state the weak duality $P(h) \leq D(h)$, which is straightforward.

Lemma 3.1 (Weak Duality)**.** *We have $P(h) \leq D(h)$ for all $h \in \mathrm{USC}_L(\mathbb{R}_+^N)$.*

Proof. Let $\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ and $\psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) \in \Psi^{\mathrm{b,a}}(h)$, and let Δ be a corresponding dynamic strategy such that $\mathrm{P\&L}^h_{\psi^{\mathrm{a}} - \psi^{\mathrm{b}}, \Delta} \geq 0$. Since $\mathbb{Q}$ is a martingale measure with $\mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_i^{\mathbb{Q}} \preceq_{\mathrm{c}} \mu_i^{\mathrm{a}}$, and $\psi_i^{\mathrm{a}}, \psi_i^{\mathrm{b}}$ are convex with linear growth, we deduce

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[h(X_1, \dots, X_N)] &\leq \mathbb{E}_{\mathbb{Q}}\Bigg[\sum_{i=1}^N \psi_i^{\mathrm{a}}(X_i) - \psi_i^{\mathrm{b}}(X_i) \Bigg] + \mathbb{E}_{\mathbb{Q}}\Bigg[\sum_{i=1}^{N-1} \Delta_i(X_1, \dots, X_i)(X_{i+1} - X_i) \Bigg] \\ &= \sum_{i=1}^N \mu_i^{\mathbb{Q}}(\psi_i^{\mathrm{a}}) - \mu_i^{\mathbb{Q}}(\psi_i^{\mathrm{b}}) \leq \sum_{i=1}^N \mu_i^{\mathrm{a}}(\psi_i^{\mathrm{a}}) - \mu_i^{\mathrm{b}}(\psi_i^{\mathrm{b}}) = \mathrm{cost}(\psi^{\mathrm{b,a}}). \end{aligned}$$

Here we have used that the discrete stochastic integral has vanishing expectation. Indeed, the negative part of the terminal value of this local martingale is $\mathbb{Q}$-integrable due to the superhedging property and the linear growth of h, and that implies that it is a true martingale [25, Theorem 2]. The result follows by taking supremum over $\mathbb{Q}$ and infimum over $\psi^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(h)$. $\square$

The nontrivial direction of the strong duality will be established in two steps. We first focus on the single-maturity case $N = 1$, where we prove strong duality from first principles via a Hahn–Banach argument. We then show the result for the general case by combining the strong duality of classical MOT with the single-maturity case.

3.1 Single Maturity

Let $N = 1$ and consider the single maturity $T_1 > 0$ as well as corresponding bid and ask marginals $\mu^{\mathrm{b}}, \mu^{\mathrm{a}} \in \mathcal{P}_1$ satisfying $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. We slightly abuse notation by writing μ^{s} for μ^{s}_1, etc., and identifying martingale measures with their marginal μ at T_1 (rather than using the full martingale measure $\delta_{x_0} \otimes \mu$). Then the primal problem (11) boils down to

$$P(h) = \sup_{\mu \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})} \mu(h), \qquad \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) = \{\mu \in \mathcal{P}_1 \mid \mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu \preceq_{\mathrm{c}} \mu^{\mathrm{a}}\}, \tag{14}$$

where $h \in \mathrm{USC}_L$ is the payoff function of a T_1-claim $h(X_1)$. On the other hand, the dual problem (13) becomes

$$D(h) = \inf_{\psi^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(h)} \mathrm{cost}(\psi^{\mathrm{b,a}}), \qquad \mathrm{cost}(\psi^{\mathrm{b,a}}) = \mu^{\mathrm{a}}(\psi^{\mathrm{a}}) - \mu^{\mathrm{b}}(\psi^{\mathrm{b}}),$$

$$\Psi^{\mathrm{b,a}}(h) = \{\psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) \in \mathrm{CVX}^2_L \mid \psi^{\mathrm{a}} - \psi^{\mathrm{b}} \geq h\}.$$

That is, $D(h)$ is the cost of the cheapest static superhedging portfolio of h, given by the difference of convex functions. Observe that $P(h) = \mu^{\mathrm{a}}(h) = D(h)$ if h is convex, and $P(h) = \mu^{\mathrm{b}}(h) = D(h)$ when h is concave. The next result extends strong duality to general payoffs.

Theorem 3.2 (Strong Duality, $N = 1$)**.** *For all $h \in \mathrm{USC}_L$, we have*

$$P(h) = \sup_{\mu \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})} \mu(h) = \inf_{\psi \in \Psi^{\mathrm{b,a}}(h)} \mathrm{cost}(\psi) = D(h).$$

Moreover, there exists a primal optimizer $\mu^\star \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$.

Proof. The weak duality $P(h) \leq D(h)$ was shown in Lemma 3.1. To show the reverse inequality, we first assume that $h \in \mathcal{C}_L(\mathbb{R}_+)$. Our aim is to construct a measure $\mu^\star \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ such that $\mu^\star(h) = D(h)$. Since we already know that $\mu^\star(h) \leq P(h) \leq D(h)$, this will imply that $P(h) = D(h)$ and $\mu^\star$ is a maximizer.

Step 1. We show that $D : \mathcal{C}_L(\mathbb{R}_+) \to \mathbb{R}$ is a sublinear functional. Clearly $D(\lambda f) = \lambda D(f)$ for all $\lambda > 0$ and $f \in \mathcal{C}_L(\mathbb{R}_+)$. Note also that $D(\lambda) = \lambda$ for all $\lambda \in \mathbb{R}$, using, e.g., $\psi \equiv (\lambda, 0) \in \Psi^{\mathrm{b,a}}(\lambda)$. Moreover, D is subadditive,

$$D(f + f') \leq D(f) + D(f') \quad \forall\ f, f' \in \mathcal{C}_L(\mathbb{R}_+).$$

Indeed, $\Psi^{\mathrm{b,a}}(f) + \Psi^{\mathrm{b,a}}(f') \subseteq \Psi^{\mathrm{b,a}}(f+f')$, which gives

$$D(f+f') = \inf_{\psi\in\Psi^{\mathrm{b,a}}(f+f')} \mathrm{cost}(\psi) \le \inf_{\psi\in\Psi^{\mathrm{b,a}}(f),\,\psi'\in\Psi^{\mathrm{b,a}}(f')} \mathrm{cost}(\psi+\psi') = D(f) + D(f')$$

since $\mathrm{cost}(\cdot)$ is linear. This completes the proof that D is sublinear.

<u>*Step 2.*</u> Consider the one-dimensional linear subspace $\mathcal{M} := \{\lambda h : \lambda \in \mathbb{R}\} \subseteq \mathcal{C}_L(\mathbb{R}_+)$ and introduce the linear functional $\ell(\lambda h) := \lambda D(h)$ on $\mathcal{M}$. Then $\ell(\lambda h) = D(\lambda h)\ \forall\lambda > 0$ due to the positive homogeneity of D. Moreover, $\ell(-\lambda h) \le D(-\lambda h)$ for all $\lambda \ge 0$, as follows from

$$\ell(\lambda h) + \ell(-\lambda h) = \ell(0) = 0 = D(0) = D(\lambda h - \lambda h) \le D(\lambda h) + D(-\lambda h).$$

Hence $\ell(f) \le D(f)$ for all $f \in \mathcal{M}$, that is, ℓ is dominated by D on $\mathcal{M}$. By the Hahn–Banach Dominated Extension Theorem, ℓ can be extended to a linear functional (again denoted ℓ) on $\mathcal{C}_L(\mathbb{R}_+)$ such that $\ell(f) \le D(f)$ for all $f \in \mathcal{C}_L(\mathbb{R}_+)$. Moreover, ℓ is a positive functional in the sense that $\ell(f) \ge 0$ whenever $f \ge 0$. Indeed, if $f \ge 0$ but $\ell(f) < 0$, then as $(0,0) \in \Psi^{\mathrm{b,a}}(-f)$, we obtain the contradiction

$$0 < -\ell(f) = \ell(-f) \le D(-f) \le \mathrm{cost}(0) = 0.$$

Finally, by the definition of $\|\cdot\|_L$, the convex function $\psi^{\mathrm{a}}(x) := \|f\|_L(1+|x|)$ dominates f. In other words, $(\psi^{\mathrm{a}}, 0) \in \Psi^{\mathrm{b,a}}(f)$, leading to

$$\ell(f) \le D(f) \le \mu^{\mathrm{a}}(\psi^{\mathrm{a}}) = (1+c)\ \|f\|_L,$$

where the constant $c = \int |x|\mu^{\mathrm{a}}(\mathrm{d}x)$ is finite due to $\mu^{\mathrm{a}} \in \mathcal{P}_1$. In summary, ℓ is a bounded, linear, positive functional on $\mathcal{C}_L(\mathbb{R}_+)$.

<u>*Step 3.*</u> Let us verify the following tightness condition: for any $\varepsilon > 0$, there exists $R > 0$ such that

$$|\ell(f)| \le \varepsilon\|f\|_L, \quad \forall f \in \mathcal{C}_L(\mathbb{R}_+) \quad \text{s.t. } f \equiv 0 \text{ on } [0,R].$$

Let $f \in \mathcal{C}_L(\mathbb{R}_+)$ satisfy $f \equiv 0$ on $[0,R]$, where $R \ge 2$. Then $|f(x)| \le 3\|f\|_L(x-R/2)^+ =: \|f\|_L g_R(x)$. Moreover, as μ^{a} has finite first moment, dominated convergence shows that $\mu^{\mathrm{a}}(g_R) \to 0$ for $R \to \infty$. By the positivity of ℓ, we deduce that

$$\ell(f) \le \ell(\|f\|_L\, g_R) \le \|f\|_L\, D(g_R) = \|f\|_L\, \mu^{\mathrm{a}}(g_R) \to 0, \quad R \to \infty.$$

The same argument applies to $-\ell(f) = \ell(-f)$, completing the proof of tightness.

<u>*Step 4.*</u> As $(\mathcal{C}_L, \|\cdot\|_L)$ is isomorphic to $(\mathcal{C}_b, \|\cdot\|_\infty)$ via $f \mapsto \frac{f}{1+|\cdot|}$ and ℓ is bounded linear (Step 2) satisfying the tightness condition (Step 3), the Riesz Representation Theorem [9, Theorem 7.10.6] implies the existence of a signed Radon measure $\mu^\star$ on $\mathbb{R}_+$ (with finite first moment), which represents ℓ via $\ell(f) = \mu^\star(f)$. Positivity of ℓ implies that $\mu^\star$ is a nonnegative measure. Moreover, $\ell(1) = 1$ due to $\ell(\pm 1) \le D(\pm 1) = \pm 1$, so that $\mu^\star$ is a probability measure. Let $f \in \mathrm{CVX}_L$, then $(f, 0) \in \Psi^{\mathrm{b,a}}(f)$ and

$$\mu^\star(f) = \ell(f) \le D(f) = \mu^{\mathrm{a}}(f).$$

Similarly, $(0, f) \in \Psi^{\mathrm{b,a}}(-f)$, leading to $\mu^\star(f) \ge \mu^{\mathrm{b}}(f)$. Thus $\mu^{\mathrm{b}}(f) \le \mu^\star(f) \le \mu^{\mathrm{a}}(f)$ for all $f \in \mathrm{CVX}_L$, showing that $\mu^\star \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$. This completes the proof that $D(h) = \ell(h) = \mu^\star(h) \le P(h)$

for the given claim $h \in \mathcal{C}_L(\mathbb{R}_+)$.

Step 5. It remains to extend the result to $h \in \mathrm{USC}_L$. Consider the sup convolution

$$h^\varepsilon(x) = \sup_{y \in \mathbb{R}_+} \left(h(y) - \frac{1}{\varepsilon}|y - x| \right), \quad \varepsilon > 0.$$

Then h^ε is finite for all $\varepsilon < \bar{\varepsilon} := \|h\|_L^{-1}$. For any such ε, recall that h^ε is Lipschitz continuous with constant $1/\varepsilon$, as follows from

$$h^\varepsilon(x) - h^\varepsilon(x') \leq \frac{1}{\varepsilon} \sup_{y \in \mathbb{R}_+} \left[|y - x| - |y - x'| \right] \leq \frac{1}{\varepsilon}|x - x'|, \quad x, x' \in \mathbb{R}_+.$$

Hence $h^\varepsilon \in \mathcal{C}_L(\mathbb{R}_+)$ and from the previous steps, $P(h^\varepsilon) = D(h^\varepsilon)$. Moreover, $h^\varepsilon \downarrow h$ as $\varepsilon \downarrow 0$. By the $\mathcal{W}_1$-compactness of the set $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ in the definition (14), this implies that $P(h^\varepsilon) \downarrow P(h)$ (as in, e.g., [14, Theorem 31]). On the other hand, $D(h^\varepsilon) \geq D(h)$ for all $\varepsilon > 0$ by monotonicity, so that $P(h) = \lim_{\varepsilon \to 0} P(h^\varepsilon) \geq D(h)$. This completes the proof that $P(h) = D(h)$. The existence of a maximizer $\mu^\star \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ follows from the $\mathcal{W}_1$-compactness of $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ and $h \in \mathrm{USC}_L$. □

Remark 3.3. Dolinsky and Soner [15] treat market friction in the static hedging instruments through an abstract pricing operator $\mathscr{P}(\cdot)$ describing their prices. This operator is used in both the dual and the primal formulation. In the primal, the marginal constraint becomes $\mu(f) \leq \mathscr{P}(f)$ for all measurable functions $f : \mathbb{R}_+ \to \mathbb{R}$ satisfying a certain growth condition. By contrast, our primal formulation (14) with the supremum over $\mu \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ makes the constraint explicit in terms of bid and ask marginals obtained from market data. While Theorem 3.2 could also be derived from the abstract duality of [15] with some additional work, we preferred to provide an elementary proof.

3.2 Multiple Maturities

We now turn to strong duality for path-dependent claims involving $N \geq 1$ maturities. Fix marginals $\mu^{\mathrm{b}}, \mu^{\mathrm{a}} \in \mathcal{P}_1^N$ satisfying Assumption 2.1. We recall the set

$$\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) = \left\{ \mathbb{Q} \mid \text{martingale measure}, \ \mu_i^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_i^{\mathbb{Q}} \preceq_{\mathrm{c}} \mu_i^{\mathrm{a}}, \ i \in [N] \right\}$$

of calibrated martingale measures and consider a payoff of the form $h(X_1, \ldots, X_N)$.

Theorem 3.4 (Strong Duality)**.** *For all $h \in \mathrm{USC}_L(\mathbb{R}_+^N)$, we have*

$$P(h) = \sup_{\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})} \mathbb{E}_{\mathbb{Q}}[h(X_1, \ldots, X_N)] = \inf_{\psi \in \Psi^{\mathrm{b,a}}(h)} \mathrm{cost}(\psi) = D(h)$$

and there exists an optimizer $\mathbb{Q}^\star \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ to the primal problem.

Proof. Set $\mathcal{Q}_i := \mathcal{Q}(\mu_i^{\mathrm{b}}, \mu_i^{\mathrm{a}})$ and $\mathcal{Q}_\otimes := \prod_{i=1}^N \mathcal{Q}_i$. Then $\mathcal{Q}_\otimes$ is convex and $\mathcal{W}_1$-compact by Lemma B.2. Next, observe that

$$P(h) = \sup_{\mu \in \mathcal{Q}_\otimes} \sup_{\mathbb{Q} \in \mathcal{Q}(\mu)} \mathbb{E}_{\mathbb{Q}}[h], \quad \text{where} \quad \mathcal{Q}(\mu) = \left\{ \mathbb{Q} \mid \text{martingale measure}, \ \mu_i^{\mathbb{Q}} = \mu_i, \ i \in [N] \right\}. \quad (15)$$

Consider $\mu \in \mathcal{Q}_\otimes$ satisfying $\mu_i \preceq_{\mathrm{c}} \mu_j$ for $1 \le i < j \le N$. We then apply the strong duality of classical MOT [3, Corollary 1.2] to obtain

$$\sup_{\mathbb{Q} \in \mathcal{Q}(\mu)} \mathbb{E}_{\mathbb{Q}}[h] = \inf_{\psi \in \Psi(h)} \sum_{i=1}^N \mu_i(\psi_i), \tag{16}$$

where $\Psi(h) = \{\psi \in \mathcal{C}_L^N \mid \exists\, \Delta \text{ s.t. } \mathrm{P\&L}^h_{\psi,\Delta} \ge 0\}$. For $\mu \in \mathcal{Q}_\otimes$ with $\mu_i \not\preceq_{\mathrm{c}} \mu_j$ for some $i < j$, we have $\mathcal{Q}(\mu) = \varnothing$ and hence the supremum on the left-hand side equals $-\infty$. On the other hand, such μ allows for calendar arbitrage, and using also the linear growth of h, we see that the right-hand side also equals $-\infty$. In summary, (16) holds for any $\mu \in \mathcal{Q}_\otimes$, and combining with (15) yields

$$P(h) = \sup_{\mu \in \mathcal{Q}_\otimes} \sup_{\mathbb{Q} \in \mathcal{Q}(\mu)} \mathbb{E}_{\mathbb{Q}}[h] = \sup_{\mu \in \mathcal{Q}_\otimes} \inf_{\psi \in \Psi(h)} \sum_{i=1}^N \mu_i(\psi_i)\,.$$

Next, we apply the minimax theorem. Consider $\mathcal{X} = \mathcal{Q}_\otimes$ equipped with the $\mathcal{W}_1$-topology and $\mathcal{Y} = \Psi(h)$ equipped with the product topology induced by $\|\cdot\|_L$. Then $\mathcal{X}, \mathcal{Y}$ are convex and $\mathcal{X}$ is compact. Moreover, $\mathcal{X} \times \mathcal{Y} \ni (\mu, \psi) \mapsto \sum_{i=1}^N \mu_i(\psi_i)$ is linear and continuous with respect to the product topology. Therefore, the minimax theorem [35, Corollary 3.3] yields

$$P(h) = \inf_{\psi \in \Psi(h)} \sup_{\mu \in \mathcal{Q}_\otimes} \sum_{i=1}^N \mu_i(\psi_i) = \inf_{\psi \in \Psi(h)} \sum_{i=1}^N \sup_{\mu_i \in \mathcal{Q}_i} \mu_i(\psi_i). \tag{17}$$

For each i, applying Theorem 3.2 to the claim $f = \psi_i \in \mathcal{C}_L$ yields

$$\sup_{\mu_i \in \mathcal{Q}_i} \mu_i(\psi_i) = \inf_{\psi_i^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(\psi_i)} \mu_i^{\mathrm{a}}(\psi_i^{\mathrm{a}}) - \mu_i^{\mathrm{b}}(\psi_i^{\mathrm{b}}).$$

Inserting this in (17) yields

$$\begin{aligned} P(h) &= \inf_{\psi \in \Psi(h)} \sum_{i=1}^N \inf_{\psi_i^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(\psi_i)} \mu_i^{\mathrm{a}}(\psi_i^{\mathrm{a}}) - \mu_i^{\mathrm{b}}(\psi_i^{\mathrm{b}}) \\ &\ge \inf_{\psi^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(h)} \sum_{i=1}^N \mu_i^{\mathrm{a}}(\psi_i^{\mathrm{a}}) - \mu_i^{\mathrm{b}}(\psi_i^{\mathrm{b}}) = D(h). \end{aligned}$$

with $\Psi^{\mathrm{b,a}}(h)$ given in (12), and now combining with weak duality of Lemma 3.1 concludes the proof of $P(h) = D(h)$. The existence of a primal optimizer again follows by compactness and semicontinuity. □

4 Convergence to Classical MOT

In this section, we study the convergence of BAMOT to classical MOT when the bid–ask spread converges to zero. That is, for each maturity, the bid and ask marginals converge to a single marginal, generating the unambiguous prices of vanillas for that maturity. The first subsection establishes consistency in the sense that the BAMOT value (and hence the optimal superhedging cost) converges to the natural limit. Section 4.2 introduces a novel metric, termed the *bid–ask distance*, which is directly tied to the bid–ask spreads of call and put options. Section 4.3 uses this metric for a quantitative analysis of the convergence to MOT. We restrict our attention to the single-maturity case $N = 1$ and establish two results: a linear rate of convergence when h can be written as a difference of Lipschitz convex functions, and a square-root rate of convergence when $h \in \mathrm{USC}_L$ with bounded variation. The general multi-maturity setting is left for future investigation.

4.1 Consistency

The following result shows the continuity of the BAMOT value for a sequence of bid and ask marginals with monotonically decreasing spread. In particular, if the spread converges to zero, meaning that bid and ask marginals converge to the same limit, it establishes the consistency of BAMOT with classical MOT. Thus, if the spread is small, the superhedging price with bid–ask spread is approximately equal to the superhedging price obtained with static hedging instruments priced with the mid marginal.

Proposition 4.1. *Let $\{(\mu_n^{\mathrm{b}}, \mu_n^{\mathrm{a}})\}_{n\in\mathbb{N}} \subset \mathcal{P}_1^N \times \mathcal{P}_1^N$ be a sequence of bid and ask marginals such that*

$$\mu_{n,i}^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_{m,i}^{\mathrm{b}} \preceq_{\mathrm{c}} \mu_{m,i}^{\mathrm{a}} \preceq_{\mathrm{c}} \mu_{n,i}^{\mathrm{a}} \quad \forall n \leq m,\ i \in [N],$$

which implies the existence of limits $\mu_\infty^{\mathrm{b}} = \lim_m \mu_m^{\mathrm{b}}$ and $\mu_\infty^{\mathrm{a}} = \lim_m \mu_m^{\mathrm{a}}$ that are again valid bid and ask marginals. Consider the corresponding BAMOT problems

$$P_n(h) := \sup_{\mathbb{Q}\in\mathcal{Q}(\mu_n^{\mathrm{b}},\mu_n^{\mathrm{a}})} \mathbb{E}_{\mathbb{Q}}[h(X_1,\dots,X_N)], \quad n \in \mathbb{N}\cup\{\infty\}$$

for a fixed payoff function $h \in USC_L(\mathbb{R}_+^N)$ that is bounded from above. Then

$$P_n(h) \to P_\infty(h).$$

In particular, if $\mu_\infty^{\mathrm{b}} = \mu_\infty^{\mathrm{a}} =: \mu_\infty$, then $P_n(h)$ converges to the value $\sup_{\mathbb{Q}\in\mathcal{Q}(\mu_\infty)} \mathbb{E}_{\mathbb{Q}}[h(X_1,\dots,X_N)]$ of the classical MOT problem associated with μ_∞.

Proof. The existence of $\lim \mu_n^{\mathrm{b}}$ can be seen from the associated call prices (or potential functions), which are monotone in n and therefore convergent. Similarly for μ_n^{a}.

Set $\mathcal{Q}_n = \mathcal{Q}(\mu_n^{\mathrm{b}}, \mu_n^{\mathrm{a}})$, so that $\mathcal{Q}_\infty = \cap_n \mathcal{Q}(\mu_n^{\mathrm{b}}, \mu_n^{\mathrm{a}})$. Since $\mathcal{Q}_n$ decreases to $\mathcal{Q}_\infty$, clearly $\{P_n(h)\}_{n\in\mathbb{N}}$ is a decreasing sequence with $\lim_{n\to\infty} P_n(h) \geq P_\infty(h)$. We show that $\lim_{n\to\infty} P_n(h) \leq P_\infty(h)$.

Suppose first that h is bounded and continuous. Let $\mathbb{Q}_n^* \in \mathcal{Q}_n$ denote a maximizer to $P_n(h)$. As $\{\mathbb{Q}_n^*\}_{n\in\mathbb{N}} \subset \mathcal{Q}_1$ and $\mathcal{Q}_1$ is weakly compact, after taking a subsequence, there exists $\mathbb{Q}_\infty^*$ such that $\mathbb{Q}_n^*$ converges weakly to $\mathbb{Q}_\infty^*$. Then $\mathbb{E}_{\mathbb{Q}_n^*}[h] \to \mathbb{E}_{\mathbb{Q}_\infty^*}[h]$ since h is bounded and continuous. Moreover, as $\{\mathcal{Q}_m\}_{m\in\mathbb{N}}$ is decreasing, $\mathbb{Q}_m^* \in \mathcal{Q}_m \subset \mathcal{Q}_n$ for all $m \geq n$. Letting $m \to \infty$ yields $\mathbb{Q}_\infty^* \in \mathcal{Q}_n$ as $\mathcal{Q}_n$ is weakly closed, and as n was arbitrary, we conclude that $\mathbb{Q}_\infty^* \in \cap_{n\in\mathbb{N}}\mathcal{Q}_n = \mathcal{Q}_\infty$. In summary, $\lim_{n\to\infty} P_n(h) = \lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}_n^*}[h] = \mathbb{E}_{\mathbb{Q}_\infty^*}[h] \leq P_\infty(h)$, completing the proof when h is bounded and continuous.

In the general case, there exist $h_k \in \mathcal{C}_b(\mathbb{R}_+^N)$ decreasing to h, and the above shows that $P_n(h_k)$ decreases to $P_\infty(h_k)$ for every k. For each $n \in \mathbb{N}\cup\{\infty\}$, the sublinear expectation P_n satisfies $P_n(h_k) \downarrow P_n(h)$ by [14, Theorem 31]. Given $\varepsilon > 0$, choose k such that $|P_\infty(h) - P_\infty(h_k)| < \varepsilon/2$ and then $n \in \mathbb{N}$ such that $|P_\infty(h_k) - P_n(h_k)| < \varepsilon/2$. By monotonicity of P_n, we have $P_n(h) \leq P_n(h_k) \leq P_\infty(h) + \varepsilon$, showing that $\lim_{n\to\infty} P_n(h) \leq P_\infty(h)$ as desired. □

4.2 The Bid–Ask Distance

Next, we introduce a metric that is directly tied to the bid–ask spreads of call and put options. It also provides a characterization of the convex order between probability measures.

Definition 4.2. Given $\mu, \nu \in \mathcal{P}_1$, introduce the directed distance

$$\vec{d}(\mu, \nu) = \sup \{(\mu - \nu)(\psi) \mid \psi \text{ convex}, \operatorname{Lip}(\psi) \leq 1\},$$

and the *bid–ask distance* given by the symmetrization

$$d(\mu, \nu) = \frac{\vec{d}(\mu, \nu) + \vec{d}(\nu, \mu)}{2}.$$

Recently, [39] (see also [1]) proposed an alternative characterization of convex order in the 2-Wasserstein space. Their criterion relies on maximal covariances, a formulation that does not directly translate to the bid–ask interpretation needed for our analysis. The bid–ask distance also relates to the relaxed L^∞ metric introduced in [36], which is constructed from an asymmetric distance to compare semicontinuous functions in a nearly uniform manner. The following proposition collects key properties of $\vec{d}$ and d.

Proposition 4.3. *Let $\mathcal{W}_1$ denote the 1-Wasserstein distance. Then the following hold.*

(i) $\vec{d}(\mu, \nu) = 0$ *if and only if* $\mu \preceq_{\mathrm{c}} \nu$.

(ii) $\vec{d}$ *is an asymmetric distance [30]; it is nonnegative, satisfies the triangle inequality, and separates points in the sense that* $\vec{d}(\mu, \nu) = \vec{d}(\nu, \mu) = 0$ *implies* $\mu = \nu$. *Consequently, d is a metric.*

(iii) If μ, ν have identical barycenters, then

$$\vec{d}(\mu, \nu) = 2 \sup_{K \geq 0} \left(\mathbb{E}_\mu[(X - K)^+] - \mathbb{E}_\nu[(X - K)^+]\right).$$

(iv) We have $d(\mu, \nu) \leq \mathcal{W}_1(\mu, \nu)$ *for all* $\mu, \nu \in \mathcal{P}_1$. *Moreover, there exist sequences* $(\mu_n), (\nu_n)$ *in* $\mathcal{P}_1$ *such that*

$$\lim_{n \to \infty} d(\mu_n, \nu_n) = 0, \quad \textit{while } \mathcal{W}_1(\mu_n, \nu_n) = 1 \quad \forall n \in \mathbb{N}.$$

In particular, d and $\mathcal{W}_1$ induce distinct topologies.

Proof. See Appendix B. □

Remark 4.4. (a) Property (iii) highlights the connection between the proposed distance and the bid–ask spreads of quoted vanilla options; if $c^{\mathrm{s}}(K,T) = \mathbb{E}_{\mu^{\mathrm{s}}}[(X - K)^+]$, $\mathrm{s} \in \{\mathrm{b}, \mathrm{a}\}$, denotes the price of the (K, T) call option on each side of market, then

$$d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) = \frac{1}{2} \vec{d}(\mu^{\mathrm{a}}, \mu^{\mathrm{b}}) = \sup_{K \geq 0} \left(c^{\mathrm{a}}(K,T) - c^{\mathrm{b}}(K,T)\right), \tag{18}$$

where the first equality follows from property (i). As each spread in (18) represents the (round-trip) transaction cost of trading these options, the distance indeed measures the level of bid–ask friction in the market.

(b) The bid–ask distance d is equivalent to the following *integral probability metric*,

$$\tilde{d}(\mu,\nu) = \sup\left\{(\mu-\nu)(\psi) \mid \psi = \psi^{\mathrm{a}} - \psi^{\mathrm{b}}\right\}, \tag{19}$$

where ψ^{a} and ψ^{b} are convex and 1-Lipschitz. Indeed, observe that $\vec{d}(\mu,\nu) \le \tilde{d}(\mu,\nu)$, so $d(\mu,\nu) \le \tilde{d}(\mu,\nu)$ as well. On the other hand, for any test function $\psi = \psi^{\mathrm{a}} - \psi^{\mathrm{b}}$ in (19),

$$(\mu-\nu)(\psi) = (\mu-\nu)(\psi^{\mathrm{a}}) + (\nu-\mu)(\psi^{\mathrm{b}}) \le \vec{d}(\mu,\nu) + \vec{d}(\nu,\mu) = 2\,d(\mu,\nu).$$

(c) The directed distance $\vec{d}$ is related to weak optimal transport through the convex Kantorovich–Rubinstein duality formula (see, e.g., [8]),

$$\sup_{\psi \text{ convex, } \mathrm{Lip}(\psi)\le 1} (\mu-\nu)(\psi) = \inf_{\mathbb{P}\in\Pi(\mu,\nu)} \mathbb{E}_{\mathbb{P}}\left[\left|\mathbb{E}_{\mathbb{P}}[X_1 \mid X_0] - X_0\right|\right], \tag{20}$$

where $(X_0, X_1) \sim \mathbb{P}$ and $\Pi(\mu,\nu)$ denotes the set of all couplings between μ and ν. The right-hand side of (20) thus provides an alternative characterization of $\vec{d}(\mu,\nu)$. As observed in [8], (20) is a quantitative analogue of Strassen's theorem; the left-hand side quantifies the extent to which convex order is violated, while the dual measures the deviation from the martingale property.

4.3 Convergence Rate

Next, we focus on the single-maturity case $N = 1$ and show how the bid–ask distance controls the discrepancy between the BAMOT price and its frictionless counterpart. When h is given by the difference of convex functions, then the rate is linear as shown next. In particular, the result applies to all multi-leg option strategies $h(x) = \int_{\mathbb{R}_+} (x-K)^+ \lambda(\mathrm{d}K)$ such as bull, bear, and butterfly spreads. The following results are presented from the primal perspective (14), noting that similar rates apply to the dual value by strong duality.

Proposition 4.5. *Let $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$ and denote the mid marginal by $\mu^{\mathrm{m}} = \frac{\mu^{\mathrm{b}}+\mu^{\mathrm{a}}}{2}$. Suppose that $h = \psi^{\mathrm{a}} - \psi^{\mathrm{b}}$ for some Lipschitz pair $(\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) \in \mathrm{CVX}^2_L$. Then,*

$$0 \le P(h) - \mu^{\mathrm{m}}(h) \le C\, d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}),$$

where $C = \mathrm{Lip}(\psi^{\mathrm{a}}) + \mathrm{Lip}(\psi^{\mathrm{b}})$.

Proof. The lower bound follows as $\mu^{\mathrm{m}} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$. On the other hand, from the definitions of μ^{m} and $\vec{d}$, we obtain after rearranging that

$$P(h) - \mu^{\mathrm{m}}(h) \le \frac{1}{2}(\mu^{\mathrm{a}} - \mu^{\mathrm{b}})(\psi^{\mathrm{a}}) + \frac{1}{2}(\mu^{\mathrm{a}} - \mu^{\mathrm{b}})(\psi^{\mathrm{b}}) \le \frac{1}{2}(\mathrm{Lip}(\psi^{\mathrm{a}}) + \mathrm{Lip}(\psi^{\mathrm{b}}))\vec{d}(\mu^{\mathrm{a}}, \mu^{\mathrm{b}}). \tag{21}$$

Since $\vec{d}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) = 0$ by $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$ and Proposition 4.3 (i), we know that $\frac{1}{2}\vec{d}(\mu^{\mathrm{a}}, \mu^{\mathrm{b}}) = d(\mu^{\mathrm{a}}, \mu^{\mathrm{b}})$, which, together with (21), concludes the upper bound. □

If h is upper semicontinuous with bounded variation, as in the case of digital options, the convergence rate is typically square-root, as shown next. We will see in Section 5.3 that this rate is in fact numerically sharp.

Proposition 4.6. *Let $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$ and assume that the mid-marginal $\mu^{\mathrm{m}} = \frac{\mu^{\mathrm{b}}+\mu^{\mathrm{a}}}{2}$ admits a density ϕ^{m} that is bounded by $M > 0$. Let $h \in \mathrm{USC}_L$ have bounded variation $V(h) < \infty$. Then,*

$$0 \le P(h) - \mu^{\mathrm{m}}(h) \le C\sqrt{d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})},$$

where $C = \sqrt{2M}V(h)$.

Proof. Fix $\mu \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ and denote its call price function by $c(K) = \mathbb{E}_{\mu}[(X-K)^+]$. Define also $c^{\mathrm{s}}(K) = \mathbb{E}_{\mu^{\mathrm{s}}}[(X-K)^+]$, $\mathrm{s} \in \{\mathrm{b}, \mathrm{m}, \mathrm{a}\}$. For technical reasons, we extend c to $\mathbb{R}$ by setting $c(K) = \mathbb{E}_{\mu}[X] - K$ when $K < 0$. Since $\mu \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$, Proposition 4.3 (iii) implies that

$$|c(K) - c^{\mathrm{m}}(K)| \le \tfrac{1}{2}\big(c^{\mathrm{a}}(K) - c^{\mathrm{b}}(K)\big) \le \frac{1}{2}d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}), \quad \forall K \ge 0.$$

In addition, $c(K) - c^{\mathrm{m}}(K) = 0$ for any $K < 0$ because $\mathbb{E}_{\mu}[X] = \mathbb{E}_{\mu^{\mathrm{m}}}[X]$ and both μ and μ^{m} are supported on $\mathbb{R}_+$. Thus,

$$\sup_{K \in \mathbb{R}} |c(K) - c^{\mathrm{m}}(K)| \le \frac{1}{2}d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}).$$

Moreover, since ϕ^{m} is bounded by M, we have $c^{\mathrm{m}} \in \mathcal{C}^{1,1}(\mathbb{R})$ and $\partial_K c^{\mathrm{m}}$ is M-Lipschitz. Applying Lemma B.3 to the convex functions $\psi = c^{\mathrm{m}}$, $\varphi = c$, and $\varepsilon = \frac{1}{2}d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ on $\mathbb{R}$, we obtain

$$\sup_{K \ge 0} \left|\partial_K^+ c(K) - \partial_K c^{\mathrm{m}}(K)\right| \le \sqrt{2M\, d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})}, \tag{22}$$

where $\partial_K^+ c$ denotes the right derivative of c. Recall the identities $\partial_K^+ c = \Phi - 1$ and $\partial_K c^{\mathrm{m}} = \Phi^{\mathrm{m}} - 1$, where Φ and Φ^{m} denote the distribution functions of μ and μ^{m}, respectively [11]. Together with (22), this yields

$$\sup_{K \ge 0} |\Phi(K) - \Phi^{\mathrm{m}}(K)| \le \sqrt{2M\, d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})}. \tag{23}$$

Since h has bounded variation, integration by parts for the Riemann–Stieltjes integral gives

$$(\mu - \mu^{\mathrm{m}})(h) = \int_{\mathbb{R}_+} h\,(\mathrm{d}\Phi - \mathrm{d}\Phi^{\mathrm{m}}) = -\int_{\mathbb{R}_+} (\Phi - \Phi^{\mathrm{m}})\,\mathrm{d}h.$$

The boundary terms vanish since $\Phi(0^-) = \Phi^{\mathrm{m}}(0^-) = 0$ and $\lim_{K\to\infty}\Phi(K) = \lim_{K\to\infty}\Phi^{\mathrm{m}}(K) = 1$. Combining this with (23) yields

$$(\mu - \mu^{\mathrm{m}})(h) \le V(h)\sup_{K \ge 0} |\Phi(K) - \Phi^{\mathrm{m}}(K)| \le V(h)\sqrt{2M\, d(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})}.$$

Taking the supremum over $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ shows the claim with $C = \sqrt{2M}V(h)$. □

5 Examples

This section presents a collection of numerical examples involving multiple illiquid claims in the market. In Section 5.1, we focus on digital call options and present a closed-form solution in a one-sided market. We further validate our theoretical findings via a numerical experiment using options data on the S&P 500 Index (SPX).[4] In Section 5.2, we compare the super- and subhedging prices obtained under BAMOT with those derived using classical MOT. Finally, in Section 5.3, we demonstrate the convergence behavior as the bid–ask spread shrinks for a risk-reversal payoff and an at-the-money digital call option. Unless stated otherwise, all numerical solutions are obtained via a discretization scheme (see Appendix C for details) and solved using linear programming. We refer to the resulting numerical solutions as the primal and dual optimizers. However, it should be emphasized that a dual optimizer for the original, undiscretized BAMOT problem need not exist.

5.1 Digital Options

This section is devoted to digital options. Beyond their long-standing use in fixed income, foreign exchange, and commodities markets, their popularity has surged recently on prediction market platforms such as Polymarket and Kalshi. We first study a special case in which the market is one-sided, that is, bid quotes are absent, leading to explicit primal and dual optimizers. We then present numerical results for a digital call option written on the S&P 500 Index (SPX) using real market data, from which we observe that pricing digital options using mid marginals can substantially underestimate the superhedging prices in the presence of bid–ask frictions.

Our analysis relies on the concept of *local concentration*. Given $\mu \in \mathcal{P}_1$, we can construct another probability measure μ' such that $\mu' \preceq_{\mathrm{c}} \mu$ as follows. Let $I_1, \dots, I_n$ be a partition of $\mathbb{R}_+$ into disjoint intervals $\{I_i\}$ such that $w_i := \mu(I_i) > 0$ for each i. Let $x_i = \frac{1}{w_i} \int_{I_i} x \, \mu(\mathrm{d}x)$ denote the barycenter of μ on I_i. Then $\mu' = \sum_{i=1}^n w_i \delta_{x_i} \in \mathcal{P}_1$, and $\mu' \preceq_{\mathrm{c}} \mu$.

5.1.1 Digital Option in a One-Sided Market

We consider a one-sided option market with a single maturity T and no available bid quotes. Hence $\mu^{\mathrm{b}} = \delta_{x_0}$, where we choose $x_0 = 1$, while the ask marginal distribution is an arbitrary atomless measure with mean x_0. We aim to find the superhedging price and construct an optimal measure for an out-of-the-money (OTM) digital call option with payoff $h(x) = \mathbb{1}_{\{x \geq K\}}$, where $K > x_0$. Then, Theorem 3.2 becomes

$$P(h) = \sup_{\mu \in \underline{\mathcal{Q}}(\mu^{\mathrm{a}})} \mu([K, \infty)) = \inf_{\psi^{\mathrm{b,a}} \in \Psi^{\mathrm{b,a}}(h)} \mu^{\mathrm{a}}(\psi^{\mathrm{a}}) - \psi^{\mathrm{b}}(x_0) = D(h). \tag{24}$$

It turns out that both primal and dual optimizers of (24) can be found explicitly. Indeed, introduce $b^{\mathrm{a}}(L) := \int_L^\infty (x - K) \mu^{\mathrm{a}}(\mathrm{d}x)$. Since $b^{\mathrm{a}}(0) = x_0 - K < 0$, $b^{\mathrm{a}}(K) \geq 0$ and by the continuity of $L \mapsto b^{\mathrm{a}}(L)$, there exists a critical strike $L^\star \leq K$ such that $b^{\mathrm{a}}(L^\star) = 0$. For our purposes, we assume that $L^\star < K$ as otherwise μ^{a} is supported on $[0, K]$, in which case the problem is trivial.

Proposition 5.1. *Introduce the convex profile* $\psi_L(x) := \frac{1}{K-L}(x - L)^+$. *Then,*

(i) $\mu^\star := \mu^{\mathrm{a}}\big|_{[0,L^\star]} + \mu^{\mathrm{a}}([L^\star, \infty))\delta_K \in \underline{\mathcal{Q}}(\mu^{\mathrm{a}})$ *is a primal optimizer of* (24).

(ii) $(\psi_{L^\star}, 0) \in \Psi^{\mathrm{b,a}}(h)$ *is a dual optimizer of* (24).

(iii) $P(h) = D(h) = \frac{c^{\mathrm{a}}(L^\star, T)}{K - L^\star} = \mu^{\mathrm{a}}([L^\star, \infty))$.

[4] Data retrieved from Bloomberg.

Proof. By local concentration, $\mu^\star \in \mathcal{Q}(\delta_{x_0}, \mu^{\mathrm{a}}) = \underline{\mathcal{Q}}(\mu^{\mathrm{a}})$. It is also clear that $(\psi_{L^\star}, 0) \in \Psi^{\mathrm{b,a}}(h)$, as illustrated in the right panel of Fig. 4. Then by definition of $\mu^\star$, $L^\star$, and ψ_L,

$$D(h) \leq \mu^{\mathrm{a}}(\psi_{L^\star}) = \frac{1}{K - L^\star} \int_{L^\star}^{\infty} (x - L^\star)\, \mu^{\mathrm{a}}(\mathrm{d}x) = \mu^{\mathrm{a}}([L^\star, \infty)) = \mu^\star(h) \leq P(h).$$

The primal and dual optimality follow by weak duality. ◻

Primal-Dual Relations Let $\sigma^\star(\cdot, T)$ denote the implied volatility skew of the optimal measure $\mu^\star$. Then, the implied volatility at the optimal strike $L^\star$ satisfies

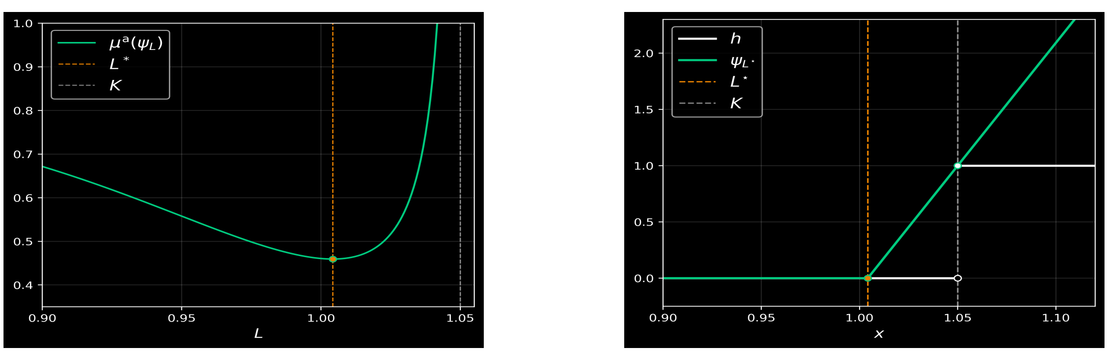
$$\sigma^\star(L^\star, T) = \sigma^{\mathrm{a}}(L^\star, T). \tag{25}$$

Indeed, the price of the $(L^\star, T)$ call under $\mu^\star$ is $c^\star(L^\star, T) = \mu^\star(h)(K - L^\star) = c^{\mathrm{a}}(L^\star, T)$. Similarly, observe that $c^\star(K, T) = c^{\mathrm{b}}(K, T) = (x_0 - K)^+ = 0$, since the optimal measure is supported on $[0, K]$. Equivalently, $\sigma^\star(K, T) = \sigma^{\mathrm{b}}(K, T) = 0$, as shown in the following example.

Example 5.2. Let $T = 1/12$ (one month), $K = 1.05$, and μ^{a} is the one-month marginal distribution in the Black–Scholes model with volatility $\sigma^{\mathrm{a}} = 20\%$ and zero interest rates. Then the superhedging price is approximately equal to $D(h) = 0.46$, with optimal lower strike $L^\star \approx 1.004$. Fig. 4 displays the cost function $L \mapsto \mu^{\mathrm{a}}(\psi_L)$ and the optimal superhedging portfolio. The superhedging price is significantly higher than the price of the digital under the ask marginal, roughly equal to 0.19. As shown in Fig. 5, different optimal measures can be constructed by varying the distribution on $[0, L^\star]$ using local concentration.

Figure 4: Superhedging of a digital call in a one-sided market.

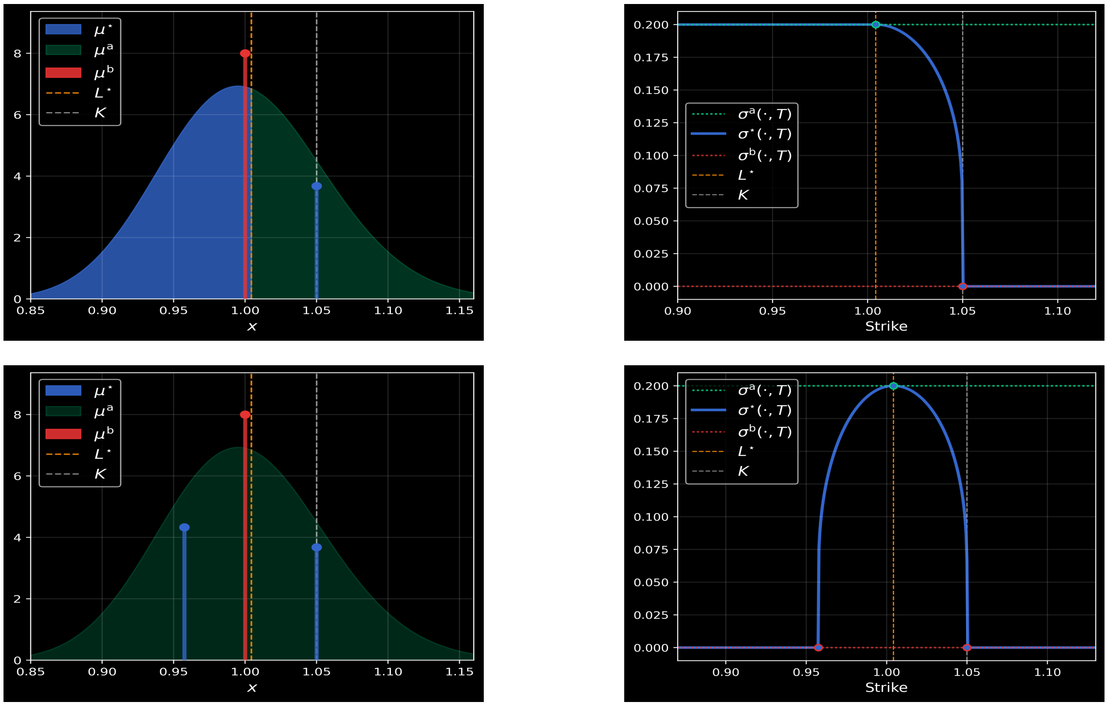
(a) Super-replication cost as function of L (b) Optimal superhedging portfolio



Connection with One-touch Options. The present example can be seen as a static analogue of Hobson's robust hedging of one-touch (American digital) options [23]. Indeed, consider the discrete optimal measure $\mu^\star$ displayed in the bottom left chart of Fig. 5. In view of $\delta_{x_0} = \mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^\star \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$, there exists a continuous-time càdlàg martingale $(X_t)_{t\in[0,1]}$ such that $X_0 = x_0$, which jumps to $X_{1/2} \sim \mu^\star$ at $t = 1/2$ and to $X_1 \sim \mu^{\mathrm{a}}$ at the terminal time. This is precisely the model constructed in [23] that achieves the superhedging price for the one-touch call option. Specifically, we have

$$\sup_{\mathbb{Q},\, X_T \sim \mu^{\mathrm{a}}} \mathbb{Q}(X_\tau \geq K) = \sup_{\mu \,\preceq_{\mathrm{c}}\, \mu^{\mathrm{a}}} \mu([K, \infty)), \tag{26}$$

Figure 5: Digital call in a one-sided market. Illustrations of different optimal measures (left) and corresponding implied volatility skews (right).



where τ is the first time X exceeds the strike K (equal to $\tau = 1$ if $X_t < K$ for all t), and $\mathbb{Q}$ is a continuous-time martingale measure. In fact, the identity still holds if the marginal constraint on the left-hand side is relaxed to $\mathrm{Law}^{\mathbb{Q}}(X_T) \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. This is because Hobson's optimal superhedging strategy utilizes only long positions in call options alongside a dynamic hedge; since no options are sold, the strategy is independent of the bid marginal. As the right-hand side corresponds to a simpler problem, our framework also provides a useful compression tool for some path-dependent robust pricing problems. In addition, by the optional sampling theorem and Strassen's theorem, (26) can be extended to connect our framework with the robust pricing of American options,

$$\sup_{\tau \le T} \sup_{\mathbb{Q},\, X_T \sim \mu^{\mathrm{a}}} \mathbb{E}_{\mathbb{Q}}[h(X_\tau)] = \sup_{\mu \,\preceq_{\mathrm{c}}\, \mu^{\mathrm{a}}} \mu(h),$$

where $h \in \mathrm{USC}_L$ and τ ranges over all stopping times taking values in $[0, T]$.

5.1.2 Digital Option on the S&P 500 Index (SPX)

We now compute the superhedging price of an out-of-the-money digital call using real market data. The bid and ask marginals are calibrated to the option chains of SPX put and call options quoted on 2025-02-27 with maturity on 2025-07-18, using four-component log-normal mixture models. The parameters are fitted using the discounted forward price $x_0 = 5861$ and are reported in Table 1. Observe that the mixture weights and component means are shared across the bid and ask marginals, while the component volatilities are calibrated so as to preserve convex order as mentioned in Section 2.1.

Figure 6: A dual optimizer, the potential function, and implied volatility skew of a primal optimizer for the digital payoff $h(x) = 100\mathbb{1}_{\{x \geq 1.05x_0\}}$ on the S&P 500 Index (SPX).

(a) Dual optimizer (b) Potential function (c) Implied volatility

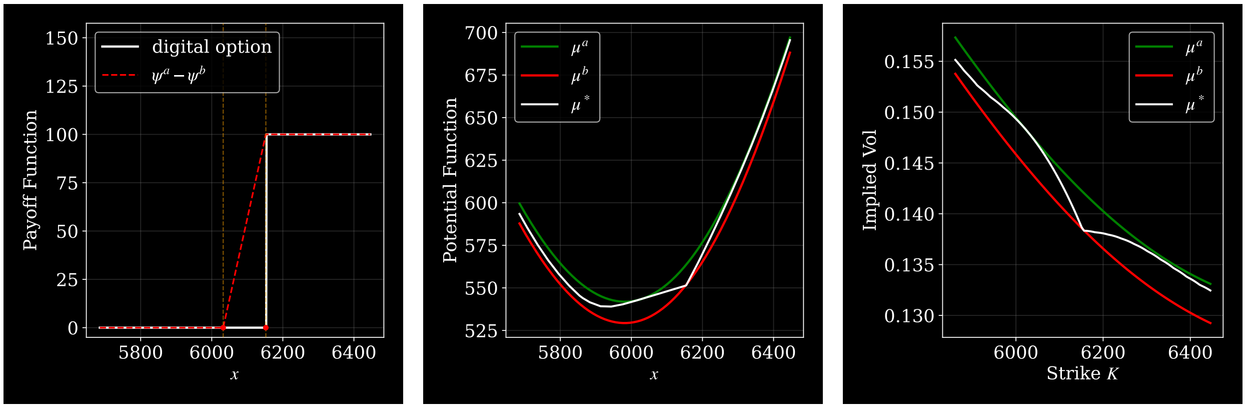


Table 1: Log-normal mixture parameters for μ^{a} and μ^{b} (values are kept up to 4 significant figures).

	Bid			**Ask**		
Index	**means**	**vols**	**weights**	**means**	**vols**	**weights**
1	6250	0.04008	0.09591	6250	0.04009	0.09591
2	6098	0.07222	0.5814	6098	0.07408	0.5814
3	5531	0.1293	0.2741	5531	0.1360	0.2741
4	4116	0.3150	0.04852	4116	0.3198	0.04852

Figure 6 presents a dual optimizer, the potential function and the implied volatility skew of a primal optimizer for the discretized BAMOT problem with $N = 1$ and payoff $h(x) = 100\,\mathbb{1}_{\{x \geq K\}}$ with $K = 1.05x_0$. Recall that the potential function of a probability measure $\mu \in \mathcal{P}_1$ is defined as $U_\mu(x) := \int |x - y|\,\mu(\mathrm{d}y)$, and that $U_\mu \leq U_\nu$ if $\mu \preceq_{\mathrm{c}} \nu$.

The computed superhedging price is 46.84, while the price obtained under the mid marginal is 37.14, resulting in a premium of $P(h) - \mu^{\mathrm{m}}(h) = 9.70$. This suggests that the common practice of pricing digital options using the mid marginal can substantially underestimate prices relative to BAMOT in the presence of bid–ask frictions.

The dual optimizer shown in Fig. 6a is a call spread with strikes $L^\star = 6032$ and $K = 1.05x_0 = 6154.05$, yielding a tradable optimal superhedging portfolio in the market. On the other hand, Fig. 6b illustrates that the potential function of the primal optimizer $\mu^\star$ is squeezed between the potential functions of the bid and ask marginals, thereby validating the feasibility of the solution. In addition, the potential function associated with $\mu^\star$ touches those of the ask and bid marginals at $L^\star, K$, respectively. This behavior is consistent with the structure of the digital call payoff, which is a step function that vanishes below the strike. Finally, Fig. 6c illustrates that the implied volatility skew of the primal optimizer lies within the region delineated by the bid and ask marginals and touches the bid and ask implied volatility skews at $L^\star, K$. To understand this, note that the potential function of $\mu^\star$ in Fig. 6b is linear between the lower strike $L^\star$ and K, implying that

$\mu^\star([L^\star, K]) = 0$. Hence, the digital call payoff and the dual optimizer $\psi^\star(x) = 100\frac{(x-L^\star)^+-(x-K)^+}{K-L^\star}$ have the same price under $\mu^\star$. Therefore,

$$P(h) = \mu^\star(h) = \mu^\star(\psi^\star) = 100 \times \frac{c^\star(L^\star,T) - c^\star(K,T)}{K - L^\star} \leq 100 \times \frac{c^{\mathrm{a}}(L^\star,T) - c^{\mathrm{b}}(K,T)}{K - L^\star} = D(h).$$

By strong duality, we conclude that $c^\star(L^\star,T) = c^{\mathrm{a}}(L^\star,T)$ and $c^\star(K,T) = c^{\mathrm{b}}(K,T)$, and the same holds for the corresponding implied volatilities.

5.2 Forward Start Option

In this section, we compare the price bounds computed under BAMOT with those obtained from MOT. Consider the case $N = 2$ and take the forward-start payoff

$$h(x_1, x_2) = (x_2 - Kx_1)^+.$$

Observe that $|x_2 - x_1| = 2(x_2 - x_1)^+ - (x_2 - x_1)$, and hence $\mathbb{E}_{\mathbb{Q}}[|X_2 - X_1|] = 2\mathbb{E}_{\mathbb{Q}}[(X_2 - X_1)^+]$ for any martingale coupling $\mathbb{Q} \in \mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$, where $\mu^{\mathrm{a}} = (\mu_1^{\mathrm{a}}, \mu_2^{\mathrm{a}})$ and $\mu^{\mathrm{b}} = (\mu_1^{\mathrm{b}}, \mu_2^{\mathrm{b}})$. Assume zero interest rates, we initialize the bid–ask marginals $\mu_i^{\mathrm{a}}, \mu_i^{\mathrm{b}}$, $i = 1, 2$, from the Black–Scholes model with parameters

$$(T_1, \sigma_1^{\mathrm{b}}) = (0.5, 0.19), \quad (T_1, \sigma_1^{\mathrm{a}}) = (0.5, 0.20), \quad (T_2, \sigma_2^{\mathrm{b}}) = (1, 0.17), \quad (T_2, \sigma_2^{\mathrm{a}}) = (1, 0.18),$$

whose density and potential functions are presented in Fig. 7a and Fig. 7b, respectively.

Fix a collection of call strikes $\{K_m\}_{m\in[M]}$ available at both $T_i, i = 1, 2$. The dual problem $D(h)$ then leads to the following discretized optimization (see Appendix C for details):

$$\begin{aligned}
\inf \quad & a + \sum_{i=1}^{2} b_i x_0 + \sum_{i=1}^{2}\sum_{m=1}^{M} \Big(c_{i,m}^{\mathrm{a}} c^{\mathrm{a}}(K_m, T_i) - c_{i,m}^{\mathrm{b}} c^{\mathrm{b}}(K_m, T_i) \Big) \\
\text{such that} \quad & a + \sum_{i=1}^{2} b_i x_i + \sum_{i=1}^{2}\sum_{m=1}^{M} (c_{i,m}^{\mathrm{a}} - c_{i,m}^{\mathrm{b}})(x_i - K_m)^+ \\
& \quad + \Delta(x_1)(x_2 - x_1) \;\geq\; (x_2 - Kx_1)^+, \qquad (x_1, x_2) \in (\mathbb{R}_+)^2, \\
& a, b_i \in \mathbb{R}, \qquad c_{i,m}^{\mathrm{a}}, c_{i,m}^{\mathrm{b}} \geq 0, \qquad i = 1, 2,\; m \in [M].
\end{aligned}$$

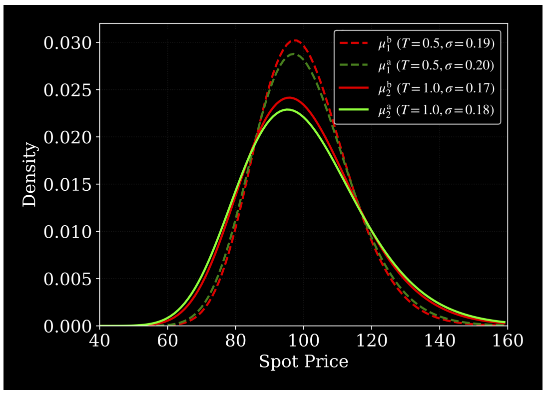
We compute and compare the super- and subhedging prices under BAMOT and MOT using mid marginals, which are constructed by averaging the bid and ask marginals, for $K \in \{0.8, 0.85, \ldots, 1.2\}$, and present the results in Fig. 7. As illustrated in Fig. 7c and Fig. 7d, BAMOT produces a wider price range than MOT. The discrepancy is most pronounced near $K = 1$ and diminishes as K approaches the boundary values.

5.3 Convergence Rates

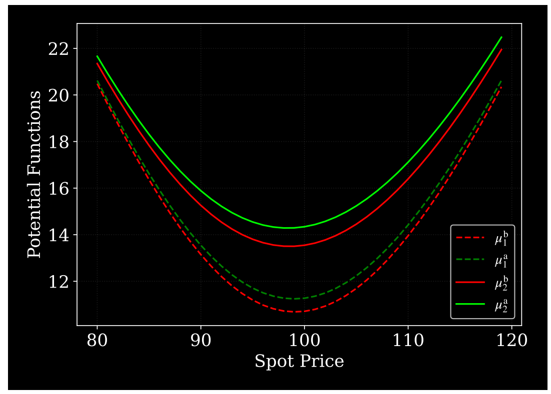
In this section, we present two results on the convergence rate of BAMOT to classical MOT when the payoff is either a risk-reversal strategy or an at-the-money (ATM) digital option. We show that the former exhibits a linear rate of convergence, while the latter admits a square-root rate.

Figure 7: (a)–(b): The left shows the density functions of $\mu_1^{\mathrm{b}}, \mu_1^{\mathrm{a}}, \mu_2^{\mathrm{b}}, \mu_2^{\mathrm{a}}$, and the right displays the call prices across strikes, confirming Assumption 2.1. (c)–(d): Super- and subhedging bounds and their differences for $h(x_1, x_2) = (x_2 - Kx_1)^+$ with $K \in \{0.8, 0.85, \ldots, 1.15, 1.2\}$, given strikes $\{60, 65, 70, \ldots, 140\}$ at maturities T_1, T_2, spot $x_0 = 100$ and zero interest rates.

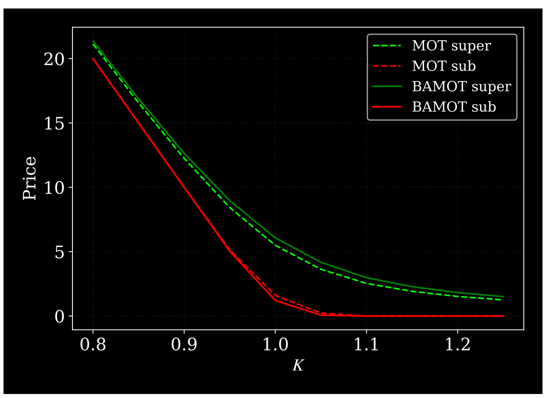
(a) Densities of $\mu_1^{\mathrm{b}}, \mu_1^{\mathrm{a}}, \mu_2^{\mathrm{b}}, \mu_2^{\mathrm{a}}$



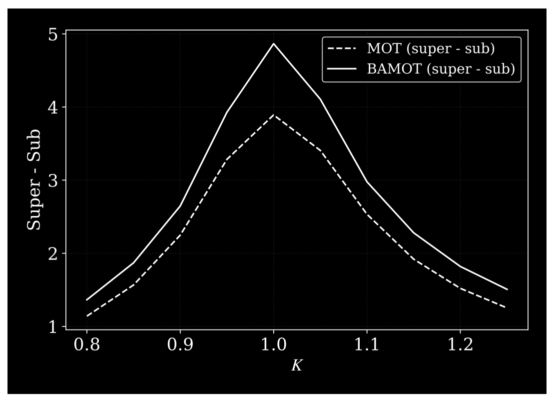
(b) Potential functions



(c) Super- and subhedging bounds



(d) Differences of robust bounds

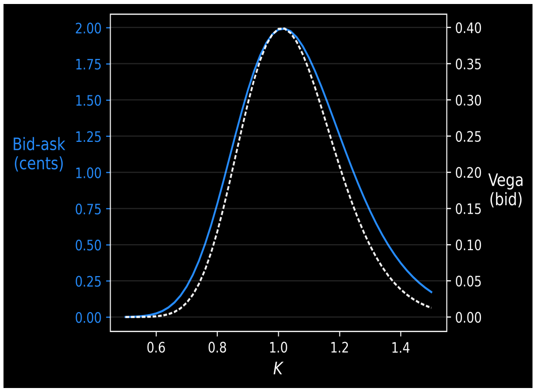


For both payoffs, we consider the setting when μ^{b} and μ^{a} are the one-year marginals, i.e., $T = 1$, in the Black–Scholes model with zero interest rates and volatilities $\sigma^{\mathrm{b}} = 15\%$ and $\sigma^{\mathrm{a}} = 20\%$, respectively. We set the spot price $x_0 = 1$ for simplicity. We then continuously deform the bid and ask marginals toward the mid marginal $\mu^{\mathrm{m}} = \frac{\mu^{\mathrm{b}}+\mu^{\mathrm{a}}}{2}$ by setting $\mu^{\mathrm{s},\gamma} = (1-\gamma)\mu^{\mathrm{s}} + \gamma\mu^{\mathrm{m}}$, $\mathrm{s} \in \{\mathrm{b}, \mathrm{a}\}$, with $\gamma \in [0, 1]$ increasing to one.

The bid–ask distances are obtained by maximizing the bid–ask spread among call options; see Remark 4.4 (a). By a first-order Taylor expansion of the Black–Scholes price $c_{\mathrm{BS}}(K, T, \sigma)$ with respect to σ, the bid–ask spread of call options is maximized where the Vega is large, namely near the money ($K \approx 1$). This behavior is confirmed in the left panel of Fig. 8, showing the spread for the initial Black–Scholes marginals $\mu^{\mathrm{b}}, \mu^{\mathrm{a}}$.

For the risk-reversal payoff $h(x) = (x - 1.05)^+ - (0.95 - x)^+$, which is a difference of two convex functions, Proposition 4.5 implies a linear convergence rate. The right panel of Figure 8 displays the payoff together with an optimal dominating profile $\psi^{\mathrm{a},\star} - \psi^{\mathrm{b},\star}$, and Fig. 9 confirms the linear rate of convergence. On the other hand, for the ATM digital option $h(x) = \mathbb{1}_{\{x \geq 1\}}$, Proposition 4.6 predicts a square-root convergence rate, and Fig. 10 illustrates that this rate is sharp.

Figure 8: Left: Bid–ask spread of one-year call options in the Black–Scholes model with $\sigma^{\mathrm{b}} = 15\%$ and $\sigma^{\mathrm{a}} = 20\%$. Comparison with Black–Scholes Vega based on bid prices. Right: Risk-reversal strategy $h(x) = (x - 1.05)^+ - (0.95 - x)^+$ and optimal dominating profile.



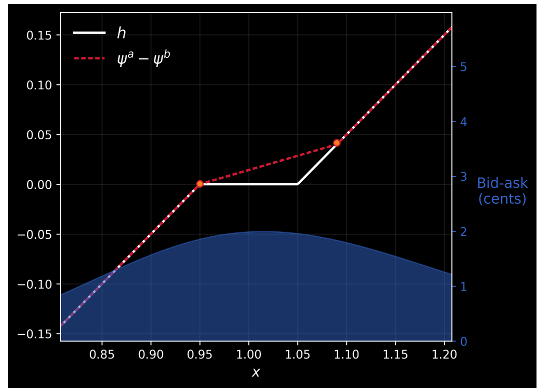
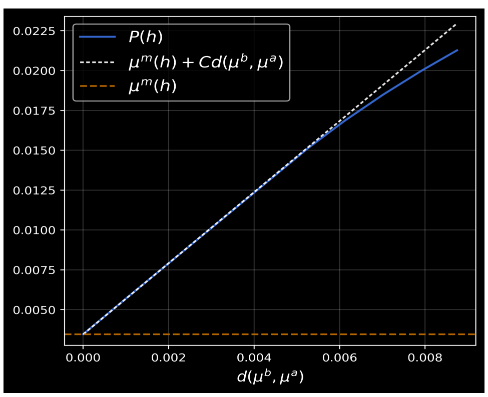


Figure 9: BAMOT convergence for a 95% − 105% risk-reversal strategy.

(a) Superhedging price $P(h)$



(b) Premium $P(h) - \mu^{\mathrm{m}}(h)$

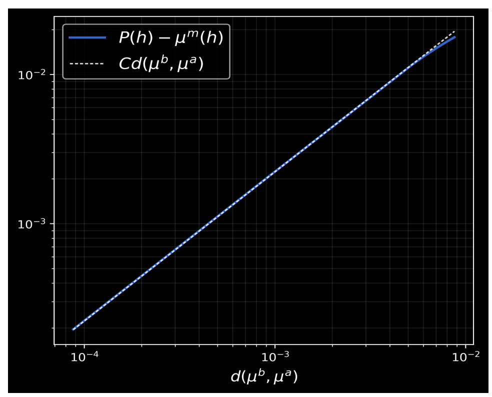
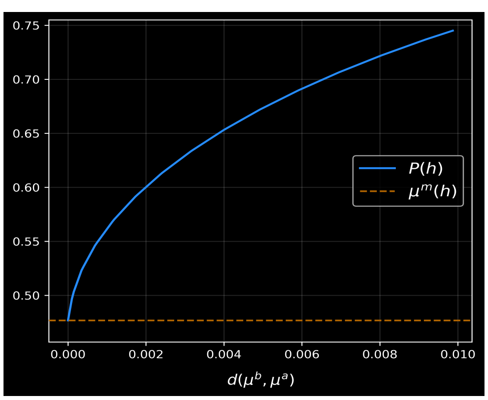
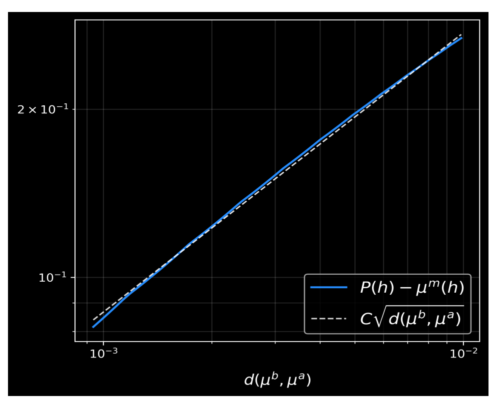


Figure 10: BAMOT convergence for an at-the-money digital option.

(a) Superhedging price $P(h)$



(b) Premium $P(h) - \mu^{\mathrm{m}}(h)$



6 Conclusion

In this work, we propose BAMOT as a practical and mathematically tractable framework for robust pricing that incorporates bid–ask spreads on vanilla options. We establish strong duality for upper semicontinuous payoff functions and convergence to classical MOT as the bid–ask spreads vanish. This opens the door to extending numerous classical MOT results to the bid–ask setting, including strong duality for more general payoffs, dual attainment and regularity of dual optimizers, and the stability under general perturbations of the bid and ask marginals. Another line of research could extend this framework to a continuous-time setting [16, 24], either by incorporating continuous price dynamics and delta hedging, or by allowing option quotes across a continuum of maturities.

A Construction of Bid and Ask Marginals from Enhanced Quotes

As noted in Section 2.1, the existence of bid and ask marginals is implicitly assumed in practice in order to produce arbitrage-free implied volatility skews. A natural question is whether these marginals always exist, which is the focus of this section. Specifically, we present a procedure for enhancing market quotes by replacing them with the "best" available super- and subhedging prices, which enables exact calibration of the ask marginals; see Fig. 11. On the other hand, our findings suggest that marginal distributions can only produce a best-fit approximation of bid market quotes.

Consider a strip of available strikes $K_0 < \ldots < K_M$ for fixed maturity T, and corresponding bid and ask quotes

$$p_m^{\mathrm{b}} \leq p_m^{\mathrm{a}}, \quad c_m^{\mathrm{b}} \leq c_m^{\mathrm{a}},$$

of put and call options struck at K_m, $m = 0, \ldots, M$. Suppose that $K_0 = 0$, with $p_0^b = p_0^a = 0$ and c_0^b, c_0^a equal to the $T-$forward price F. If the bid or ask price is unavailable for some strike, we impute it according to

$$p_m^{\mathrm{b}} = 0, \quad p_m^{\mathrm{a}} = K_m, \quad c_m^{\mathrm{b}} = 0, \quad c_m^{\mathrm{a}} = F,$$

using the bounds $(K_m - X_T)^+ \leq K_m$ and $(X_T - K_m)^+ \leq X_T$ for puts and calls, respectively. Assume the market quotes are free of static arbitrage, ensuring that the family of calibrated models

$$\mathcal{Q}_M^{\mathrm{b,a}} = \big\{\mathbb{Q} \mid \text{martingale measure}, \, v_m^{\mathrm{b}} \leq v_m^{\mathbb{Q}} \leq v_m^{\mathrm{a}} \ \forall\, m\big\}, \qquad \begin{cases} v_m^{\mathbb{Q}} = \mathbb{E}_{\mathbb{Q}}[(K_m - X_T)^+], & v_m^{\mathrm{s}} = p_m^{\mathrm{s}}, \\ v_m^{\mathbb{Q}} = \mathbb{E}_{\mathbb{Q}}[(X_T - K_m)^+], & v_m^{\mathrm{s}} = c_m^{\mathrm{s}}, \end{cases}$$

for $s \in \{\mathrm{b}, \mathrm{a}\}$, is non-empty. This section explains how $\mathcal{Q}_M^{\mathrm{b,a}}$ relates to the admissible set $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ in (10) with $N = 1$. Specifically, we construct bid and ask marginals $\mu^{\mathrm{b}}, \mu^{\mathrm{a}}$ such that the latter precisely recovers the ask quotes following a novel quotes enhancement procedure. A similar procedure is applied to the bid prices, facilitating the calibration of the bid marginal to market quotes.

A.1 Combining Put and Call Options

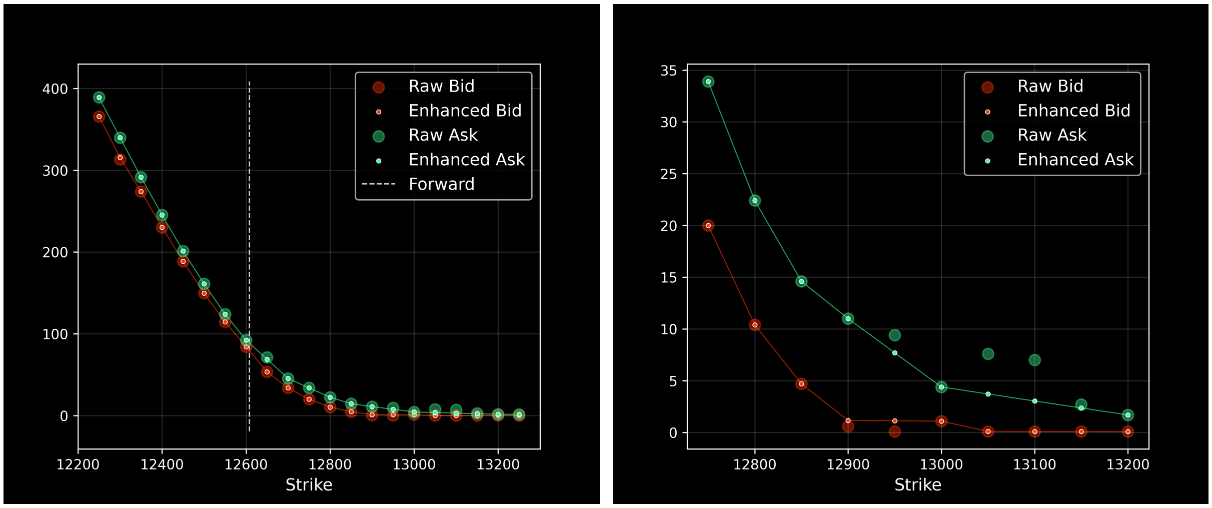
Recall that $c_0^{\mathrm{b}} = c_0^{\mathrm{a}} = F$, i.e., the forward contract has zero bid–ask spread. Then in view of put-call parity, we can replace the call quotes by

$$c_m^{\mathrm{b}} \leftarrow \max\big(c_m^{\mathrm{b}}, \, p_m^{\mathrm{b}} + F - K_m\big), \quad c_m^{\mathrm{a}} \leftarrow \min\big(c_m^{\mathrm{a}}, \, p_m^{\mathrm{a}} + F - K_m\big).$$

Then the put quotes no longer carry any additional information and can be discarded.

Figure 11: Quotes enhancement of 10-17-2025 call options on the Swiss Market Index, as of 10-09-2025.

(a) Global view (b) Zoom-in View



A.2 Quotes Enhancement

Each quote may be further enhanced using other call options as hedging instruments. To wit, consider the following super- and subhedging problems,

$$\underline{c}^{\mathrm{a}}_m = \inf_{\lambda \in \Lambda^+_m} \mathrm{cost}^+(\lambda), \quad \Lambda^+_m = \Big\{(\lambda^{\mathrm{a}}, \lambda^{\mathrm{b}}) \in (\mathbb{R}^{M+1}_+)^2 \;\Big|\; \sum_{\ell=0}^{M} (\lambda^{\mathrm{a}}_\ell - \lambda^{\mathrm{b}}_\ell)(x - K_\ell)^+ \geq (x - K_m)^+ \;\; \forall\, x \geq 0\Big\},$$

$$\overline{c}^{b}_m = \sup_{\lambda \in \Lambda^-_m} \mathrm{cost}^-(\lambda), \quad \Lambda^-_m = \Big\{(\lambda^{\mathrm{b}}, \lambda^{\mathrm{a}}) \in (\mathbb{R}^{M+1}_+)^2 \;\Big|\; \sum_{\ell=0}^{M} (\lambda^{\mathrm{b}}_\ell - \lambda^{\mathrm{a}}_\ell)(x - K_\ell)^+ \leq (x - K_m)^+ \;\; \forall\, x \geq 0\Big\},$$

with the cost functions $\mathrm{cost}^\pm(\lambda) = \pm \sum_{\ell=0}^{M} (\lambda^{\mathrm{a}}_\ell c^{\mathrm{a}}_\ell - \lambda^{\mathrm{b}}_\ell c^{\mathrm{b}}_\ell)$. By construction, we have that $c^{\mathrm{b}}_m \leq \overline{c}^{\mathrm{b}}_m$ and $\underline{c}^{\mathrm{a}}_m \leq c^{\mathrm{a}}_m$ for all m. Hence, the above quotes enhancement procedure generates tighter bid–ask spreads. Fig. 11 illustrates this for vanilla options on the Swiss Market Index (SMI) expiring on 10-17-2025, as of 10-09-2025. As can be seen, the bid and ask enhanced quotes do not cross and become non-increasing in strike. Moreover, while the bid quotes exhibit mixed convexity—being convex in some regions and concave in others—the ask quotes appear to be convex in the strike. We formalize these empirical observations in the next result.

Proposition A.1. *The enhanced bid and ask quotes satisfy the following properties:*

(i) (Consistency) $\overline{c}^{\mathrm{b}}_m \leq \underline{c}^{\mathrm{a}}_m$ *for all* m.

(ii) (Monotonicity) $F = \underline{c}^{\mathrm{a}}_0 \geq \ldots \geq \underline{c}^{\mathrm{a}}_M$*, and* $F = \overline{c}^{\mathrm{b}}_0 \geq \ldots \geq \overline{c}^{\mathrm{b}}_M$.

(iii) (Convexity, ask quotes) For all $m_- < m < m_+$,

$$\underline{c}^{\mathrm{a}}_m \leq \gamma \underline{c}^{\mathrm{a}}_{m_-} + (1-\gamma)\underline{c}^{\mathrm{a}}_{m_+} \quad \text{when} \quad K_m = \gamma K_{m_-} + (1-\gamma) K_{m_+}. \tag{27}$$

Proof. *(i)* Suppose that $\bar{c}^{\mathrm{b}}_m > \underline{c}^{\mathrm{a}}_m$ for some $m \leq M$. Then there exists $\lambda^{\pm} \in \Lambda^{\pm}_m$ such that $\mathrm{cost}^-(\lambda^-) > \mathrm{cost}^+(\lambda^+)$. Consequently,

$$\sum_{\ell=0}^{M}(\lambda^{\mathrm{b},-}_\ell - \lambda^{\mathrm{a},-}_\ell)(x-K_\ell)^+ \leq (x-K_m)^+ \leq \sum_{\ell=0}^{M}(\lambda^{\mathrm{a},+}_\ell - \lambda^{\mathrm{b},+}_\ell)(x-K_\ell)^+.$$

Then the portfolio associated with $\lambda := (\lambda^{\mathrm{a},+}+\lambda^{\mathrm{a},-}, \lambda^{\mathrm{b},+}+\lambda^{\mathrm{b},-})$ yields a nonnegative payoff for a negative price, contradicting our assumption that the options market is arbitrage-free.

(ii) Fix $\ell < m$. As $(x-K_\ell)^+ \geq (x-K_m)^+$, any superhedging portfolio of $(x-K_\ell)^+$ also superhedges $(x-K_m)^+$. Hence $\Lambda^+_\ell \subseteq \Lambda^+_m$ and $\underline{c}^{\mathrm{a}}_\ell \geq \underline{c}^{\mathrm{a}}_m$ follows. Similarly, $\Lambda^-_m \subseteq \Lambda^-_\ell$, which yields $\bar{c}^{\mathrm{b}}_\ell \geq \bar{c}^{\mathrm{b}}_m$.

(iii) For any $\varepsilon > 0$, let $\lambda_{m_-} \in \Lambda^+_{m_-}$, $\lambda_{m_+} \in \Lambda^+_{m_+}$ be $\frac{\varepsilon}{2}$-approximate optimizers to $\underline{c}^{\mathrm{a}}_{m_-}$ and $\underline{c}^{\mathrm{a}}_{m_+}$, respectively. Since $K_m = \gamma K_{m_-} + (1-\gamma)K_{m_+}$, we have $\gamma(x-K_{m_-})^+ + (1-\gamma)(x-K_{m_+})^+ \geq (x-K_m)^+$ by the convexity of $K \mapsto (x-K)^+$. Thus, $\lambda_m := \gamma\lambda_{m_-} + (1-\gamma)\lambda_{m_+} \in \Lambda^+_m$. Therefore,

$$\underline{c}^{\mathrm{a}}_m \leq \mathrm{cost}^+(\lambda_m) = \gamma \mathrm{cost}^+(\lambda_{m_-}) + (1-\gamma)\mathrm{cost}^+(\lambda_{m_+}) \leq \gamma \underline{c}^{\mathrm{a}}_{m_-} + (1-\gamma)\underline{c}^{\mathrm{a}}_{m_+} + \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, the result follows.

□

From Appendix A.1 and the consistency property in Proposition A.1, the set of calibrated models can be equivalently described as

$$\mathcal{Q}^{\mathrm{b,a}}_M = \{\mathbb{Q} \mid \text{martingale measure}, \bar{c}^{\mathrm{b}}_m \leq c^{\mathbb{Q}}_m \leq \underline{c}^{\mathrm{a}}_m \ \ \forall\, m\}.$$

The quote enhancement mechanism described above ensures the existence of an ask marginal that precisely recovers the filtered ask quotes. A construction is provided in the next section for completeness, keeping in mind that the parametric approach in Section 2.1 yields more realistic implied volatility skews and is thus preferred in practice. Moreover, the following approach shows the existence of a generalized bid marginal, potentially involving negative mass, see Remark A.3.

A.3 Existence of an Ask Marginal

Let $\underline{M} = \sup\{1 \leq m \leq M \mid \underline{c}^{\mathrm{a}}_m < \underline{c}^{\mathrm{a}}_{m-1}\}$.[5] When $\underline{M} < M$, then the enhanced ask quotes $\underline{c}^{\mathrm{a}}_m = \underline{c}^{\mathrm{a}}_{\underline{M}}$ for all $m > \underline{M}$, meaning that the cheapest superhedging of the call at K_m is given by the call at $K_{\underline{M}}$. We can therefore truncate the quotes to $\underline{c}^{\mathrm{a}}_0 \geq \ldots \geq \underline{c}^{\mathrm{a}}_{\underline{M}}$, as rational agents would always prefer $(X_T - K_{\underline{M}})^+$ over $(X_T - K_m)^+$ for $m > \underline{M}$, or be indifferent if the support of the risk-neutral distribution is contained in $[0, K_{\underline{M}}]$.

Write M in lieu of $\underline{M}$. By the maximality of M, we have $\underline{c}^{\mathrm{a}}_M < \underline{c}^{\mathrm{a}}_{M-1}$, hence $\partial^+\underline{c}^{\mathrm{a}}(K_M) := \frac{\underline{c}^{\mathrm{a}}_M - \underline{c}^{\mathrm{a}}_{M-1}}{K_M - K_{M-1}} < 0$. Let $K_{M+1} := K_M - \underline{c}^{\mathrm{a}}_M/\partial^+\underline{c}^{\mathrm{a}}(K_M) \geq K_M$ and set $\underline{c}^{\mathrm{a}}_{M+1} = 0$. Define

$$\partial^+\underline{c}^{\mathrm{a}}(K) := \sum_{m=0}^{M} \mathbb{1}_{[K_m, K_{m+1})}(K)\frac{\underline{c}^{\mathrm{a}}_{m+1} - \underline{c}^{\mathrm{a}}_m}{K_{m+1} - K_m}.$$

[5] If the set is empty, then $\underline{c}^{\mathrm{a}}_0 = \ldots = \underline{c}^{\mathrm{a}}_M = F$, and also $c^{\mathrm{a}}_0 = \ldots = c^{\mathrm{a}}_M = F$, leading to a degenerate market.

Then $\partial^+\underline{c}^{\mathrm{a}} = 0$ beyond K_{M+1}, and the convexity property (27) implies that

$$\underline{c}^{\mathrm{a}}(K) := F + \int_0^K \partial^+\underline{c}^{\mathrm{a}}(L)\,\mathrm{d}L, \qquad \underline{c}^{\mathrm{a}}(0) = F, \qquad \forall\, K \geq 0,$$

is a piecewise linear convex function that interpolates the enhanced ask quotes $\{\underline{c}^{\mathrm{a}}_m\}_{m\in[M]}$. We then introduce the measure

$$\mu^{\mathrm{a}}((L,K]) := \partial^+\underline{c}^{\mathrm{a}}(K) - \partial^+\underline{c}^{\mathrm{a}}(L),\ L \leq K, \quad \mu^{\mathrm{a}}(\{0\}) = 1 + \partial^+\underline{c}^{\mathrm{a}}(0) < 1, \tag{28}$$

which is nonnegative as $\partial^+\underline{c}^{\mathrm{a}}$ is nondecreasing in strike. We finally show that μ^{a} is a probability measure and matches all enhanced ask quotes, thus corresponding to a feasible ask measure.

Proposition A.2. *The construction μ^{a} in* (28) *is a discrete probability measure on $0 = K_0 < \ldots < K_{M+1}$, such that $\mathbb{E}_{\mu^{\mathrm{a}}}[(X-K_m)^+] = \underline{c}^{\mathrm{a}}_m$ for all m. In particular, its barycenter coincides with the forward $F = \underline{c}^{\mathrm{a}}_0$.*

Proof. Clearly, μ^{a} is supported on $\{K_m\}_{m=0}^{M+1}$. Moreover, its total mass is given by

$$\mu^{\mathrm{a}}(\mathbb{R}_+) = \mu^{\mathrm{a}}(\{0\}) + \mu^{\mathrm{a}}((0,K_{M+1}]) = 1 + \partial^+\underline{c}^{\mathrm{a}}(0) + \partial^+\underline{c}^{\mathrm{a}}(K_{M+1}) - \partial^+\underline{c}^{\mathrm{a}}(0) = 1.$$

Since $\partial^+\underline{c}^{\mathrm{a}}$ is nondecreasing and $\partial^+\underline{c}^{\mathrm{a}}(0) \geq -1$ (noting that $\underline{c}^{\mathrm{a}}_1 - \underline{c}^{\mathrm{a}}_0 \geq (F-K_1)^+ - F \geq -K_1$ and $K_0 = 0$), μ^{a} is nonnegative. Hence μ^{a} is a probability measure. Finally, by construction and integration by parts,

$$\mathbb{E}_{\mu^{\mathrm{a}}}[(X-K_m)^+] = \int_{K_m}^{K_{M+1}} (x-K_m)\mu^{\mathrm{a}}(\mathrm{d}x) = -\int_{K_m}^{K_{M+1}} \partial^+\underline{c}^{\mathrm{a}}(x)dx = \underline{c}^{\mathrm{a}}(K_m) = \underline{c}^{\mathrm{a}}_m.$$

□

Remark A.3. A parallel construction applies to the enhanced bid quotes, leading to an exact fit. However, because the curve formed by the enhanced bid quotes is not necessarily convex, the resulting object is generally a *signed* measure. The latter nonetheless yields nonnegative prices for all call options and by extension, for all nonnegative convex payoffs. This points toward a generalization of convex order to signed measures, a subject we leave for future research.

B Technical Results and Proofs

The first two lemmas are standard results related to the convex order. Given $\mu \in \mathcal{P}_1$, we introduce the sublevel set

$$\underline{\mathcal{Q}}(\mu) := \left\{\mu' \in \mathcal{P}_1 \,\middle|\, \mu' \preceq_{\mathrm{c}} \mu\right\}.$$

Observe that $\underline{\mathcal{Q}}(\mu) \subseteq \mathcal{P}_1(A)$ whenever $A := \mathrm{supp}(\mu)$ is an interval in $\mathbb{R}_+$.

Lemma B.1. *For all $f \in \mathcal{C}_L$, the functional $\underline{\mathcal{Q}}(\mu^{\mathrm{a}}) \ni \mu \mapsto \int f\,\mathrm{d}\mu$ is weakly continuous.*

Proof. Let $\underline{\mathcal{Q}}(\mu^{\mathrm{a}}) \ni \mu_n \overset{w}{\to} \mu$, then for any $R > 0$, we have

$$\left|\int f\,\mathrm{d}\mu_n - \int f\,\mathrm{d}\mu\right| \leq \left|\int_{[0,R]} f\,\mathrm{d}(\mu_n - \mu)\right| + \left|\int_{\mathbb{R}_+\setminus[0,R]} f\,\mathrm{d}\mu_n\right| + \left|\int_{\mathbb{R}_+\setminus[0,R]} f\,\mathrm{d}\mu\right|.$$

The first term converges to 0 by the definition of weak convergence. We now control the second term. Since $|f(x)| \leq \|f\|_L(1+x)$ and $x\mathbb{1}_{\{x>R\}} \leq 2(x-R/2)^+$ for all $x \in \mathbb{R}_+$,

$$\begin{aligned}\left|\int_{\mathbb{R}_+\setminus[0,R]} f \,\mathrm{d}\mu_n\right| &\leq \|f\|_L \int_{\mathbb{R}_+\setminus[0,R]} (1+x)\,\mu_n(\mathrm{d}x) \\ &\leq \|f\|_L\, \mu_n\left(\mathbb{R}_+ \setminus [0,R]\right) + \|f\|_L \int 2(x-R/2)^+\mu_n(\mathrm{d}x) \\ &\leq \frac{\|f\|_L}{R}\int x\mu^{\mathrm{a}}(\mathrm{d}x) + 2\|f\|_L \int (x-R/2)^+\mu^{\mathrm{a}}(\mathrm{d}x),\end{aligned}$$

where the last line follows from Markov's inequality and the convexity of $x \mapsto (x-R/2)^+$, $x \mapsto x$. Since $(x-R/2)^+ \to 0$ pointwise as $R \to \infty$, $(x-R/2)^+ \leq x$ and $\mu^{\mathrm{a}} \in \mathcal{P}_1$, we conclude by the dominated convergence theorem that

$$\limsup_{R\to\infty} \sup_n \left|\int_{\mathbb{R}_+\setminus[0,R]} f\,\mathrm{d}\mu_n\right| = 0.$$

The third term can be controlled using similar arguments. □

Lemma B.2. *Let $\mu^{\mathrm{a}}, \mu^{\mathrm{b}} \in \mathcal{P}_1$ satisfy $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. Then the sets $\underline{\mathcal{Q}}(\mu^{\mathrm{a}})$ and $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ are convex and compact with respect to $\mathcal{W}_1$.*

Proof. Convexity is easily verified. Fix a sequence (μ_n) in $\underline{\mathcal{Q}}(\mu^{\mathrm{a}})$ converging weakly to some limit μ. By (7) and Lemma B.1, we derive that $\mu \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. Thus, $\underline{\mathcal{Q}}(\mu^{\mathrm{a}})$ is weakly closed.

Next, we show that (μ_n) contains a subsequence that converges weakly to some $\mu \in \underline{\mathcal{Q}}(\mu^{\mathrm{a}})$. First, take $\psi(x) = |x|$, and observe that $\sup_{\mu\in\underline{\mathcal{Q}}(\mu^{\mathrm{a}})} \mu(\psi) = \mu^{\mathrm{a}}(\psi)$. Hence, applying Markov's inequality implies the tightness of (μ_n). By Prokhorov's theorem (see, e.g., [9, Theorem 8.6.2]) and by the weak closedness of $\underline{\mathcal{Q}}(\mu^{\mathrm{a}})$, there exists a subsequence (μ_n), up to re-indexing, such that $\mu_n \overset{w}{\to} \mu \in \underline{\mathcal{Q}}(\mu^{\mathrm{a}})$. Moreover, the mapping $\underline{\mathcal{Q}}(\mu^{\mathrm{a}}) \ni \mu \to \mu(\psi)$ is weakly continuous by Lemma B.1, implying that $\mu_n \to \mu$ with respect to $\mathcal{W}_1$.

Moving on to the general case, note that $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}}) \subseteq \underline{\mathcal{Q}}(\mu^{\mathrm{a}})$ for all $\mu^{\mathrm{b}} \preceq_{\mathrm{c}} \mu^{\mathrm{a}}$. It is thus sufficient to show the closedness of $\mathcal{Q}(\mu^{\mathrm{b}}, \mu^{\mathrm{a}})$ under $\mathcal{W}_1$-topology, which follows from the Kantorovich–Rubinstein theorem [37, Theorem 1.14] and (7). □

Proof of Proposition 4.3. (i) If $\mu \preceq_{\mathrm{c}} \nu$, then clearly $\vec{d}(\mu,\nu) \leq 0$. Moreover, choosing $\psi(x) = x$ yields

$$0 \geq \vec{d}(\mu,\nu) \geq \mathbb{E}_\mu[X] - \mathbb{E}_\nu[X] = 0,$$

recalling that measures in convex order have identical barycenters. Conversely, suppose that $\vec{d}(\mu,\nu) = 0$. Then $\mathbb{E}_\mu[X] = \mathbb{E}_\nu[X]$ and $\mathbb{E}_\mu[(K-X)^+] \leq \mathbb{E}_\nu[(K-X)^+]\ \forall\ K \geq 0$, so we conclude from (7) that $\mu \preceq_{\mathrm{c}} \nu$.

(ii) The triangle inequality is clear. Also, $\vec{d}$ is nonnegative since $\psi \equiv 0$ is convex and 1-Lipschitz. Finally, suppose that $\vec{d}(\mu,\nu) = \vec{d}(\nu,\mu) = 0$. Then $\mu \preceq_{\mathrm{c}} \nu$ and $\nu \preceq_{\mathrm{c}} \mu$, which implies that $\mu = \nu$.

(iii) Let $\vec{d}_C(\mu,\nu) := \sup\{\mathbb{E}_\mu[(X-K)^+] - \mathbb{E}_\nu[(X-K)^+] \mid K \geq 0\}$. Let ψ convex with $\mathrm{Lip}(\psi) \leq 1$. From Carr–Madan's formula, ψ can be decomposed as

$$\psi(x) = \psi(0) + \psi'(0)x + \int_{\mathbb{R}_+} (x-K)^+\lambda^\psi(\mathrm{d}K), \tag{29}$$

where ψ' is the right derivative of ψ and λ^ψ the nonnegative Radon measure on $\mathbb{R}_+$ induced by ψ'. Note that since $\mathrm{Lip}(\psi) \leq 1$ and ψ' is nondecreasing, we have $\lambda^\psi(\mathbb{R}_+) \leq 2$. In addition, $(\mu-\nu)(1) = (\mu-\nu)(x) = 0$ as $\mu, \nu \in \mathcal{P}_1$ with identical barycenters. Therefore, by (29), we have

$$(\mu-\nu)(\psi) = \int \left(\int (x-K)^+ \lambda^\psi(\mathrm{d}K) \right) (\mu-\nu)(\mathrm{d}x). \tag{30}$$

Applying Fubini's theorem to (30) yields

$$\begin{aligned}
(\mu-\nu)(\psi) &= \int_{\mathbb{R}_+} \left(\mathbb{E}_\mu[(X-K)^+] - \mathbb{E}_\nu[(X-K)^+] \right) \lambda^\psi(\mathrm{d}K) \\
&\leq \sup_{K \geq 0} \left(\mathbb{E}_\mu[(X-K)^+] - \mathbb{E}_\nu[(X-K)^+] \right) \lambda^\psi(\mathbb{R}_+) \\
&\leq 2\vec{d}_C(\mu,\nu).
\end{aligned}$$

Next, we show that $2\vec{d}_C(\mu,\nu) \leq \vec{d}(\mu,\nu)$. Fix $K \geq 0$, take $\psi(x) = |x-K|$. Then, using $|x-K| = 2(x-K)^+ + (K-x)$, we derive that

$$\vec{d}(\mu,\nu) \geq (\mu-\nu)(\psi) = 2\left(\mathbb{E}_\mu[(X-K)^+] - \mathbb{E}_\nu[(X-K)^+] \right).$$

Taking supremum over $K \geq 0$ on both sides, we conclude that $\vec{d}(\mu,\nu) \geq 2\vec{d}_C(\mu,\nu)$, as desired.

(iv) Recall that $\mathcal{W}_1(\mu,\nu) = \sup\{(\mu-\nu)(\varphi) \mid \|\varphi\|_{\mathrm{Lip}} \leq 1\}$ by Kantorovich–Rubinstein theorem. Then $\vec{d}(\mu,\nu) \leq \mathcal{W}_1(\mu,\nu)$ for all $\mu,\nu \in \mathcal{P}_1$, hence $d(\mu,\nu) \leq \mathcal{W}_1(\mu,\nu)$. Let us move on to the second statement. For ease of presentation, we construct sequences or discrete probability measures on $\mathbb{R}$ satisfying the desired property. Given their compact support, these measures can be shifted to $\mathbb{R}_+$, leading to valid sequences in $\mathcal{P}_1$. Let $\xi = \frac{1}{2}(\delta_{-1} + \delta_1)$ be the Rademacher distribution. For each $n \geq 1$, introduce

$$\mu_n = c_n \sum_{|m| \leq n} \delta_{2m}, \; c_n = \frac{1}{2n+1}, \quad \nu_n = \mu_n * \xi := c_n \sum_{|m| \leq n} \frac{\delta_{2m-1} + \delta_{2m+1}}{2}.$$

Observe that ν_n is supported on the odd integers $-2n-1, -2n+1, \ldots, 2n+1$, with probability c_n except at the endpoints where the probability is halved. We first verify that $\mathcal{W}_1(\mu_n,\nu_n) = 1$ irrespective of n. Indeed, consider the butterfly spreads portfolio

$$\varphi(x) = \sum_{|m| \leq n} B(x, 2m), \quad B(x,K) = (x-K+1)^+ - 2(x-K)^+ + (x-K-1)^+.$$

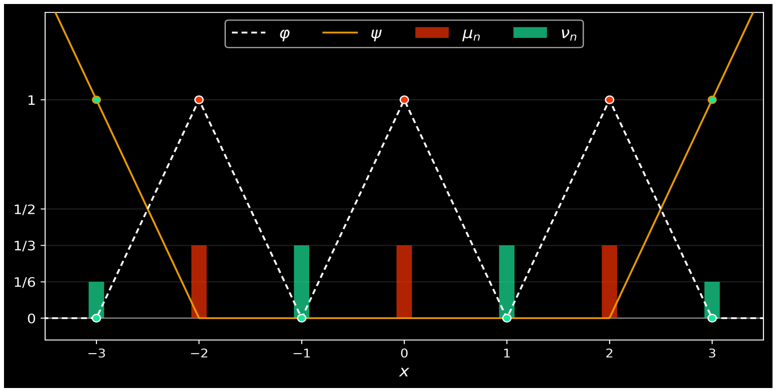
Then $\varphi(2m) = 1$ for all $|m| \leq n$, while φ vanishes on the support of ν_n. It is easy to check that φ is 1-Lipschitz, then we conclude from Kantorovich–Rubinstein theorem that

$$\mathcal{W}_1(\mu_n,\nu_n) \geq (\mu_n - \nu_n)(\varphi) = \mu_n(\varphi) = \frac{1}{2n+1} \sum_{|m| \leq n} \varphi(2m) = 1.$$

Next, consider the coupling $\mathbb{Q}_n = \mathrm{Law}(X_0, X_1)$, where $X_0 \sim \mu_n$ and $X_1 = X_0 + Z$ with $Z \sim \xi$ independent of X_0. Then $X_1 \sim \nu_n$, and $\mathbb{Q}_n$ is a coupling between μ_n and ν_n. Hence,

$$\mathcal{W}_1(\mu_n,\nu_n) \leq \mathbb{E}_{\mathbb{Q}_n}[|X_1 - X_0|] = \mathbb{E}_{\mathbb{Q}_n}[|Z|] = 1,$$

Figure 12: Discrete measures μ_n, ν_n, $n = 1$, and dual optimizers of $\mathcal{W}_1(\mu_n, \nu_n)$ and $\vec{d}(\nu_n, \mu_n)$ (φ and ψ, respectively) in the proof of Proposition 4.3 (iv).



which shows that $\mathcal{W}_1(\mu_n, \nu_n) = 1$.

Moving to the bid–ask distance, observe that $\mu_n \preceq_c \nu_n$ as ξ is centered. This can also be seen from the above martingale coupling and Strassen's theorem. Hence $\vec{d}(\mu_n, \nu_n) = 0$ by the second property. Flipping the roles of μ_n, ν_n, observe that $\vec{d}(\nu_n, \mu_n) = c_n = \frac{1}{2n+1}$. Indeed, for any 1-Lipschitz convex function ψ, we have after rearranging that

$$\frac{1}{c_n}(\nu_n - \mu_n)(\psi) = \sum_{m=-n+1}^{n} \psi(2m-1) - \frac{1}{2}[\psi(2m-2) + \psi(2m)]$$
$$+ \frac{1}{2}[\psi(-2n-1) - \psi(-2n) + \psi(2n+1) - \psi(2n)].$$

The first term is nonpositive using the (midpoint) convexity of ψ, while the second is bounded by

$$\frac{1}{2}\big[|\psi(-2n-1) - \psi(-2n)| + |\psi(2n+1) - \psi(2n)|\big] \leq 1,$$

since ψ is 1-Lipschitz. As ψ is arbitrary, we conclude that $\vec{d}(\nu_n, \mu_n) \leq c_n$, where the upper bound is achieved, e.g., by the strangle $\psi(x) = (-2n - x)^+ + (x - 2n)^+$; see Fig. 12. As $\lim_{n\to\infty} c_n = 0$, the claim follows. □

Lemma B.3. *Let* $\psi \in \mathrm{CVX}(\mathbb{R}) \cap \mathcal{C}^1(\mathbb{R})$ *with* ψ' *M-Lipschitz for some* $M > 0$. *Then for any* $\varphi \in \mathrm{CVX}(\mathbb{R})$ *satisfying*

$$\sup_{x\in\mathbb{R}} |(\psi - \varphi)(x)| \leq \varepsilon,$$

we have

$$\sup_{x\in\mathbb{R}} \sup_{\gamma\in\partial\varphi(x)} |\gamma - \psi'(x)| \leq 2\sqrt{M\varepsilon},$$

where $\partial\varphi(x)$ *denotes the subdifferential of* φ *at* x.

Proof. Fix x, let $\gamma \in \partial\varphi(x)$. As ψ' is M-Lipschitz, [33, Theorem 9.22] implies that

$$\psi'(x)\delta - \frac{M}{2}\delta^2 \leq \psi(x+\delta) - \psi(x) \leq \psi'(x)\delta + \frac{M}{2}\delta^2, \quad \delta > 0. \tag{31}$$

On the other hand, from $\sup_{x\in\mathbb{R}} |(\psi - \varphi)(x)| \leq \varepsilon$ and the convexity of φ, we derive that

$$\psi(x+\delta) + \varepsilon \geq \varphi(x+\delta) \geq \varphi(x) + \gamma\delta \geq \psi(x) - \varepsilon + \gamma\delta.$$

Subtracting $\psi(x)$ and applying the upper bound in (31), we arrive at

$$\psi'(x)\delta + \frac{M}{2}\delta^2 + \varepsilon \geq \gamma\delta - \varepsilon.$$

Rearranging the terms gives

$$\gamma - \psi'(x) \leq \frac{2\varepsilon}{\delta} + \frac{M}{2}\delta. \tag{32}$$

Minimizing the right-hand side of (32) over $\delta > 0$ yields the optimal $\delta^* = 2\sqrt{\varepsilon/M}$, which gives the upper bound

$$\gamma - \psi'(x) \leq 2\sqrt{M\varepsilon}.$$

When $\delta < 0$, repeating the argument for $\psi(x-\delta) - \psi(x)$ shows the lower bound

$$\gamma - \psi'(x) \geq -2\sqrt{M\varepsilon},$$

as desired. □

C Discretized Problems

Suppose the spot price of the underlying asset is x_0, and options at strikes $\{K_m\}_{m\in[M]}$ are available at time T_i, $i \in [N]$. By the same argument as in [3], we could restrict the dual optimization to

$$\Psi^{\mathrm{b,a}}_{\mathcal{S}}(h) = \left\{\psi^{\mathrm{b,a}} = (\psi^{\mathrm{a}}, \psi^{\mathrm{b}}) \in (\mathcal{S})^{2N} \;\middle|\; \exists\ \Delta \text{ s.t. } \mathrm{P\&L}^{h}_{\psi^{\mathrm{a}}-\psi^{\mathrm{b}},\Delta} \geq 0\right\}$$

where

$$\mathcal{S} := \left\{\psi : \mathbb{R}_+ \to \mathbb{R} \;\middle|\; \psi(x) = a + bx + \sum_{m=1}^{M} c_m(x-K_m)^+,\ a, b \in \mathbb{R},\ c_m \geq 0\right\}.$$

Introducing the restricted superhedging problem as

$$D_{\mathrm{fin}} := \inf_{\psi^{\mathrm{b,a}}\in\Psi^{\mathrm{b,a}}_{\mathcal{S}}(h)} \sum_{i=1}^{N} \mu^{\mathrm{a}}_i(\psi^{\mathrm{a}}_i) - \mu^{\mathrm{b}}_i(\psi^{\mathrm{b}}_i).$$

Write $c^{\mathrm{s}}(K_m, T_i)$, $i \in [N]$, $m \in [M]$, for the call option price with strike K_m and maturity T_i under the marginal μ^{s}_i, $\mathrm{s} \in \{\mathrm{a}, \mathrm{b}\}$. As $\mathbb{E}_{\mu^{\mathrm{b}}_i}[X] = \mathbb{E}_{\mu^{\mathrm{a}}_i}[X] = x_0$ for all $i \in [N]$, the optimization becomes

$$\begin{aligned}
&\min a + \sum_{i=1}^{N} b_i x_0 + \sum_{i=1}^{N}\sum_{m=1}^{M} c^{\mathrm{a}}_{i,m} c^{\mathrm{a}}(K_m, T_i) - \sum_{i=1}^{N}\sum_{m=1}^{M} c^{\mathrm{b}}_{i,m} c^{\mathrm{b}}(K_m, T_i)\\
\text{s.t.}\quad & a + \sum_{i=1}^{N} b_i x_i + \sum_{i=1}^{N}\sum_{m=1}^{M}\left(c^{\mathrm{a}}_{i,m} - c^{\mathrm{b}}_{i,m}\right)(x_i - K_m)^+\\
&+ \sum_{i=1}^{N-1} \Delta(x_1, \ldots, x_i)(x_{i+1} - x_i) \geq h(x_1, \ldots, x_N), \quad (x_1, \ldots, x_N) \in (\mathbb{R}_+)^N,\\
& a, b_i \in \mathbb{R}, c^{\mathrm{a}}_{i,m}, c^{\mathrm{b}}_{im} \geq 0.
\end{aligned}$$

The problem is then solved by linear programming. For completeness, we present the discretized version of the primal problem (14) in the single-maturity case $N = 1$, which is used to find a primal optimizer in Section 5.1.2. To that end, we construct a finite grid $\{y_i\}_{i\in[I]}$ on the support of μ^{a}. Given the precomputed call option prices $\{c^{\mathrm{s}}(K_m)\}_{m\in[M],\,\mathrm{s}\in\{\mathrm{a},\mathrm{b}\}}$, the discretized problem reads:

$$
\begin{aligned}
&\max_p \sum_{i=1}^I p_i\, h(y_i) \\
\text{s.t.}\quad &\sum_{i=1}^I p_i = 1, \qquad p_i \geq 0, \quad \sum_{i=1}^I p_i y_i = x_0, \\
&\sum_{i=1}^I p_i (y_i - K_m)^+ - c^{\mathrm{a}}(K_m) \leq 0, \\
&\sum_{i=1}^I p_i (y_i - K_m)^+ - c^{\mathrm{b}}(K_m) \geq 0, \qquad m \in [M].
\end{aligned}
$$

References

[1] B. Acciaio and G. Pammer. A short proof of the characterisation of convex order using the 2-Wasserstein distance. *arXiv:2207.02096*, 2022.

[2] E. Bayraktar and Y. Zhang. Fundamental theorem of asset pricing under transaction costs and model uncertainty. *Math. Oper. Res.*, 41(3):1039–1054, 2016.

[3] 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.

[4] M. Beiglböck, B. Jourdain, W. Margheriti, and G. Pammer. Approximation of martingale couplings on the line in the adapted weak topology. *Probab. Theory Related Fields*, 183(1-2):359–413, 2022.

[5] M. Beiglböck, B. Jourdain, W. Margheriti, and G. Pammer. Stability of the weak martingale optimal transport problem. *Ann. Appl. Probab.*, 33(6B):5382–5412, 2023.

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

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

[8] M. Beiglböck, G. Pammer, L. Riess, and S. Schrott. The fundamental theorem of weak optimal transport. *arXiv:2501.16316*, 2025.

[9] V. I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[10] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.*, 25(2):823–859, 2015.

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

[12] D. Brigo and F. Mercurio. Lognormal-mixture dynamics and calibration to market volatility smiles. *Int. J. Theor. Appl. Finance*, 5(4):427–446, 2002.

[13] P. Cheridito, M. Kupper, and L. Tangpi. Duality formulas for robust pricing and hedging in discrete time. *SIAM J. Financial Math.*, 8(1):738–765, 2017.

[14] L. Denis, M. Hu, and S. Peng. Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. *Potential Anal.*, 34(2):139–161, 2011.

[15] Y. Dolinsky and H. M. Soner. Robust hedging with proportional transaction costs. *Finance Stoch.*, 18(2):327–347, 2014.

[16] Y. Dolinsky and H. M. Soner. Martingale optimal transport in the Skorokhod space. *Stochastic Processes and their Applications*, 125(10):3893–3931, 2015.

[17] B. Dupire. A market design to trade bundles of securities and applications to options. *Bachelier Colloquium, Presentation Slides*, 2026.

[18] A. Fahim and Y.-J. Huang. Model-independent superhedging under portfolio constraints. *Finance Stoch.*, 20(1):51–81, 2016.

[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] G. Guo and J. Obłój. Computational methods for martingale optimal transport problems. *The Annals of Applied Probability*, 29(6):3311–3347, 2019.

[21] P. Henry-Labordère. Automated option pricing: numerical methods. *Int. J. Theor. Appl. Finance*, 16(8):1350042, 27, 2013.

[22] P. Henry-Labordère. *Model-free Hedging: A Martingale Optimal Transport Viewpoint*. Chapman and Hall/CRC Financial Mathematics Series. CRC Press, Boca Raton, 2017.

[23] 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.*, pages 267–318. Springer, Berlin, 2011.

[24] Z. Hou and J. Obłój. Robust pricing-hedging dualities in continuous time. *Finance Stoch.*, 22(3):511–567, 2018.

[25] 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.

[26] E. Jouini and H. Kallal. Martingales and arbitrage in securities markets with transaction costs. *J. Econom. Theory*, 66(1):178–197, 1995.

[27] N. Juillet. Peacocks parametrised by a partially ordered set. In *Séminaire de Probabilités XLVIII*, volume 2168 of *Lecture Notes in Math.*, pages 13–32. Springer, Cham, 2016.

[28] D. B. Madan and W. Schoutens. Two processes for two prices. *International Journal of Theoretical and Applied Finance*, 17(01):1450005, 2014.

[29] D. B. Madan, W. Schoutens, and K. Wang. Modeling the bid and ask prices of options. *The Journal of Computational Finance*, 26(4):1–36, 2023.

[30] A. C. G. Mennucci. On asymmetric distances. *Anal. Geom. Metr. Spaces*, 1:200–231, 2013.

[31] A. Neufeld and J. Sester. On the stability of the martingale optimal transport problem: a set-valued map approach. *Statist. Probab. Lett.*, 176:Paper No. 109131, 7, 2021.

[32] A. Neufeld and J. Sester. A deep learning approach to data-driven model-free pricing and to martingale optimal transport. *IEEE Trans. Inform. Theory*, 69(5):3172–3189, 2023.

[33] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1998.

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

[35] M. Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.

[36] H. M. Soner and V. Tissot-Daguette. Stopping times of boundaries: relaxation and continuity. *SIAM J. Control Optim.*, 63(4):2835–2855, 2025.

[37] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, 2009.

[38] J. Wiesel. Continuity of the martingale optimal transport problem on the real line. *Ann. Appl. Probab.*, 33(6A):4645–4692, 2023.

[39] J. Wiesel and E. Zhang. An optimal transport-based characterization of convex order. *Depend. Model.*, 11(1):Paper No 20230102, 15, 2023.

[40] Z. Zhou, J. Blanchet, and P. W. Glynn. Distributionally robust martingale optimal transport. *arXiv preprint arXiv:2106.07191*, 2021.