The Annals of Applied Probability
2019, Vol. 29, No. 6, 3348–3386
https://doi.org/10.1214/19-AAP1482

ROBUST PRICING AND HEDGING AROUND THE GLOBE

BY SEBASTIAN HERRMANN AND FLORIAN STEBEGG

University of Michigan and Columbia University

We consider the martingale optimal transport duality for càdlàg processes with given initial and terminal laws. Strong duality and existence of dual optimizers (robust semistatic superhedging strategies) are proved for a class of payoffs that includes American, Asian, Bermudan and European options with intermediate maturity. We exhibit an optimal superhedging strategy for which the static part solves an auxiliary problem and the dynamic part is given explicitly in terms of the static part.

1. Introduction. This paper studies the robust pricing and superhedging of derivative securities with a payoff of the form

$$F(X, A) = f\bigg(\int_{[0,T]} X_t \,\mathrm{d}A_t\bigg). \tag{1.1}$$

Here f is a nonnegative Borel function, X is a càdlàg price process (realized on the Skorokhod space), and A is chosen by the buyer from a given set $\mathcal{A}$ of exercise rights. More precisely, $\mathcal{A}$ is a set of so-called averaging processes, that is, nonnegative and nondecreasing adapted càdlàg processes A with $A_T \equiv 1$. Setting $\mathcal{A} = \{\mathbf{1}_{[\![\tau,T]\!]} : \tau \text{ a } [0,T]\text{-valued stopping time}\}$ or $\mathcal{A} = \{t \mapsto t/T\}$ reduces (1.1) to the relevant special cases of American- or Asian-style derivatives, respectively,

$$f(X_\tau) \quad \text{or} \quad f\bigg(\frac{1}{T}\int_0^T X_t \,\mathrm{d}t\bigg). \tag{1.2}$$

Other relevant examples are Bermudan options and European options with intermediate maturity (cf. Examples 3.3–3.4).

Robust pricing problem. Let μ and ν be probability measures on $\mathbb{R}$. We denote by $\mathcal{M}(\mu,\nu)$ the set of (continuous-time) martingale couplings between μ and ν, that is, probability measures P under which X is a martingale with marginal distributions $X_0 \overset{P}{\sim} \mu$ and $X_T \overset{P}{\sim} \nu$. The value of the *primal problem*

$$\mathbf{S} := \sup_{P\in\mathcal{M}(\mu,\nu)} \sup_{A\in\mathcal{A}} E^P\big[F(X,A)\big] \tag{1.3}$$

Received September 2017; revised October 2018.
MSC2010 subject classifications. 60G44, 49N05, 91G20.
Key words and phrases. Robust superhedging, semistatic strategies, martingale optimal transport, duality.

can be interpreted as the maximal model-based price for F over all models which are consistent with the given marginals.

If $\mathcal{A}$ is a singleton, then (1.3) is a so-called (continuous-time) martingale optimal transport problem. This problem was introduced (for general payoffs) by Beiglböck, Henry-Labordère and Penkner [6] in a discrete-time setting and by Galichon, Henry-Labordère and Touzi [15] in continuous time; cf. the survey [26].

Semistatic superhedging problem. The formal dual problem to (1.3) has a natural interpretation as a superhedging problem.[1] Loosely speaking, a semistatic superhedge is a triplet (φ, ψ, H) consisting of functions φ, ψ and a suitable process H such that for every $A \in \mathcal{A}$, the superhedging inequality holds:

$$\varphi(X_0) + \psi(X_T) + \int_0^T H_{t-}^A \, \mathrm{d}X_t \geq F(X, A) \quad \text{pathwise.} \tag{1.4}$$

Here the strategy $H = H^A$ may depend in an adapted way on A (cf. Section 3.2 for a precise formulation). For the example of an American-style payoff, this means that at the chosen exercise time τ, the buyer communicates her decision to exercise to the seller, who can then adjust the dynamic part of his hedging strategy (cf. [3], Section 3). The left-hand side in (1.4) is the payoff of a static position in two European-style derivatives on X plus the final value of a self-financing dynamic trading strategy in X. The inequality (1.4) says that the final value of this semistatic portfolio dominates the payoff F for every choice of exercise right and "all" price paths. The initial cost to set up a semistatic superhedge (φ, ψ, H) equals the price $\mu(\varphi) + \nu(\psi)$ of the static part.[2] The formal *dual problem* to (1.3),

$$\mathbf{I} := \inf\{\mu(\varphi) + \nu(\psi) : (\varphi, \psi, H) \text{ is a semistatic superhedge}\}, \tag{1.5}$$

amounts to finding the cheapest semistatic superhedge (if it exists) and its initial cost, the so-called robust superhedging price.

Main objectives and relaxation of the dual problem. We are interested in *strong duality*, that is, $\mathbf{S} = \mathbf{I}$, and *dual attainment*, that is, the existence of a dual minimizer.

Dual attainment requires a suitable relaxation of the dual problem. Indeed, for the discrete-time martingale optimal transport problem, Beiglböck, Henry-Labordère and Penkner [6] show strong duality for upper semicontinuous payoffs but provide a counterexample that shows that dual attainment can fail even if the payoff function is bounded and continuous. Beiglböck, Nutz and Touzi [8] achieve

[1] The primal problem (1.3) can be viewed as an optimization over finite measures P with three constraints: two marginal constraints and the martingale constraint. Its formal dual problem is the Lagrange dual problem where suitable functions φ and ψ and a suitable process H are used as Lagrange multipliers for the marginal and martingale constraints, respectively.

[2] We use the common notation $\mu(\varphi)$ for the integral of φ against μ.

strong duality and dual attainment for general payoffs and marginals in the one-step case by relaxing the dual problem in two aspects. First, they only require the superhedging inequality to hold in the quasi-sure sense, that is, outside a set which is a null set under every one-step martingale coupling between μ and ν. The reason is that the marginal constraints may introduce barriers on the real line which (almost surely) cannot be crossed by *any* martingale with these marginals; this was first observed by Hobson [20] (see also Cox [11] and Beiglböck and Juillet [7]). These barriers partition the real line into intervals and the marginal laws into so-called irreducible components. Then strong duality and dual attainment can be reduced to proving the same results for each irreducible component [8, 20]. Second, Beiglböck, Nutz and Touzi [8] extend the meaning of the expression $\mu(\varphi) + \nu(\psi)$ to certain situations where both individual integrals are infinite. For example, it can happen that $\mu(\varphi) = -\infty$ and $\nu(\psi) = \infty$, but the price $E^P[\varphi(X_0) + \psi(X_T)]$ of the combined static part is well defined, finite and invariant under the choice of $P \in \mathcal{M}(\mu, \nu)$. In this situation, this price is still denoted by $\mu(\varphi) + \nu(\psi)$. We employ both relaxations for the precise definition of the dual problem in Section 3.2.

In continuous time, Dolinsky and Soner [13, 14] show strong duality for uniformly continuous payoffs satisfying a certain growth condition. They use the integration by parts formula to define the stochastic integral $\int_0^T H_{t-}\,\mathrm{d}X_t$ pathwise for finite variation integrands H. However, dual attainment cannot be expected in this class in general. For our payoffs (1.1), we need to allow integrands that are of finite variation whenever they are bounded but can become arbitrarily large or small on certain price paths. As the integrands are not of finite variation on these paths, the meaning of the pathwise integral needs to be extended appropriately.

For the purpose of the Introduction, we discuss our results and methodology in a nonrigorous fashion, ignoring all aspects relating to the relaxation of the dual problem.

Main results. We prove strong duality and dual attainment for payoffs of the form (1.1) under mild conditions on f and $\mathcal{A}$ for irreducible marginals (Theorem 3.9); all results can be extended to general marginals along the lines of [8], Section 7. The key idea is the reduction of the primal and dual problems to simpler auxiliary problems, which do not depend on the set $\mathcal{A}$ of exercising rights. In particular, our results cover American-style derivatives $f(X_\tau)$ for Borel-measurable f and Asian-style derivatives $f(\frac{1}{T}\int_0^T X_t\,\mathrm{d}t)$ for lower semicontinuous f and show that both derivatives have (perhaps surprisingly) the same robust superhedging prices and structurally similar semistatic superhedges.

Methodology. Our methodology relies on two crucial observations which allow us to bound the primal problem from below and the dual problem from above by simpler auxiliary primal and dual problems, respectively. To obtain a primal lower bound, we show that for any law θ which is in convex order between μ and

ν, there is a sequence $(P_n)_{n\geq 1} \subset \mathcal{M}(\mu,\nu)$ such that the law of $\int_{[0,T]} X_t \,\mathrm{d}A_t$ under P_n converges weakly to θ if A is a suitable averaging process. This allows us to bound $\mathbf{S}$ from below by the value of the auxiliary primal problem

$$\widetilde{\mathbf{S}} := \sup_{\mu \leq_c \theta \leq_c \nu} \theta(f).$$

(The converse inequality also holds (cf. Lemma 4.1), so that in fact $\mathbf{S} = \widetilde{\mathbf{S}}$.)

Regarding the dual upper bound, we prove (modulo technicalities) that if φ is concave and ψ is convex such that $\varphi + \psi \geq f$, then (φ, ψ, H) is a semistatic superhedge, where the dynamic part H is given explicitly in terms of φ and ψ by

$$H_t := \varphi'(X_0) - \int_{[0,t]} \{\varphi'(X_0) + \psi'(X_s)\} \,\mathrm{d}A_s. \tag{1.6}$$

This allows us to bound $\mathbf{I}$ from above by the value of the auxiliary dual problem

$$\widetilde{\mathbf{I}} := \inf\{\mu(\varphi) + \nu(\psi) : \varphi \text{ concave}, \psi \text{ convex}, \varphi + \psi \geq f\}.$$

As a consequence, strong duality and dual attainment for $\mathbf{S}$ and $\mathbf{I}$ follow from the same assertions for the simpler problems $\widetilde{\mathbf{S}}$ and $\widetilde{\mathbf{I}}$, which we prove by adapting the techniques of [8]. Moreover, our reduction of the dual problem implies that if (φ, ψ) is optimal for $\widetilde{\mathbf{I}}$, then it is also the static part of an optimal semistatic superhedge and the dynamic part H can be computed ex post via (1.6). This dramatically decreases the complexity of the superhedging problem: the optimization over two functions *and a process* satisfying an inequality constraint *on the Skorokhod space* is reduced to an optimization over two functions satisfying an inequality constraint on $\mathbb{R}$.

Our methodology reveals that many derivatives have the same robust superhedging prices and semi-static superhedges. Indeed, $\widetilde{\mathbf{I}}$ and $\widetilde{\mathbf{S}}$ do not depend on the set $\mathcal{A}$ of exercise rights granted to the buyer, and this independence transfers to $\mathbf{S}$ and $\mathbf{I}$ under mild conditions on f and $\mathcal{A}$. For example, if f is lower semicontinuous, then the Asian-style derivative $f(\frac{1}{T}\int_0^T X_t \,\mathrm{d}t)$, the American-style derivative $f(X_\tau)$, and the European-style derivative $f(X_{T'})$ (for a fixed maturity $T' \in (0,T)$) all have the same robust superhedging price (Remark 3.10). This invariance breaks down when more than two marginals are given.

Related literature. Much of the extant literature on robust superhedging in semistatic settings is concerned with strong duality and dual attainment. The results vary in their generality and explicitness as well as their precise formulation. The semistatic setting, where call options are available as additional hedging instruments, dates back to Hobson's seminal paper [16] on the lookback option.[3] Many other specific exotic derivatives (mostly without special exercise rights) have been analyzed in this framework in the past two decades; cf. the survey [17].

[3] We note that the superhedging strategies described in [16] are actually dynamic in the call options.

Securities with special exercise rights have been studied in the context of American-style derivatives in discrete-time settings. Bayraktar, Huang and Zhou [3] obtain a duality result for a somewhat different primal problem (cf. [3], Theorem 3.1) and show that duality may fail in their setting if they formulate their primal problem in analogy to the present paper; see also [4] for related results with portfolio constraints. Hobson and Neuberger [19] (based on Neuberger's earlier manuscript [23]) resolve this issue by adopting a weak formulation for the primal problem: instead of optimizing only over martingale measures on a fixed filtered path space, the optimization there runs over filtered probability spaces supporting a martingale and thereby allows richer information structures and hence more stopping times. We also refer to [1, 5, 18] for recent developments in this regard. We note that all these papers permit significant restrictions on the set of possible price paths (e.g., binomial trees) while we allow all càdlàg paths. This difference may be the reason why strong duality holds in our setting without any relaxation of the primal problem.

The case of an Asian-style payoff as in (1.2) has been studied in the case of a Dirac initial law μ. For convex or concave f, Stebegg [25] shows strong duality and dual attainment. For nonnegative Lipschitz f, Cox and Källblad [12] obtain a PDE characterization of the maximal model-based price for finitely supported ν. Bayraktar, Cox and Stoev [2] provide a similar, but not identical PDE for the corresponding pricing problem for American-style payoffs as in (1.2). A consequence of our main duality result is that the Asian and American pricing problems are the same, so that both these PDEs have the same (viscosity) solution.

Organization of the paper. The remainder of the article is organized as follows. In Section 2, we recall basic results on the convex order and potential functions, introduce the generalized integral of [8] and its relevant properties and present the extension of the pathwise definition of the stochastic integral for finite variation integrands. Section 3 introduces the robust pricing and semistatic superhedging problems and presents our duality result. The duality between the auxiliary problems, the structure of their optimizers and their relation to the robust pricing and superhedging problem are treated in Section 4. In Section 5, we provide simple geometric constructions of primal and dual optimizers for risk reversals and butterfly spreads. Finally, some counterexamples are collected in Section 6.

2. Preliminaries. Fix a time horizon T and let $\Omega = D([0, T]; \mathbb{R})$ be the space of real-valued càdlàg paths on $[0, T]$. We endow Ω with the Skorokhod topology and denote by $\mathcal{F}$ the corresponding Borel σ-algebra, by $X = (X_t)_{t\in[0,T]}$ the canonical process on Ω, and by $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ the (raw) filtration generated by X. Unless otherwise stated, all probabilistic notions requiring a filtration pertain to $\mathbb{F}$.

For any process $Y = (Y_t)_{t\in[0,T]}$ on Ω, we set $Y_{0-} = 0$, so that the jump of Y at time 0 is $\Delta Y_0 = Y_0$.

2.1. *Martingale measures and convex order.* Let μ and ν be finite[4] measures on $\mathbb{R}$ with finite first moment. We denote by $\Pi(\mu,\nu)$ the set of (continuous-time) couplings of μ and ν, that is, finite measures P on Ω such that $P\circ X_0^{-1}=\mu$ and $P\circ X_T^{-1}=\nu$. If, in addition, the canonical process X is a martingale under P (defined in the natural way if P is not a probability measure), then we write $P\in\mathcal{M}(\mu,\nu)$ and say that P is a (continuous-time) martingale coupling between μ and ν.

We also consider discrete-time versions of these notions. To wit, we denote by $\Pi^d(\mu,\nu)$ the set of finite measures Q on $\mathbb{R}^2$ with marginal distributions μ and ν and by $\mathcal{M}^d(\mu,\nu)$ the subset of measures Q under which the canonical process on $\mathbb{R}^2$ is a martingale (in its own filtration). The sets $\Pi^d(\mu,\theta,\nu)$ and $\mathcal{M}(\mu,\theta,\nu)$ of finite measures on $\mathbb{R}^3$ with prescribed marginal distributions are defined analogously.

We write $\mu\le_c\nu$ if μ and ν are in *convex order* in the sense that $\mu(\varphi)\le\nu(\varphi)$ holds for any convex function $\varphi:\mathbb{R}\to\mathbb{R}$. In this case, μ and ν have the same mass and the same barycenter $\operatorname{bary}(\mu):=\frac{1}{\mu(\mathbb{R})}\int x\mu(\mathrm{d}x)$.

The potential function $u_\mu:\mathbb{R}\to[0,\infty]$ of μ is defined as

$$u_\mu(x):=\int|x-y|\mu(\mathrm{d}y). \tag{2.1}$$

We refer to [7], Section 4.1, for basic properties of potential functions. In particular, the following relationship between the convex order, potential functions and martingale measures is well known.

PROPOSITION 2.1. *Let μ and ν be finite measures with finite first moments and $\mu(\mathbb{R})=\nu(\mathbb{R})$. Then the following are equivalent*: (*i*) $\mu\le_c\nu$, (*ii*) $u_\mu\le u_\nu$, (*iii*) $\mathcal{M}^d(\mu,\nu)\neq\varnothing$ *and* (*iv*) $\mathcal{M}(\mu,\nu)\neq\varnothing$.

An analogous result holds for three marginals μ,θ,ν, the corresponding potential functions and the set $\mathcal{M}^d(\mu,\theta,\nu)$.

We recall the following definition from [8], Definition 2.2 (see also [7], Definition A.3).

DEFINITION 2.2. A pair of finite measures $\mu\le_c\nu$ is called *irreducible* if the set $I=\{u_\mu<u_\nu\}$ is connected and $\mu(I)=\mu(\mathbb{R})$. In this situation, let J be the union of I and any endpoints of I that are atoms of ν; then (I,J) is the *domain* of $\mu\le_c\nu$.

We work with irreducible $\mu\le_c\nu$ for the remainder of this article.

[4] As in [8], using finite measures as opposed to probability measures turns out to be useful.

2.2. *Generalized integral.* Let $\mu \preceq_c \nu$ be irreducible with domain (I, J). Beiglböck and Juillet [7], Section A.3, and Beiglböck, Nutz, and Touzi [8], Section 4, appropriately extend the meaning of the expression $\mu(\varphi) + \nu(\psi)$ to the case where the individual integrals are not necessarily finite. We present here a slight extension of their work in order to deal with intermediate laws $\mu \preceq_c \theta \preceq_c \nu$ for which the pairs $\mu \preceq_c \theta$ and $\theta \preceq_c \nu$ may not be irreducible.

For the rest of this article, whenever we write $\mu \preceq_c \nu$ for any two measures μ and ν, we implicitly assume that both measures are finite and have a finite first moment. Throughout this section, we fix $\mu \preceq_c \theta_1 \preceq_c \theta_2 \preceq_c \nu$.

DEFINITION 2.3. Let $\chi : J \to \mathbb{R}$ be concave. Denote by $-\chi''$ the second derivative measure of the convex function $-\chi$ and by $\Delta\chi$ the possible jumps of χ at the endpoints of I. We set

$$(\theta_1 - \theta_2)(\chi) := \frac{1}{2}\int_I (u_{\theta_1} - u_{\theta_2})\,\mathrm{d}\chi'' + \int_{J\setminus I} |\Delta\chi|\,\mathrm{d}(\theta_2 - \theta_1) \in [0, \infty]. \tag{2.2}$$

The right-hand side is well defined in $[0, \infty]$ because $u_{\theta_1} \le u_{\theta_2}$ on I and $\theta_1(\{b\}) \le \theta_2(\{b\})$ for $b \in J \setminus I$.

If $\theta_1 = \mu$ and $\theta_2 = \nu$, then (2.2) coincides with equation (4.2) in [8] because μ is concentrated on I. As in [8], there is an alternative representation of $(\theta_1 - \theta_2)(\chi)$ in terms of an iterated integral with respect to a disintegration of a (one-step) martingale coupling on $\mathbb{R}^2$:

LEMMA 2.4. *Let $\chi : J \to \mathbb{R}$ be concave and let $Q \in \mathcal{M}^d(\theta_1, \theta_2)$. For any disintegration $Q = \theta_1 \otimes \kappa$, we have*

$$(\theta_1 - \theta_2)(\chi) = \int_J \left[\chi(x_1) - \int_J \chi(x_2)\kappa(x_1, \mathrm{d}x_2)\right]\theta_1(\mathrm{d}x_1).$$

PROOF. The proof of [8], Lemma 4.1, does not use that $\mu \preceq_c \nu$ is irreducible. Moreover, for $\bar\chi : J \to \mathbb{R}$ concave and continuous, the same arguments as in the proof of [8], Lemma 4.1, yield

$$\frac{1}{2}\int_I (u_{\theta_1} - u_{\theta_2})\,\mathrm{d}\bar\chi'' = \int_J \left[\bar\chi(x_1) - \int_J \bar\chi(x_2)\kappa(x_1, \mathrm{d}x_2)\right]\theta_1(\mathrm{d}x_1). \tag{2.3}$$

(Note that $\int_J \bar\chi(x_2)\kappa(x_1, \mathrm{d}x_2) = \bar\chi(x_1)$ for boundary points $x_1 \in J \setminus I$ because κ is a martingale kernel concentrated on J.)

For a general concave $\chi : J \to \mathbb{R}$, we write $\chi = \bar\chi - |\Delta\chi|\mathbf{1}_{J\setminus I}$ with $\bar\chi$ continuous. Then we can replace $\bar\chi$ with χ on the left-hand side of (2.3) and the integrand on the right-hand side reads as

$$\chi + |\Delta\chi|\mathbf{1}_{J\setminus I} - \int_J \chi(x_2)\kappa(\cdot, \mathrm{d}x_2) - \int_{J\setminus I} |\Delta\chi(x_2)|\kappa(\cdot, \mathrm{d}x_2).$$

Integrating this against θ_1 and using Fubini's theorem yields

$$\int_J \left[\chi(x_1) - \int_J \chi(x_2)\kappa(x_1, \mathrm{d}x_2)\right]\theta(\mathrm{d}x_1) + \int_{J\setminus I} |\Delta\chi|\,\mathrm{d}\theta_1 - \int_{J\setminus I} |\Delta\chi|\,\mathrm{d}\theta_2.$$

Together with (2.3), this proves the claim. □

It can be shown as in [8] that $(\theta_1 - \theta_2)(\chi) = \theta_1(\chi) - \theta_2(\chi)$ if at least one of the individual integrals is finite.

We can now define the integral $\theta_1(\varphi) + \theta_2(\psi)$ as in [8], Definition 4.7.

DEFINITION 2.5. Let $\varphi : J \to \overline{\mathbb{R}}$ and $\psi : J \to \overline{\mathbb{R}}$ be Borel functions. If there exists a concave function $\chi : J \to \mathbb{R}$ such that $\varphi - \chi \in L^1(\theta_1)$ and $\psi + \chi \in L^1(\theta_2)$, we say that χ is a *concave moderator* for (φ, ψ) with respect to $\theta_1 \leq_c \theta_2$ and set

$$\theta_1(\varphi) + \theta_2(\psi) := \theta_1(\varphi - \chi) + \theta_2(\psi + \chi) + (\theta_1 - \theta_2)(\chi) \in (-\infty, \infty]. \tag{2.4}$$

As in [8], the expression $\theta_1(\varphi) + \theta_2(\psi)$ defined in (2.4) does not depend on the choice of the concave moderator.

DEFINITION 2.6. We write $L^c(\theta_1, \theta_2)$ for the space of pairs of Borel functions $\varphi, \psi : J \to \overline{\mathbb{R}}$ which admit a concave moderator χ with respect to $\theta_1 \leq_c \theta_2$ such that $(\theta_1 - \theta_2)(\chi) < \infty$.

We next present additional properties of the notions introduced above.

LEMMA 2.7. *Let* $(\varphi, \psi) \in L^c(\theta_1, \theta_2)$.

(i) *φ is finite on atoms of θ_1. If φ is concave, then $\varphi < \infty$ on J and $\varphi > -\infty$ on the interior of the convex hull of the support of θ_1.*

(ii) *ψ is finite on atoms of θ_2. If ψ is convex, then $\psi > -\infty$ on J and $\psi < \infty$ on the interior of the convex hull of the support of θ_2.*

(iii) *If $a, b : \mathbb{R} \to \mathbb{R}$ are affine, then $(\varphi + a, \psi + b) \in L^c(\theta_1, \theta_2)$ and*

$$\theta_1(\varphi + a) + \theta_2(\psi + b) = \{\theta_1(\varphi) + \theta_2(\psi)\} + \theta_1(a) + \theta_2(b).$$

PROOF. We only prove (iii). Let χ be a concave moderator for (φ, ψ) with respect to $\theta_1 \leq_c \theta_2$. Then $\varphi - \chi \in L^1(\theta_1)$, $\psi + \chi \in L^1(\theta_2)$ and $(\theta_1 - \theta_2)(\chi) < \infty$. Being affine, a and b are θ_1- and θ_2-integrable. It follows that χ is also a concave moderator for $(\varphi + a, \psi + b)$ with respect to $\theta_1 \leq_c \theta_2$ and that $(\varphi + a, \psi + b) \in L^c(\theta_1, \theta_2)$. The last assertion is a direct computation. □

REMARK 2.8. Recall that I is the interior of the convex hull of the support of ν and that J is the union of I and any endpoints of I that are atoms of ν. Hence, Lemma 2.7(ii) shows in particular, that if $(\varphi, \psi) \in L^c(\theta_1, \nu)$ with ψ convex, then ψ is finite on J.

We conclude this section with a number of calculation rules for the integrals defined above when φ is concave and ψ is convex.

LEMMA 2.9. *Let $\mu \leq_c \theta_1 \leq_c \theta_2 \leq_c \theta_3 \leq_c \nu$ (where the pair $\mu \leq_c \nu$ is irreducible) and let $(\varphi, \psi) \in L^c(\theta_1, \theta_3)$ be such that φ is concave and finite, ψ is convex and finite and $\varphi + \psi$ is bounded from below by a concave θ_3-integrable function.*

(i) *φ and $-\psi$ are concave moderators for (φ, ψ) with respect to $\theta_1 \leq_c \theta_3$.*
(ii) $(\varphi, \psi) \in L^c(\theta_1, \theta_2) \cap L^c(\theta_2, \theta_3)$.
(iii) $\theta_1(\varphi) + \theta_2(\psi) \leq \theta_1(\varphi) + \theta_3(\psi)$.
(iv) $\theta_2(\varphi) + \theta_3(\psi) \leq \theta_1(\varphi) + \theta_3(\psi)$.

PROOF. Denote by ξ a concave θ_3-integrable lower bound for $\varphi + \psi$. By the concavity of ξ, we have $\theta_1(\xi) \geq \theta_2(\xi) \geq \theta_3(\xi) > -\infty$, so that ξ is also θ_1- and θ_2-integrable.

(i) Regarding the concave moderator property of φ, it suffices to show that $\varphi + \psi$ is θ_3-integrable. Denote by φ' the left-derivative of the concave function φ on I. Then for $(x_1, x_3) \in I \times J$,

$$\xi(x_3) \leq \varphi(x_3) + \psi(x_3) \leq \varphi(x_1) + \psi(x_3) + \varphi'(x_1)(x_3 - x_1). \tag{2.5}$$

Fix any $Q \in \mathcal{M}^d(\theta_1, \theta_3)$. Then (2.5) also holds Q-a.e. on $J \times J$ (setting $\varphi' = 0$ on $J \setminus I$, e.g.); this uses that any mass in a point of $J \setminus I$ stays put under a martingale transport plan. Since ξ is θ_3-integrable, the negative part of the right-hand side in (2.5) is Q-integrable. Then it can be argued as in [8], Remark 4.10, that the right-hand side in (2.5) is Q-integrable. It follows that $\varphi + \psi$ is θ_3-integrable.

Regarding the assertion about $-\psi$, it suffices to show that $\varphi + \psi$ is θ_1-integrable. We have

$$\begin{aligned} \xi(x_1) &\leq \varphi(x_1) + \psi(x_1) \\ &= \big[\varphi(x_1) + \psi(x_3) + \varphi'(x_1)(x_3 - x_1)\big] \\ &\quad + \big[\psi(x_1) - \psi(x_3) - \varphi'(x_1)(x_3 - x_1)\big] \quad Q\text{-a.e. on } J \times J. \end{aligned} \tag{2.6}$$

By the above, the first term on the right-hand side is Q-integrable. Thus, the negative part of the second term is also Q-integrable. Hence, we may integrate the second term iteratively using Fubini's theorem as in [8], Remark 4.10. The Q-integral equals $-(\theta_1 - \theta_3)(-\psi) \leq 0$. In particular, the right-hand side in (2.6) is Q-integrable. It follows that $\varphi + \psi$ is θ_1-integrable.

(ii)–(iv) We know from the above that $\varphi + \psi$ is θ_3-integrable. It follows that φ is a concave moderator for (φ, ψ) with respect to $\theta_2 \leq_c \theta_3$. Because $\theta_1 \leq_c \theta_2$, we have that $u_{\theta_1} \leq u_{\theta_2}$ and $\theta_1(\{b\}) \leq \theta_2(\{b\})$ for $b \in J \setminus I$. Thus, $(\theta_2 - \theta_3)(\varphi) \leq$

$(\theta_1 - \theta_3)(\varphi) < \infty$ (cf. Definition 2.3). Hence, $(\varphi, \psi) \in L^c(\theta_2, \theta_3)$ and

$$\begin{aligned}\theta_2(\varphi) + \theta_3(\psi) &= \theta_2(\varphi - \varphi) + \theta_3(\varphi + \psi) + (\theta_2 - \theta_3)(\varphi)\\ &\le \theta_1(\varphi - \varphi) + \theta_3(\varphi + \psi) + (\theta_1 - \theta_3)(\varphi)\\ &= \theta_1(\varphi) + \theta_3(\psi).\end{aligned}$$

One can show similarly that $(\varphi, \psi) \in L^c(\theta_1, \theta_2)$ and that $\theta_1(\varphi) + \theta_2(\psi) \le \theta_1(\varphi) + \theta_3(\psi)$. $\square$

2.3. *Pathwise stochastic integration.* For any $\mathbb{F}$-adapted càdlàg process H of finite variation, the integral $H_- \bullet X_T = \int_{(0,T]} H_{t-}\,\mathrm{d}X_t$ can be defined *pathwise*, that is, for each $\omega \in \Omega$ individually, via integration by parts as follows:

$$H_- \bullet X_T := X_T H_T - X_0 H_0 - \int_{(0,T]} X_t\,\mathrm{d}H_t, \tag{2.7}$$

where the integral on the right-hand side is the pathwise Lebesgue–Stieltjes integral. Setting $H_{0-} = 0$, so that $\Delta H_0 = H_0$, we can recast (2.7) as

$$H_- \bullet X_T = (X_T - X_0)H_0 + \int_{(0,T]} (X_T - X_t)\,\mathrm{d}H_t. \tag{2.8}$$

For any martingale measure P, if the (standard) stochastic integral of H_- with respect to X exists, then it is P-indistinguishable from the pathwise stochastic integral.

We need to give a sensible meaning to the integral $H_- \bullet X_T$ for certain integrands H which are not necessarily of finite variation, but may diverge in finite time.

EXAMPLE 2.10. The following example motivates our extension of the pathwise stochastic integral for finite variation integrands. Let $\mu = \delta_0$ and $\nu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. Then $\mu \le_c \nu$ are irreducible with domain $(I, J) = ((-1, 1), [-1, 1])$. Consider a payoff function f which is convex on $[-1, 1]$ and has infinite (one-sided) derivatives at -1 and 1, for example, $f(x) = 1 - \sqrt{1 - x^2}\mathbf{1}_{[-1,1]}(x)$, a semistatic superhedge for the Asian-style derivative $f(\frac{1}{T}\int_0^T X_t\,\mathrm{d}t)$ can be derived as follows. By Jensen's inequality and the convexity of f, for every path of X that evolves in $[-1, 1]$,

$$\begin{aligned}f\left(\frac{1}{T}\int_0^T X_t\,\mathrm{d}t\right) &\le \int_0^T f(X_t)\,\frac{\mathrm{d}t}{T} \le \int_0^T \big(f(X_T) - f'(X_t)(X_T - X_t)\big)\,\frac{\mathrm{d}t}{T}\\ &= f(X_T) - \int_0^T (X_T - X_t) f'(X_t)\,\frac{\mathrm{d}t}{T}.\end{aligned}$$

Comparing this with (2.8), a semistatic superhedge for the Asian-style derivative is obtained from a European-style derivative with payoff $f(X_T)$ maturing at T and

a dynamic trading strategy H with $H_0 = 0$ and dynamics $\mathrm{d}H_t = -f'(X_t)\frac{\mathrm{d}t}{T}$. Then H is of finite variation whenever X stays away from the boundaries of $(-1, 1)$. But, as X approaches -1 or 1, the derivative $f'(X_t)$ becomes arbitrarily large (in absolute value), and H may fail to be of finite variation. It turns out, however, that the integral $\int_0^T (X_T - X_t) f'(X_t) \frac{\mathrm{d}t}{T}$ is still well defined on these paths. The reason is that when paths of X come arbitrarily close to 1, say, then for any martingale coupling $P \in \mathcal{M}(\mu, \nu)$, $X_T = 1$ P-a.s. on these paths (because $J = [-1, 1]$), so that $X_T - X_t$ becomes small and counteracts the growth of $f'(X_t)$.

We shall define a pathwise stochastic integral for $\mathbb{F}$-adapted càdlàg integrators X and integrands $\hat{H}_-$ of the form

$$\hat{H}_t = h_0 + \int_{(0,t]} h_s \, \mathrm{d}Y_s \tag{2.9}$$

for an $\mathbb{F}$-adapted càdlàg process $Y = (Y_t)_{t\in[0,T]}$ of finite variation and an $\mathbb{F}$-adapted process $h = (h_t)_{t\in[0,T]}$—even in certain situations where the right-hand side of (2.9) is not finite. The idea is to formally substitute (2.9) into (2.8), formally use the associativity of Lebesgue–Stieltjes integrals, and then employ the resulting expression as a definition for a pathwise stochastic integral. We first introduce a set of integrands for this integral.

DEFINITION 2.11. Let $\Omega' \subset D([0, T]; \mathbb{R})$. We denote by $L(\Omega')$ the set of pairs (h, Y) consisting of an $\mathbb{F}$-adapted process h and an $\mathbb{F}$-adapted càdlàg process Y of finite variation such that the process $((X_T - X_t)h_t)_{t\in[0,T]}$ is $\mathrm{d}Y$-integrable on $(0, T]$ for each path in Ω'.

If Y is a $\mathbb{F}$-adapted càdlàg process of finite variation, then $(1, Y) \in L(\Omega')$ for any $\Omega' \subset D([0, T]; \mathbb{R})$ (because any càdlàg function is bounded on compact intervals).

We fix a set $\Omega' \subset D([0, T]; \mathbb{R})$ for the rest of this section.

DEFINITION 2.12. For $H = (h, Y) \in L(\Omega')$, we set

$$H \diamond X_T := (X_T - X_0)h_0 + \int_{(0,T]} (X_T - X_t)h_t \, \mathrm{d}Y_t \quad \text{on } \Omega'. \tag{2.10}$$

We note that the Lebesgue–Stieltjes integral on the right-hand side of (2.10) is well defined and finite by the definition of $L(\Omega')$. The following result shows that for pathwise bounded h, $H \diamond X_T$ coincides with $\hat{H}_- \bullet X_T$ for $\hat{H}$ as in (2.9). This motivates the interpretation of $H \diamond X_T$ as the gains from trading in X according to a self-financing trading strategy $\hat{H}_-$.

PROPOSITION 2.13. *Let $H = (h, Y) \in L(\Omega')$ and $\omega \in \Omega'$. If the function $t \mapsto h_t(\omega)$ is bounded on $[0, T]$, then*

$$(H \diamond X_T)(\omega) = (\hat{H}_- \bullet X_T)(\omega),$$

where $\hat{H} = h_0 + \int_{(0,\cdot]} h \, \mathrm{d}Y$.

If we set $h_0 = Y_0$ and $h_t = 1$ for $t \in (0, T]$ for an $\mathbb{F}$-adapted càdlàg process Y of finite variation, then $H = (h, Y) \in L(\Omega)$ and by Proposition 2.13,

$$H \diamond X_T = Y_- \bullet X_T \quad \text{on } \Omega.$$

So the integral $H \diamond X_T$ embeds all pathwise stochastic integrals $Y_- \bullet X_T$.

PROOF OF PROPOSITION 2.13. Since $h(\omega)$ is bounded on $[0, T]$, $\hat{H}_t(\omega) = h_0(\omega) + \int_{(0,t]} h_s(\omega)\,\mathrm{d}Y_s(\omega)$ is a well-defined càdlàg finite variation function on $[0, T]$. Thus, by (2.8),

$$\hat{H}_- \bullet X_T = (X_T - X_0)h_0 + \int_{(0,T]} (X_T - X_s)h_s\,\mathrm{d}Y_s = H \diamond X_T. \qquad \square$$

3. Robust pricing and superhedging problems. Throughout this section, we fix an irreducible pair $\mu \le_c \nu$ with domain (I, J) and a Borel function $f : \mathbb{R} \to \overline{\mathbb{R}}$ which is bounded from below by a ν-integrable concave function.

3.1. *Pricing problem.* Our pricing and hedging duality applies to a wide range of exotic derivatives including American options, fixed strike Asian options, Bermudan options and European options with an intermediate maturity. We now describe this class of derivative securities.

DEFINITION 3.1. A nonnegative $\mathbb{F}$-adapted nondecreasing càdlàg process $A = (A_t)_{t\in[0,T]}$ is called an *averaging process* if $A_T(\omega) = 1$ for every $\omega \in \Omega$. If, in addition, $A_0(\omega) = 0$ and $\Delta A_T(\omega) = 0$ for each $\omega \in \Omega$, then A is called an *interior averaging process*. If, in addition, there is $t \in (0, T)$ such that $A_t(\omega) = 0$ for each $\omega \in \Omega$, then A is called a *strictly interior averaging process*.

Recall that we set $A_{0-} = 0$ and note that for each $\omega \in \Omega$, $A(\omega)$ can be identified with a Borel probability measure on $[0, T]$. If A is an interior averaging process, then this probability measure is supported on $(0, T)$, and if A is a strictly interior averaging process then its support is (uniformly in ω) contained in $[t, T)$ for some $t \in (0, T)$.

Given a nonempty set $\mathcal{A}$ of averaging processes, we consider a derivative security whose payoff at time T is

$$f\Big(\int_{[0,T]} X_t\,\mathrm{d}A_t\Big), \tag{3.1}$$

where $A \in \mathcal{A}$ is chosen by the buyer and the seller observes $(A_s)_{s\in[0,t]}$ at time t. Then the *robust model-based price* is defined as

$$\mathbf{S}_{\mu,\nu}(f, \mathcal{A}) = \sup_{P\in\mathcal{M}(\mu,\nu)} \sup_{A\in\mathcal{A}} E^P\Big[f\Big(\int_{[0,T]} X_t\,\mathrm{d}A_t\Big)\Big]. \tag{3.2}$$

In other words, $\mathbf{S}_{\mu,\nu}(f, \mathcal{A})$ is the highest model-based price of the derivative security (3.1) among all martingale models which are consistent with the given marginal distributions.

REMARK 3.2. One can show that for each $P \in \mathcal{M}(\mu, \nu)$ and each averaging process A, the law of $\int_{[0,T]} X_t \, \mathrm{d}A_t$ under P is in convex order between μ and ν; cf. Lemma 4.1. Because f is by assumption bounded from below by a ν-integrable concave function, the expectations in (3.2) are well defined.

Important special cases are obtained for specific choices of $\mathcal{A}$.

EXAMPLE 3.3 (No special exercise rights). Setting $\mathcal{A} = \{A\}$ deprives the buyer of any special exercise rights and reduces (3.2) to the more familiar robust pricing problem

$$\sup_{P \in \mathcal{M}(\mu,\nu)} E^P[F]$$

for the derivative security $F = f(\int_{[0,T]} X_t \, \mathrm{d}A_t)$.

(i) *Asian options.* Setting $A_t = t/T$ recovers the Asian-style derivative $f(\frac{1}{T}\int_0^T X_t \, \mathrm{d}t)$; this includes fixed strike Asian puts and calls, but not floating strike Asian options. This robust pricing problem is analyzed in [12].

(ii) *European options.* Setting $A_t = \mathbf{1}_{[T',T]}(t)$ yields a European-style payoff $f(X_{T'})$ with an intermediate maturity $T' \in (0, T)$.

EXAMPLE 3.4 (Special exercise rights). Fix a nonempty set $\mathcal{T}$ of $[0, T]$-valued $\mathbb{F}$-stopping times, and consider $\mathcal{A} = \{\mathbf{1}_{[\![\tau,T]\!]} : \tau \in \mathcal{T}\}$. Then (3.2) reduces to

$$\sup_{P \in \mathcal{M}(\mu,\nu)} \sup_{\tau \in \mathcal{T}} E^P\big[f(X_\tau)\big]. \tag{3.3}$$

(i) *American options.* If $\mathcal{T}$ consists of all $[0, T]$-valued $\mathbb{F}$-stopping times, then (3.3) is the robust American option pricing problem analyzed in [2].

(ii) *Bermudan options.* Bermudan options with exercise dates $0 \le T_1 < \cdots < T_n \le T$ are covered by choosing $\mathcal{T}$ to be the set of $\{T_1, \ldots, T_n\}$-valued $\mathbb{F}$-stopping times.

3.2. *Superhedging problem.* In the case of robust semistatic superhedging of American options, it is well known that a pricing-hedging duality can in general only hold if the seller of the option can adjust the dynamic part of his trading strategy after the option has been exercised; cf. [3], Section 3. In other words, the buyer has to communicate her decision of exercising to the seller at the time of exercising. The analog in our setting is that the seller observes A_t at time t. That is, his trading strategy can be "adapted" to the averaging process chosen by the buyer.

To make this precise, let $\hat{\Omega}$ be the cartesian product of Ω and the set of nonnegative, nondecreasing, càdlàg functions $a : [0, T] \to [0, 1]$ with $a(T) = 1$. As

$\hat{\Omega}$ is a subspace of the Skorokhod space $D([0,T];\mathbb{R}\times[0,1])$, we can equip it with the subspace Skorokhod topology and denote by $\hat{\mathcal{F}}$ the corresponding Borel σ-algebra. We write $\hat{\mathbb{F}}=(\hat{\mathcal{F}}_t)_{t\in[0,T]}$ for the (raw) filtration generated by the canonical process on $\hat{\Omega}$. For any process Z on $\hat{\Omega}$ and any averaging process A (on Ω), we define the process Z^A on Ω by

$$Z_t^A(\omega)=Z_t\big(\omega,A(\omega)\big),\quad \omega\in\Omega.$$

Note that if Z is $\hat{\mathbb{F}}$-adapted, then Z^A is $\mathbb{F}$-adapted, and if Z is càdlàg or of finite variation, then so is Z^A.

Next, we define a suitable set of paths for the hedging problem. Let $\Omega_{\mu,\nu}\subset\Omega$ denote the subset of paths which start in I, evolve in J, and are "captured" if they approach the boundary ∂J:

$$\begin{aligned}\Omega_{\mu,\nu}:=\big\{&\omega\in\Omega:\omega_0\in I,\ \omega_t\in J \text{ for all } t\in(0,T],\\ &\text{if } \omega_{t-}\in\partial J, \text{ then } \omega_u=\omega_{t-} \text{ for all } u\in[t,T], \text{ and}\\ &\text{if } \omega_t\in\partial J, \text{ then } \omega_u=\omega_t \text{ for all } u\in[t,T]\big\}.\end{aligned}\tag{3.4}$$

One can show that every martingale coupling between μ and ν is concentrated on $\Omega_{\mu,\nu}$:[5]

LEMMA 3.5. *$\Omega_{\mu,\nu}\in\mathcal{F}$ and $P[\Omega_{\mu,\nu}]=P[\Omega]$ for every $P\in\mathcal{M}(\mu,\nu)$.*

We are now ready to define the trading strategies for the robust superhedging problem.

DEFINITION 3.6. A *semistatic trading strategy* is a triplet (φ,ψ,H) consisting of a pair of functions $(\varphi,\psi)\in L^c(\mu,\nu)$ and a pair $H=(h_t,Y_t)_{t\in[0,T]}$ of $\hat{\mathbb{F}}$-adapted processes on $\hat{\Omega}$ such that

$$H^A:=\big(h^A,Y^A\big)\in L(\Omega_{\mu,\nu})\quad \text{for every averaging process } A.\tag{3.5}$$

The portfolio value at time T of a semistatic trading strategy is given by the sum of the static part with payoffs $\varphi(X_0)$ and $\psi(X_T)$ and the gains $H^A\diamond X_T$ of the dynamic part:

$$\varphi(X_0)+\psi(X_T)+H^A\diamond X_T.\tag{3.6}$$

The initial cost to set up this position is equal to the initial price of the static part:

$$\mu(\varphi)+\nu(\psi).\tag{3.7}$$

[5]The fact that μ and ν are concentrated on I and J, respectively, together with the martingale property implies that P-a.e. path has the first two properties in (3.4). The other two properties can be shown similar to the fact that nonnegative supermartingales are almost surely captured in zero (cf. [21], Lemma 7.31).

We now turn our attention to semistatic trading strategies which dominate the payoff (3.1) of our derivative security for each path in $\Omega_{\mu,\nu}$ and every averaging process in $\mathcal{A}$.

DEFINITION 3.7. A semistatic trading strategy (φ, ψ, H) is called a *semistatic superhedge (for f and $\mathcal{A}$)* if for every $A \in \mathcal{A}$,

$$f\left(\int_{[0,T]} X_t \,\mathrm{d}A_t\right) \le \varphi(X_0) + \psi(X_T) + H^A \diamond X_T \quad \text{on } \Omega_{\mu,\nu} \tag{3.8}$$

and

$$E^P\big[\varphi(X_0) + \psi(X_T) + H^A \diamond X_T\big] \le \mu(\varphi) + \nu(\psi), \quad P \in \mathcal{M}(\mu,\nu). \tag{3.9}$$

The set of semistatic superhedges for f and $\mathcal{A}$ is denoted by $\mathcal{D}_{\mu,\nu}(f,\mathcal{A})$.

The requirement (3.9) is an admissibility condition. It demands that for every $P \in \mathcal{M}(\mu,\nu)$, the portfolio value, consisting of both the static and the dynamic part, is a one-step P-supermartingale between the time at which the static part is set up and time T. In other words, the expectation of the terminal portfolio value (3.6) is less than or equal to the initial portfolio value (3.7).

We define the *robust superhedging price (for f and $\mathcal{A}$)* as the "minimal" initial capital required to set up a semistatic superhedge for f and $\mathcal{A}$:[6]

$$\mathbf{I}_{\mu,\nu}(f,\mathcal{A}) = \inf_{(\varphi,\psi,H)\in\mathcal{D}_{\mu,\nu}(f,\mathcal{A})} \big\{\mu(\varphi) + \nu(\psi)\big\}. \tag{3.10}$$

3.3. *Weak and strong duality.* Weak duality between the robust pricing and hedging problems is an immediate consequence of their definitions.

LEMMA 3.8 (Weak duality). *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be Borel and bounded from below by a ν-integrable concave function and let $\mathcal{A}$ be a nonempty set of averaging processes. Then*

$$\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) \le \mathbf{I}_{\mu,\nu}(f,\mathcal{A}).$$

PROOF. Let $P \in \mathcal{M}(\mu,\nu)$, $A \in \mathcal{A}$, and $(\varphi,\psi,H) \in \mathcal{D}_{\mu,\nu}(f,\mathcal{A})$ (there is nothing to show if this set is empty). Taking P-expectations in (3.8) and using (3.9) shows that $E^P[f(\int_{[0,T]} X_t \,\mathrm{d}A_t)] \le \mu(\varphi) + \nu(\psi)$.

This proves the claim as P, A, and (φ,ψ,H) were arbitrary. □

With an additional mild assumption on either $\mathcal{A}$ or f, we obtain strong duality and the existence of dual minimizers.

[6]We use the convention $\inf \varnothing = \infty$.

THEOREM 3.9. *Let $\mu \leq_c \nu$ be irreducible, let $f : \mathbb{R} \to [0, \infty]$ be Borel, and let $\mathcal{A}$ be a set of averaging processes. Suppose that one of the following two conditions holds*:

- *f is lower semicontinuous and $\mathcal{A}$ contains an interior averaging process*;
- *$\mathcal{A}$ contains a strictly interior averaging process.*

Then

$$\mathbf{S}_{\mu,\nu}(f, \mathcal{A}) = \mathbf{I}_{\mu,\nu}(f, \mathcal{A}) \in [0, \infty]$$

and this value is independent of $\mathcal{A}$ as long as one of the two conditions above holds. Moreover, if $\mathbf{I}_{\mu,\nu}(f, \mathcal{A}) < \infty$, then there exists an optimizer $(\varphi, \psi, H) \in \mathcal{D}_{\mu,\nu}(f, \mathcal{A})$ for $\mathbf{I}_{\mu,\nu}(f, \mathcal{A})$.

REMARK 3.10.

(i) For fixed f, the robust model-based price $\mathbf{S}_{\mu,\nu}(f, \mathcal{A})$ is invariant under the choice of the set $\mathcal{A}$ (as long as the assumptions of Theorem 3.9 hold). In particular, American, Bermudan and European options with intermediate maturity (cf. Examples 3.3–3.4) all have the same robust model-based price (because the corresponding sets $\mathcal{A}$ all contain a strictly interior averaging process). If f is lower semicontinuous, this extends to the Asian-style option of Example 3.3(i). If more than two marginals are given, then the robust model-based prices of these derivatives typically differ; see Example 6.3.

(ii) Derivatives of the form (3.1) that depend distinctly on X_0 and/or X_T such as $f(\frac{1}{2}(X_0 + X_T))$ are not covered by Theorem 3.9 ($\mathcal{A}$ does not contain an *interior* averaging process). In these cases, the robust model-based price is still bounded above by the corresponding robust model-based price of, say, the European-style derivative $f(X_{T/2})$. However, the inequality is typically strict; see Example 6.4.

REMARK 3.11.

(i) Theorem 3.9 can be extended to nonirreducible marginals along the lines of [8], Section 7.

(ii) Strong duality continues to hold if we restrict ourselves to finite variation strategies; cf. Remark 4.15 for an outline of the argument. It is an open question whether there is (in general), a dual minimizer (φ, ψ, H) with a dynamic part H of finite variation.

We defer the proof of Theorem 3.9 to the end of Section 4.4. The idea is as follows. We bound the pricing problem from below and the hedging problem from above by auxiliary maximization and minimization problems, respectively, and show that strong duality holds between those two auxiliary problems. Then all four problems have equal value and in particular strong duality for the pricing and

hedging problems holds. Moreover, we show that the auxiliary dual problem admits a minimizer and that every element in the dual space of the auxiliary problem gives rise to a semistatic superhedge with the same cost. Then, in particular, the minimizer of the auxiliary dual problem yields an optimal semistatic superhedge for f and $\mathcal{A}$ (which is independent of $\mathcal{A}$).

4. Auxiliary problems. Throughout this section, we fix an irreducible pair $\mu \leq_c \nu$ with domain (I, J) and a function $f : \mathbb{R} \to \overline{\mathbb{R}}$ which is bounded from below by a ν-integrable concave function.

The auxiliary primal and dual problems are formally derived in Section 4.1. They are rigorously introduced in Sections 4.2–4.3 and proved to be lower and upper bounds of the robust model-based price and the robust superhedging price, respectively. Their strong duality is proved in Section 4.4. Finally, structural properties of primal and dual optimizers of the auxiliary problems are studied in Section 4.5.

4.1. *Motivation.* The key property of payoffs of the form (3.1) is that the law of $\int_{[0,T]} X_t \, \mathrm{d}A_t$ under $P \in \mathcal{M}(\mu, \nu)$ is in convex order between μ and ν. In this section, we explain this observation and how it can be used to estimate the robust pricing problem from below and the robust superhedging problem from above.

Let $P \in \mathcal{M}(\mu, \nu)$ and let τ be a $[0, T]$-valued $\mathbb{F}$-stopping time. An application of the optional stopping theorem and Jensen's inequality shows that for any convex function ψ,

$$\mu(\psi) = E^P\big[\psi(X_0)\big] = E^P\big[\psi\big(E^P[X_\tau \mid \mathcal{F}_0]\big)\big] \leq E^P\big[\psi(X_\tau)\big] \quad \text{and}$$
$$\nu(\psi) = E^P\big[\psi(X_T)\big] \geq E^P\big[\psi\big(E^P[X_T \mid \mathcal{F}_\tau]\big)\big] = E^P\big[\psi(X_\tau)\big],$$

so that the law of X_τ under P is in convex order between μ and ν.

Using a time change argument and again Jensen's inequality and the optional stopping theorem, it can be shown that this property generalizes to the random variable $\int_{[0,T]} X_t \, \mathrm{d}A_t$ for an averaging process A.

LEMMA 4.1. *Let $P \in \mathcal{M}(\mu, \nu)$ and let A be an averaging process. Then the law of $\int_{[0,T]} X_t \, \mathrm{d}A_t$ under P is in convex order between μ and ν.*

In the sequel, we write $\mathbf{S} = \mathbf{S}_{\mu,\nu}(f, \mathcal{A})$ and $\mathbf{I} = \mathbf{I}_{\mu,\nu}(f, \mathcal{A})$ for brevity. Lemma 4.1 implies that

$$\mathbf{S} \leq \sup_{\mu \leq_c \theta \leq_c \nu} \theta(f) =: \widetilde{\mathbf{S}}.$$

We show in Section 4.2 that also the converse inequality holds under mild assumptions on f and $\mathcal{A}$. Thus, $\mathbf{S} = \widetilde{\mathbf{S}}$ and one is led to expect that $\mathbf{I} = \widetilde{\mathbf{I}}$ for a suitable dual problem $\widetilde{\mathbf{I}}$ to $\widetilde{\mathbf{S}}$.

Let us thus formally derive the Lagrange dual problem for $\widetilde{\mathbf{S}}$. Dualizing the constraint $\mu \leq_c \theta \leq_c \nu$ suggests to consider the Lagrangian

$$L(\theta, \psi_1, \psi_2) := \theta(f) + \big(\theta(\psi_1) - \mu(\psi_1)\big) + \big(\nu(\psi_2) - \theta(\psi_2)\big), \tag{4.1}$$

where *convex* functions ψ_1, ψ_2 are taken as Lagrange multipliers.[7] Then the Lagrange dual problem is

$$\widetilde{\mathbf{I}} = \inf_{\psi_1,\psi_2} \sup_{\theta} L(\theta, \psi_1, \psi_2) = \inf_{\psi_1,\psi_2} \sup_{\theta} \big\{\theta(f + \psi_1 - \psi_2) - \mu(\psi_1) + \nu(\psi_2)\big\},$$

where the infima are taken over convex functions and the suprema are taken over finite measures. Viewing the finite measure θ as a Lagrange multiplier for the constraint $f \leq -\psi_1 + \psi_2$ and relabeling $\varphi = -\psi_1$ and $\psi = \psi_2$, we obtain

$$\widetilde{\mathbf{I}} = \inf\big\{\mu(\varphi) + \nu(\psi) : \varphi \text{ concave, } \psi \text{ convex, and } \varphi + \psi \geq f\big\}. \tag{4.2}$$

In the precise definition of $\widetilde{\mathbf{I}}$ in Section 4.3, $\mu(\varphi) + \nu(\psi)$ is understood in the generalized sense of Definition 2.5 and the inequality $\varphi + \psi \geq f$ is required to hold on J. We then show that each feasible element (φ, ψ) for $\widetilde{\mathbf{I}}$ entails an element $(\varphi, \psi, H) \in \mathcal{D}_{\mu,\nu}(f, \mathcal{A})$ (Proposition 4.7), which implies that $\mathbf{I} \leq \widetilde{\mathbf{I}}$.

Combining the above with the weak duality inequality (Lemma 3.8) yields

$$\widetilde{\mathbf{S}} = \mathbf{S} \leq \mathbf{I} \leq \widetilde{\mathbf{I}}.$$

Hence, strong duality and dual attainment for the robust pricing and superhedging problems reduce to the same assertions for the simpler auxiliary problems, which are proved in Section 4.4.

4.2. *Auxiliary primal problem.* Consider the *auxiliary primal problem*

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \sup_{\mu \leq_c \theta \leq_c \nu} \theta(f), \tag{4.3}$$

where $\theta(f)$ is understood as the outer integral if f is not Borel-measurable. Under suitable conditions on f and $\mathcal{A}$, the primal value $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ is a lower bound for the robust model-based price (3.2).

PROPOSITION 4.2. *Let $\mathcal{A}$ be a set of averaging processes. Suppose that one of the following two sets of conditions holds*:

(i) *$\mathcal{A}$ contains an interior averaging process and f is lower semicontinuous and bounded from below by a ν-integrable concave function $\varphi : J \to \overline{\mathbb{R}}$;*

(ii) *$\mathcal{A}$ contains a strictly interior averaging process and f is Borel.*

[7] Note that the last two terms in (4.1) are nonnegative for all convex ψ_1, ψ_2 if and only if the primal constraint $\mu \leq_c \theta \leq_c \nu$ holds.

Then

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f) \le \mathbf{S}_{\mu,\nu}(f, \mathcal{A}).$$

The proof of Proposition 4.2 is given at the end of this section. It is based on the following construction of measures in $\mathcal{M}(\mu, \nu)$ under which the law of $\int_{[0,T]} X_t \, \mathrm{d}A_t$ equals (approximately or exactly) a given θ. This construction also highlights the importance of $\mathcal{A}$ containing an *interior* averaging process, which does not put any mass on the times 0 and T at which the marginal distributions of X are given; see Example 6.4 for a counterexample.

LEMMA 4.3. *Let $\mu \le_c \theta \le_c \nu$.*

(i) *There is a sequence $(P_n)_{n\ge1} \subset \mathcal{M}(\mu, \nu)$ such that*

$$\mathcal{L}^{P_n}\left(\int_{[0,T]} X_t \, \mathrm{d}A_t\right) \xrightarrow{n\to\infty} \theta \quad \textit{weakly}$$

for every interior averaging process A.

(ii) *If A is a strictly interior averaging process, then there is $P \in \mathcal{M}(\mu, \nu)$ (depending on A) such that $\mathcal{L}^P(\int_{[0,T]} X_t \, \mathrm{d}A_t) = \theta$.*

PROOF. (i) By the two-step adaptation of Proposition 2.1, there exists a measure $Q \in \mathcal{M}^d(\mu, \theta, \nu)$. For all n large enough, let $\iota^n : \mathbb{R}^3 \to \Omega$ be the embedding of $\mathbb{R}^3$ in Ω which maps (y_1, y_2, y_3) to the piecewise constant path

$$[0, T] \ni t \mapsto y_1 \mathbf{1}_{[0,\frac{1}{n})}(t) + y_2 \mathbf{1}_{[\frac{1}{n},T)}(t) + y_3 \mathbf{1}_{\{T\}}(t) \tag{4.4}$$

(which jumps (at most) at times $\frac{1}{n}$ and T), and denote by $P_n := Q \circ (\iota^n)^{-1}$ the associated pushforward measure. Then $P_n \in \mathcal{M}(\mu, \nu)$ by the corresponding properties of Q. Moreover, denoting the canonical process on $\mathbb{R}^3$ by (Y_1, Y_2, Y_3) and setting $A^n = A \circ \iota^n$ for an interior averaging process A, we have

$$\begin{aligned}\int_{[0,T]} (\iota^n)_t \, \mathrm{d}A^n_t - Y_2 &= Y_1 A^n_{\frac{1}{n}-} + Y_2(A^n_{T-} - A^n_{\frac{1}{n}-}) + Y_3 \Delta A^n_T - Y_2 A_T \\ &= (Y_1 - Y_2) A^n_{\frac{1}{n}-} + (Y_3 - Y_2) \Delta A^n_T \\ &= (Y_1 - Y_2) A^n_{\frac{1}{n}-} \quad \text{on } \mathbb{R}^3,\end{aligned} \tag{4.5}$$

where we use the properties $A_T = 1$ and $\Delta A_T = 0$ of an interior averaging process.

By construction, the law of $\int_{[0,T]} (\iota^n)_t \, \mathrm{d}A^n_t$ under Q coincides with the law of $\int_{[0,T]} X_t \, \mathrm{d}A_t$ under P_n and the law of Y_2 under Q is θ. It thus suffices to prove that the right-hand side in (4.5) converges to zero in $L^1(Q)$ as $n \to \infty$. To this end, note that $|Y_1 - Y_2| \le |Y_1| + |Y_2|$ is Q-integrable because μ and θ have finite first moments. Thus, by dominated convergence, it is enough to show that $A^n_{\frac{1}{n}-} \to 0$

pointwise as $n \to 0$. So fix $(y_1, y_2, y_3) \in \mathbb{R}^3$. Since A is $\mathbb{F}$-adapted, $A_{\frac{1}{n}-}(\omega)$ only depends on the values of the path ω on the interval $[0, \frac{1}{n})$. In view of the embedding (4.4), this means that

$$A^n_{\frac{1}{n}-}(y_1, y_2, y_3) = A_{\frac{1}{n}-}\big(\iota^n(y_1, y_2, y_3)\big) = A_{\frac{1}{n}-}(y_1 \mathbf{1}_{[0,T]}),$$

where $y_1 \mathbf{1}_{[0,T]}$ denotes the constant path at y_1. Hence, the asserted pointwise convergence follows from the fact that $A_0 = 0$ and A is right continuous.

(ii) If A is a strictly interior averaging process, then the last expression in (4.5) is identically zero for n large enough and setting $P = P_n$ gives the desired result. □

REMARK 4.4. Part (i) of Lemma 4.3 remains true if we restrict ourselves to martingale measures with almost surely continuous paths. The analog of part (ii) for continuous martingales requires the additional assumption that there exists $t < T$ such that $A_t \equiv 1$.

The main ingredient for this assertion is [9], Theorem 11: for every discrete time-martingale $\{Y_n\}_{n\geq 0}$, there is a continuous-time martingale $\{Z_t\}_{t\geq 0}$ with continuous sample paths such that the processes $\{Y_n\}_{n\geq 0}$ and $\{Z_n\}_{n\geq 0}$ have the same (joint) distribution.

PROOF OF PROPOSITION 4.2. Let $\mu \leq_c \theta \leq_c \nu$. Assume first that condition (ii) holds and let A be a strictly interior averaging process. Then by Lemma 4.3(ii), there is $P \in \mathcal{M}(\mu, \nu)$ such that $\mathcal{L}^P(\int_{[0,T]} X_t \, dA_t) = \theta$. Hence,

$$\theta(f) = E^P\left[f\left(\int_{[0,T]} X \, dA\right)\right] \leq \mathbf{S}_{\mu,\nu}(f, \mathcal{A}).$$

As θ was arbitrary, the claim follows.

Next, assume instead that condition (i) holds and let A be an interior averaging process and φ as in condition (i). By Lemma 4.3(i), there is a sequence $(P_n)_{n\in\mathbb{N}} \subset \mathcal{M}(\mu, \nu)$ such that $\theta_n := \mathcal{L}^{P_n}(\int_{[0,T]} X_t \, dA_t) \to \theta$ weakly. Define $f_k = f \vee (-k)$, $k \geq 1$. Then f_k is bounded from below and lower semicontinuous, so $\liminf_{n\to\infty} \theta_n(f_k) \geq \theta(f_k)$ by the Portmanteau theorem.

Fix $\varepsilon > 0$. Choose first k large enough such that $\nu((\varphi + k)^-) \leq \frac{\varepsilon}{2}$ and then N large enough such that $\theta_n(f_k) - \theta(f_k) \geq -\frac{\varepsilon}{2}$ for all $n \geq N$. Using that $0 \leq f_k - f \leq (\varphi + k)^-$ and that $(\varphi + k)^-$ is convex, we obtain for $n \geq N$,

$$\begin{aligned}\theta_n(f) - \theta(f) &= \theta_n(f - f_k) + \big(\theta_n(f_k) - \theta(f_k)\big) + \theta(f_k - f) \\ &\geq -\theta_n\big((\varphi + k)^-\big) - \frac{\varepsilon}{2} \geq -\nu\big((\varphi + k)^-\big) - \frac{\varepsilon}{2} \geq -\varepsilon.\end{aligned}$$

Thus, $\liminf_{n\to\infty} \theta_n(f) \geq \theta(f)$. Now the claim follows from

$$\theta(f) \leq \liminf_{n\to\infty} \theta_n(f) = \liminf_{n\to\infty} E^{P_n}\left[f\left(\int_{[0,T]} X_t \, dA_t\right)\right] \leq \mathbf{S}_{\mu,\nu}(f, \mathcal{A}).$$ □

4.3. *Auxiliary dual problem.* Consider the *auxiliary dual problem*

$$\widetilde{\mathbf{I}}_{\mu,\nu}(f) = \inf_{(\varphi,\psi)\in\widetilde{\mathcal{D}}_{\mu,\nu}(f)} \{\mu(\varphi) + \nu(\psi)\}, \tag{4.6}$$

where $\widetilde{\mathcal{D}}_{\mu,\nu}(f)$ denotes the set of $(\varphi, \psi) \in L^c(\mu, \nu)$ with concave $\varphi : J \to \overline{\mathbb{R}}$ and convex $\psi : J \to \overline{\mathbb{R}}$ such that $\varphi + \psi \geq f$ on J.

The dual value $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$ is an upper bound for the robust superhedging price (3.10).

PROPOSITION 4.5. *Let $f : \mathbb{R} \to [0, \infty]$ be Borel. Then $\mathbf{I}_{\mu,\nu}(f, \mathcal{A}) \leq \widetilde{\mathbf{I}}_{\mu,\nu}(f)$.*

Proposition 4.5 follows immediately from the next result (Proposition 4.7) which shows that every $(\varphi, \psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ gives rise to a semistatic superhedge for f and $\mathcal{A}$. More precisely, the semistatic superhedge is of the form (φ, ψ, H) and the dynamic part H can be explicitly written in terms of the "derivatives" of φ and ψ.

Given a convex function $\psi : J \to \mathbb{R}$, a Borel function $\psi' : I \to \mathbb{R}$ is called a *subderivative* of ψ if for every $x_0 \in I$, $\psi'(x_0)$ belongs to the subdifferential of ψ at x_0, that is,

$$\psi(x) - \psi(x_0) \geq \psi'(x_0)(x - x_0), \quad x \in J.$$

Symmetrically, for a concave function $\varphi : J \to \mathbb{R}$, a Borel function $\varphi' : I \to \mathbb{R}$ is called a *superderivative* of φ if $-\varphi'$ is a subderivative of $-\varphi$.

REMARK 4.6. If $(\varphi, \psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ and $f > -\infty$ on J, then φ and ψ are both finite (so that sub and superderivatives are well defined). Indeed, we already know from Remark 2.8 that ψ is finite on J. Moreover, $\varphi < \infty$ on J by Lemma 2.7(i) and if $f > -\infty$ on J, then $\varphi \geq f - \psi > -\infty$, so that also φ is finite on J.

PROPOSITION 4.7. *Let $f : \mathbb{R} \to [0, \infty]$ be Borel and let $(\varphi, \psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$. Denoting the canonical process on $\hat{\Omega}$ by (X, A), define the $\hat{\mathbb{F}}$-adapted process $h = (h_t)_{t\in[0,T]}$ (on $\hat{\Omega}$) by*

$$\begin{aligned} h_0 &= \varphi'(X_0)(1 - A_0) - \psi'(X_0)A_0, \\ h_t &= -\varphi'(X_0) - \psi'(X_t), \quad t \in (0, T], \end{aligned} \tag{4.7}$$

where φ' is any superderivative of φ, ψ' is any subderivative of ψ, and we set $\varphi' = \psi' = 0$ on $\mathbb{R} \setminus I$. Set $H = (h, A)$. Then $(\varphi, \psi, H) \in \mathcal{D}_{\mu,\nu}(f, \mathcal{A})$ for any nonempty set $\mathcal{A}$ of averaging processes.

The proof of Proposition 4.7, given at the end of this section, relies on the following two technical lemmas. The definition of $\Omega_{\mu,\nu}$ in (3.4) is crucial for the first one. We recall that (real-valued) càdlàg functions are bounded on compact intervals.

LEMMA 4.8. *Let ψ and ψ' be as in Proposition 4.7. For each $\omega \in \Omega_{\mu,\nu}$, the function $[0,T] \ni t \mapsto (\omega_T - \omega_t)\psi'(\omega_t)$ is bounded.*

PROOF. Fix $\omega \in \Omega_{\mu,\nu}$ and write $I = (l,r)$ with $l, r \in \overline{\mathbb{R}}$. We consider three cases: (i) $J = I$, (ii) $J = [l,r)$, and (iii) $J = [l,r]$. The case $(l,r]$ is symmetric to (ii).

(i) Suppose that $J = I = (l,r)$. We claim that ω evolves in a compact (and hence strict) subset of I. Suppose for the sake of contradiction hat $\inf_{t\in[0,T]} \omega_t = l \in [-\infty,\infty)$. Then there is a sequence $(t_n)_{n\in\mathbb{N}} \subset [0,T]$ such that $\lim_{n\to\infty} \omega_{t_n} = l$. Passing to a subsequence if necessary, this sequence may be chosen to be either (strictly) increasing or nonincreasing to a limit $t^\star \in [0,T]$. Then, as ω is càdlàg, $\omega_{t^\star -} = l$ or $\omega_{t^\star} = l$. But then $\omega_{t^\star} = l$ in any case by the definition of $\Omega_{\mu,\nu}$, a contradiction to $\omega_{t^\star} \in J = I$. Thus, $\inf_{t\in[0,T]} \omega_t > l$ and symmetrically $\sup_{t\in[0,T]} \omega_t < r$. This proves the claim. It follows that $(\omega_T - \omega_t)\psi'(\omega_t)$ is bounded over $t \in [0,T]$ because the subderivative ψ' is bounded on compact subsets of I.

(ii) Suppose that $J = [l,r)$, i.e., ν has an atom in $l > -\infty$. If ω evolves in I, then we can argue as in (i). We may thus assume that $t^\star := \inf\{t \in [0,T] : \omega_t = l\} \in (0,T]$. Then, as ω is càdlàg and by the definition of $\Omega_{\mu,\nu}$, we have $\omega_u = l$ for all $u \in [t^\star, T]$). In particular, $\omega_T = l$ and it is enough to show that $[0,t^\star) \ni t \mapsto (\omega_T - \omega_t)\psi'(\omega_t)$ is bounded.

We can argue similarly as in (i) that $r' := \sup_{t\in[0,T]} \omega_t < r$, so that the path ω evolves in the compact interval $[l,r']$. Because ψ is convex and finite on (l,r), ψ' is bounded from above on $[l,r']$. It follows that $t \mapsto (\omega_T - \omega_t)\psi'(\omega_t)$ is bounded from below on $[0,t^\star)$. To show that this function is also bounded from above, we observe that by the convexity of ψ,

$$(\omega_T - \omega_t)\psi'(\omega_t) \le \psi(\omega_T) - \psi(\omega_t) = \psi(l) - \psi(\omega_t). \tag{4.8}$$

Now $\psi(l)$ is finite because ν has an atom at l, and ψ is bounded from below on $[l,r']$ because it is finite and convex on the compact interval $[l,r']$. Using this in (4.8) shows the assertion.

(iii) Suppose that $J = [l,r]$, that is, ν has atoms at $l > -\infty$ and $r < \infty$. As in (ii), we may assume that ω hits one of the endpoints of J before T. By symmetry, we may assume that ω hits l. By definition of $\Omega_{\mu,\nu}$, the path ω is then bounded away from the right endpoint r (otherwise it would be captured in r), that is, $\sup_{t\in[0,T]} \omega_t < r$. Now the same argument as in (ii) proves the assertion. □

The second technical lemma is an adaptation of [8], Remark 4.10, to our setting. It is used to show the admissibility condition (3.9) of the semistatic trading strategy in Proposition 4.7.

LEMMA 4.9. *Let $(\varphi,\psi) \in L^c(\mu,\nu)$ and let $g_0, g_1 : J \to \mathbb{R}$ be Borel. Let τ be a $[0,T]$-valued $\mathbb{F}$-stopping time such that*

$$\varphi(X_0) + \psi(X_T) + g_0(X_0)(X_\tau - X_0) + g_1(X_\tau)(X_T - X_\tau) \tag{4.9}$$

is bounded from below on $\Omega_{\mu,\nu}$. Then for all $P \in \mathcal{M}(\mu,\nu)$,

$$E^P\big[\varphi(X_0)+\psi(X_T)+g_0(X_0)(X_\tau-X_0)+g_1(X_\tau)(X_T-X_\tau)\big]=\mu(\varphi)+\nu(\psi).$$

PROOF. Let χ be a concave moderator for (φ,ψ) with respect to $\mu \le_c \nu$ and let θ be the law of X_τ. By optional stopping, $\mu \le_c \theta \le_c \nu$. We expand (4.9) to

$$\begin{aligned}(\varphi-\chi)(X_0)+(\psi+\chi)(X_T)+\big[&\chi(X_0)-\chi(X_T)\\&+g_0(X_0)(X_\tau-X_0)+g_1(X_\tau)(X_T-X_\tau)\big],\end{aligned}\tag{4.10}$$

and observe that the first two terms are P-integrable. Then the assumed lower bound yields that the last term has a P-integrable negative part. We can therefore apply Fubini's theorem and evaluate its integral iteratively. To this end, let Q be the law of (X_0,X_τ,X_T) on the canonical space $\mathbb{R}^3$ with a disintegration

$$\mathrm{d}Q=\mu(\mathrm{d}x_0)\otimes\kappa_0(x_0,\mathrm{d}x_1)\otimes\kappa_1(x_0,x_1,\mathrm{d}x_2)$$

for martingale kernels κ_0 and κ_1. In view of the definition of $\mu(\varphi)+\nu(\psi)$ in (2.4), we have to show that the P-expectation of the last term in (4.10) is $(\mu-\nu)(\chi)$.

To this end, we observe that for $\mu\otimes\kappa_0$-a.e. $(x_0,x_1)\in J^2$,

$$\begin{aligned}&\int_J\big[\chi(x_0)-\chi(x_2)+g_0(x_0)(x_1-x_0)+g_1(x_1)(x_2-x_1)\big]\kappa_1(x_0,x_1,\mathrm{d}x_2)\\&=\int_J\big[\chi(x_0)-\chi(x_2)+g_0(x_0)(x_1-x_0)\big]\kappa_1(x_0,x_1,\mathrm{d}x_2)\\&=\chi(x_0)-\int_J\chi(x_2)\kappa_1(x_0,x_1,\mathrm{d}x_2)+g_0(x_0)(x_1-x_0).\end{aligned}\tag{4.11}$$

Integrating the left-hand side of (4.11) against $\mu\otimes\kappa_0$ gives the P-expectation of the last term in (4.10). It thus remains to show that the corresponding integral of the right-hand side equals $(\mu-\nu)(\chi)$. Integrating the right-hand side of (4.11) first against $\kappa_0(x_0,\mathrm{d}x_1)$ yields for μ-a.e. $x_0\in J$,

$$\chi(x_0)-\int_J\chi(x_2)\kappa(x_0,\mathrm{d}x_2),\tag{4.12}$$

where $\kappa(x_0,\cdot)=\int_J\kappa_1(x_1,\cdot)\kappa_0(x_0,\mathrm{d}x_1)$ is again a martingale kernel. Finally, the integral of (4.12) against μ is

$$\int_J\left[\chi(x_0)-\int_J\chi(x_2)\kappa(x_0,\mathrm{d}x_2)\right]\mu(\mathrm{d}x_0).$$

Noting that $\mu\otimes\kappa$ is a disintegration of a one-step martingale measure on $\mathbb{R}^2$ with marginals μ and ν, the last term equals $(\mu-\nu)(\chi)$ by Lemma 2.4. □

PROOF OF PROPOSITION 4.7. First, we show that (φ,ψ,H) is a semistatic trading strategy. As h and A are clearly $\hat{\mathbb{F}}$-adapted and $(\varphi,\psi)\in L^c(\mu,\nu)$ by assumption, it remains to check condition (3.5) (with Y^A replaced by A). So fix an

averaging process A and note that $H^A = (h^A, A)$. The only nontrivial part in proving $H^A \in L(\Omega_{\mu,\nu})$ is to show that $(X_T - X_t)h_t^A$ is $\mathrm{d}A$-integrable on $(0, T]$ for each path in $\Omega_{\mu,\nu}$. To this end, note that $\varphi'(X_0)$ and $\psi'(X_0)$ are finite because $X_0 \in I$. It thus suffices to show that $(X_T - X_t)\psi'(X_t)$ is bounded on $[0, T]$ for each path in $\Omega_{\mu,\nu}$; this is the content of Lemma 4.8.

Second, we show the superhedging property (3.8). Fix an averaging process A and a path in $\Omega_{\mu,\nu}$. To ease the notation, we write h instead of h^A in the following. Note, however, that h^A has the same formal expression as h in (4.7), but with A being the fixed averaging process (and not the second component of the canonical process on $\hat{\Omega}$).

Using the definitions of $H \diamond X_T$ and h as well as the fact that $A_0 = \Delta A_0$, we obtain

$$\begin{aligned} H \diamond X_T &= (X_T - X_0)h_0 + \int_{(0,T]} (X_T - X_t)h_t \,\mathrm{d}A_t \\ &= (X_T - X_0)\varphi'(X_0) - \int_{[0,T]} (X_T - X_t)\big(\varphi'(X_0) + \psi'(X_t)\big)\,\mathrm{d}A_t. \end{aligned}$$

Then, using that $\mathrm{d}A$ is a probability measure on $[0, T]$, the concavity of φ and the convexity of ψ, and Jensen's inequality, we can estimate

$$\begin{aligned} H \diamond X_T &= \int_{[0,T]} \varphi'(X_0)(X_t - X_0)\,\mathrm{d}A_t - \int_{[0,T]} \psi'(X_t)(X_T - X_t)\,\mathrm{d}A_t \\ &\geq \varphi'(X_0)\bigg(\int_{[0,T]} X_t \,\mathrm{d}A_t - X_0\bigg) - \int_{[0,T]} \big(\psi(X_T) - \psi(X_t)\big)\,\mathrm{d}A_t \\ &\geq \varphi\bigg(\int_{[0,T]} X_t \,\mathrm{d}A_t\bigg) - \varphi(X_0) - \psi(X_T) + \int_{[0,T]} \psi(X_t)\,\mathrm{d}A_t \\ &\geq \varphi\bigg(\int_{[0,T]} X_t \,\mathrm{d}A_t\bigg) - \varphi(X_0) - \psi(X_T) + \psi\bigg(\int_{[0,T]} X_t \,\mathrm{d}A_t\bigg). \end{aligned} \tag{4.13}$$

Rearranging terms and using that $\varphi + \psi \geq f$ on J, we find

$$\varphi(X_0) + \psi(X_T) + H \diamond X_T \geq f\bigg(\int_{[0,T]} X_t \,\mathrm{d}A_t\bigg).$$

Third, we show the admissibility condition (3.9). Fix an averaging process A and $P \in \mathcal{M}(\mu, \nu)$. Define the family of $\mathbb{F}$-stopping times C_s, $s \in (0, 1)$, by

$$C_s = \inf\big\{t \in [0, T] : A_t > s\big\}$$

and note that $0 \leq C_s \leq T$ for $s \in (0, 1)$ because $A_T = 1$. Then using the family C_s as a time change (cf. [24], Proposition 0.4.9) for the integral in (4.13) yields

$$\begin{aligned} &\varphi(X_0) + \psi(X_T) + H \diamond X_T \\ &\quad = \int_0^1 \big\{\varphi(X_0) + \psi(X_T) + \varphi'(X_0)(X_{C_s} - X_0) \\ &\qquad - \psi'(X_{C_s})(X_T - X_{C_s})\big\}\,\mathrm{d}s. \end{aligned} \tag{4.14}$$

Now, suppose that the integrand in (4.14) is bounded from below, uniformly over $s \in (0,1)$ and $\omega \in \Omega_{\mu,\nu}$. Then by Lemma 4.9, the P-expectation of the integrand equals $\mu(\varphi) + \nu(\psi)$ for each $s \in (0,1)$. Using this together with Tonelli's theorem and (4.14) gives

$$E^P\big[\varphi(X_0) + \psi(X_T) + H \diamond X_T\big] = \mu(\varphi) + \nu(\psi),$$

so that (3.9) holds.

It remains to show that the integrand in (4.14) is uniformly bounded from below. This follows from concavity of φ and convexity of ψ together with the fact that $\varphi + \psi \geq f \geq 0$ on J:

$$\begin{aligned}&\varphi(X_0) + \psi(X_T) + \varphi'(X_0)(X_t - X_0) - \psi'(X_t)(X_T - X_t)\\ &\quad\geq \varphi(X_t) + \psi(X_t) \geq f(X_t) \geq 0, \quad t \in [0,T].\end{aligned}$$

This completes the proof. □

4.4. *Duality.* We now turn to the duality between the auxiliary problems $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ and $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$.

THEOREM 4.10. *Let $\mu \leq_c \nu$ be irreducible with domain (I,J) and let $f : \mathbb{R} \to [0,\infty]$.*

(i) *If f is upper semianalytic, then $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \widetilde{\mathbf{I}}_{\mu,\nu}(f) \in [0,\infty]$.*
(ii) *If $\widetilde{\mathbf{I}}_{\mu,\nu}(f) < \infty$, then there exists a dual minimizer $(\varphi,\psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$.*

A couple of remarks are in order.

REMARK 4.11. We only state the duality for one irreducible component. One can formulate and prove the full duality for arbitrary marginals $\mu \leq_c \nu$ in analogy to [8], Section 7. We omit the details in the interest of brevity.

REMARK 4.12. The lower bound on f in Theorem 4.10 can be relaxed. Indeed, suppose that $f : \mathbb{R} \to \overline{\mathbb{R}}$ is upper semianalytic and bounded from below by an affine function g.

We first consider the primal problem. Because g is affine and any $\mu \leq_c \theta \leq_c \nu$ has the same mass and barycenter as μ,

$$\theta(f-g) = \theta(f) - \theta(g) = \theta(f) - \mu(g).$$

Thus,

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f-g) = \widetilde{\mathbf{S}}_{\mu,\nu}(f) - \mu(g). \tag{4.15}$$

Regarding the dual problem, we note that $(\varphi,\psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f-g)$ if and only if $(\varphi+g,\psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ and that by Lemma 2.7(iii),

$$\mu(\varphi) + \nu(\psi) = \big\{\mu(\varphi+g) + \nu(\psi)\big\} - \mu(g).$$

Hence,

$$\widetilde{\mathbf{I}}_{\mu,\nu}(f-g)=\widetilde{\mathbf{I}}_{\mu,\nu}(f)-\mu(g). \tag{4.16}$$

Because $f-g$ is nonnegative, the left-hand sides of (4.15)–(4.16) coincide by Theorem 4.10(i). Therefore, $\widetilde{\mathbf{S}}_{\mu,\nu}(f)=\widetilde{\mathbf{I}}_{\mu,\nu}(f)\in(-\infty,\infty]$.

Moreover, if $\widetilde{\mathbf{I}}_{\mu,\nu}(f)<\infty$, then also $\widetilde{\mathbf{I}}_{\mu,\nu}(f-g)<\infty$ and a dual minimizer $(\varphi,\psi)\in\widetilde{\mathcal{D}}_{\mu,\nu}(f-g)$ for $\widetilde{\mathbf{I}}_{\mu,\nu}(f-g)$ exists by Theorem 4.10(ii). Now the above shows that $(\varphi+g,\psi)\in\widetilde{\mathcal{D}}_{\mu,\nu}(f)$ is a dual minimizer for $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$.

The proof of Theorem 4.10 is based on several preparatory results. We start with the crucial closedness property of the dual space in the spirit of [8], Proposition 5.2.

PROPOSITION 4.13. *Let $\mu\le_c\nu$ be irreducible with domain (I,J), let f, $f_n: J\to[0,\infty]$ be such that $f_n\to f$ pointwise, and let $(\varphi_n,\psi_n)\in\widetilde{\mathcal{D}}_{\mu,\nu}(f_n)$ with $\sup_n\{\mu(\varphi_n)+\nu(\psi_n)\}<\infty$. Then there is $(\varphi,\psi)\in\widetilde{\mathcal{D}}_{\mu,\nu}(f)$ such that $\mu(\varphi)+\nu(\psi)\le\liminf_{n\to\infty}\{\mu(\varphi_n)+\nu(\psi_n)\}$.*

PROOF. Let $h_n=\varphi_n'\colon I\to\mathbb{R}$ be a superderivative of the concave function φ_n. As

$$\varphi_n(x)+\psi_n(y)+h_n(x)(y-x)\ge\varphi_n(y)+\psi_n(y)\ge f_n(y)\ge 0,\quad (x,y)\in I\times J,$$

(φ_n,ψ_n,h_n) is in the dual space $\mathcal{D}^c_{\mu,\nu}(0)$ of [8]. Hence, following the line of reasoning in the proof of [8], Proposition 5.2, (which is based on Komlos' lemma; we recall that convex combinations of convex (concave) functions are again convex (concave)), we may assume without loss of generality that

$$\varphi_n\to\bar\varphi\quad\mu\text{-a.e.}\quad\text{and}\quad\psi_n\to\bar\psi\quad\nu\text{-a.e.}$$

for some $(\bar\varphi,\bar\psi)\in L^c(\mu,\nu)$. Moreover, the arguments in [8] also show that $\mu(\bar\varphi)+\nu(\bar\psi)\le\liminf_{n\to\infty}\{\mu(\varphi_n)+\nu(\psi_n)\}$.

Now, define the functions $\varphi,\psi: J\to\overline{\mathbb{R}}$ by $\varphi:=\liminf_{n\to\infty}\varphi_n$ and $\psi:=\limsup_{n\to\infty}\psi_n$. Then φ is convex, ψ is concave, $\varphi=\bar\varphi$ μ-a.e., and $\psi=\bar\psi$ ν-a.e. In particular, $(\varphi,\psi)\in L^c(\mu,\nu)$ and $\mu(\varphi)+\nu(\psi)\le\liminf_{n\to\infty}\{\mu(\varphi_n)+\nu(\psi_n)\}$. Furthermore, as $\varphi_k+\psi_k\ge f_k$ on J, we have for each n that

$$\inf_{k\ge n}\varphi_k+\sup_{k\ge n}\psi_k\ge\inf_{k\ge n}(\varphi_k+\psi_k)\ge\inf_{k\ge n}f_k\quad\text{on } J.$$

Sending $n\to\infty$ gives $\varphi+\psi\ge f$. In summary, $(\varphi,\psi)\in\widetilde{\mathcal{D}}_{\mu,\nu}(f)$. □

We proceed to show strong duality for bounded upper semicontinuous functions.

LEMMA 4.14. *Let $f:\mathbb{R}\to[0,\infty]$ be bounded and upper semicontinuous. Then $\widetilde{\mathbf{S}}_{\mu,\nu}(f)=\widetilde{\mathbf{I}}_{\mu,\nu}(f)$.*

The proof is based on a Hahn–Banach separation argument similar to [8], Lemma 6.4.

PROOF. We first show the weak duality inequality. Let $\mu \le_c \theta \le_c \nu$ and $(\varphi, \psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$. In particular, $\varphi + \psi$ is bounded from below. Then by Lemma 2.9(iii)–(iv),

$$\theta(f) \le \theta(\varphi + \psi) = \theta(\varphi) + \theta(\psi) \le \theta(\varphi) + \nu(\psi) \le \mu(\varphi) + \nu(\psi), \tag{4.17}$$

and the inequality $\widetilde{\mathbf{S}}_{\mu,\nu}(f) \le \widetilde{\mathbf{I}}_{\mu,\nu}(f)$ follows.

The converse inequality is based on a Hahn–Banach argument, so let us introduce a suitable space. By the de la Vallée-Poussin theorem, there is an increasing convex function $\zeta_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ of superlinear growth such that $x \mapsto \zeta_\nu(|x|)$ is ν-integrable. Now, set $\zeta(x) = 1 + \zeta_\nu(|x|)$, $x \in \mathbb{R}$, and denote by C_ζ the space of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ such that f/ζ vanishes at infinity. We endow C_ζ with the norm $\|f\|_\zeta := \|f/\zeta\|_\infty$. With this notation, the same arguments as in the proof of [8], Lemma 6.4, show that the dual space C_ζ^* of continuous linear functionals on C_ζ can be represented by finite signed measures.

Fix $f \in C_\zeta$. Then

$$-\zeta(x)\|f\|_\zeta \le f(x) \le \zeta(x)\|f\|_\zeta, \quad x \in J. \tag{4.18}$$

Because ζ_ν is convex and $x \mapsto \zeta_\nu(|x|)$ is ν-integrable, we have $\theta(\zeta) \le \nu(\zeta) < \infty$ for all $\mu \le_c \theta \le_c \nu$. This together with (4.18) shows that $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ is finite. Thus, adding a suitable constant to f, we may assume that $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = 0$. For the following Hahn–Banach argument, we consider the convex cone

$$K := \{g \in C_\zeta : \widetilde{\mathbf{I}}_{\mu,\nu}(g) \le 0\}.$$

Proposition 4.13 implies that K is closed.

Suppose for the sake of contradiction that $\widetilde{\mathbf{I}}_{\mu,\nu}(f) > 0$. Then, by the Hahn–Banach theorem, K and f can be strictly separated by a continuous linear functional on C_ζ. That is, there is a finite signed measure ρ such that $\rho(f) > 0$ and $\rho(g) \le 0$ for all $g \in K$. For any compactly supported nonnegative continuous function $g \in C_\zeta$, we have $\widetilde{\mathbf{I}}_{\mu,\nu}(-g) \le 0$. That is, $-g \in K$, and hence $\rho(-g) \le 0$. This shows that ρ is a (nonnegative) finite measure. Multiplying ρ by a positive constant if necessary, we may assume that ρ has the same mass as μ and ν. Next, let ψ be convex and of linear growth. Then $\psi\,\nu(\mathbb{R}) - \nu(\psi) \in K$ and $-\psi\,\mu(\mathbb{R}) + \mu(\psi) \in K$. Using that $\rho \le 0$ for these two functions yields $\mu(\psi) \le \rho(\psi) \le \nu(\psi)$. We conclude that $\mu \le_c \rho \le_c \nu$. But now $\rho(f) > 0$ contradicts $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = 0$. Thus, $\widetilde{\mathbf{I}}_{\mu,\nu}(f) \le \widetilde{\mathbf{S}}_{\mu,\nu}(f)$.

Finally, let f be bounded and upper semicontinuous and choose $f_n \in C_b(\mathbb{R}) \subseteq C_\zeta$ such that $f_n \searrow f$. By the above, we have $\widetilde{\mathbf{S}}_{\mu,\nu}(f_n) = \widetilde{\mathbf{I}}_{\mu,\nu}(f_n)$ for all n. We show below that $\lim_{n\to\infty} \widetilde{\mathbf{S}}_{\mu,\nu}(f_n) = \widetilde{\mathbf{S}}_{\mu,\nu}(f)$. Using this and the monotonicity of $\widetilde{\mathbf{I}}_{\mu,\nu}$, we obtain

$$\widetilde{\mathbf{I}}_{\mu,\nu}(f) \le \lim_{n\to\infty} \widetilde{\mathbf{I}}_{\mu,\nu}(f_n) = \lim_{n\to\infty} \widetilde{\mathbf{S}}_{\mu,\nu}(f_n) = \widetilde{\mathbf{S}}_{\mu,\nu}(f) \le \widetilde{\mathbf{I}}_{\mu,\nu}(f).$$

So strong duality holds for bounded upper semicontinuous functions.

It remains to argue that $\lim_{n\to\infty} \widetilde{\mathbf{S}}_{\mu,\nu}(f_n) = \widetilde{\mathbf{S}}_{\mu,\nu}(f)$. We show more generally that $\widetilde{\mathbf{S}}_{\mu,\nu}$ is continuous along decreasing sequences of bounded upper semicontinuous functions. So let $f_n \searrow f$ be a convergent sequence of bounded upper semicontinuous functions. Fix $\varepsilon > 0$ and set $\ell := \lim_{n\to\infty} \widetilde{\mathbf{S}}_{\mu,\nu}(f_n)$. Then for each n, $\ell \le \mathbf{S}_{\mu,\nu}(f_n) < \infty$, and thus the set

$$A_n := \{\mu \le_c \theta \le_c \nu : \theta(f_n) \ge \ell - \varepsilon\}$$

is nonempty. Moreover, each A_n is a closed subset of the weakly compact set $\{\theta : \mu \le_c \theta \le_c \nu\}$ and $A_{n+1} \subseteq A_n$. Therefore, there exists a θ' in the intersection $\bigcap_{n\ge1} A_n$. We then obtain by monotone convergence that

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f) \ge \theta'(f) = \lim_{n\to\infty} \theta'(f_n) \ge \ell - \varepsilon.$$

This implies that $\widetilde{\mathbf{S}}_{\mu,\nu}(f) \ge \ell$ as ε was arbitrary. The converse inequality follows from the monotonicity of $\widetilde{\mathbf{S}}_{\mu,\nu}$. This completes the proof. □

PROOF OF THEOREM 4.10. (i) This is a consequence of Lemma 4.14 and a capacitability argument that is almost verbatim to [8], Section 6. The same arguments can be found in [22]. We therefore omit these elaborations.

(ii) Applying Proposition 4.13 to the constant sequence $f_n = f$ and a minimizing sequence $(\varphi_n, \psi_n) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ of $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$ yields a dual minimizer. □

We are now in a position to prove the duality between the robust pricing and superhedging problems.

PROOF OF THEOREM 3.9. By Proposition 4.2, Lemma 3.8 and Proposition 4.5,

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f) \le \mathbf{S}_{\mu,\nu}(f,\mathcal{A}) \le \mathbf{I}_{\mu,\nu}(f,\mathcal{A}) \le \widetilde{\mathbf{I}}_{\mu,\nu}(f),$$

and Theorem 4.10 shows that $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \widetilde{\mathbf{I}}_{\mu,\nu}(f)$. Hence,

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \mathbf{S}_{\mu,\nu}(f,\mathcal{A}) = \mathbf{I}_{\mu,\nu}(f,\mathcal{A}) = \widetilde{\mathbf{I}}_{\mu,\nu}(f). \tag{4.19}$$

In particular, the quantities in (4.19) are all independent of the choice of $\mathcal{A}$ (as long as one of the two conditions of Theorem 3.9 holds).

If $\mathbf{I}_{\mu,\nu}(f,\mathcal{A}) < \infty$, then $\widetilde{\mathbf{I}}_{\mu,\nu}(f) < \infty$, and hence there is an optimizer $(\varphi,\psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ for $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$. Then Proposition 4.7 provides an $H = (h, A)$ such that $(\varphi,\psi,H) \in \mathcal{D}_{\mu,\nu}(f,\mathcal{A})$. By (4.19) and the definition of $\mathbf{I}_{\mu,\nu}(f,\mathcal{A})$, (φ,ψ,H) is an optimizer for $\mathbf{I}_{\mu,\nu}(f,\mathcal{A})$. □

REMARK 4.15. Strong duality (without dual attainment) for the robust pricing and superhedging problems continues to hold if we restrict ourselves to trading strategies whose dynamic part is of finite variation.

First, observe that the process $\hat{H}$ defined by (2.9) is of finite variation when h is bounded. Recalling the definition (4.7) of h in Proposition 4.7, we see that h is bounded on $\{\omega : \omega_T \in J^\circ\}$ as these paths are bounded in a compact subset of J, on which ψ' is bounded. This will more generally hold for almost all paths if ψ' is uniformly bounded on J. Therefore, strong duality (and dual attainment in strategies of finite variation) holds if J is open.

Second, consider the case $J = [a, b)$ for some $-\infty < a < b \leq \infty$. Suppose that the assumptions of Theorem 3.9 hold and that $\mathbf{I}_{\mu,\nu}(f, \mathcal{A}) < \infty$, and let $(\varphi, \psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ be a dual auxiliary optimizer. Then $\psi(a) < \infty$ as ν has an atom at a (cf. Lemma 2.7). If $\psi'(a) > -\infty$, then the same argument as above shows that the dynamic trading strategy constructed in Proposition 4.7 is of finite variation. If $\psi'(a) = -\infty$, then we construct a sequence of functions

$$\psi_k(x) := \begin{cases} \psi(x) & \text{for } x \geq a + \dfrac{1}{k}, \\ \psi(a) + k(x-a)\left(\psi\left(a + \dfrac{1}{k}\right) - \psi(a)\right) & \text{for } x < a + \dfrac{1}{k} \end{cases}$$

that approximates ψ by linear interpolation on the interval $[a, a + \frac{1}{k}]$. We then have $\psi_k \searrow \psi$ and $\mu(\varphi) + \nu(\psi_k) \searrow \mu(\varphi) + \nu(\psi) = \widetilde{\mathbf{I}}_{\mu,\nu}(f)$ as $k \to \infty$. Since $\psi_k'(a) > -\infty$, the associated process $H^{(k)}$ is of finite variation almost surely. The cases $J = (a, b]$ and $J = [a, b]$ are analogous.

4.5. *Structure of primal and dual optimizers.* If a primal optimizer to the auxiliary problem exists, we can derive some necessary properties for the dual optimizer.

PROPOSITION 4.16. *Let $\mu \leq_c \nu$ be irreducible with domain (I, J) and let $f : \mathbb{R} \to [0, \infty]$ be Borel. Suppose that $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \widetilde{\mathbf{I}}_{\mu,\nu}(f)$, that $\mu \leq_c \theta \leq_c \nu$ is an optimizer for $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$, and that $(\varphi, \psi) \in \widetilde{\mathcal{D}}_{\mu,\nu}(f)$ is an optimizer for $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$. Then*:

(i) $\varphi + \psi = f$ *θ-a.e.*,
(ii) *φ is affine on the connected components of $\{u_\mu < u_\theta\}$*,
(iii) *ψ is affine on the connected components of $\{u_\theta < u_\nu\}$*,
(iv) *φ does not have a jump at a finite endpoint b of J if $\theta(\{b\}) > 0$ and*
(v) *ψ does not have a jump at a finite endpoint b of J if $\theta(\{b\}) < \nu(\{b\})$.*

PROOF. As in the proof of Lemma 4.14, we obtain (cf. (4.17)) that

$$\theta(f) \leq \theta(\varphi + \psi) \leq \mu(\varphi) + \nu(\psi).$$

By the absence of a duality gap as well as the optimality of θ and (φ, ψ), all inequalities are equalities:

$$\theta(f) = \theta(\varphi + \psi) = \mu(\varphi) + \nu(\psi). \tag{4.20}$$

Now (i) follows from the first equality in (4.20) and the fact that $\varphi + \psi \geq f$ on J. Rearranging the second equality, we can write

$$\begin{aligned} 0 &= \{\mu(\varphi) + \nu(\psi)\} - \theta(\varphi + \psi) \\ &= \{\mu(\varphi) + \nu(\psi)\} - \{\theta(\varphi) + \nu(\psi)\} + \{\theta(\varphi) + \nu(\psi)\} - \theta(\varphi + \psi). \end{aligned}$$

Using the definition (2.4) of the first three expressions (using φ as a concave moderator for the first two terms and $-\psi$ for the third; cf. Lemma 2.9(i)), we obtain

$$\begin{aligned} 0 &= (\mu - \nu)(\varphi) - (\theta - \nu)(\varphi) + (\theta - \nu)(-\psi) \\ &= (\mu - \theta)(\varphi) + (\theta - \nu)(-\psi), \end{aligned} \tag{4.21}$$

where the last equality is a direct consequence of the definitions of $(\mu - \nu)(\varphi)$ and $(\theta - \nu)(\varphi)$ (cf. (2.2)). Both terms on the right-hand side of (4.21) are nonnegative by definition, and hence must vanish:

$$0 = (\mu - \theta)(\varphi) = \int_I (u_\mu - u_\theta)\, \mathrm{d}\varphi'' + \int_{J \setminus I} |\Delta\varphi|\, \mathrm{d}\theta$$

and similarly for $(\theta - \nu)(-\psi)$. This implies that $\varphi'' = 0$ on $\{u_\mu < u_\theta\}$ (which is assertion (ii)) and that $|\Delta\varphi| = 0$ for every endpoint of J on which θ has an atom (which is assertion (iv)). The proofs of (iii) and (v) are similar. □

The next result shows that for upper semicontinuous f, there is a maximizer for $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ which is maximal with respect to the convex order. We omit the proof in the interest of brevity.

PROPOSITION 4.17. *Let $\mu \leq_c \nu$ be irreducible and let $f : \mathbb{R} \to [0, \infty]$ be upper semicontinuous and bounded from above by a convex, continuous and ν-integrable function. Furthermore, fix a strictly convex function $g : \mathbb{R} \to \mathbb{R}$ with linear growth, and consider the "secondary" optimization problem*

$$\sup_{\theta \in \Theta(f)} \theta(g), \tag{4.22}$$

where $\Theta(f) := \{\theta : \mu \leq_c \theta \leq_c \nu \text{ and } \theta(f) \geq \widetilde{\mathbf{S}}_{\mu,\nu}(f)\}$ is the set of optimizers of the auxiliary primal problem.

(i) *$\Theta(f)$ is nonempty, convex and weakly compact and (4.22) admits an optimizer.*

(ii) *Any optimizer θ of (4.22) has the following properties:*

- *θ is maximal in $\Theta(f)$ with respect to the convex order.*
- *If O is an open interval such that $O \subseteq \{u_\theta < u_\nu\}$ and $f|_O$ is convex, then $\theta(O) = 0$.*
- *If K is an interval such that $K^\circ \subseteq \{u_\mu < u_\theta\}$, $f|_K$ is strictly concave, and $\theta(K) > 0$, then $\theta|_K$ is concentrated in a single atom.*

The following example shows that the set optimizers for $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ can have multiple maximal or minimal elements with respect to the convex order; there is in general no greatest or least element for this partially ordered set.

EXAMPLE 4.18. Let $\mu = \delta_0$ and $\nu = \frac{1}{3}(\delta_{-1} + \delta_0 + \delta_1)$ and let f be piecewise linear with $f(-1) = f(1) = 3$, $f(-1/2) = f(1/2) = 2$ and $f(0) = 0$. We claim that there is no greatest or least primal optimizer.

We construct candidate primal and dual optimizers as follows. On the primal side, set $\theta_1 = \frac{2}{3}\delta_{-\frac{1}{2}} + \frac{1}{3}\delta_1$ and $\theta_2 = \frac{1}{3}\delta_{-1} + \frac{2}{3}\delta_{\frac{1}{2}}$. On the dual side, set $\varphi \equiv 0$ and let ψ be the convex function that interpolates linearly between $\psi(-1) = \psi(1) = 3$ and $\psi(0) = 1$. Direct computations yield $\theta_1(f) = \theta_2(f) = \frac{7}{3} = \nu(\psi)$ which shows that θ_1 and θ_2 are primal optimizers and that (φ, ψ) is a dual optimizer.

First, we show that there is no primal optimizer which dominates both θ_1 and θ_2 in convex order. Indeed, one can check that $u_\nu = \max(u_{\theta_1}, u_{\theta_2})$, so that ν is the only feasible primal element which dominates both θ_1 and θ_2 in convex order. But $\nu(f) = 2 < \frac{7}{3}$ and, therefore, ν is not optimal.

Second, we show that there is no primal optimizer which is dominated by both θ_1 and θ_2. Indeed, one can check that $\{u_\mu < \min(u_{\theta_1}, u_{\theta_2})\} = (-\frac{1}{2}, \frac{1}{2})$, so that every feasible primal element that is dominated by both θ_1 and θ_2 must be concentrated on $[-\frac{1}{2}, \frac{1}{2}]$. But $f \le 2$ on $[-\frac{1}{2}, \frac{1}{2}]$, so that no primal optimizer can be concentrated on this interval.

We conclude this section with an example that shows that primal attainment does not hold in general if f is not upper semicontinuous.

EXAMPLE 4.19. Let $\mu = \delta_0$, $\nu = \frac{1}{2}(\delta_{-1} + \delta_1)$, and set $f(x) := |x|\mathbf{1}_{(-1,1)}(x)$. Then $\mu \le_c \nu$ is irreducible with domain $((-1, 1), [-1, 1])$. Considering the sequence $\theta_n := \frac{1}{2}(\delta_{-1+\frac{1}{n}} + \delta_{1-\frac{1}{n}})$, one can see that $\widetilde{\mathbf{S}}_{\mu,\nu}(f) \ge 1$. But there is no $\mu \le_c \theta \le_c \nu$ such that $\theta(f) \ge 1$ because $f < 1$ on $[-1, 1]$.

5. Examples. Two common payoff functions are risk reversals and butterfly spreads. In this section, we provide solutions to the auxiliary primal and dual problems for these payoffs. Throughout this section, we fix irreducible marginals $\mu \le_c \nu$ and denote their common total mass and first moment by m_0 and m_1, respectively.

5.1. *Risk reversals.* The payoff function of a risk reversal is of the form

$$f(x) = -(a - x)_+ + (x - b)_+,$$

for fixed $a < b$. The following result provides a simple geometric construction of the primal and dual optimizers in terms of the potential functions of μ and ν.[8] We

[8]The authors thank David Hobson for the idea of this construction.

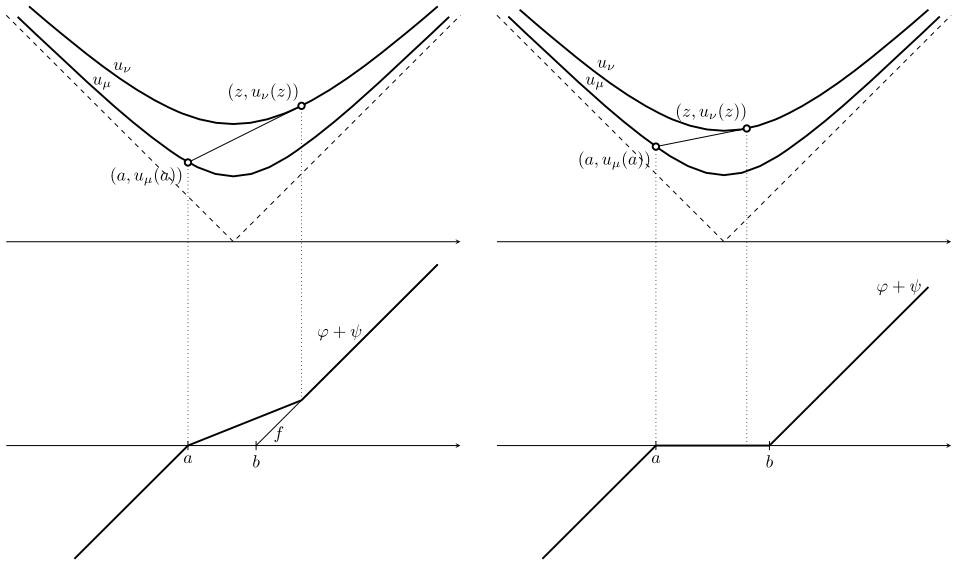

FIG. 1. *Construction of the potential functions of the optimal intermediate laws* θ (*top*) *and the dual optimizers* $\varphi + \psi$ (*bottom*) *for a risk reversal as described in Proposition* 5.1; $z > b$ *in the left panel and* $z < b$ *in the right panel.*

recall that any convex function u lying between the potential functions u_μ and u_ν is the potential function of a measure θ which is in convex order between μ and ν (cf., e.g., [10]).

PROPOSITION 5.1. *Consider the line through the point* $(a, u_\mu(a))$ *of maximal slope lying below* (*or on*) *the graph of* u_ν; *cf. Figure* 1. *This line is either* (*i*) *a tangent line to the graph of* u_ν *with a tangent point* $(z, u_\nu(z))$ *for some* $z \in (a, \infty)$ *or* (*ii*) *the asymptote line for the graph of* u_ν *near* $+\infty$.[9]

In case (*i*), *define the concave function* φ *and the convex function* ψ *by*

$$\varphi(x) = -\alpha(x - a)_+,$$
$$\psi(x) = x - a + \alpha\big(x - (z \vee b)\big)_+,$$

where $\alpha = (b-a)/((z \vee b) - a)$. *Moreover, let* u *be the unique convex function that coincides with* u_μ *on* $(-\infty, a]$ *and with* u_ν *on* $[z, \infty)$ *and is affine on* $[a, z]$ (*i.e.*, u *coincides on* $[a, z]$ *with the tangent line considered above*). *Denote by* θ *the unique measure with potential function* $u_\theta = u$. *In case* (*ii*), *set* $\varphi(x) = 0$, $\psi(x) = x - a$ *and* $\theta = \mu$.

[9]Note that case (ii) can only happen when $(a, \mu(a))$ lies on the increasing part of the dashed potential function in Figure 1. In particular, in this case, μ is concentrated on the left of a.

Then θ is an optimizer for the auxiliary primal problem $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$, (φ,ψ) is an optimizer for the auxiliary dual problem $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$, and the common optimal optimal value is given in terms of the potential functions of μ and ν by

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f)=\widetilde{\mathbf{I}}_{\mu,\nu}(f)=\begin{cases} m_1-\dfrac{a+b}{2}m_0+\dfrac{b-a}{2}\dfrac{u_\nu(z\vee b)-u_\mu(a)}{(z\vee b)-a} & \text{in case } (i),\\ m_1-am_0 & \text{in case } (ii).\end{cases}$$

PROOF. We first note that θ and (φ,ψ) are admissible elements for the auxiliary primal and dual problems, respectively. Indeed, by construction, u_θ is convex and lies between u_μ and u_θ. Thus, the associated measure θ satisfies $\mu\le_c\theta\le_c\nu$. Moreover, a straightforward computation shows that $\varphi+\psi\ge f$, and $(\varphi,\psi)\in L^c(\mu,\nu)$ by Lemma 2.9.

By the weak duality inequality (4.17) (this also holds if f is bounded from below by an affine function; cf. Remark 4.12), $\theta(f)\le\mu(\varphi)+\nu(\psi)$ holds for any admissible primal and dual elements. It thus suffices to show that $\theta(f)=\mu(\varphi)+\nu(\psi)$ for our particular choices for θ and (φ,ψ).

Case (i): Using the identity $(t-s)_+=\frac{1}{2}(|t-s|+t-s)$, the integrals $\theta(f)$, $\mu(\varphi)$ and $\nu(\psi)$ can be expressed in terms of the potential functions of μ, θ and ν as follows:

$$\begin{aligned}\theta(f)&=\frac{1}{2}\big(u_\theta(b)-u_\theta(a)\big)+m_1-\frac{a+b}{2}m_0,\\ \mu(\varphi)&=-\frac{\alpha}{2}\big(u_\mu(a)+m_1-am_0\big),\\ \nu(\psi)&=m_1-am_0+\frac{\alpha}{2}\big(u_\nu(z\vee b)+m_1-(z\vee b)m_0\big).\end{aligned}$$

Substituting $\alpha=(b-a)/((z\vee b)-a)$ and simplifying gives

$$\begin{aligned}\mu(\varphi)+\nu(\psi)&=m_1+\frac{1}{2}\big((a-(z\vee b))\alpha-2a\big)m_0-\frac{\alpha}{2}\big(u_\mu(a)-u_\nu(z\vee b)\big)\\ &=m_1-\frac{a+b}{2}m_0+\frac{b-a}{2}\frac{u_\nu(z\vee b)-u_\mu(a)}{(z\vee b)-a}.\end{aligned}$$

Hence,

$$\theta(f)-\big(\mu(\varphi)+\nu(\psi)\big)=\frac{b-a}{2}\left(\frac{u_\theta(b)-u_\theta(a)}{b-a}-\frac{u_\nu(z\vee b)-u_\mu(a)}{(z\vee b)-a}\right),$$

and it suffices to show that the two quotients inside the brackets are equal. To this end, we distinguish two cases. On the one hand, if $z\le b$, it is enough to observe that $u_\theta(a)=u_\mu(a)$ and $u_\theta(b)=u_\nu(b)$ by the construction of $u=u_\theta$. On the other hand, if $z\ge b$, then the two quotients are the same because $u=u_\theta$ is affine on $[a,z]\supset[a,b]$ and coincides with u_μ at a and with u_ν at z.

Case (ii): On the one hand, since $\theta=\mu$ and μ is concentrated on the left of a, we have $\theta(f)=\mu(f)=m_1-am_0$. On the other hand, $\mu(\varphi)+\nu(\psi)=\int(x-a)\nu(\mathrm{d}x)=m_1-am_0$. □

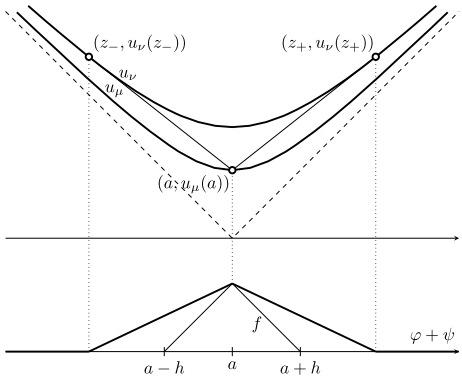

FIG. 2. *Construction of the potential function of the optimal intermediate law* θ (*top*) *and the dual optimizer* $\varphi + \psi$ (*bottom*) *for a butterfly spread as described in Proposition* 5.2.

5.2. *Butterfly spreads*. The payoff function of a butterfly spread is of the form

$$f(x) = \big(x - (a-h)\big)_+ - 2(x-a)_+ + \big(x - (a+h)\big)_+,$$

for fixed a and $h > 0$. We have the following analog to Proposition 5.1; we omit the proof.

PROPOSITION 5.2. *Consider the two lines* l_+, l_- *through the point* $(a, u_\mu(a))$ *of maximal and minimal slope, respectively, lying below* (*or on*) *the graph of* u_ν; *cf. Figure* 2. *We distinguish the cases* $(i+)$ l_+ *is a tangent line with tangent point* $(z_+, u_\nu(z_+))$, $(ii+)$ l_+ *is an asymptote*, $(i-)$ l_- *is a tangent line with tangent point* $(z_-, u_\nu(z_-))$, $(ii-)$ l_- *is an asymptote. In case* $(ii\pm)$, *we set* $z_\pm = \pm\infty$.

Let u *be the convex function that coincides with* u_ν *on* $(-\infty, z_-] \cup [z_+, \infty)$ *and is affine on* $[z_-, a]$ *and on* $[a, z_+]$, *and define the concave function* φ *and the convex function* ψ *by*

$$\varphi(x) = -(\alpha+\beta)(x-a)_+,$$
$$\psi(x) = \alpha\big(x - \big(z_- \wedge (a-h)\big)\big)_+ + \beta\big(x - \big(z_+ \vee (a+h)\big)\big)_+,$$

where $\alpha = \frac{h}{a-(z_-\wedge(a-h))}$ *and* $\beta = \frac{h}{(z_+\vee(a+h))-a}$. *Here, in the asymptote cases* $(ii\pm)$, φ, ψ *need to be interpreted as the limiting functions that arise as* $z_\pm \to \pm\infty$.[10]

Then the intermediate law θ *with potential function* $u_\theta = u$ *is an optimizer for the auxiliary primal problem* $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$, (φ, ψ) *is an optimizer for the auxiliary dual problem* $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$, *and the common optimal optimal value is given in terms of the potential functions of* μ *and* ν *by*

$$\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \widetilde{\mathbf{I}}_{\mu,\nu}(f) = \frac{h}{2}(s_+ + s_-),$$

[10]For instance, if $z_- = -\infty$ and $z_+ < \infty$, then $\varphi(x) = -\beta(x-a)_+$ and $\psi(x) = h + \beta(x - (z_+ \vee (a+h)))_+$.

where

$$s_+ = \begin{cases} \dfrac{u_\nu(z_+ \vee (a+h)) - u_\mu(a)}{(z_+ \vee (a+h)) - a} & \text{in case } (i+), \\ m_0 & \text{in case } (ii+), \end{cases}$$

$$s_- = \begin{cases} \dfrac{u_\nu(z_- \wedge (a-h)) - u_\mu(a)}{a - (z_- \wedge (a-h))} & \text{in case } (i-), \\ -m_0 & \text{in case } (ii-). \end{cases}$$

6. Counterexamples. In this section, we give four counterexamples. Example 6.1 shows that strong duality for the auxiliary problems may fail for general (not necessarily irreducible) marginals if the dual elements φ, ψ are required to be *globally* concave and convex, respectively. Example 6.2 shows that strong duality may fail if the dual elements φ and ψ are required to be μ- and ν-integrable, respectively. Example 6.3 shows that the robust model-based prices of Asian- and American-style derivatives are typically not equivalent when more than two marginals are given. Example 6.4 shows that the equality $\mathbf{S}_{\mu,\nu}(f, \mathcal{A}) = \widetilde{\mathbf{S}}_{\mu,\nu}(f)$ may fail when the assumptions of Proposition 4.2 are violated.

EXAMPLE 6.1 (Duality gap with globally convex/concave dual elements). Let $\mu = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, let ν be the uniform distribution on $(-2, 2)$, and set $f(x) := |x|^{-\frac{1}{2}}$, $x \in \mathbb{R}$ (with $f(0) = \infty$).

First, we show that $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ is finite. Fix any $\mu \leq_c \theta \leq_c \nu$. Computing the potential functions u_μ and u_ν shows that $\mu \leq_c \nu$ and that $\{u_\mu < u_\nu\} = I_1 \cup I_2$ with $I_1 = (-2, 0)$ and $I_2 = (0, 2)$. Because ν does not have an atom at the common boundary 0 of I_1 and I_2, also θ cannot have an atom at 0. Thus, we can write $\theta = \theta_1 + \theta_2$ with

$$\frac{1}{2}\delta_{-1} \leq_c \theta_1 \leq_c \nu|_{I_1} \quad \text{and} \quad \frac{1}{2}\delta_1 \leq_c \theta_2 \leq_c \nu|_{I_2}.$$

Since f is convex when restricted to I_1 or I_2, we have

$$\theta(f) = \theta_1(f) + \theta_2(f) \leq \nu|_{I_1}(f) + \nu|_{I_2}(f) = \nu(f) < \infty.$$

It follows $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \nu(f) < \infty$.

Second, let φ be concave and ψ be convex such that $\varphi + \psi \geq f$. We show that then necessarily $\mu(\varphi) + \nu(\psi) = \infty$. To this end, we may assume that $\varphi < \infty$ on $\operatorname{supp}(\mu) = \{-1, 1\}$. Then $\varphi < \infty$ everywhere by concavity. Thus, evaluating $\varphi + \psi \geq f$ at 0 implies that $\psi(0) = \infty$. Therefore, $\psi = \infty$ on $(-\infty, 0]$ or on $[0, \infty)$ by the convexity of ψ. In both cases, we have $\mu(\varphi) + \nu(\psi) = \infty$.

EXAMPLE 6.2 (Duality gap with individually integrable dual elements). We consider the marginals

$$\mu := C \sum_{n \geq 1} n^{-3} \mu_n \quad \text{and} \quad \nu := C \sum_{n \geq 1} n^{-3} \nu_n,$$

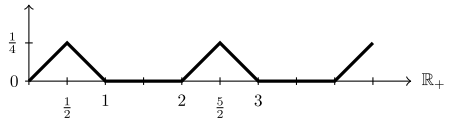

FIG. 3. *The function f in Example 6.2.*

where $C := (\sum_{n\geq 1} n^{-3})^{-1}$, $\mu_n := \delta_n$ and $\nu_n := \frac{1}{3}(\delta_{n-1} + \delta_n + \delta_{n+1})$ for $n \geq 1$. These are the same marginals as in [8], Example 8.5, where it is shown that $\mu \leq_c \nu$ is irreducible with domain $((0,\infty),[0,\infty))$. We now let $f : \mathbb{R}_+ \to [0,1]$ be the piecewise linear function through the points given by $f(n) = 0$ and $f(2n + \frac{1}{2}) = \frac{1}{4}$ for $n \geq 0$; cf. Figure 3.

We proceed to construct candidates for optimizers for $\widetilde{\mathbf{S}}_{\mu,\nu}(f)$ and $\widetilde{\mathbf{I}}_{\mu,\nu}(f)$. For the primal problem, define the sequence $(\bar{\theta}_n)_{n\geq 1}$ by

$$\bar{\theta}_n = \begin{cases} \frac{1}{3}(\delta_{n-1} + 2\delta_{n+\frac{1}{2}}) & \text{for } n \text{ even}, \\ \frac{1}{3}(2\delta_{n-\frac{1}{2}} + \delta_{n+1}) & \text{for } n \text{ odd}, \end{cases}$$

and set $\bar{\theta} := C\sum_{n\geq 1} n^{-3}\bar{\theta}_n$. One can check that $\mu_n \leq_c \bar{\theta}_n \leq_c \nu_n$ and compute $\bar{\theta}_n(f) = \frac{1}{6}$. Hence, $\mu \leq_c \bar{\theta} \leq_c \nu$ (by linearity of potential functions in the measure) and $\bar{\theta}(f) = \frac{1}{6}$.

We now turn to the dual problem. Let $\bar{\varphi}$ and $\bar{\psi}$ be the unique concave and convex functions, respectively, with second derivative measures

$$-\bar{\varphi}'' = \sum_{n\geq 0} \delta_{2n+\frac{1}{2}} \quad \text{and} \quad \bar{\psi}'' = \frac{1}{2}\sum_{n\geq 1} \delta_n$$

and $\bar{\varphi}(0) = \bar{\psi}(0) = 0$, $\bar{\varphi}'(0) = f'(0) = \frac{1}{2}$, and $\bar{\psi}'(0) = 0$. The "initial conditions" are chosen such that $f(0) = \bar{\varphi}(0) + \bar{\psi}(0)$ and $f'(0) = \bar{\varphi}'(0) + \bar{\psi}'(0)$ and the choice of the second derivative measures ensures that φ and ψ pick up the negative and positive curvature of f, respectively. Thus, $\bar{\varphi} + \bar{\psi} = f$ on $\mathbb{R}_+$ by construction. We proceed to compute $\mu(\bar{\varphi}) + \nu(\bar{\psi})$ in the sense of Definition 2.3. (The individual integrals are infinite because $\bar{\varphi}$ and $\bar{\psi}$ have quadratic growth while μ and ν have no second moments.) To this end, we note that $\bar{\varphi} + \bar{\psi} = f$ vanishes on the support of ν. This implies that $\bar{\varphi}$ is a concave moderator for $(\bar{\varphi}, \bar{\psi})$ with respect to $\mu \leq_c \nu$. We can then compute

$$\begin{aligned} \mu(\bar{\varphi}) + \nu(\bar{\psi}) &= \mu(\bar{\varphi} - \bar{\varphi}) + \nu(\bar{\psi} + \bar{\varphi}) + (\mu - \nu)(\bar{\varphi}) \\ &= (\mu - \nu)(\bar{\varphi}) \\ &= C\sum_{n\geq 1} n^{-3}(\mu_n - \nu_n)(\bar{\varphi}). \end{aligned}$$

Fix $n \geq 1$. Because $\bar{\varphi}$ is continuous, we have

$$(\mu_n - \nu_n)(\bar{\varphi}) = \frac{1}{2}\int_I (u_{\mu_n} - u_{\nu_n})\,\mathrm{d}\bar{\varphi}''. \tag{6.1}$$

The difference $u_{\mu_n} - u_{\nu_n}$ vanishes outside $(n-1, n+1)$ and on this interval, $\bar{\varphi}''$ is concentrated on either $n - \frac{1}{2}$ (if n is odd) or on $n + \frac{1}{2}$ (if n is even) with mass 1. Therefore, the right-hand side of (6.1) collapses to $\frac{1}{2}(u_{\mu_n} - u_{\nu_n})(n \pm \frac{1}{2}) = \frac{1}{6}$. It follows that $\mu(\bar{\varphi}) + \nu(\bar{\psi}) = \frac{1}{6} = \bar{\theta}(f)$. Hence, by (weak) duality, $\bar{\theta}$ and $(\bar{\varphi}, \bar{\psi})$ are primal and dual optimizers, respectively.

We are now in a position to argue that no dual optimizer lies in $L^1(\mu) \times L^1(\nu)$. Suppose for the sake of contradiction that $(\varphi, \psi) \in L^1(\mu) \times L^1(\nu)$ is a dual optimizer and note that $\operatorname{supp}(\bar{\theta}) = \{0.5, 1, 2, 2.5, 3, \ldots\}$. We have $\varphi + \psi = f$ $\bar{\theta}$-a.e. by Proposition 4.16(i). One can show that the following modifications of (φ, ψ) do not affect its optimality nor the individual integrability of φ and ψ; we omit the tedious details. First, ψ is replaced by its piecewise linear interpolation at the atoms of ν. Second, φ is replaced by its piecewise linear interpolation at the kinks of f. Third, a suitable convex function is added to φ and subtracted from ψ (preserving their concavity and convexity, resp.) such that the second derivative measures $-\varphi''$ and ψ'' become singular.

Because $\varphi + \psi = f$ on $\operatorname{supp}(\bar{\theta})$ and both sides are piecewise linear, we conclude that $\varphi + \psi = f$ holds on $[\frac{1}{2}, \infty)$. As $-\varphi''$ and ψ'' are singular, φ and ψ must then account for the negative and positive curvature of f, respectively. It follows that both φ and ψ have quadratic growth. Since μ and ν do not have a second moment, we conclude that $\mu(\varphi) = -\infty$ and $\nu(\psi) = \infty$, a contradiction.

EXAMPLE 6.3 (Different robust model-based prices for Asian- and American-style derivatives for multiple marginals). For $n \geq 2$ given marginals $\mu_0 \leq_c \mu_1 \leq_c \cdots \leq_c \mu_n$ corresponding to the time points $0, 1, \ldots, n$ (say), the robust model-based price $\mathbf{S}_{\mu_0,\ldots,\mu_n}(f, \mathcal{A})$ is defined analogously. But this robust model-based price now depends nontrivially on $\mathcal{A}$, as the following example shows. Fix a strictly convex function f. On the one hand, if $\mathcal{A}$ corresponds to American-style derivatives, then one can check that $\mathbf{S}_{\mu_0,\ldots,\mu_n}(f, \mathcal{A}) = \mu_n(f)$. On the other hand, for Asian-style derivatives, that is, $\mathcal{A}' = \{t \mapsto t/n\}$, Jensen's inequality yields

$$f\left(\frac{1}{n}\int_0^n X_t\,\mathrm{d}t\right) \leq \frac{1}{n}\sum_{i=0}^{n-1} f\left(\int_i^{i+1} X_t\,\mathrm{d}t\right),$$

so that

$$\mathbf{S}_{\mu_0,\ldots,\mu_n}(f, \mathcal{A}') \leq \frac{1}{n}\sum_{i=1}^{n} \mu_i(f) \leq \mu_n(f).$$

For a generic choice of marginals, both inequalities are strict. Hence, the robust model-based price of an Asian-style derivative with a strictly convex payoff function is typically smaller than that of the corresponding American-style derivative.

EXAMPLE 6.4 (Necessity of the assumptions of Proposition 4.2).

(i) We show that $\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) = \widetilde{\mathbf{S}}_{\mu,\nu}(f)$ may fail if $\mathcal{A}$ does not contain an interior averaging process. Set $\mathcal{A} = \{A\} = \{t \mapsto \frac{1}{2} + \frac{1}{2}\mathbf{1}_{\{t=T\}}\}$, so that $\int_0^T X_t\,dA_t = (X_0 + X_T)/2$, and consider $f(x) = x^2$. Then, using the martingale property of X under any $P \in \mathcal{M}(\mu,\nu)$, one can check that $\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) = (3\mu(f) + \nu(f))/4$, whereas $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = \nu(f)$ since f is convex. Now, choose μ and ν such that $\mu(f) < \nu(f)$ (f is strictly convex). Then $\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) < \widetilde{\mathbf{S}}_{\mu,\nu}(f)$.

(ii) We show that $\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) = \widetilde{\mathbf{S}}_{\mu,\nu}(f)$ may fail if $\mathcal{A}$ contains an interior averaging process but f is not lower semicontinuous. Set $\mathcal{A} = \{t \mapsto t/T\}$ and $f(x) = \mathbf{1}_{\{|x|\geq 1\}}$, and choose $\mu = \delta_0$ and $\nu = (\delta_1 + \delta_{-1})/2$. On the one hand, since $\nu(f) = 1$ and $f \leq 1$, we have $\widetilde{\mathbf{S}}_{\mu,\nu}(f) = 1$. On the other hand, we claim that $\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) = 0$. To this end, fix $P \in \mathcal{M}(\mu,\nu)$. Since P-a.e. path of X starts in 0, evolves in $[-1,1]$, and is right continuous, $|\frac{1}{T}\int_0^T X_t\,\mathrm{d}t| < 1$ P-a.s. Thus, $E^P[f(\frac{1}{T}\int_0^T X_t\,\mathrm{d}t)] = 0$. Since $P \in \mathcal{M}(\mu,\nu)$ was arbitrary, $\mathbf{S}_{\mu,\nu}(f,\mathcal{A}) = 0$.

Acknowledgments. The authors thank David Hobson, Sigrid Källblad, Marcel Nutz and Yavor Stoev for stimulating discussions.

REFERENCES

[1] AKSAMIT, A., DENG, S., OBŁÓJ, J. and TAN, X. (2019). The robust pricing-hedging duality for American options in discrete time financial markets. *Math. Finance* **29** 861–897. MR3974286

[2] BAYRAKTAR, E., COX, A. M. G. and STOEV, Y. (2018). Martingale optimal transport with stopping. *SIAM J. Control Optim.* **56** 417–433. MR3763082

[3] BAYRAKTAR, E., HUANG, Y.-J. and ZHOU, Z. (2015). On hedging American options under model uncertainty. *SIAM J. Financial Math.* **6** 425–447. MR3356981

[4] BAYRAKTAR, E. and ZHOU, Z. (2017). On arbitrage and duality under model uncertainty and portfolio constraints. *Math. Finance* **27** 988–1012. MR3705160

[5] BAYRAKTAR, E. and ZHOU, Z. (2017). Super-hedging American options with semi-static trading strategies under model uncertainty. *Int. J. Theor. Appl. Finance* **20** 1750036, 10. MR3695458

[6] BEIGLBÖCK, M., HENRY-LABORDÈRE, P. and PENKNER, F. (2013). Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17** 477–501. MR3066985

[7] BEIGLBÖCK, M. and JUILLET, N. (2016). On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44** 42–106. MR3456332

[8] BEIGLBÖCK, M., NUTZ, M. and TOUZI, N. (2017). Complete duality for martingale optimal transport on the line. *Ann. Probab.* **45** 3038–3074. MR3706738

[9] CHACON, R. V. (1977). Potential processes. *Trans. Amer. Math. Soc.* **226** 39–58. MR0501374

[10] CHACON, R. V. and WALSH, J. B. (1976). One-dimensional potential embedding. In *Séminaire de Probabilités X. Lecture Notes in Math.* **511** 19–23. Springer, Berlin. MR0445598

[11] COX, A. M. G. (2008). Extending Chacon–Walsh: Minimality and generalised starting distributions. In *Séminaire de Probabilités XLI. Lecture Notes in Math.* **1934** 233–264. Springer, Berlin. MR2483735

[12] COX, A. M. G. and KÄLLBLAD, S. (2017). Model-independent bounds for Asian options: A dynamic programming approach. *SIAM J. Control Optim.* **55** 3409–3436. MR3719009
[13] DOLINSKY, Y. and SONER, H. M. (2014). Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Related Fields* **160** 391–427. MR3256817
[14] DOLINSKY, Y. and SONER, H. M. (2015). Martingale optimal transport in the Skorokhod space. *Stochastic Process. Appl.* **125** 3893–3931. MR3373308
[15] GALICHON, A., HENRY-LABORDÈRE, P. and TOUZI, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.* **24** 312–336. MR3161649
[16] HOBSON, D. (1998). Robust hedging of the lookback option. *Finance Stoch.* **2** 329–347.
[17] HOBSON, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris–Princeton Lectures on Mathematical Finance* 2010. *Lecture Notes in Math.* **2003** 267–318. Springer, Berlin. MR2762363
[18] HOBSON, D. and NEUBERGER, A. (2016). More on the hedging of American options under model uncertainty. Preprint. Available at arXiv:1604.02274v1.
[19] HOBSON, D. and NEUBERGER, A. (2017). Model uncertainty and the pricing of American options. *Finance Stoch.* **21** 285–329. MR3590709
[20] HOBSON, D. G. (1998). The maximum maximum of a martingale. In *Séminaire de Probabilités, XXXII. Lecture Notes in Math.* **1686** 250–263. Springer, Berlin. MR1655298
[21] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169
[22] KELLERER, H. G. (1984). Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** 399–432. MR0761565
[23] NEUBERGER, A. (2007). Bounds on the American option. Preprint. Available at SSRN:966333.
[24] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357
[25] STEBEGG, F. (2014). Model-independent pricing of Asian options via optimal martingale transport. Preprint. Available at arXiv:1412.1429v1.
[26] TOUZI, N. (2014). Martingale inequalities, optimal martingale transport, and robust superhedging. In *Congrès SMAI* 2013. *ESAIM Proc. Surveys* **45** 32–47. EDP Sci., Les Ulis. MR3451815

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
530 CHURCH STREET
ANN ARBOR, MICHIGAN 48109
USA
E-MAIL: sherrma@umich.edu

DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
1255 AMSTERDAM AVENUE
NEW YORK, NEW YORK 10027
USA
E-MAIL: florian.stebegg@columbia.edu