Decisions in Economics and Finance
https://doi.org/10.1007/s10203-023-00432-y

On entropy martingale optimal transport theory

Alessandro Doldi[1] · **Marco Frittelli**[2] · **Emanuela Rosazza Gianin**[3]

Received: 5 September 2023 / Accepted: 20 December 2023

Abstract
In this paper, we give an overview of (nonlinear) pricing-hedging duality and of its connection with the theory of entropy martingale optimal transport (EMOT), recently developed, and that of convex risk measures. Similarly to Doldi and Frittelli (Finance Stoch 27(2):255–304, 2023), we here establish a duality result between a convex optimal transport and a utility maximization problem. Differently from Doldi and Frittelli (Finance Stoch 27(2):255–304, 2023), we provide here an alternative proof that is based on a compactness assumption. Subhedging and superhedging can be obtained as applications of the duality discussed above. Furthermore, we provide a dual representation of the generalized optimized certainty equivalent associated with indirect utility.

Keywords Martingale optimal transport · Entropy optimal transport · Pricing-hedging duality · Robust finance · Pathwise finance

Mathematics Subject Classification 49Q25 · 49J45 · 60G46 · 91G80 · 90C46

JEL Classification C61 · G13

All authors are members of GNAMPA-INDAM.

✉ Emanuela Rosazza Gianin
emanuela.rosazza1@unimib.it

Alessandro Doldi
alessandro.doldi@unifi.it

Marco Frittelli
marco.frittelli@unimi.it

1 Dipartimento di Scienze per l'Economia e l'Impresa (DISEI), Università degli Studi di Firenze, Via delle Pandette, 9, 50127 Florence, Italy

2 Dipartimento di Matematica, Università degli Studi di Milano, Via Saldini 50, 20133 Milan, Italy

3 Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milan, Italy

Published online: 26 January 2024

1 Introduction

Nonlinear option pricing and risk measures are two widely explored topics in Mathematical Finance. In Sect. 2, we provide an overview of nonlinear pricing-hedging duality and its connection with entropy martingale optimal transport (EMOT) and risk measures. We proceed step-by-step, starting from the classical pricing theory and its relation with coherent/convex risk measures via the subhedging/indifference pricing, going through the recent theory of model uncertainty and pathwise finance and its links with martingale optimal transport. We further discuss the duality between the EMOT problem and the subhedging price, already established in Doldi and Frittelli (2023) and here proved under stronger assumptions but with an alternative and simpler proof.

We synthetically anticipate the discussion in Sect. 2 in the following table, and we point out that in *this paper (as well as in* Doldi and Frittelli (2023)*) we develop the duality theory sketched in the last line of the table and provide its financial interpretation*.

In row 1 (resp. 2), one finds, in the column "functional form", the definition of a coherent (resp. convex) risk measure, while in row 3 (resp. 4) the definition of the subreplication (resp. indifference buyer) price is sketched. The last four rows treat the optimal transport problems, in their various formulations: the genuine optimal

Table 1 c is the contingent claim to be evaluated; $\mathcal{A}$ is the set of acceptable random variables defined on the probability space $(\Omega, \mathcal{F}, P)$; $\Pi(\Omega)$ is the set of all probabilities on Ω; $\mathcal{P}(P) = \{Q \in \Pi(\Omega) \mid Q \ll P\}$; $\mathrm{Mart}(\Omega)$ is the set of all martingale probabilities on Ω; $\mathcal{M}(P) = \mathrm{Mart}(\Omega) \cap \mathcal{P}(P)$; the functions $\alpha_{\mathcal{A}}$, α_U, $\mathcal{D}_U$ in the last column are penalty functions over probability measures; $I^{\Delta}(X)$ is the stochastic integral of the trading strategy Δ with respect to the underlying price process X; $\Pi(Q_1, Q_2) = \{Q \in \Pi(\Omega)$ with given marginals $(Q_1, Q_2)\}$; $\mathrm{Mart}(Q_1, Q_2) = \{Q \in \mathrm{Mart}(\Omega)$ with given marginals $(Q_1, Q_2)\}$; $\mathrm{Meas}(\Omega)$ is the set of all positive finite measures on Ω; the functions ψ and φ represent options to be used in static hedging; $\mathrm{Sub}(c)$ is the set of static parts of semistatic subhedging strategies for c (see Eq. (54)); U is a concave proper utility functional; and S^U is the associated generalized optimized certainty equivalent

		Functional form	Sublinear	Convex
1	- Coherent R.M	$-\inf\{m \in \mathbb{R} \mid c + m \in \mathcal{A}\}$, $\mathcal{A}$ cone	$\inf_{Q \in \mathcal{Q} \subseteq \mathcal{P}(P)} E_Q[c]$	
2	- Convex R.M	$-\inf\{m \in \mathbb{R} \mid c + m \in \mathcal{A}\}$, $\mathcal{A}$ convex		$\inf_{Q \in \mathcal{P}(P)} (E_Q[c] + \alpha_{\mathcal{A}}(Q))$
3	Subreplic. price	$\sup\left\{m \in \mathbb{R} \mid \exists \Delta : m + I^{\Delta}(X) \leq c\right\}$	$\inf_{Q \in \mathcal{M}(P)} E_Q[c]$	
4	Indiff. price	$\sup\{m \in \mathbb{R} \mid U(c - m) \geq U(0)\}$		$\inf_{Q \in \mathcal{M}(P)} (E_Q[c] + \alpha_U(Q))$
5	O.T	$\sup_{\varphi + \psi \leq c} \left(E_{Q_1}[\varphi] + E_{Q_2}[\psi]\right)$	$\inf_{Q \in \Pi(Q_1, Q_2)} E_Q[c]$	
6	E.O.T	$\sup_{\varphi + \psi \leq c} U(\varphi, \psi)$		$\inf_{Q \in \mathrm{Meas}(\Omega)} (E_Q[c] + \mathcal{D}_U(Q))$
7	M.O.T	$\sup_{[\varphi, \psi] \in \mathrm{Sub}(c)} \left(E_{Q_1}[\varphi] + E_{Q_2}[\psi]\right)$	$\inf_{Q \in \mathrm{Mart}(Q_1, Q_2)} E_Q[c]$	
8	E.M.O.T	$\sup_{[\varphi, \psi] \in \mathrm{Sub}(c)} S^U(\varphi, \psi)$		$\inf_{Q \in \mathrm{Mart}(\Omega)} (E_Q[c] + \mathcal{D}_U(Q))$

transport (OT), the entropic optimal transport (EOT), the martingale optimal transport (MOT) and the entropic martingale optimal transport (EMOT). The associate definitions are sketched in the column "functional form." In the last two columns, one finds the corresponding dual representation. While in row 5 (OT) the supremum is taken for the sum of two option prices (the sum of two expectations), in row 6 the supremum is taken for the nonlinear evaluation functional U. A similar difference is between rows 7 and 8. Observe that differently from rows 1, 2, 5, 6, in rows 3, 4, 7, 8, the financial market is present and martingale measures are involved in the dual formulation. In rows 1, 2, 3, 4, we illustrate the classical setting, where the conditions in the functional form hold P-a.s., while in the last four rows optimal transport is applied to treat the robust versions, where the inequalities holds for all elements of Ω.

To ease readability, a summary of the main symbols and notations used in this paper can be found in Sect. A.4.

2 Review on risk measures and the pricing-hedging duality

In this section, we provide a brief review of the theory of risk measures and of the pricing-hedging duality.

2.1 Risk measures and the pricing-hedging duality: the classical setup

The notion of subhedging price is one of the most analyzed concepts in financial mathematics. Although specular considerations can be done for the superhedging price, in this introduction we focus on the subhedging price. We are assuming a discrete-time market model with zero interest rate. It may be convenient for the reader to have at hand the summary described in Table 1. In the classical setup of stochastic securities market models, one considers an adapted stochastic process $X = (X_t)_t$, $t = 0, ..., T$, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$, representing the price of some underlying asset. Let $\mathcal{P}(P)$ be the set of all probability measures on Ω that are absolutely continuous with respect to P, $\mathrm{Mart}(\Omega)$ be the set of all probability measures on Ω under which X is a martingale and $\mathcal{M}(P) = \mathcal{P}(P) \cap \mathrm{Mart}(\Omega)$. We also let $\mathcal{H}$ be the class of admissible integrands and $I^\Delta := I^\Delta(X)$ be the stochastic integral of X with respect to $\Delta \in \mathcal{H}$. Under reasonable assumptions on $\mathcal{H}$, the equality

$$E_Q\left[I^\Delta(X)\right] = 0 \tag{1}$$

holds for all $Q \in \mathcal{M}(P)$ and, as well known, all linear pricing functionals compatible with no arbitrage are expectations $E_Q[\cdot]$ under some probability $Q \in \mathcal{M}(P)$ such that $Q \sim P$.

We denote with p the **subhedging price** of a contingent claim $Z := c(X_T)$ written on the payoff X_T of the underlying asset. If we let $\mathcal{L}(P) \subseteq L^0(\Omega, \mathcal{F}_T, P)$ be the space of random payoffs, then $p : \mathcal{L}(P) \to \mathbb{R}$ is defined by

$$p(Z) := \sup\left\{m \in \mathbb{R} \mid \exists \Delta \in \mathcal{H} \text{ s.t. } m + I^\Delta(X) \le Z,\ P-\text{a.s.}\right\}. \tag{2}$$

The subhedging price is independent from the preferences of the agents, but it depends on the reference probability measure via the class of P-null events. It satisfies the following two key properties:

(CA) Cash Additivity on $\mathcal{L}(P)$: $p(Z+k) = p(Z) + k$, for all $k \in \mathbb{R}$, $Z \in \mathcal{L}(P)$.
(IA) Integral Additivity on $\mathcal{L}(P)$: $p(Z + I^{\Delta}) = p(Z)$, for all $\Delta \in \mathcal{H}$, $Z \in \mathcal{L}(P)$.

When a functional p satisfies (CA), then Z, k and $p(Z)$ must be expressed in the same monetary unit and this allows for the *monetary* interpretation of p, as the price of the contingent claim. This will be one of the key features that we will require also in the novel definition of the nonlinear subhedging value. The (IA) property and $p(0) = 0$ imply that the p price of any stochastic integral $I^{\Delta}(X)$ is equal to zero, as in (1).

Since the seminal works of El Karoui and Quenez (1995), Karatzas (1997), Delbaen and Schachermayer (1994), it was discovered that, under the no arbitrage assumption, the dual representation of the subhedging price p is

$$p(Z) = \inf_{Q \in \mathcal{M}(P)} E_Q[Z]. \tag{3}$$

More or less in the same period, the concept of a **coherent risk measure** was introduced in the pioneering work by Artzner et al. (1999). A coherent risk measure $\rho : \mathcal{L}(P) \to \mathbb{R}$ determines the minimal capital required to make acceptable a financial position and its dual formulation is assigned by

$$-\rho(Y) = \inf_{Q \in \mathcal{Q} \subseteq \mathcal{P}(P)} E_Q[Y], \tag{4}$$

where Y is a random variable representing future profit and loss and $\mathcal{Q} \subseteq \mathcal{P}(P)$. Coherent risk measures ρ are convex, cash additive, monotone and positively homogeneous. We take the liberty to label both the representations in (3) and (4) as the "*sublinear case*".

In the study of incomplete markets, the concept of the (buyer) **indifference price** p^b, originally introduced by Hodges and Neuberger (1989), received, in the early 2000, increasing consideration (see Frittelli (2000), Rouge and El Karoui (2000), Delbaen et al. (2002), Bellini and Frittelli (2002)) as a tool to assess, *consistently with the no arbitrage principle*, the value of nonreplicable contingent claims, and not just to determine an upper bound (the superhedging price) or a lower bound (the subhedging price) for the price of the claim. Differently from the notion of subhedging, p^b is based on some concave increasing utility function $u : \mathbb{R} \to [-\infty, +\infty)$ of the agent. By defining the indirect utility function

$$U(w_0) := \sup_{\Delta \in \mathcal{H}} E_P[u(w_0 + I^{\Delta}(X))],$$

where $w_0 \in \mathbb{R}$ is the initial wealth, the indifference price p^b is defined as

$$p^b(Z) := \sup \{ m \in \mathbb{R} \mid U(Z - m) \geq U(0) \}.$$

Under suitable assumptions, the dual formulation of p^b is

$$p^b(Z) = \inf_{Q \in \mathcal{M}(P)} \left\{ E_Q[Z] + \alpha_u(Q) \right\}, \tag{5}$$

and the penalty term $\alpha_u : \mathcal{M}(P) \to [0, +\infty]$ is associated with the particular utility function u appearing in the definition of p^b via the Fenchel conjugate of u. We observe that in case of the exponential utility function $u(x) = 1 - \exp(-x)$, the penalty is $\alpha_{\exp}(Q) := H(Q, P) - \min_{Q \in \mathcal{M}(P)} H(Q, P)$, where

$$H(Q, P) := \int F\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \mathrm{d}P, \quad \text{if } Q \ll P \text{ and } F(y) = y \ln(y),$$

is the relative entropy. In this case, the penalty $\alpha_{\exp}$ is a divergence functional, similarly, e.g., to those considered in (11). Observe that the functional p^b is concave, monotone increasing and satisfies both properties (CA) and (IA), but it is not necessarily linear on the space of all contingent claims. As recalled in the conclusion of Frittelli (2000), "there is no reason why a price functional defined on the whole space of bundles and consistent with no arbitrage should be linear also outside the space of marketed bundles".

It was exactly the particular form (5) of the indifference price that suggested to Frittelli and Rosazza Gianin (2002) to introduce the concept of **convex risk measure** (also independently introduced by Föllmer and Schied (2002)), as a map $\rho : \mathcal{L}(P) \to \mathbb{R}$ that is convex, cash additive and monotone decreasing. Under good continuity properties, the Fenchel–Moreau theorem shows that any convex risk measure admits the following representation

$$-\rho(Y) = \inf_{Q \in \mathcal{P}(P)} \left\{ E_Q[Y] + \alpha(Q) \right\} \tag{6}$$

for some penalty $\alpha : \mathcal{P}(P) \to [0, +\infty]$. We will then label functional in the form (5) or (6) as the "*convex case*". As a consequence of the cash additivity property, in the dual representations (5) or (6) the infimum is taken with respect to *probability measures*, namely with respect to normalized nonnegative elements in the dual space, which in this case can be taken as $L^1(P)$. Differently from the indifference price p^b, convex risk measures do not necessarily take into account the presence of the stochastic security market, as reflected by the absence of any reference to martingale measures in the dual formulation (6) and (4), in contrast to (5) and (3). The discussion and comparison regarding convex/coherent risk measures are summarized in rows 1–2 of Table 1, while rows 3–4 compare the subheding price with the indifference price.

2.2 Pathwise finance

In the classical setting of nonlinear pricing recalled before, it was implicitly assumed that a reference probability P was fixed and known a priori. The financial crises in 2008, however, somehow inspired and motivated an increasing interest in the case where uncertainty in the selection of a reference probability occurs. The classical notions of arbitrage and pricing-hedging duality have been therefore investigated in this new framework. Two main approaches have been adopted to deal with uncertainty in P. One approach consisted in replacing the single reference probability P with a family of—a priori nondominated—probability measures, leading to the theory of quasi-sure stochastic analysis (see Bayraktar and Zhang (2016), Bayraktar and Zhou (2017), Bouchard and Nutz (2015), Cohen (2012), Denis and Martini (2006), Peng (2019), Soner et al. (2011)). An alternative approach, even more radical, developed a probability-free, pathwise, theory of financial markets, see Acciaio et al. (2016), Burzoni et al. (2016), Burzoni et al. (2017), Burzoni et al. (2019), Riedel (2015). In such framework, optimal transport theory became a very powerful tool to prove pathwise pricing-hedging duality results with relevant contributions by many authors (Beiglböck et al. (2013), Davis et al. (2014), Dolinsky and Soner (2014) and Dolinsky and Soner (2015), Galichon et al. (2014), Henry-Labordère (2013), Henry-Labordère et al. (2016); Hou and Obłój (2018), Tan and Touzi (2013), Bartl et al. (2019), Cheridito et al. (2020) and Cheridito et al. (2017), Guo and Obłój (2019), Backhoff-Veraguas and Pammer (2020), Neufeld and Sester (2021), Sester (2023) and Sester (2023)).

These contributions mainly deal with what we labeled above as the sublinear case, while our main interest in this paper is to develop the convex case theory, as explained below.

From now on, we will abandon the classical setup described above and work without a reference probability measure. We consider a finite horizon $T \in \mathbb{N}$, $T \geq 1$, and

$$\Omega := K_0 \times \cdots \times K_T$$

for $K_0, \ldots, K_T$ subsets of $\mathbb{R}$ and assume that K_0 is a singleton, that is, $K_0 = \{x_0\}$, $x_0 \in \mathbb{R}$. We let $X_0, \ldots, X_T$ be the canonical projections $X_t : \Omega \to K_t$, for $t = 0, 1, ..., T$. We denote

$$\mathrm{Mart}(\Omega) := \{\text{Martingale probability measures for the canonical process of } \Omega\},$$

and when μ is a measure defined on the Borel σ-algebra of $(K_0 \times \cdots \times K_T)$, its marginals will be denoted with $\mu_0, \ldots, \mu_T$. We consider a contingent claim $c : \Omega \to (-\infty, +\infty]$ which is now allowed to depend on the whole path, and for hedging purposes, we will adopt semistatic trading strategies. In other words, in addition to dynamic trading in X via the admissible integrands $\Delta \in \mathcal{H}$, we may invest in "vanilla" options $\varphi_t : K_t \to \mathbb{R}$. For modeling purposes, we take vector subspaces $\mathcal{E}_t \subseteq \mathcal{C}_b(K_t)$ for $t = 0, \ldots, T$, where $\mathcal{C}_b(K_t)$ is the space of real-valued, continuous, bounded functions on K_t. For each t, $\mathcal{E}_t$ represents the set of options to be used for hedging, say affine combinations of options with same maturity t but different strikes. The key assumption in the robust, optimal transport-based formulation is that the marginals $(\widehat{Q}_0, \widehat{Q}_1, ..., \widehat{Q}_T)$ of the underlying price process X are known. This assumption can be justified (see the seminal papers by Breeden and Litzenberger (1978) and Hobson (1998), as well as the many contributions by Hobson (2011), Cox and Obłój (2011a), Cox and Obłój (2011b), Cox and Wang (2013), Henry-Labordère et al. (2016), Brown et al. (2001), Hobson and Klimmek (2013)) by assuming the knowledge of a sufficiently large number of plain vanilla options maturing at each intermediate date, implying then the possibility of calibration.

Thus,

$$\mathcal{M}(\widehat{Q}_0, \widehat{Q}_1, ...\widehat{Q}_T) := \left\{ Q \in \mathrm{Mart}(\Omega) \mid X_t \sim_Q \widehat{Q}_t \text{ for each } t = 0, \dots, T \right\}$$

represents the set of arbitrage-free pricing measures that are compatible with the observed prices of the options. In this framework, the set of admissible trading strategies and of the corresponding stochastic integrals are, respectively, given by

$$\mathcal{H} := \{\Delta = [\Delta_0, \dots, \Delta_{T-1}] \mid \Delta_t \in \mathcal{C}_b(K_0 \times \cdots \times K_t;\mathbb{R})\} \tag{7}$$

$$\mathcal{I} := \left\{ I^\Delta(x) = \sum_{t=0}^{T-1} \Delta_t(x_0, \dots, x_t)(x_{t+1} - x_t) \mid \Delta \in \mathcal{H} \right\} \tag{8}$$

and the subhedging duality, obtained in Beiglböck et al. (2013) Th. 1.1, takes the form:

$$\begin{aligned} &\inf_{Q \in \mathcal{M}(\widehat{Q}_0, \widehat{Q}_1, ...\widehat{Q}_T)} E_Q[c] \\ &= \sup\left\{ \sum_{t=0}^{T} E_{\widehat{Q}_t}[\varphi_t] \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi_t(x_t) + I^\Delta(x) \le c(x)\ \forall x \in \Omega \right\}, \end{aligned} \tag{9}$$

where the RHS of (9) is known as the **robust subhedging price** of c. Comparing (9) with the duality between (2) and (3), we observe that: (i) the $P-$a.s. inequality in (2) has been replaced by an inequality that holds for all $x \in \Omega$; (ii) in (9) the infimum of the price of the contingent claim c is taken under all martingale measure compatible with the option prices, with no reference to the probability P; and (iii) in (9) static hedging with options is allowed.

As can be seen from the LHS of (9), this case falls into the category labeled above as the *sublinear case*, and the purpose of this paper (as well as of Doldi and Frittelli (2023)) is to investigate the *convex case*, in the robust setting, using the tools from entropy optimal transport (EOT) recently developed in Liero et al. (2018).

Let us first describe the financial interpretation of the problems that we are going to study.

2.3 The dual problem

Differently from the pricing theory in finance where the problem $\inf_{Q \in \mathcal{M}(\widehat{Q}_0, \widehat{Q}_1, ...\widehat{Q}_T)} E_Q[c]$ in the LHS of (9) is a dual problem, in **martingale optimal transport** (MOT) it represents the primal problem (called henceforth *sublinear case* of MOT). In Liero et al. (2018), the primal **entropy optimal transport** (EOT) problem takes the form

$$\inf_{\mu \in \mathrm{Meas}(\Omega)} \left(E_Q [c] + \sum_{t=0}^{T} \mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t) \right), \tag{10}$$

where $\mathrm{Meas}(\Omega)$ is the set of all positive finite measures μ on Ω, and $\mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t)$ is a divergence in the form:

$$\mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t) := \int_{K_t} F_t \left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\widehat{Q}_t} \right) \mathrm{d}\widehat{Q}_t, \text{ if } \mu_t \ll \widehat{Q}_t; \tag{11}$$

otherwise, $\mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t) := +\infty$. We label with $F := (F_t)_{t=0,\dots,T}$ the family of divergence functions $F_t : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ appearing in (11). Problem (10) represents the *convex case* of OT theory. Notice that in the EOT primal problem (10) the typical constraint that μ has prescribed marginals $(\widehat{Q}_0, \widehat{Q}_1, \dots \widehat{Q}_T)$ has been relaxed thanks to the introduction of the divergence functional $\mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t)$, which penalizes those measures μ that are "far" from some reference marginals $(\widehat{Q}_0, \widehat{Q}_1, \dots \widehat{Q}_T)$. We are then naturally led to the study of the convex case of MOT, i.e., to the entropy martingale optimal transport (EMOT) problem

$$\inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q [c] + \sum_{t=0}^{T} \mathcal{D}_{F_t, \widehat{Q}_t}(Q_t) \right) \tag{12}$$

having also a clear financial interpretation. The marginals are not any more fixed a priori, to capture the fact that the available information might not be enough to detect them with satisfactory precision. So the infimum is taken over *all martingale probability measures*, but those that are far from some estimate $(\widehat{Q}_0, \widehat{Q}_1, \dots \widehat{Q}_T)$ are appropriately penalized through $\mathcal{D}_{F_t, \widehat{Q}_t}$. Of course, when $\mathcal{D}_{F_t, \widehat{Q}_t}(\cdot) = \delta_{\widehat{Q}_t}(\cdot)$, we recover the sublinear MOT problem, where only martingale probability measures with fixed marginals are allowed. Observe that in addition to the martingale property, the elements $Q \in \mathrm{Mart}(\Omega)$ in (12) are required to be probability measures, while in the EOT (10) theory all positive finite measure are allowed. As it was recalled after equation (6), this normalization feature of the dual elements ($\mu(\Omega) = 1$) is not surprising when one deals with dual problems of primal problems with a cash additive objective functional as, for example, in the theory of coherent and convex risk measures.

Potentially, we could push our smoothing argument above even further: In place of the functionals $\mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t)$, $t = 0, \dots, T$, we might as well consider more general marginal penalizations, not necessarily in the divergence form (11), yielding the problem

$$\mathfrak{D}(c) := \inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q [c] + \sum_{t=0}^{T} \mathcal{D}_t(Q_t) \right). \tag{13}$$

These penalizations $\mathcal{D}_0, \dots, \mathcal{D}_T$ will be better specified later.

We point out that an additional entropic term has been added to optimal transport problems since the seminal work of Cuturi (2013) (see also the survey/monograph

Peyré and Cuturi (2019)). On this topic, we also cite Nutz and Wiesel (2011), Bernton et al. (2021), Ghosal et al. (2021), De March and Henry-Labordère (2020), Henry-Labordère (2019), Blanchet et al. (2020). We also point out that in all the works stemming from Cuturi (2013), the exact matching of the marginals is still required. In this paper such constraint is absent to take into account uncertainty regarding the marginals themselves. In Table 1, rows 5–6, one may compare optimal transport with entropy optimal transport, while in rows 5–7 one may compare optimal transport with martingale optimal transport.

3 Toward entropy martingale optimal transport

3.1 The primal problem: the nonlinear subhedging value

We provide the financial interpretation of the primal problem which will yield the EMOT problem in (12) as its dual. It is convenient to reformulate the robust subhedging price in the RHS of (9) in a more general setting.

Definition 3.1 Consider a measurable function $c : \Omega \to \mathbb{R}$ representing a (possibly path dependent) option, the set $\mathcal{V}$ of hedging instruments and a suitable pricing functional $\pi : \mathcal{V} \to \mathbb{R}$. Then the robust subhedging value of c is defined by

$$\Pi_{\pi,\mathcal{V}}(c) := \sup \{\pi(v) \mid v \in \mathcal{V} \text{ s.t. } v \leq c\}.$$

In the classical setting, functionals of this form (and even with a more general formulation) are known as general capital requirement, see for example Frittelli and Scandolo (2006). We stress, however, that in Definition 3.1 the inequality $v \leq c$ holds for all elements in Ω with no reference to a probability measure whatsoever. The novelty in this definition is that a priori π may not be linear and it is crucial to understand which evaluating functional π we may use. For our discussion, we assume that the vector subspaces $\mathcal{E}_t \subseteq \mathcal{C}_b(K_t)$ satisfies $\mathbb{R} \subseteq \mathcal{E}_t$, for $t = 0, \ldots, T$. We let $\mathcal{E} := \mathcal{E}_0 \times \cdots \times \mathcal{E}_T$, and the Minkowski sum $\mathcal{V} := \mathcal{E}_0 + \cdots + \mathcal{E}_T + \mathcal{I}$, meant as a vector subspace of the class of continuous functions on Ω. Suppose we took a linear pricing rule $\pi : \mathcal{V} \to \mathbb{R}$ defined via a $\widehat{Q} \in \text{Mart}(\Omega)$ by

$$\pi(v) := E_{\widehat{Q}}\left[\sum_{t=0}^{T} \varphi_t + I^{\Delta}\right] \stackrel{(i)}{=} E_{\widehat{Q}}\left[\sum_{t=0}^{T} \varphi_t\right] \stackrel{(ii)}{=} \sum_{t=0}^{T} E_{\widehat{Q}_t}[\varphi_t], \tag{14}$$

where we used (1) and the fact that $\widehat{Q}_t$ is the marginal of $\widehat{Q}$. In this case, we would trivially obtain for the robust subhedging value of c

$$\Pi_{\pi,\mathcal{V}}(c) = \sup\{\pi(v) \mid v \in \mathcal{V} \text{ s.t. } v \le c\} \tag{15}$$

$$= \sup\left\{\sum_{t=0}^{T} E_{\widehat{Q}_t}[\varphi_t] \mid \exists\Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T}\varphi_t(x_t) + I^\Delta(x) \le c(x)\ \forall x \in \Omega\right\}$$

$$= \sup\left\{m \in \mathbb{R} \mid \exists\Delta \in \mathcal{H}, \varphi \in \mathcal{E}, \text{ s.t. } m - \sum_{t=0}^{T} E_{\widehat{Q}_t}[\varphi_t] + \sum_{t=0}^{T}\varphi_t + I^\Delta \le c\right\}$$

$$= \sup\left\{m \in \mathbb{R} \mid \exists\Delta \in \mathcal{H}, \varphi \in \mathcal{E}, \text{ with } E_{\widehat{Q}_t}[\varphi_t] = 0 \text{ s.t. } m + \sum_{t=0}^{T}\varphi_t + I^\Delta \le c\right\}, \tag{16}$$

where in the last equality we replaced φ_t with $(E_{\widehat{Q}_t}[\varphi_t] - \varphi_t) \in \mathcal{E}_t$, which satisfies:

$$E_{\widehat{Q}_t}\left[E_{\widehat{Q}_t}[\varphi_t] - \varphi_t\right] = 0. \tag{17}$$

Interpretation: $\Pi_{\pi,\mathcal{V}}(c)$ is the supremum amount $m \in \mathbb{R}$ for which we may buy zero-cost portfolios of options φ_t and dynamic strategies $\Delta \in \mathcal{H}$ such that $m + \sum_{t=0}^{T}\varphi_t + I^\Delta \le c$, where the values of both the portfolios of options and the stochastic integrals are computed as the expectation under the same martingale measure ($\widehat{Q}$ for the integral I^Δ; its marginals $\widehat{Q}_t$ for each option φ_t).

However, as mentioned above when presenting the indifferent price p^b, there is a priori no reason why one has to allow only linear functional in the evaluation of $v \in \mathcal{V}$. *We thus generalize the expression for* $\Pi_{\pi,\mathcal{V}}(c)$ *by considering valuation functionals* $S : \mathcal{V} \to \mathbb{R}$ and $S_t : \mathcal{E}_t \to \mathbb{R}$ *more general than* $E_{\widehat{Q}}[\cdot]$ *and* $E_{\widehat{Q}_t}[\cdot]$.

Nonetheless, in order to be able to repeat the same key steps we used in (15)–(16) and therefore to keep the same interpretation, we shall impose that such functionals S and S_t satisfy the property in (17) and the two properties (i) and (ii) in Eq. (14), that is:

(a) $S_t(\varphi_t + k) = S_t(\varphi_t) + k$ and $S_t(0) = 0$, for all $\varphi_t \in \mathcal{C}_b(K_t), k \in \mathbb{R}, t = 0, \dots, T$.

(b) $S\left(\sum_{t=0}^{T}\varphi_t + I^\Delta(x)\right) = S\left(\sum_{t=0}^{T}\varphi_t\right)$ for all $\Delta \in \mathcal{H}$ and $\varphi \in \mathcal{E}$.

(c) $S\left(\sum_{t=0}^{T}\varphi_t\right) = \sum_{t=0}^{T} S_t(\varphi_t)$ for all $\varphi \in \mathcal{E}$.

We immediately recognize that (a) is the cash additivity (CA) property on $\mathcal{C}_b(K_t)$ of the functional S_t and (b) implies the integral additivity (IA) property on $\mathcal{V}$. As a consequence, repeating the same steps in (15)–(16), we will obtain as primal problem the **nonlinear subhedging value** of c :

$$
\begin{aligned}
\mathfrak{P}(c) &= \sup\{S(v) \mid v \in \mathcal{V} \text{ s.t. } v \le c\} \\
&= \sup\left\{\sum_{t=0}^{T} S_t(\varphi_t) \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi_t(x_t) + I^{\Delta}(x) \le c(x)\ \forall x \in \Omega\right\} \qquad (18)\\
&= \sup\left\{m \in \mathbb{R} \mid \exists \Delta \in \mathcal{H}, \varphi \in \mathcal{E},\ \text{with } S_t(\varphi_t) = 0 \text{ s.t. } m + \sum_{t=0}^{T} \varphi_t + I^{\Delta} \le c\right\}, \qquad (19)
\end{aligned}
$$

to be compared with (16).

Interpretation: $\mathfrak{P}(c)$ is the supremum amount $m \in \mathbb{R}$ for which we may buy zero-value portfolios of options φ_t and dynamic strategies $\Delta \in \mathcal{H}$ such that $m + \sum_{t=0}^{T} \varphi_t + I^{\Delta} \le c$, where the values of both the portfolios of options and the stochastic integrals are computed with the same functional S.

3.2 Stock additivity

Before further elaborating on these issues, let us introduce the concept of stock additivity, which is the natural counterpart of properties (IA) and (CA) when we are evaluating hedging instruments depending solely on the value of the underlying stock X at some fixed date $t \in \{0, \dots, T\}$. Let Id_t be the identity function on K_t

$$\mathrm{Id}_t : x_t \mapsto x_t .$$

We recall that the set of hedging instruments is denoted by $\mathcal{E}_t \subseteq \mathcal{C}_b(K_t)$ and we will suppose that $\mathrm{Id}_t \in \mathcal{E}_t$ (that is, we can use units of stock at time t for hedging) and that $\mathbb{R} \subseteq \mathcal{E}_t$ (that is, deterministic amounts of cash can be used for hedging as well).

Definition 3.2 A functional $p_t : \mathcal{E}_t \to \mathbb{R}$ is stock additive on $\mathcal{E}_t$ if $p_t(0) = 0$ and

$$p_t(\varphi_t + \alpha_t \mathrm{Id}_t + \lambda_t) = p_t(\varphi_t) + \alpha_t x_0 + \lambda_t \quad \forall \varphi_t \in \mathcal{E}_t, \lambda_t \in \mathbb{R}, \alpha_t \in \mathbb{R} .$$

We now clarify the role of stock additive functionals in our setup. Suppose that $S_t : \mathcal{E}_t \to \mathbb{R}$ are stock additive on $\mathcal{E}_t$, $t = 0, \dots, T$. It can be shown (see Lemma A.3) that if there exist $\varphi, \psi \in \mathcal{E}_0 \times \dots \times \mathcal{E}_T$ and $\Delta \in \mathcal{H}$ such that $\sum_{t=0}^{T} \varphi_t = \sum_{t=0}^{T} \psi_t + I^{\Delta}$ then

$$\sum_{t=0}^{T} S_t(\varphi_t) = \sum_{t=0}^{T} S_t(\psi_t).$$

This allows us to define a functional $S : \mathcal{V} = \mathcal{E}_0 + \dots + \mathcal{E}_T + \mathcal{I} \to \mathbb{R}$ by

$$S(\upsilon) := \sum_{t=0}^{T} S_t(\varphi_t), \quad \text{for } \upsilon = \sum_{t=0}^{T} \varphi_t + I^{\Delta}. \qquad (20)$$

Then S is a well-defined, integral additive functional on $\mathcal{V}$, and $S, S_0, \dots, S_T$ satisfy the properties (a), (b), (c). There is a natural way to produce a variety of stock additive functionals, as explained in Example 3.3.

Example 3.3 Consider a martingale measure $\widehat{Q} \in \mathrm{Mart}(\Omega)$, a concave nondecreasing utility function $u_t : \mathbb{R} \to [-\infty, +\infty)$, satisfying $u_t(0) = 0$ and $u_t(x_t) \le x_t \ \forall x_t \in \mathbb{R}$, and define

$$U_{\widehat{Q}_t}(\varphi_t) := \sup_{\alpha\in\mathbb{R},\ \lambda\in\mathbb{R}} \left(\int_\Omega u_t\left(\varphi_t(x_t) + \alpha x_t + \lambda\right) \mathrm{d}\widehat{Q}_t(x_t) - (\alpha x_0 + \lambda) \right). \tag{21}$$

If we take $S_t(\varphi_t) = U_{\widehat{Q}_t}(\varphi_t)$, then, as shown in Lemma 5.2, the stock additivity property is satisfied for these functionals. Two relevant examples of $S_t = U_{\widehat{Q}_t}$ are those corresponding to linear or exponential utility functions (see Sect. 5.1). For a linear utility $u_t(x) = x$, we get

$$U_{\widehat{Q}_t}(\varphi_t) = E_{\widehat{Q}_t}[\varphi_t].$$

For the exponential utility $u_t(x) = 1 - e^{-x}$, we obtain

$$U_{\widehat{Q}_t}(\varphi_t) = E_{Q_{\widehat{\alpha}}}[\varphi_t(x_t)] + H(Q_{\widehat{\alpha}}, \widehat{Q}_t) = \min_{Q\in\mathrm{Mart}(K_0\times K_t)} \left(E_Q[\varphi_t(x_t)] + H(Q, \widehat{Q}_t)\right) \tag{22}$$

where $\widehat{\alpha} \in \mathbb{R}$ satisfies the martingale condition:

$$E_{Q_{\widehat{\alpha}}}[\mathrm{Id}_t] = x_0 \tag{23}$$

for

$$\frac{dQ_\alpha}{d\widehat{Q}} := \frac{\exp(-\varphi_t(x_t) - \alpha x_t)}{\widehat{E}[\exp(-\varphi_t(x_t) - \alpha x_t)]}, \tag{24}$$

using the notation $\widehat{E} = E_{\widehat{Q}}$, $H(Q, \widehat{Q})$ for the relative entropy and $\mathrm{Mart}(K_0 \times K_t) = \{Q \in \mathrm{Prob}(K_t) \mid E_Q[\mathrm{Id}_t] = x_0\}$ (see Propositions 5.4 and 5.5 for details).

When we consider stock additive functionals $S_0, \dots, S_T$ that induce the functional S as explained in (20), we can focus our attention to the optimization problem (18) or (19), that will be referred to as our primal problem. We mention at this point that different formulations of nonlinear subhedging prices can be already found in the literature, see Föllmer and Schied (2016), Cheridito et al. (2017), Pennanen and Perkkiö (2019). We refer to Doldi and Frittelli (2023) Sect. 2.3 for further discussion of this related literature.

3.3 Entropy martingale optimal transport duality

It was proved in Doldi and Frittelli (2023) Theorem 3.4 that under fairly general assumptions if

$$\mathcal{D}_t(Q_t) := \sup_{\varphi_t \in \mathcal{E}_t} \left(S_t(\varphi_t) - \int_{K_t} \varphi_t \, \mathrm{d}Q_t \right) \quad \text{for } Q_t \in \mathrm{Prob}(K_t), \quad t = 0, \ldots, T,$$

then $\mathfrak{D}(c) = \mathfrak{P}(c)$, namely

$$\begin{aligned} &\inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q[c] + \sum_{t=0}^{T} \mathcal{D}_t(Q_t) \right) \\ &= \sup \left\{ \sum_{t=0}^{T} S_t(\varphi_t) \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi_t(x_t) + I^{\Delta}(x) \le c(x) \; \forall x \in \Omega \right\}. \end{aligned} \tag{25}$$

In the particular case of $S_0, \ldots, S_T$ induced by utility functions, as explained in Example 3.3, this yields the duality

$$\begin{aligned} &\inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q[c] + \sum_{t=0}^{T} \mathcal{D}_{F_t, \widehat{Q}_t}(Q_t) \right) \\ &= \sup \left\{ \sum_{t=0}^{T} U_{\widehat{Q}_t}(\varphi_t) \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi_t(x_t) + I^{\Delta}(x) \le c(x) \; \forall x \in \Omega \right\}. \end{aligned} \tag{26}$$

The functions F_t appearing in $\mathcal{D}_{F_t, \widehat{Q}_t}$, defined in (11), are associated with the utility functions u_t appearing in $U_{\widehat{Q}_t}$ via the conjugacy relation:

$$F_t(y) := v_t^*(y) = \sup_{x_t \in \mathbb{R}} \{x_t y - v(y)\} = \sup_{x_t \in \mathbb{R}} \{u_t(x_t) - x_t y\}, \quad y \in \mathbb{R},$$

where $v(y) := -u(-y)$. Thus, depending on which utility function u is selected in the primal problem in the RHS of (26) to evaluate the options through $U_{\widehat{Q}_t}$, the penalization term $\mathcal{D}_{F_t, \widehat{Q}_t}$ in (26) has a particular form induced by $F_t = v_t^*$. In the special case of linear utility functions $u_t(x_t) = x_t$, we recover the sublinear MOT theory. Indeed, in this case, $v_t^*(y) = +\infty$, for all $y \neq 1$ and $v_t^*(1) = 0$, so that

$$\mathcal{D}_{F_t, \widehat{Q}_t}(\cdot) = \delta_{\widehat{Q}_t}(\cdot),$$

and thus, we obtain the robust pricing-hedging duality (9) of the classical MOT. For the exponential utility $u_t(x) = 1 - e^{-x}$, one can verify that $\mathcal{D}_{F_t, \widehat{Q}_t}(Q_t) = H(Q_t, \widehat{Q}_t)$ (see (59)).

In this work, we focus on the case of **compact** *underlying space which allows us to provide an alternative, simpler proof of the duality* (25).

To achieve this, we first need to present a more general setting (which will lead to the duality (28)), and then, we show how to recover (25) from (28). Following Doldi and Frittelli (2023), we start by introducing two general functionals U and $\mathcal{D}_U$ that are associated through a Fenchel–Moreau-type relation (see (31)). The functional $U:\mathcal{E}\to[-\infty,+\infty)$ is defined on the vector space $\mathcal{E}\subseteq\mathcal{C}_b(\Omega;\mathbb{R}^{T+1})$ consisting of continuous and bounded functions defined on some Polish space Ω and with values in $\mathbb{R}^{T+1}$, while $\mathcal{D}_U:\mathrm{ca}(\Omega)\to(-\infty,+\infty]$ with $\mathrm{ca}(\Omega)$ being the set of all finite signed Borel measures on Ω. As better discussed later, we can think at $\mathcal{E}$ as the set of financial instruments that can be used for hedging, while at U as the evaluation functional of the hedging instruments.

The map U is not necessarily cash additive. In order to turn U in a cash additive functional, we then rely on the notion of the optimized certainty equivalent (OCE) that was introduced in Ben-Tal and Teboulle (1986) and further analyzed in Ben Tal and Ben-Tal and Teboulle (2007). We introduce the generalized optimized certainty equivalent associated with U as the functional $S^U:\mathcal{E}\to[-\infty,+\infty]$ defined by

$$S^U(\varphi):=\sup_{\xi\in\mathbb{R}^{T+1}}\left(U(\varphi+\xi)-\sum_{t=0}^{T}\xi_t\right),\quad \varphi\in\mathcal{E}. \tag{27}$$

As it is easily recognized, any OCE is, except for the sign, a particular convex risk measure, and so, it is cash additive. The cash additivity $S^U(\varphi+\xi)=S^U(\varphi)+\sum_{t=0}^{T}\xi_t$ of the map S^U will ensure that in the problem (10) the elements $\mu\in\mathrm{Meas}(\Omega)$ are normalized, i.e., are probability measures. A family of examples of S^U can be built by considering $U(\varphi)=\sum_{t=0}^{T}U_{\widehat{Q}_t}(\varphi_t)$, with $\widehat{Q}_t$ being the marginal of $\widehat{Q}\in\mathrm{Mart}(\Omega)$ and $U_{\widehat{Q}_t}$ as in Example 3.3. In particular, for a linear utility function $u_t(x)=x$, $U(\varphi)=S^U(\varphi)=\sum_{t=0}^{T}E_{\widehat{Q}_t}[\varphi_t]$ since U already satisfies cash additivity.

In Theorem 4.4, we then prove the following duality

$$\inf_{Q\in\mathrm{Mart}(\Omega)}\left(E_Q[c(X)]+\mathcal{D}_U(Q)\right)=\sup_{\varphi\in\mathcal{S}_{\mathrm{sub}}(c)}S^U(\varphi) \tag{28}$$

where

$$\mathcal{S}_{\mathrm{sub}}(c):=\left\{\varphi\in\mathrm{dom}(U)\mid\exists\Delta\in\mathcal{H}\text{ s.t. }\sum_{t=0}^{T}\varphi_t(x_t)+I^{\Delta}(x)\le c(x)\ \ \forall x\in\Omega\right\}. \tag{29}$$

Observe that

$$\sup_{\varphi\in\mathcal{S}_{\mathrm{sub}}(c)}S^U(\varphi)$$
$$=\sup_{\Delta\in\mathcal{H}}\left\{\sup\left\{S^U(\varphi)\mid\varphi\in\mathrm{dom}(U),\ \sum_{t=0}^{T}\varphi_t(x_t)+I^{\Delta}(x)\le c(x)\ \ \forall x\in\Omega\right\}\right\},$$

which is equal to the RHS of (43). Referring to the class of the above examples and for a fixed $\widehat{Q}$, we have for a linear utility function $\mathcal{D}_U(Q)=\sum_{t=0}^{T}\delta_{\widehat{Q}_t}(Q)$ where

$\delta_A = \infty \mathbf{1}_{A^c}$, corresponding to the classical MOT problem, while for an exponential utility $\mathcal{D}_U(Q) = \sum_{t=0}^{T} H(Q_t, \widehat{Q}_t)$.

Note that the aforementioned Theorem 4.4, which in principle could be considered as a corollary of Doldi and Frittelli (2023) Theorem 2.4 (see Doldi and Frittelli (2023) Corollary 2.5), is obtained by a different technique, which allows for avoiding much of the technicalities involved in the noncompact case. Indeed, we first show a Kantorovich-type duality result for a generalization of the EOT problem (10) (see Theorem 4.3). We then deduce (28) following closely Beiglböck et al. (2013) by means of a minimax argument. The duality (25) is then deduced in Sect. 4.3, Theorem 4.8. In Sect. 5.1, we provide an example of application, obtaining a nonlinear pricing-hedging duality for the stock additive pricing functionals of the type in Example 3.3. Finally, in Sect. 5.2 we show the flexibility of our previous result beyond subhedging and superhedging dualities. Indeed, we prove a dual robust representation of the generalized optimized certainty equivalent associated with the indirect utility function. A recap on the EMOT problem is contained in row 8 of Table 1.

4 Main results

The main duality in Theorem 4.4 is obtained applying a preparatory result stated in Theorem 4.3 that we now illustrate.

4.1 A generalized optimal transport duality

We introduce some notations used in the sequel. For unexplained concepts on measure theory, we refer to Appendix A.1. We let Ω be a Polish Space, $\mathcal{B}(\Omega)$ its Borel sigma algebra and define the following sets:

$$\begin{aligned}
\mathrm{ca}(\Omega) &:= \{\gamma : \mathcal{B}(\Omega) \to (-\infty, +\infty) \mid \gamma \text{ is a finite signed Borel measure on } \Omega\}, \\
\mathrm{Meas}(\Omega) &:= \{\mu : \mathcal{B}(\Omega) \to [0, +\infty) \mid \mu \text{ is a nonnegative finite Borel measure on } \Omega\}, \\
\mathrm{Prob}(\Omega) &:= \{Q : \mathcal{B}(\Omega) \to [0, 1] \mid Q \text{ is a probability Borel measure on } \Omega\}. \\
\mathcal{C}_b(\Omega, \mathbb{R}^M) &:= (\mathcal{C}_b(\Omega))^M = \{\varphi : \Omega \to \mathbb{R}^M \mid \varphi \text{ is bounded and continuous on } \Omega\}.
\end{aligned}$$

We let $\mathcal{E} \subseteq \mathcal{C}_b(\Omega; \mathbb{R}^{M+1})$ be a vector subspace, $U : \mathcal{E} \to [-\infty, +\infty)$ be a proper concave functional and set

$$\mathrm{dom}(U) := \{\varphi \in \mathcal{E} \mid U(\varphi) > -\infty\} \tag{30}$$

and

$$V(\varphi) := -U(-\varphi).$$

We define $\mathcal{D} : \mathrm{ca}(\Omega) \to (-\infty, +\infty]$ by

$$\begin{aligned}\mathcal{D}(\gamma) &:= \sup_{\substack{\varphi\in\mathcal{E}\\ \varphi=(\varphi_1,\dots,\varphi_M)}} \left(U(\varphi) - \sum_{m=0}^{M} \int_\Omega \varphi_m \mathrm{d}\gamma \right) \\ &= \sup_{\varphi\in\mathcal{E}} \left(\sum_{m=0}^{M} \int_\Omega \varphi_m \mathrm{d}\gamma - V(\varphi) \right), \quad \gamma \in \mathrm{ca}(\Omega). \end{aligned} \tag{31}$$

$\mathcal{D}$ is a convex functional and is $\sigma(\mathrm{ca}(\Omega), \mathcal{E})$- lower semicontinuous, even if we do not require that U is $\sigma(\mathcal{E}, \mathrm{ca}(\Omega))$-upper semicontinuous.

The following assumption will hold throughout all the paper without further mention.

Standing Assumption 4.1 *$\mathcal{D}$ is proper, i.e.,* $\mathrm{dom}(\mathcal{D}) = \{\gamma \in \mathrm{ca}(\Omega) \mid \mathcal{D}(\gamma) < +\infty\} \neq \emptyset$.

Remark 4.2 So far, we have introduced the functionals U and $\mathcal{D}$ by fixing U and then defining $\mathcal{D}$ by duality. As already discussed in Doldi and Frittelli (2023), an alternative approach could be to do the converse. Let $\mathcal{D} : \mathrm{ca}(\Omega) \to (-\infty, +\infty]$ be a proper convex functional that is $\sigma(\mathrm{ca}(\Omega), \mathcal{E})$ -lower semicontinuous for some vector subspace $\mathcal{E} \subseteq \mathcal{C}_b(\Omega, \mathbb{R}^{M+1})$. Set now V the Fenchel–Moreau (convex) conjugate of $\mathcal{D}$, i.e.,

$$V(\varphi) := \sup_{\gamma\in\mathrm{ca}(\Omega)} \left(\sum_{m=0}^{M} \int_\Omega \varphi_m \,\mathrm{d}\,\gamma - \mathcal{D}(\gamma) \right), \quad \varphi \in \mathcal{E}, \tag{32}$$

and

$$U(\varphi) := -V(-\varphi), \quad \varphi \in \mathcal{E}. \tag{33}$$

By the Fenchel–Moreau theorem, it holds that $\mathcal{D}$ if of the form (31).

Theorem 4.3 *Let $c : \Omega \to (-\infty, +\infty]$ be proper lower semicontinuous with compact sublevel sets and assume the following condition on U holds*[1]*:*

$$\exists \text{ a sequence } (k^n)_n \subseteq \mathbb{R}^{M+1} \text{ with } \limsup_n \sum_{m=0}^{M} k_m^n = +\infty \text{ and } U(-k^n) > -\infty \ \forall n\,. \tag{34}$$

Then

$$\inf_{\mu\in\mathrm{Meas}(\Omega)} \left(\int_\Omega c \,\mathrm{d}\mu + \mathcal{D}(\mu) \right) = \sup_{\varphi\in\Phi(c)} U(\varphi)\,,$$

where

$$\Phi(c) := \left\{ \varphi \in \mathrm{dom}(U) \mid \sum_{m=0}^{M} \varphi_m(x) \leq c(x) \ \forall x \in \Omega \right\}. \tag{35}$$

[1] The condition (34) is a finiteness requirement. It is straightforwardly satisfied if the functional U is real-valued on the whole $\mathcal{E}$.

Proof We start applying (31) to get that

$$\int_\Omega c\,\mathrm{d}\mu + \mathcal{D}(\mu) = \int_\Omega c\,\mathrm{d}\mu + \sup_{\varphi\in\mathcal{E}}\left(U(\varphi) - \sum_{m=0}^{M}\int_\Omega \varphi_m\,\mathrm{d}\mu\right).$$

We then consider $\mathcal{L} : \mathrm{Meas}(\Omega) \times \mathrm{dom}(U) \to (-\infty, +\infty]$ defined by

$$\mathcal{L}(\mu,\varphi) := \int_\Omega \left(c - \sum_{m=0}^{M}\varphi_m\right)\mathrm{d}\,\mu + U(\varphi),$$

and we set $M := \{\mu \in \mathrm{Meas}(\Omega) \mid \int_\Omega c\,\mathrm{d}\mu < +\infty\}$. We observe that $\mathcal{L}$ is real-valued on $M \times \mathrm{dom}(U)$, and for any $\mu \in \mathrm{Meas}(\Omega)\backslash M$, we have $\mathcal{L}(\mu,\varphi) = +\infty$ for all $\varphi \in \mathrm{dom}(U)$ (since c is bounded from below). We also see that setting $\mathcal{C} := \mathrm{dom}(U)$

$$\inf_{\mu\in\mathrm{Meas}(\Omega)}\left(\int_\Omega c\mathrm{d}\mu + \mathcal{D}(\mu)\right) = \inf_{\mu\in\mathrm{Meas}(\Omega)}\sup_{\varphi\in\mathcal{C}}\mathcal{L}(\mu,\varphi) = \inf_{\mu\in M}\sup_{\varphi\in\mathcal{C}}\mathcal{L}(\mu,\varphi). \tag{36}$$

The aim is now to interchange sup and inf in RHS of (36), using Theorem A.5.

To this end, without loss of generality we can assume $\alpha := \sup_{\varphi\in\mathcal{C}}\inf_{\mu\in\mathrm{Meas}(\Omega)}\mathcal{L}(\mu,\varphi) < +\infty$ and we have to find $\varphi \in \mathcal{C}$ and $C > \alpha$ such that the sublevel set $\mu_C := \{\mu \in \mathrm{Meas}(\Omega) \mid \mathcal{L}(\mu,\varphi) \le C\}$ is weakly compact. The functional c is proper, lower continuous and has compact sublevel sets; hence, it attains a minimum on Ω. Therefore, for any $\varepsilon > 0$ we can choose, by Assumption (34), a deterministic vector $\varphi \in \mathcal{C}$ having all components φ_m equal to some constant $-k_m^n < 0$, such that $\varphi \in \mathrm{dom}(U)$ and

$$\inf_{x\in\Omega}\left(c(x) - \sum_{m=0}^{M}\varphi_m(x)\right) > \varepsilon > 0.$$

For such choice of φ and for a sufficiently big constant $C > \alpha$, there exists another constant $D := C - U(\varphi) \ge 0$, independent of μ, such that

$$\mu_C = \left\{\mu \in \mathrm{Meas}(\Omega) \mid \int_\Omega \left(c - \sum_{m=0}^{M}\varphi_m\right)\mathrm{d}\mu \le D\right\} \tag{37}$$

$$= \left\{\mu \in \mathrm{Meas}(\Omega) \mid \int_\Omega \left(c - \sum_{m=0}^{M}\varphi_m - \varepsilon\right)\mathrm{d}\mu + \varepsilon\mu(\Omega) \le D\right\}. \tag{38}$$

Consequently, the set μ_C is:

1. Nonempty, as the measure $\mu \equiv 0$ belongs to μ_C.
2. Weakly closed. Indeed, for each $\varphi \in \mathcal{C}$ the function $c - \varphi$ is lower semicontinuous on Ω, and so, it is the pointwise supremum of bounded continuous functions $(c_n)_n \subseteq \mathcal{C}_b(\Omega)$. For each n, $\mu \mapsto \int_\Omega c_n\,\mathrm{d}\mu$ is weakly continuous on $\mathrm{Meas}(\Omega)$, by definition. Hence, by monotone convergence theorem the map

$\mu \mapsto \int_\Omega \left(c - \sum_{m=0}^M \varphi_m\right) \mathrm{d}\mu$ is the pointwise supremum of weakly continuous functions and is then lower semicontinuous with respect to the weak topology. We conclude that for each $\varphi \in \mathcal{C}$ the functional $\mathcal{L}(\,\cdot\,, \varphi)$ is weakly lower semicontinuous, and has closed sublevel sets. This implies that in particular μ_C is weakly closed, using (37).

3. Bounded: having a sequence of measures in μ_C with unbounded total mass would result in a contradiction with the constraint in (38), taking into account that $c - \sum_{m=0}^M \varphi_m - \varepsilon \geq 0$ and $\varepsilon > 0$.
4. Tight: let $0 \leq f := c - \sum_{m=0}^M \varphi_m - \varepsilon$. Since $\varepsilon\mu(\Omega) \geq 0$ for all $\mu \in \mathrm{Meas}(\Omega)$, by (37) the inclusion $\mu_C \subseteq \{\mu \in \mathrm{Meas}(\Omega) \mid \int_\Omega f \mathrm{d}\mu \leq D\}$ holds. Now it is easy to check that for all $\mu \in \mu_C$ and $\alpha > 0$

$$D \geq \int_\Omega f \mathrm{d}\mu \geq \int_{f>\alpha} f \mathrm{d}\mu \geq \alpha\mu(\{f > \alpha\})\,.$$

   Observing that the sublevels of f are compact, by lower semicontinuity of c and compactness of its sublevel sets, we see that $\{f > \alpha\}$ are complementaries of compact subsets of Ω and can be taken with arbitrarily small measure, just by increasing α, uniformly in $\mu \in \mu_C$. Thus, tightness follows.
5. A subset of M.

These properties in turn yield $\sigma(\,\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$-compactness by Prohorov theorem (e.g., Föllmer and Schied (2016) Theorem A.45). As a consequence, by Item 5, μ_C is compact in the relative topology induced on M, namely $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))|_M$. We now may apply Theorem A.5. Indeed, $\mathcal{L}$ is real-valued on $M \times \mathcal{C}$. Items 1 and 2 of Theorem A.5 are fulfilled for: $A = M$ endowed with the topology $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))|_M$; $B = \mathcal{C}$; and C taken as above. We only justify explicitly lower semicontinuity $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))|_M$ for Item 1, which can be obtained arguing as in Item 2. Hence, we may interchange sup and inf in RHS of (36), obtaining

$$\inf_{\mu\in M} \sup_{\varphi\in\mathcal{C}} \mathcal{L}(\mu, \varphi) = \sup_{\varphi\in\mathcal{C}} \inf_{\mu\in M} \mathcal{L}(\mu, \varphi) = \sup_{\varphi\in\mathcal{C}} \inf_{\mu\in\mathrm{Meas}(\Omega)} \mathcal{L}(\mu, \varphi) \tag{39}$$

where the last equality follows from the fact that $\mathcal{L}(\mu, \varphi) = +\infty$ on the complementary of M in $\mathrm{Meas}(\Omega)$ for every $\varphi \in \mathcal{C}$. It is now easy to check that for every $\varphi \in \mathcal{C}$

$$\inf_{\mu\in Meas} \mathcal{L}(\mu, \varphi) = \begin{cases} U(\varphi) & \text{if } \sum_{m=0}^M \varphi_m(x) \leq c(x)\,\forall\, x \in \Omega \\ -\infty & \text{otherwise;} \end{cases}$$

thus,

$$\sup_{\varphi\in\mathcal{C}} \inf_{\mu\in\mathrm{Meas}(\Omega)} \mathcal{L}(\mu, \varphi) = \sup_{\varphi\in\Phi(c)} U(\varphi)$$

which concludes the proof, given (36) and (39). ◻

4.2 The entropy martingale optimal transport duality

In order to describe a suitable theory to develop the entropy optimal transport duality in a dynamic setting, in this section we will adopt a particular product structure of the set Ω.

To this end, in addition to the notations already introduced at the beginning of Sect. 4.1, we consider a finite horizon $T \in \mathbb{N}$, $T \geq 1$, and

$$\Omega := K_0 \times \cdots \times K_T \tag{40}$$

for $K_0, \ldots, K_T \subseteq \mathbb{R}$, with $K_0 = \{x_0\}$, $x_0 \in \mathbb{R}$. We denote with $X_0, \ldots, X_T$ the canonical projections $X_t : \Omega \to K_t$, and we set $X = [X_0, \ldots, X_T] : \Omega \to \mathbb{R}^{T+1}$, to be considered as discrete-time stochastic process X representing the price of an underlying asset. We denote with:

$$\mathrm{Mart}(\Omega) := \{\text{Martingale probability measures for the canonical process of } \Omega\}. \tag{41}$$

When $\mu \in \mathrm{Meas}(K_0 \times \cdots \times K_T)$, its marginals will be denoted with: $\mu_0, \ldots, \mu_T$.

We recall, respectively, from (7) and (8), that $\mathcal{H}$ is the set of admissible trading strategies and $\mathcal{I}$ is the set of elementary stochastic integral. We take $\mathcal{E} = \mathcal{E}_0 \times \cdots \times \mathcal{E}_T$ where $\mathcal{E}_t \subseteq \mathcal{C}_b(K_0 \times \cdots \times K_t)$ is a vector subspace, for every $t = 0, \ldots, T$. Then $\mathcal{E}$ is clearly a vector subspace of $\mathcal{C}_b(\Omega; \mathbb{R}^{T+1})$, and in the stochastic process interpretation, its elements are processes adapted to the natural filtration of the process X.

We suppose that $U : \mathcal{E} \to [-\infty, +\infty)$ is proper and concave, $\mathcal{D} : \mathrm{Meas}(\Omega) \to (-\infty, +\infty]$ is defined in (31), and as in (27),

$$S^U(\varphi) := \sup_{\xi \in \mathbb{R}^{T+1}} \left(U(\varphi + \xi) - \sum_{t=0}^{T} \xi_t \right), \quad \varphi \in \mathcal{E}.$$

The following result establishes the duality in (43). The result here presented can be obtained as a corollary of Theorem 2.4 of Doldi and Frittelli (2023) where, differently from below, there is no assumption on the compactness of the sets $K_0, \ldots, K_T$. We provide, however, a proof that is different from Doldi and Frittelli (2023) and simpler, yet holding under the compactness assumption.

Theorem 4.4 *Assume that $\Omega := K_0 \times \cdots \times K_T$ for compact sets $K_0, \ldots, K_T \subseteq \mathbb{R}$, that $c : \Omega \to (-\infty, +\infty]$ is lower semicontinuous, that $\mathcal{D} : \mathrm{Meas}(\Omega) \to (-\infty, +\infty]$ is lower bounded on $\mathrm{Meas}(\Omega)$ and proper. Suppose also U satisfies (34), and that*

$$\mathcal{N} := \left\{ \mu \in \mathrm{Meas}(\Omega) \cap \mathrm{dom}(\mathcal{D}) \mid \int_\Omega c \, \mathrm{d}\mu < +\infty \right\} \neq \emptyset$$
$$\textit{and} \quad \mathrm{dom}(U) + \mathbb{R}^{T+1} \subseteq \mathrm{dom}(U). \tag{42}$$

Then the following holds:

$$\inf_{Q\in\mathrm{Mart}(\Omega)}\left(E_Q\left[c(X)\right]+\mathcal{D}(Q)\right)=\sup_{\Delta\in\mathcal{H}}\ \sup_{\varphi\in\Phi_\Delta(c)} S^U(\varphi) \tag{43}$$

where for each $\Delta\in\mathcal{H}$

$$\Phi_\Delta(c) := \left\{\varphi\in\mathrm{dom}(U)\mid \sum_{t=0}^{T}\varphi_t(x_t)+\sum_{t=0}^{T-1}\Delta_t(x_0,\dots,x_t)(x_{t+1}-x_t)\le c(x)\ \forall\, x\in\Omega\right\}. \tag{44}$$

Proof The first part of the proof is inspired by Beiglböck et al. (2013) Equations (3.4)–(3.3)–(3.2)–(3.1).

$$\inf_{Q\in\mathrm{Mart}(\Omega)}\left(E_Q\left[c(X)\right]+\mathcal{D}(Q)\right) \tag{45}$$

$$=\inf_{Q\in\mathrm{Mart}(\Omega)}\sup_{\Delta\in\mathcal{H}}\left(E_Q\left[c(X)-\sum_{t=0}^{T-1}\Delta_t(X_0,\dots,X_t)(X_{t+1}-X_t)\right]+\mathcal{D}(Q)\right) \tag{46}$$

$$=\inf_{Q\in\mathrm{Prob}(\Omega)}\sup_{\Delta\in\mathcal{H}}\left(E_Q\left[c(X)-\sum_{t=0}^{T-1}\Delta_t(X_0,\dots,X_t)(X_{t+1}-X_t)\right]+\mathcal{D}(Q)\right) \tag{47}$$

$$=\inf_{Q\in\mathrm{Prob}(\Omega)}\sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}}}\left(E_Q\left[c(X)-\sum_{t=0}^{T-1}\Delta_t(X_0,\dots,X_t)(X_{t+1}-X_t)+\lambda\right]-\lambda+\mathcal{D}(Q)\right) \tag{48}$$

$$=\inf_{\mu\in\mathrm{Prob}(\Omega)}\sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}}}\left(\int_\Omega\left[c(x)-I^\Delta(x)+\lambda\right]\mathrm{d}\mu(x)-\lambda+\mathcal{D}(\mu)\right) \tag{49}$$

$$=\inf_{\mu\in\mathrm{Meas}(\Omega)}\sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}}}\left(\int_\Omega\left[c-I^\Delta+\lambda\right]\mathrm{d}\mu-\lambda+\mathcal{D}(\mu)\right) \tag{50}$$

$$=\inf_{\mu\in\mathrm{Meas}(\Omega)}\sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}}\left(\int_\Omega\left[c-I^\Delta+\sum_{t=0}^{T}\lambda_t\right]\mathrm{d}\mu-\sum_{t=0}^{T}\lambda_t+\mathcal{D}(\mu)\right). \tag{51}$$

The above equality chain is justified as follows: (45) = (46) is trivial; (46) = (47) follows using the same argument as in Beiglböck et al. (2013) Lemma 2.3, which yields that the inner supremum explodes to $+\infty$ unless Q is a martingale measure on Ω; (47) = (48) and (48) = (49) are trivial; (49) = (50) follows observing that the inner supremum over $\lambda\in\mathbb{R}$ explodes to $+\infty$ unless $\mu(\Omega)=1$; (50) = (51) is trivial.

We define now $\mathcal{K} : \mathrm{Meas}(\Omega) \times (\mathcal{H} \times \mathbb{R}^M) \to (-\infty, +\infty]$ as

$$\mathcal{K}(\mu, \Delta, \lambda) := \int_\Omega \left[c - I^\Delta + \sum_{t=0}^T \lambda_t \right] \mathrm{d}\mu - \sum_{t=0}^T \lambda_t + \mathcal{D}(\mu) .$$

From (42), we observe that $\mathcal{K}$ is real-valued on $\mathcal{N} \times (\mathcal{H} \times \mathbb{R}^{T+1})$ and that $K(\mu, \Delta, \lambda) = +\infty$ if $\mu \in \mathrm{Meas}(\Omega)\backslash\mathcal{N}$, for all $(\Delta, \lambda) \in \mathcal{H} \times \mathbb{R}^{T+1}$. This, together with our previous computations, provides

$$\begin{aligned}&\inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q [c(X)] + \mathcal{D}(Q) \right) \\ &= \inf_{\mu \in \mathrm{Meas}(\Omega)} \sup_{\substack{\Delta \in \mathcal{H} \\ \lambda \in \mathbb{R}^{T+1}}} \mathcal{K}(\mu, \Delta, \lambda) = \inf_{\mu \in \mathcal{N}} \sup_{\substack{\Delta \in \mathcal{H} \\ \lambda \in \mathbb{R}^{T+1}}} \mathcal{K}(\mu, \Delta, \lambda). \end{aligned} \tag{52}$$

As in the proof of Theorem 4.3, we wish to apply the minimax theorem A.5 in order to interchange inf and sup in RHS of (52), and without loss of generality, we can assume that $\alpha := \sup_{\substack{\Delta \in \mathcal{H} \\ \lambda \in \mathbb{R}^{T+1}}} \inf_{\mu \in \mathcal{N}} \mathcal{K}(\mu, \Delta, \lambda) < +\infty$. The functional $\mathcal{K}$ is real-valued on $\mathcal{N} \times (\mathcal{H} \times \mathbb{R}^{T+1})$ and convexity in Item 1, concavity in 2 of Theorem A.5 are clearly satisfied. We have to find $\Delta \in \mathcal{H}$, $\lambda \in \mathbb{R}^{T+1}$ and $C > \alpha$ such that the sublevel set $M_C := \{\mu \in \mathrm{Meas}(\Omega) \mid \mathcal{K}(\mu, \Delta, \lambda) \le C\}$ is weakly compact. Fix a $\varepsilon > 0$. As the functional c is lower semicontinuous on the compact Ω, it is lower bounded on Ω and we can take $\Delta = 0$ and λ sufficiently big in such a way that $\inf_{x \in \Omega}(c(x) + \sum_{t=0}^T \lambda_t) > \varepsilon$. For such a choice of (Δ, λ), we have that M_C satisfies

$$\begin{aligned} M_C \subseteq \Bigg\{ \mu \in \mathrm{Meas}(\Omega) \mid &\int_\Omega \left[c + \sum_{t=0}^T \lambda_t - \varepsilon \right] \mathrm{d}\mu(x) + \varepsilon \mu(\Omega) \\ &\le C + \lambda - \inf_{\mu \in \mathrm{Meas}(\Omega)} \mathcal{D}(\mu) =: D \Bigg\} \end{aligned}$$

where $D \in \mathbb{R}$ since $\mathcal{D}(\cdot)$ is lower bounded by hypothesis. By (42) and for large enough C, the set M_C is nonempty, and the same arguments in Items 2, 3 and 4 of the proof of Theorem 4.3 can be applied to conclude that the set M_C is $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$-compact. Moreover, we see that $M_C \subseteq \mathcal{N}$; hence, it is also compact in the topology $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))|_\mathcal{N}$. We finally verify $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))|_\mathcal{N}$-lower semicontinuity of $\mathcal{K}(\cdot, \Delta, \lambda)$ on $\mathcal{N}$ for every $(\Delta, \lambda) \in (\mathcal{H} \times \mathbb{R}^{T+1})$. To see this, observe that arguing as in Item 2 of the proof of Theorem 4.3 we get that $\mu \mapsto \int_\Omega \left[c - I^\Delta + \sum_{t=0}^T \lambda_t \right] \mathrm{d}\mu - \sum_{t=0}^T \lambda_t$ is $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))|_\mathcal{N}$-lower semicontinuous, while $\mathcal{D}$ is by definition $\sigma(\mathrm{ca}(\Omega), \mathcal{E})|_\mathcal{N}$ lower semicontinuous (being supremum of linear functionals each continuous in such a topology). Since sum of lower semicontinuous functions is lower semicontinuous, the desired lower semicontinuity of $\mathcal{K}(\cdot, \Delta, \lambda)$ follows. All the hypotheses of Theorem A.5 are now verified, and we may then interchange sup and inf in RHS of (52) and obtain

$$\inf_{\mu\in\mathcal{N}} \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \mathcal{K}(\mu,\Delta,\lambda) = \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \inf_{\mu\in\mathcal{N}} \mathcal{K}(\mu,\Delta,\lambda) \overset{(\star)}{=} \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \inf_{\mu\in\mathrm{Meas}(\Omega)} \mathcal{K}(\mu,\Delta,\lambda)$$

$$= \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \inf_{\mu\in\mathrm{Meas}(\Omega)} \left(\int_\Omega \left[c - I^\Delta + \sum_{t=0}^T \lambda_t\right] \mathrm{d}\mu + \mathcal{D}(\mu)\right) - \sum_{t=0}^T \lambda_t, \tag{53}$$

where in $(\star)$ we used the fact that $\mathcal{K}(\mu,\Delta,\lambda) = +\infty$ on the complementary of $\mathcal{N}$ in $\mathrm{Meas}(\Omega)$, for every $(\Delta,\lambda)\in\mathcal{H}\times\mathbb{R}^{T+1}$.

We apply now Theorem 4.3 to the inner infimum with the cost functional $c - I^\Delta + \sum_{t=0}^T \lambda_t$, observing that, since we are assuming $\mathrm{dom}(U) + \mathbb{R}^{T+1} = \mathrm{dom}(U)$ (see (42)), the condition (34) is satisfied. We get that

$$(53) = \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \sup_{\varphi\in\Phi_{\Delta,\lambda}(c)} \left(U(\varphi) - \sum_{t=0}^T \lambda_t\right)$$

where $\Phi_{\Delta,\lambda}(c)$, which depends on $\Delta,\lambda\in\mathcal{H}\times\mathbb{R}^{T+1}$, is defined according to (35) by

$$\Phi_{\Delta,\lambda}(c) = \left\{\varphi\in\mathrm{dom}(U),\ \sum_{t=0}^T \varphi_t(x) \le c(x) - I^\Delta(x) + \sum_{t=0}^T \lambda_t \ \ \forall x\in\Omega\right\}.$$

From (42), $(\varphi_t - \lambda_t)_t \in \mathrm{dom}(U)$ and we can absorb λ in φ obtaining $\Phi_{\Delta,\lambda}(c) = \Phi_\Delta(c) + \lambda$, $\forall\lambda\in\mathbb{R}^{T+1}$, $\Delta\in\mathcal{H}$, with $\Phi_\Delta(c)$ given in (44), so that

$$(53) = \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \sup_{\varphi\in\Phi_\Delta(c)} \left(U(\varphi+\lambda) - \sum_{t=0}^T \lambda_t\right)$$

$$= \sup_{\Delta\in\mathcal{H}} \sup_{\varphi\in\Phi_\Delta(c)} \sup_{\lambda\in\mathbb{R}^{T+1}} \left(U(\varphi+\lambda) - \sum_{t=0}^T \lambda_t\right).$$

We now recognize the expression in (27) and we conclude that

$$\inf_{\mu\in\mathrm{Meas}(\Omega)} \sup_{\substack{\Delta\in\mathcal{H}\\ \lambda\in\mathbb{R}^{T+1}}} \mathcal{K}(\mu,\Delta,\lambda) = (53) = \sup_{\Delta\in\mathcal{H}} \sup_{\varphi\in\Phi_\Delta(c)} S^U(\varphi),$$

and consequently, recalling our minimax argument,

$$\inf_{Q\in\mathrm{Mart}(\Omega)} \left(E_Q[c(X)] + \mathcal{D}(Q)\right) \overset{\text{Eq.(52)}}{=} (53) = \sup_{\Delta\in\mathcal{H}} \sup_{\varphi\in\Phi_\Delta(c)} S^U(\varphi).$$

$\square$

Remark 4.5 (i) The assumptions of Theorem 4.4 are reasonably weak and are satisfied, for example, if: $\mathrm{dom}(U) = \mathcal{E}$, there exists a $\widehat{\mu} \in \mathrm{Meas}(\Omega) \cap \partial U(0)$ such that $c \in L^1(\widehat{\mu})$, and c is lower semicontinuous. Indeed, for all $\mu \in \mathrm{Meas}(\Omega)$, $\mathcal{D}(\mu) \geq U(0) - 0 > -\infty$. Clearly $\mathrm{dom}(U) + \mathbb{R}^{T+1} = \mathrm{dom}(U)$. Finally, $\widehat{\mu} \in \mathcal{N}$, because $c \in L^1(\widehat{\mu})$ and $-\infty < U(0) \leq \mathcal{D}(\widehat{\mu}) \leq 0$, by definition of $\mathcal{D}$.

(ii) The step (46) = (47) is the crucial point where compactness of the sets $K_0, \ldots, K_T \subseteq \mathbb{R}$ is necessary for a smooth argument, since integrability of the underlying stock process is in this case automatically satisfied for all $Q \in \mathrm{Prob}(\Omega)$, not only for $Q \in \mathrm{Mart}(\Omega)$. Also, compactness is key in guaranteeing that the cost functional $c - I^{\Delta} + \sum_t \lambda_t$ is bounded from below, in order to apply Theorem 4.3.

The following result guarantees that, under suitable assumptions, the infimum in (43) is attained. A similar result can be also found in Corollary 2.5 of Doldi and Frittelli (2023).

Proposition 4.6 *Suppose that LHS of* (43) *is finite and that* $\mathcal{D}|_{\mathrm{Meas}(\Omega)}$ *is* $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$*-lower semicontinuous. Then, under the same the assumptions of Theorem* 4.4*, the problem in LHS of* (43) *admits an optimum.*

Proof Similarly to what we argued in Item 2 of the proof of Theorem 4.3, the map $\mu \mapsto \int_\Omega c\, \mathrm{d}\mu$ is $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$-lower semicontinuous, and we deduce the lower semicontinuity of

$$Q \mapsto \mathcal{J}(Q) := E_Q[c] + \sum_{t=0}^{T} \mathcal{D}(Q), \quad Q \in \mathrm{Mart}(\Omega).$$

Moreover, for C big enough the sublevel $\{Q \in \mathrm{Mart}(\Omega) \mid \mathcal{J}(Q) \leq C\}$ is nonempty (since we are assuming LHS of (43) is finite); hence, $\mathcal{J}$ is proper on $\mathrm{Mart}(\Omega)$. Since K_0, ..., K_T are compact, $\mathrm{Prob}(\Omega)$ is $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$-compact (see Aliprantis and Border (2006) Theorem 15.11), and $\mathrm{Mart}(\Omega)$ is $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$-closed because, arguing as in Beiglböck et al. (2013) Lemma 2.3,

$$\mathrm{Mart}(\Omega) = \bigcap_{\Delta \in \mathcal{H}} \left\{ Q \in \mathrm{Prob}(\Omega) \mid \int_\Omega \left(\sum_{t=0}^{T-1} \Delta_t(x_0, \ldots, x_t)(x_{t+1} - x_t) \right) \mathrm{d}Q(x) \leq 0 \right\}.$$

We conclude that $\mathrm{Mart}(\Omega)$ is $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega))$-compact, and $\mathcal{J}$ is lower semicontinuous and proper on it; hence, it attains a minimum. □

In view of our discussion of Sect. 5, we now rephrase the results of Theorem 4.4 as formulated in Corollary 4.7. In particular, this reformulation will come in handy when dealing with subhedging and superhedging dualitites in Corollary 5.3.

For a given proper concave $U : \mathcal{E} \to \mathbb{R}$, we consider the corresponding $V(\cdot) = -U(-\cdot)$ and define

$$S_V(\varphi) := \inf_{\lambda \in \mathbb{R}^{T+1}} \left(V(\varphi + \lambda) - \sum_{t=0}^{T} \lambda_t \right) = -S^U(-\varphi), \quad \varphi \in \mathrm{dom}(V),$$

where $\mathrm{dom}(V) := \{\varphi \in \mathcal{E} \mid V(\varphi) < +\infty\} = -\mathrm{dom}(U)$ and S^U is defined as in (27).

Furthermore, given two functionals $c : \Omega \to (-\infty, +\infty]$, $d : \Omega \to [-\infty, +\infty)$ we introduce the sets

$$\mathcal{S}_{\mathrm{sub}}(c) := \left\{ \varphi \in \mathrm{dom}(U) \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi(x_t) + I^{\Delta}(x) \leq c(x) \ \ \forall x \in \Omega \right\} \tag{54}$$

$$\mathcal{S}_{\mathrm{sup}}(d) := \left\{ \varphi \in \mathrm{dom}(V) \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi(x_t) + I^{\Delta}(x) \geq d(x) \ \ \forall x \in \Omega \right\} \tag{55}$$

and observe that $\mathcal{S}_{\mathrm{sup}}(\Psi) = -\mathcal{S}_{\mathrm{sub}}(-\Psi)$.

Corollary 4.7 *Suppose that the assumptions in Theorem 4.4 are satisfied, that $d : \Omega \to [-\infty, +\infty)$ is upper semicontinuous and that $\{\mu \in \mathrm{Meas}(\Omega) \cap \mathrm{dom}(\mathcal{D}) \mid \int_\Omega d \, \mathrm{d}\mu > -\infty\} \neq \emptyset$. Then the following hold*

$$\inf_{Q \in \mathrm{Mart}(\Omega)} \big(E_Q[c(X)] + \mathcal{D}(Q)\big) = \sup_{\varphi \in \mathcal{S}_{sub}(c)} S^U(\varphi), \tag{56}$$

$$\sup_{Q \in \mathrm{Mart}(\Omega)} \big(E_Q[d(X)] - \mathcal{D}(Q)\big) = \inf_{\varphi \in \mathcal{S}_{sup}(d)} S_V(\varphi). \tag{57}$$

Proof Equation (56) is an easy rephrasing of the corresponding (43). As to (57), we observe that for $c := -d$ we get from (56)

$$\begin{aligned} &\sup_{\varphi \in \mathcal{S}_{\mathrm{sub}}(-d)} S^U(\varphi) \\ &= \inf_{Q \in \mathrm{Mart}(\Omega)} \big(E_Q[-d(X)] + \mathcal{D}(Q)\big) = - \sup_{Q \in \mathrm{Mart}(\Omega)} \big(E_Q[d(X)] - \mathcal{D}(Q)\big). \end{aligned}$$

Observing that

$$\mathcal{S}_{\mathrm{sup}}(d) = -\mathcal{S}_{\mathrm{sub}}(-d)$$

and that $S_V(\cdot) = -S^U(-\cdot)$ on $\mathrm{dom}(V)$ we get $\sup_{\varphi \in \mathcal{S}_{\mathrm{sub}}(-d)} S^U(\varphi) = -\inf_{\varphi \in \mathcal{S}_{\mathrm{sup}}(d)} S_V(\varphi)$. This completes the proof. ☐

4.3 Duality in an additive setting

Differently from Sect. 4.1, we will now assume an additive structure of U and $\mathcal{D}$. In the whole Sect. 4.3, we consider for each $t = 0, \dots, T$ a vector subspace $\mathcal{E}_t \subseteq \mathcal{C}_b(K_t)$ such that $\mathbb{R} \subseteq \mathcal{E}_t$ and set $\mathcal{E} = \mathcal{E}_0 \times \cdots \times \mathcal{E}_T$. Note that this automatically implies that $\mathcal{E} + \mathbb{R}^{T+1} = \mathcal{E}$. Furthermore, $\mathcal{E}$ can be seen as a subspace of $\mathcal{C}_b(\Omega, \mathbb{R}^{T+1})$ once $\mathcal{E}_0, \dots, \mathcal{E}_T$ can be interpreted as subspaces of $\mathcal{C}_b(\Omega)$.

The following result provides the form of the penalization in an additive setup and the duality $\mathfrak{P}(c) = \mathfrak{D}(c)$.

Theorem 4.8 *Suppose that, for each $t = 0, \dots, T$, $\mathcal{E}_t \subseteq \mathcal{C}_b(K_t)$ is a vector subspace satisfying* $\mathrm{Id}_t \in \mathcal{E}_t$ *and* $\mathbb{R} \subseteq \mathcal{E}_t$ *and that $S_t : \mathcal{E}_t \to \mathbb{R}$ is a concave, cash additive functional null in* 0. *Consider for every $t = 0, \dots, T$ the penalizations*

$$\mathcal{D}_t(Q_t) := \sup_{\varphi_t \in \mathcal{E}_t} \left(S_t(\varphi_t) - \int_{K_t} \varphi_t \, \mathrm{d}Q_t \right) \quad \textit{for } Q_t \in \mathrm{Prob}(K_t),$$

and set $\mathcal{D}(Q) := \sum_{t=0}^{T} \mathcal{D}_t(Q_t)$. Let $c : \Omega \to (-\infty, +\infty]$ be lower semi-continuous and let $\mathfrak{D}(c)$ and $\mathfrak{P}(c)$ be defined, respectively, in (13) *and* (18). *If* $\mathcal{N} := \left\{\mu \in \mathrm{Meas}(\Omega) \cap \mathrm{dom}(\mathcal{D}) \mid \int_\Omega c \, \mathrm{d}\mu < +\infty\right\} \neq \emptyset$ *then*

$$\mathfrak{P}(c) = \mathfrak{D}(c).$$

Proof Set $\mathcal{E} = \mathcal{E}_0 \times \cdots \times \mathcal{E}_T$ and $U(\varphi) := \sum_{t=0}^{T} S_t(\varphi_t)$, for $\varphi \in \mathcal{E}$, and let $\mathcal{D}$ defined as in (31) for $M = T$. For any $\mu \in \mathrm{Meas}(\Omega)$, we have $\mathcal{D}(\mu) \geq \sum_{t=0}^{T} S_t(0) - 0 = 0$; hence, $\mathcal{D}$ is lower bounded on $\mathrm{Meas}(\Omega)$. Observe that $\mathrm{dom}(U) = \mathcal{E}$, which implies $\mathrm{dom}(U) + \mathbb{R}^{T+1} = \mathrm{dom}(U)$, and that we are in Setup 3.2 of Doldi and Frittelli (2023). By Doldi and Frittelli (2023) Lemma 3.3 $S^U(\varphi) = \sum_{t=0}^{T} S_t^{U_t}(\varphi_t) = \sum_{t=0}^{T} S_t(\varphi_t)$, since $S_0, \dots, S_T$ are Cash Additive, and $\mathcal{D}$ coincides on $\mathrm{Mart}(\Omega)$ with the penalization term $Q \mapsto \sum_{t=0}^{T} \mathcal{D}_t(Q_t)$, as provided in the statement of this theorem. Since all the assumptions of Theorem 4.4 are fulfilled, we can apply Corollary 4.7, which yields exactly $\mathfrak{D}(c) = \mathfrak{P}(c)$. ☐

Assumption 4.9 We consider concave, upper semicontinuous nondecreasing functions $u_0, \dots, u_T : \mathbb{R} \to [-\infty, +\infty)$ with $u_0(0) = \cdots = u_T(0) = 0$, $u_t(x) \leq x \ \forall x \in \mathbb{R}$ (that is $1 \in \partial u_0(0) \cap \cdots \cap \partial u_T(0)$). For each $t = 0, \dots, T$ we define $v_t(x) := -u_t(-x)$, $x \in \mathbb{R}$, and its convex conjugate

$$v_t^*(y) := \sup_{x \in \mathbb{R}} (xy - v_t(x)) = \sup_{x \in \mathbb{R}} (u_t(x) - xy), \qquad y \in \mathbb{R}. \tag{58}$$

By Fenchel–Moreau theorem, we recall that $v_t(y) = v_t^{**}(y) = \sup_{x \in \mathbb{R}} (xy - v_t^*(y))$ for all $y \in \mathbb{R}$ and that v_t^* is convex, lower semicontinuous and lower bounded on $\mathbb{R}$.

Example 4.10 Assumption 4.9 is satisfied by a wide range of functions. Just to mention a few with various peculiar features, we might take u_t of the following forms: $u_t(x) = 1 - \exp(-x)$, whose convex conjugate is given by $v_t^*(y) = -\infty$ for $y < 0$, $v_t^*(0) = 0$, $v_t^*(y) = (y \log(y) - y + 1)$ for $y > 0$; $u_t(x) = \alpha x 1_{(-\infty,0]}(x)$ for $\alpha \geq 1$, so that $v_t^*(y) = +\infty$ for $y < 0$, $v_t^*(y) = 0$ for $y \in [0, \alpha]$, $v_t^*(y) = +\infty$ for $y > \alpha$; $u_t(x) = \log(x+1)$ for $x > -1$, $u_t(x) = -\infty$ for $x \leq -1$, so that $v_t^*(y) = +\infty$ for $y \leq 0$, $v_t^*(y) = y - \log(y) - 1$ for $y > 0$; $u_t(x) = -\infty$ for $x \leq -1$, $u_t(x) = \frac{x}{x+1}$ for $x > -1$ so that $v_t^*(y) = -\infty$ for $y < 0$, $v_t^*(y) = y - 2\sqrt{y} + 1$ for $y \geq 0$; $u_t(x) = -\infty$ for $x < 0$, $u_t(x) = 1 - \exp(-x)$ for $x \geq 0$, so that $v_t^*(y) = +\infty$ for $y < 0$, $v_t^*(y) = y \log(y) - y + 1$ for $0 \leq y \leq 1$, $v_t^*(y) = 0$ for $y > 1$.

Fix $\widehat{\mu}_t \in \mathrm{Meas}(K_t)$. For $\mu \in \mathrm{Meas}(K_t)$, we define

$$\mathcal{D}_{v_t^*,\widehat{\mu}_t}(\mu) := \begin{cases} \int_{K_t} v_t^*\left(\frac{\mathrm{d}\mu}{\mathrm{d}\widehat{\mu}_t}\right)\mathrm{d}\widehat{\mu}_t & \text{if } \mu \ll \widehat{\mu}_t \\ +\infty & \text{otherwise.} \end{cases} \tag{59}$$

The next proposition, whose proofs can be found in Liero et al. (2018), Theorem 2.7 and Remark 2.8, or (Doldi and Frittelli 2023), provides the dual representation of the divergence terms.

Proposition 4.11 *Take $u_0, \dots, u_T$ satisfying Assumption 4.9, and suppose* $\mathrm{dom}(u_0) = \cdots = \mathrm{dom}(u_T) = \mathbb{R}$. *Let $\widehat{\mu}_t \in \mathrm{Meas}(K_t)$ and $v_t(\cdot) := -u_t(-\cdot)$, $t = 0, \dots, T$. Then*

$$\mathcal{D}_{v_t^*,\widehat{\mu}_t}(\mu) = \sup_{\varphi_t \in \mathcal{C}_b(K_t)} \left(\int_{K_t} \varphi_t(x_t)\,\mathrm{d}\mu(x_t) - \int_{K_t} v_t(\varphi_t(x_t))\,\mathrm{d}\widehat{\mu}_t(x_t) \right). \tag{60}$$

5 Applications of the main theorems of Sect. 4

In this section, we suppose the following requirements are fulfilled:

Standing Assumption 5.1 $\Omega := K_0 \times \cdots \times K_T$ *for compact sets $K_0, \dots, K_T \subseteq \mathbb{R}$ and $K_0 = \{x_0\}$; the functional $c : \Omega \to (-\infty, +\infty]$ is lower semicontinuous and $d : \Omega \to [-\infty, +\infty)$ is upper semicontinuous;* $\mathrm{Mart}(\Omega) \neq \emptyset$; $\widehat{Q} \in \mathrm{Mart}(\Omega)$ *is a given probability measure with marginals $\widehat{Q}_0, \dots, \widehat{Q}_T$; $c, d \in L^1(\widehat{Q})$.*

5.1 Subhedging and superhedging

As it will become clear from the proofs, in all the results in Sect. 5.1 the functional U is real-valued on the whole $\mathcal{E}$, that is, $\mathrm{dom}(U) = \mathcal{E}$. Thus, we will exploit Theorem 4.4 and Corollary 4.7, in particular (54) and (55), in the case $\mathrm{dom}(U) = \mathrm{dom}(V) = \mathcal{E}$.

We recall from (21) that

$$\begin{aligned} &U_{\widehat{Q}_t}(\varphi_t) \\ &\quad= \sup_{\alpha,\lambda\in\mathbb{R}} \left(\int_{K_t} u_t(\varphi_t(x_t) + \alpha \mathrm{Id}_t(x_t) + \lambda)\mathrm{d}\widehat{Q}_t(x_t) - (\alpha x_0 + \lambda) \right), \quad \varphi_t \in \mathcal{C}_b(K_t), \\ &V_{\widehat{Q}_t}(\varphi_t) = -U_{\widehat{Q}_t}(-\varphi_t) \\ &\quad= \inf_{\alpha,\lambda\in\mathbb{R}} \left(\int_{K_t} v_t(\varphi_t(x_t) + \alpha \mathrm{Id}_t(x_t) + \lambda)\mathrm{d}\widehat{Q}_t(x_t) + (\alpha x_0 + \lambda) \right). \end{aligned} \tag{61}$$

In Sects. 4.1 and 4.2 of Doldi and Frittelli (2023), a counterpart to (61) was investigated, without the additional supremum over α above. Thus, the following results generalize those in Doldi and Frittelli (2023), establishing at the same time stock additivity of U and of V.

We observe that Assumption 4.9 does **not** impose that the functions u_t are real-valued on the whole $\mathbb{R}$. Nevertheless, for the functionals $U_{\widehat{Q}_t}, V_{\widehat{Q}_t}$ we have:

Lemma 5.2 *Under Assumption 4.9, for each $t = 0, \dots, T$*

1. *$U_{\widehat{Q}_t}$ and $V_{\widehat{Q}_t}$ are real-valued on $\mathcal{C}_b(K_t)$ and null in 0.*
2. *$U_{\widehat{Q}_t}$ and $V_{\widehat{Q}_t}$ are concave and convex, respectively, and both nondecreasing.*
3. *$U_{\widehat{Q}_t}$ and $V_{\widehat{Q}_t}$ are stock additive on $\mathcal{C}_b(K_t)$, namely for every $\alpha_t, \lambda_t \in \mathbb{R}$ and $\varphi_t \in \mathcal{C}_b(K_t)$*

$$U_{\widehat{Q}_t}(\varphi_t + \alpha_t \mathrm{Id}_t + \lambda_t) = U_{\widehat{Q}_t}(\varphi_t) + \alpha_t x_0 + \lambda_t,$$
$$V_{\widehat{Q}_t}(\varphi_t + \alpha_t \mathrm{Id}_t + \lambda_t) = V_{\widehat{Q}_t}(\varphi_t) + \alpha_t x_0 + \lambda_t.$$

Proof Since $V_{\widehat{Q}_t}(\varphi_t) = -U_{\widehat{Q}_t}(-\varphi_t)$, w.l.o.g. we prove the claims only for $U_{\widehat{Q}_t}$. Clearly $U_{\widehat{Q}_t}(\varphi_t) > -\infty$, as we may choose $\lambda_t \in \mathbb{R}$ so that $(\varphi_t + 0\mathrm{Id}_t + \lambda_t) \in \mathrm{dom}(u) \supseteq [0, +\infty)$. Furthermore,

$$\begin{aligned} U_{\widehat{Q}_t}(\varphi_t) &\overset{1\in\partial U_t(0)}{\leq} \sup_{\alpha,\lambda\in\mathbb{R}} \left(\int_{K_t} (\varphi_t + \alpha \mathrm{Id}_t + \lambda)\, \mathrm{d}\widehat{Q}_t - (\alpha x_0 + \lambda) \right) \\ &\overset{\widehat{Q}\in\mathrm{Mart}(\Omega)}{=} \sup_{\alpha,\lambda\in\mathbb{R}} \left(\int_{K_t} \varphi_t \, \mathrm{d}\widehat{Q}_t + (\alpha x_0 + \lambda - \alpha x_0 - \lambda) \right) \leq \|\varphi_t\|_\infty . \end{aligned}$$

Finally, $0 = \int_{K_t} u(0)\, \mathrm{d}\widehat{Q}_t \leq U_{\widehat{Q}_t}(0) \leq \|0\|_\infty$.

Item 2: trivial from the definitions. Item 3: we see that

$$\begin{aligned} &U_{\widehat{Q}_t}(\varphi_t + \alpha_t \mathrm{Id}_t + \lambda_t) \\ &\quad = \sup_{\alpha,\lambda\in\mathbb{R}} \left(\int_{K_t} u_t\left(\varphi_t(x_t) + (\alpha + \alpha_t)x_t + (\lambda + \lambda_t)\right) \mathrm{d}\widehat{Q}_t(x_t) - (\alpha x_0 + \lambda) \right) \\ &= \sup_{\alpha,\lambda\in\mathbb{R}} \left(\int_{K_t} u_t\left(\varphi_t(x_t) + (\alpha + \alpha_t)x_t + (\lambda + \lambda_t)\right) \mathrm{d}\widehat{Q}_t(x_t) - ((\alpha_t + \alpha)x_0 + (\lambda_t + \lambda)) \right) \\ &\quad + \alpha_t x_0 + \lambda_t, \end{aligned}$$

in which we recognize the definition of $U_{\widehat{Q}_t}(\varphi_t) + \alpha_t x_0 + \lambda_t$. □

As in Beiglböck et al. (2013), **in the next two Corollaries** we suppose that the elements in $\mathcal{E}_t$ represent portfolios obtained combining deterministic amounts, units of the underlying stock at time t (x_t), and call options with maturity t, that is, $\mathcal{E}_t$ consists of all the functions in $\mathcal{C}_b(K_t)$ with the following form:

$$\varphi_t(x_t) = a + b x_t + \sum_{n=1}^{N} c_n (x_t - K_n)^+, \text{ for } a, b, c_n, k_n \in \mathbb{R},\ x_t \in K_t$$

and take $\mathcal{E} = \mathcal{E}_0 \times \cdots \times \mathcal{E}_T$. As shown in the proof, one could as well take $\mathcal{E} = \mathcal{C}_b(K_0) \times \cdots \times \mathcal{C}_b(K_T)$ preserving validity of (62), (63), (68) and (69). As for Lemma 5.2, the following result extends Corollary 4.3 of Doldi and Frittelli (2023) to cover the stock additive case: Indeed, the definition of $U_{\widehat{Q}_t}(\varphi_t)$ used in the next corollary is different from the one in Corollary 4.3 of Doldi and Frittelli (2023).

Corollary 5.3 *Take $u_0, \dots, u_T$ satisfying Assumption 4.9, and suppose* $\mathrm{dom}(u_0) = \dots = \mathrm{dom}(u_T) = \mathbb{R}$. *Then the following equalities hold:*

$$\inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q[c(X)] + \sum_{t=0}^{T} \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q_t) \right) = \sup \left\{ \sum_{t=0}^{T} U_{\widehat{Q}_t}(\varphi_t) \mid \varphi \in \mathcal{S}_{sub}(c) \right\} \tag{62}$$

$$\sup_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q[d(X)] - \sum_{t=0}^{T} \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q_t) \right) = \inf \left\{ \sum_{t=0}^{T} V_{\widehat{Q}_t}(\varphi_t) \mid \varphi \in \mathcal{S}_{sup}(d) \right\} \tag{63}$$

Proof We prove (62), since (63) can be obtained in a similar fashion. Set $U(\varphi) = \sum_{t=0}^{T} U_{\widehat{Q}_t}(\varphi_t)$ for $\varphi \in \mathcal{E}$. We observe that $\mathcal{E}_t$ consists of all piecewise linear functions on K_t, which are norm dense in $\mathcal{C}_b(K_t)$. By Lemma 5.2 for each $t = 0, \dots, T$, the monotone concave functional $\varphi_t \mapsto U_{\widehat{Q}_t}(\varphi_t)$ is actually well defined, finite-valued, concave and nondecreasing on the whole $\mathcal{C}_b(K_t)$. Hence, by the extended Namioka–Klee theorem (see Biagini and Frittelli (2009)) it is norm continuous on $\mathcal{C}_b(K_t)$ and we can take $\mathcal{E} = \mathcal{C}_b(K_t) \times \dots \times \mathcal{C}_b(K_T)$ in place of $\mathcal{E}_0 \times \dots \times \mathcal{E}_T$ in the RHS of (62) and prove equality to LHS in this more comfortable case (notice that $\mathcal{S}_{\mathrm{sub}}(c)$ depends on $\mathcal{E}$). We also observe that in this case we are in Setup of Sect. 4.3. Define $\mathcal{D}$ as in (31) with $M = T$. Using the facts that if $\varphi_t \in \mathcal{E}_t, \alpha, \lambda \in \mathbb{R}$ then $(\varphi_t + \alpha \mathrm{Id}_t + \lambda) \in \mathcal{E}_t$, that $Q \in$Mart(Ω) and that $v_t(\cdot) := -u_t(-\cdot)$ one may easily check that

$$\begin{aligned}
\mathcal{D}(Q) &:= \sup_{\varphi \in \mathcal{E}} \left(U(\varphi) - \sum_{t=0}^{T} \int_{K_t} \varphi_t \, \mathrm{d}Q_t \right) \\
&= \sup_{\varphi \in \mathcal{E}} \left(\sum_{t=0}^{T} \int_{K_t} u_t(\varphi_t(x_t)) \, \mathrm{d}\widehat{Q}_t(x_t) - \sum_{t=0}^{T} \int_{K_t} \varphi_t \, \mathrm{d}Q_t \right) \\
&= \sum_{t=0}^{T} \sup_{\psi_t \in \mathcal{E}_t} \left(\int_{K_t} \psi_t \, \mathrm{d}Q_t - \int_{K_t} v_t(\psi_t(x_t)) \, \mathrm{d}\widehat{Q}_t(x_t) \right) \\
&= \sum_{t=0}^{T} \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q_t), \quad \forall Q \in \mathrm{Mart}(\Omega)
\end{aligned} \tag{64}$$

where the last equality follows from Proposition 4.11 Eq. (60). The standing assumption 4.1 is satisfied. Indeed, from Assumption 4.9 we have $v_0^*(1), \dots, v_T^*(1) < +\infty$; hence, $\mathcal{D}_{v_t^*, \widehat{Q}_t}(\widehat{Q}_t) = \int_{K_t} v_t^* \left(\frac{\mathrm{d}\widehat{Q}_t}{\mathrm{d}\widehat{Q}_t} \right) \mathrm{d}\widehat{Q}_t < +\infty$, and therefore, $\widehat{Q} \in \mathrm{dom}(\mathcal{D})$. Recalling that $c \in L^1(\widehat{Q})$, this in turns yields $\widehat{Q} \in \mathcal{N} = \left\{ \mu \in \mathrm{Meas}(\Omega) \cap \mathrm{dom}(\mathcal{D}) \mid \int_\Omega c \, \mathrm{d}\mu < +\infty \right\}$. Moreover, by Lemma 5.2 Item 1, $\mathrm{dom}(U) = \mathcal{E}$, and for every

$\mu \in \mathrm{Meas}(\Omega)$ $\mathcal{D}(\mu) \geq U(0) - 0 = 0$; hence, $\mathcal{D}$ is lower bounded on the whole $\mathrm{Meas}(\Omega)$. We conclude that U and $\mathcal{D}$ satisfy the assumptions of Theorem 4.4.

Using Lemma 3.3 of Doldi and Frittelli (2023) and the fact that $U_{\widehat{Q}_0}, \ldots, U_{\widehat{Q}_T}$ are cash additive, we get $S^U(\varphi) = \sum_{t=0}^T S^{U_{\widehat{Q}_t}}(\varphi_t) = \sum_{t=0}^T U_{\widehat{Q}_t}(\varphi_t) = U(\varphi)$, and by Corollary 4.7 Eq. (56), we obtain

$$\inf_{Q \in \mathrm{Mart}(\Omega)} \left(E_Q\left[c(X)\right] + \sum_{t=0}^{T} \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q_t) \right) = \sup \left\{ \sum_{t=0}^{T} U_{\widehat{Q}_t}(\varphi_t) \mid \varphi \in \mathcal{S}_{\mathrm{sub}}(c) \right\}.$$

□

We stress the fact that in Corollary 5.3 we assume that all the functions $u_0, \ldots, u_T$ are real-valued on the whole $\mathbb{R}$.

Proposition 5.4 *The following dual representations hold:*

$$\begin{aligned} V_{\widehat{Q}_t}(\varphi_t) &= \max_{Q \in \mathrm{Mart}(K_0 \times K_t)} \left(E_Q[\varphi_t] - \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q) \right), \\ U_{\widehat{Q}_t}(\varphi_t) &= \min_{Q \in \mathrm{Mart}(K_0 \times K_t)} \left(E_Q[\varphi_t] + \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q) \right), \end{aligned} \tag{65}$$

where, with a slight abuse of notation and consistently with (41), $\mathrm{Mart}(K_0 \times K_t) = \{Q \in \mathrm{Prob}(K_t) \mid E_Q[\mathrm{Id}_t] = x_0\}$.

Proof We prove the dual representation for $V_{\widehat{Q}_t}$ as the other one follows by a change of signs. By extended Namioka–Klee theorem (Biagini and Frittelli (2009)), together with the compactness of the underlying canonical space we have that

$$V_{\widehat{Q}_t}(\varphi_t) = \max_{\gamma \in (\mathcal{C}_b(K_t))^*} \left(\int_{K_t} \varphi_t \mathrm{d}\gamma - (V_{\widehat{Q}_t})^*(\gamma) \right)$$

where the conjugate $(V_{\widehat{Q}_t})^*(\gamma)$ is given as usual by $(V_{\widehat{Q}_t})^*(\gamma) = \sup_{\psi_t \in \mathcal{C}_b(K_t)} \Big(\int_{K_t} \psi_t \mathrm{d}\gamma - V_{\widehat{Q}_t}(\psi_t) \Big)$. We show that $(V_{\widehat{Q}_t})^*(\gamma) < +\infty$ implies $\gamma \in \mathrm{Mart}(K_0 \times K_t)$. By standard monotonicity and cash additivity arguments (see, e.g., Föllmer and Schied (2016) Remark 4.18), it can be seen that $(V_{\widehat{Q}_t})^*(\gamma)$ implies that γ is a nonnegative normalized element of the dual space $(\mathcal{C}_b(K_t))^*$. Since K_t is compact, γ is then identified with an element of $\mathrm{Prob}(K_t)$. We show that the martingale property must hold: For any $\gamma \in \mathrm{Prob}(K_t)$,

$$\begin{aligned} (V_{\widehat{Q}_t})^*(\gamma) &\geq \sup_{\alpha_t \in \mathbb{R}} \left(\int_{K_t} \alpha_t \mathrm{Id}_t \mathrm{d}\gamma - V_{\widehat{Q}_t}(\alpha_t \mathrm{Id}_t) \right) \\ &\overset{L.5.2.3}{\geq} \sup_{\alpha_t \in \mathbb{R}} \left(\int_{K_t} \alpha_t \mathrm{Id}_t \mathrm{d}\gamma - \alpha_t x_0 \right) = \sup_{\alpha_t \in \mathbb{R}} \alpha_t \left(\int_{K_t} \mathrm{Id}_t \mathrm{d}\gamma - x_0 \right). \end{aligned}$$

Now, the last term in RHS is finite if and only if $\int_{K_t} \mathrm{Id}_t \mathrm{d}\gamma = x_0$. To conclude, observe that arguing as we did to obtain (64), we can also show here that for any $Q \in \mathrm{Mart}(K_0 \times K_t)$ we have

$$(V_{\widehat{Q}_t})^*(Q) = \sup_{\psi_t \in \mathcal{C}_b(K_t)} \left(\int_{K_t} \psi_t \,\mathrm{d}Q_t - \int_{K_t} v_t(\psi_t)\,\mathrm{d}\widehat{Q}_t \right) = \mathcal{D}_{v_t^*, \widehat{Q}_t}(Q_t)$$

by Proposition 4.11 Eq. (60). □

Proposition 5.5 *For the exponential utility function $u_t(x) = 1 - e^{-x}$, there exists $\widehat{\alpha} \in \mathbb{R}$ satisfying (23) and for which (22) holds.*

Proof One can verify that $\mathcal{D}_{v_t^*, \widehat{Q}_t}(Q_t) = H(Q_t, \widehat{Q}_t)$ (see (59)). Hence, from (65),

$$U_{\widehat{Q}_t}(\varphi_t) = \min_{Q \in \mathrm{Mart}(K_0 \times K_t)} \left(E_Q[\varphi_t] + H(Q, \widehat{Q}_t) \right).$$

Set $\widehat{E}[\cdot] := E_{\widehat{Q}_t}[\cdot]$ and

$$f(\alpha, \lambda) := \widehat{E}[u_t(\varphi_t + \alpha \mathrm{Id}_t + \lambda)] - (\alpha x_0 + \lambda), \quad \alpha \in \mathbb{R}, \lambda \in \mathbb{R}.$$

The first-order conditions for $U_{\widehat{Q}_t}(\varphi_t) := \sup_{\alpha,\lambda} f(\alpha, \lambda)$ are:

$$\widehat{E}[\mathrm{Id}_t \exp(-\varphi_t - \alpha \mathrm{Id}_t - \lambda)] = x_0, \tag{66}$$

$$\widehat{E}[\exp(-\varphi_t - \alpha \mathrm{Id}_t - \lambda)] = 1. \tag{67}$$

Assuming there exists $\widehat{\alpha} \in \mathbb{R}$ satisfying (23), and taking $\widehat{\lambda} = \log \widehat{E}(\exp(-\varphi_t - \widehat{\alpha}\mathrm{Id}_t))$, we easily get that (66) and (67) are satisfied, and we compute

$$U_{\widehat{Q}_t}(\varphi_t) = \sup_{\alpha,\lambda} f(\alpha, \lambda) = f(\widehat{\alpha}, \widehat{\lambda}) = -\widehat{\alpha} x_0 - \widehat{\lambda} = -\widehat{\alpha} x_0 - \ln(\widehat{E}[\exp(-\varphi_t(x_t) - \widehat{\alpha} x_t)]).$$

From the definition of $\frac{\mathrm{d}Q_{\widehat{\alpha}}}{\mathrm{d}\widehat{Q}_t}$ in (24), we obtain:

$$\widehat{E}[\exp(-\varphi_t(x_t) - \widehat{\alpha} x_t)] = \frac{\exp(-\varphi_t(x_t) - \widehat{\alpha} x_t)}{\frac{\mathrm{d}Q_{\widehat{\alpha}}}{\mathrm{d}\widehat{Q}_t}},$$

so that

$$U_{\widehat{Q}_t}(\varphi_t) = -\widehat{\alpha} x_0 - \ln(\widehat{E}[\exp(-\varphi_t(x_t) - \widehat{\alpha} x_t)]) = -\widehat{\alpha} x_0 + \varphi_t(x_t) + \widehat{\alpha} x_t + \ln\left(\frac{\mathrm{d}Q_{\widehat{\alpha}}}{\mathrm{d}\widehat{Q}_t}\right).$$

Since $U_{\widehat{Q}_t}(\varphi_t) = f(\widehat{\alpha}, \widehat{\lambda}) \in \mathbb{R}$, we conclude:

$$U_{\widehat{Q}_t}(\varphi_t) = E_{Q_{\widehat{\alpha}}}[U_{\widehat{Q}_t}(\varphi_t)] = E_{Q_{\widehat{\alpha}}}[\varphi_t(x_t)] + E_{Q_{\widehat{\alpha}}}\left[\ln\left(\frac{\mathrm{d}Q_{\widehat{\alpha}}}{\mathrm{d}\widehat{Q}_t}\right)\right],$$

as $E_{Q_{\widehat{\alpha}}}[\mathrm{Id}_t] = x_0$, which concludes the proof of (22) because $H(Q_{\widehat{\alpha}}, \widehat{Q}_t) := E_{Q_{\widehat{\alpha}}}\left[\ln\left(\frac{\mathrm{d}Q_{\widehat{\alpha}}}{\mathrm{d}\widehat{Q}_t}\right)\right]$. We are only left with proving that such an $\widehat{\alpha} \in \mathbb{R}$ exists. Since $\widehat{Q} \in \mathrm{Mart}(\Omega)$, $E_{\widehat{Q}}[\mathrm{Id}_t] = x_0$ and we must have $x_0 \in [\operatorname{ess\,inf}_{\widehat{Q}}(\mathrm{Id}_t), \operatorname{ess\,sup}_{\widehat{Q}}(\mathrm{Id}_t)]$. If either $x_0 = \operatorname{ess\,inf}_{\widehat{Q}}(\mathrm{Id}_t)$ or $x_0 = \operatorname{ess\,sup}_{\widehat{Q}}(\mathrm{Id}_t)$, then we must have $\widehat{Q}_t = \delta_{\{x_0\}}$ (the latter being a Dirac delta at the point x_0) by the martingale property and a solution to (23) exists trivially. Otherwise, assume $x_0 \in (\operatorname{ess\,inf}_{\widehat{Q}}(\mathrm{Id}_t), \operatorname{ess\,sup}_{\widehat{Q}}(\mathrm{Id}_t))$. By Lemma A.4 we have that $\lim_{\alpha\to+\infty} E_{Q_{\widehat{\alpha}}}[\mathrm{Id}_t] = \operatorname{ess\,inf}_{\widehat{Q}}(\mathrm{Id}_t)$ and $\lim_{\alpha\to-\infty} E_{Q_{\widehat{\alpha}}}[\mathrm{Id}_t] = \operatorname{ess\,sup}_{\widehat{Q}}(\mathrm{Id}_t)$. Furthermore, by dominated convergence theorem $\alpha \mapsto E_{Q_{\widehat{\alpha}}}[\mathrm{Id}_t]$ is continuous, and the existence of a solution $\widehat{\alpha}$ for (23) follows. □

We now take $u_t(x) = x$ for each $t = 0, \dots, T$, and get $U_{\widehat{Q}_t}(\varphi_t) = V_{\widehat{Q}_t}(\varphi_t) = E_{\widehat{Q}_t}[\varphi_t]$. Hence, with an easy computation we have

$$\mathcal{D}_{v_t^*,\widehat{Q}_t}(Q_t) = \begin{cases} 0 & \text{if } Q_t \equiv \widehat{Q}_t \\ +\infty & \text{otherwise.} \end{cases} \quad \text{for all } Q \in \mathrm{Mart}(\Omega).$$

Recalling that $\mathrm{Mart}(\widehat{Q}_1, \dots, \widehat{Q}_T) = \{Q \in \mathrm{Mart}(\Omega) \mid Q_t \equiv \widehat{Q}_t \ \forall t = 0, \dots, T\}$, from Corollary 5.3 we can recover the following result of Beiglböck et al. (2013) (under the more stringent compactness assumption).

Corollary 5.6 *(*Beiglböck et al. 2013 Theorem 1.1 and Corollary 1.2*) The following equalities hold:*

$$\inf_{Q\in\mathrm{Mart}(\widehat{Q}_1,\dots,\widehat{Q}_T)} E_Q[c] = \sup\left\{\sum_{t=0}^{T} E_{\widehat{Q}_t}[\varphi_t] \mid \varphi \in \mathcal{S}_{sub}(c)\right\} \tag{68}$$

$$\sup_{Q\in\mathrm{Mart}(\widehat{Q}_1,\dots,\widehat{Q}_T)} E_Q[d] = \inf\left\{\sum_{t=0}^{T} E_{\widehat{Q}_t}[\varphi_t] \mid \varphi \in \mathcal{S}_{sup}(d)\right\} \tag{69}$$

5.2 Dual representation for generalized OCE associated with the indirect utility function

We now explore the versatility of Corollary 4.7, which can be used beyond the semistatic subhedging and superhedging problems in Sect. 5.1. Note that in Sect. 5.1 we chose for static hedging portfolios the sets $\mathcal{E}_t, t = 0, \dots, T$ consisting of deterministic amounts, units of underlying stock at time t and call options with the same maturity t but different strike prices. This affected the primal problem in the fact that the penalty $\mathcal{D}$ turned out to depend solely on the (one dimensional) marginals of $\widehat{Q}$. Nonetheless, Theorem 4.4 allows to choose for each $t = 0, \dots, T$ a subspace $\mathcal{E}_t \subseteq \mathcal{C}_b(K_0 \times \cdots \times K_t)$, potentially allowing to consider also Asian and path dependent options in the sets $\mathcal{E}_t$. We expect that this would translate in the penalty $\mathcal{D}$ depending no more only on the one dimensional marginals of $\widehat{Q}$. The study of these less restrictive, yet technically more complex cases is left for future research.

In the following, we will treat a slightly different problem, which, however, helps understanding how also the extreme case $\mathcal{E}_t = \mathcal{C}_b(K_0 \times \cdots \times K_t)$, $t = 0, \ldots, T$, is of interest.

Theorem 4.4 yields the following dual robust representation of the generalized optimized certainty equivalent associated with the indirect utility function. We stress here the fact that, again, $\widehat{Q} \in \mathrm{Mart}(\Omega)$ is a fixed martingale measure, but we will not focus anymore on its marginals only, as will become clear in the following.

Theorem 5.7 *Take $u : \mathbb{R} \to \mathbb{R}$ such that $u_0 = \ldots, u_T := u$ satisfy Assumption 4.9 and let v^* be defined in (58) with u in place of u_t. Let $\widehat{U} : \mathcal{C}_b(\Omega) \to \mathbb{R}$ be the associated indirect utility*

$$\widehat{U}(\varphi) := \sup_{\Delta \in \mathcal{H}} \int_\Omega u(\varphi + I^\Delta)\, \mathrm{d}\widehat{Q}$$

and $S^{\widehat{U}}$ be the associated optimized certainty equivalent defined according to (27), namely

$$S^{\widehat{U}}(\varphi) := \sup_{\xi \in \mathbb{R}} \left(\widehat{U}(\varphi + \xi) - \xi \right) \quad \varphi \in \mathcal{C}_b(\Omega)\,.$$

Then for every $c \in \mathcal{C}_b(\Omega)$

$$S^{\widehat{U}}(c) = \inf_{Q \in \mathrm{Mart}(\Omega)} \left(\int_\Omega c\, \mathrm{d}Q + \mathcal{D}_{\widehat{Q}}(Q) \right)$$

where for $\mu \in \mathrm{Meas}(\Omega)$

$$\mathcal{D}_{\widehat{Q}}(\mu) := \begin{cases} \int_\Omega v^* \left(\frac{\mathrm{d}\mu}{\mathrm{d}\widehat{Q}} \right) \mathrm{d}\widehat{Q} & \text{if } \mu \ll \widehat{Q} \\ +\infty & \text{otherwise} \end{cases}.$$

Proof Take $\mathcal{E}_t = \mathcal{C}_b(K_0 \times \cdots \times K_t)$ for $t = 0, \ldots, T$. Define, for $\psi \in \mathcal{E} = \mathcal{E}_0 \times \ldots \times \mathcal{E}_T$, $U(\psi) := \widehat{U}\left(\sum_{t=0}^T \psi_t\right)$. Clearly, $U(\psi) > -\infty$ for any $\psi \in \mathcal{E}$, and since $\widehat{Q} \in \mathrm{Mart}(\Omega)$ and $u(x) \le x$ for all $x \in \mathbb{R}$ we also have $U(\psi) \le \sum_{t=0}^T \|\varphi_t\|_\infty < +\infty$. Moreover, it is easy to verify that defining $\mathcal{D}$ as in (31) for any $Q \in \mathrm{Mart}(\Omega)$ we have

$$\mathcal{D}(Q) := \sup_{\psi \in \mathcal{E}} \left(U(\psi) - \int_\Omega \left(\sum_{t=0}^T \psi_t \right) \mathrm{d}Q \right) = \sup_{\varphi \in \mathcal{C}_b(\Omega)} \left(\int_\Omega u(\varphi)\, \mathrm{d}\widehat{Q} - \int_\Omega \varphi\, \mathrm{d}Q \right)$$

and arguing as in Proposition 4.11, we get $\mathcal{D}(Q) = \mathcal{D}_{\widehat{Q}}(Q)$. From the fact that $u(x) \le x$ for every $x \in \mathbb{R}$, we have $v^*(1) < +\infty$; hence, from Assumption 5.1 $\widehat{Q} \in \mathrm{dom}(\mathcal{D})$. This and $c \in L^1(\widehat{Q})$ in turns yields $\widehat{Q} \in \mathcal{N}$ (see (42)). Moreover, $\mathrm{dom}(U) = \mathcal{E}$, and by definition of $\mathcal{D}$ for any $\mu \in \mathrm{Meas}(\Omega)$, we have $\mathcal{D}(\mu) \ge U(0) - 0 = 0$; hence, $\mathcal{D}$ is lower bounded on the whole $\mathrm{Meas}(\Omega)$. We conclude that U and $\mathcal{D}$ satisfy the assumptions of Theorem 4.4. We then get

$$\inf_{Q\in \text{Mart}(\Omega)} \left(E_Q\left[c(X)\right] + \mathcal{D}_{\widehat{Q}}(Q) \right) = \inf_{Q\in \text{Mart}(\Omega)} \left(E_Q\left[c(X)\right] + \mathcal{D}(Q) \right)$$
$$= \sup_{\Delta\in\mathcal{H}} \sup_{\psi\in\Phi_\Delta(c)} S^U(\psi).$$

Observe now that S^U satisfies

$$S^U(\psi) := \sup_{\lambda\in\mathbb{R}^{T+1}} \left(U(\psi+\lambda) - \sum_{t=0}^T \lambda_t \right) = \sup_{\lambda\in\mathbb{R}^{T+1}} \left(\widehat{U}\left(\sum_{t=0}^T \psi_t + \sum_{t=0}^T \lambda_t \right) - \sum_{t=0}^T \lambda_t \right)$$
$$= \sup_{\xi\in\mathbb{R}} \left(\widehat{U}\left(\sum_{t=0}^T \psi_t + \xi \right) - \xi \right) =: S^{\widehat{U}}\left(\sum_{t=0}^T \psi_t \right).$$

$S^{\widehat{U}} : \mathcal{C}_b(\Omega) \to \mathbb{R}$ is (IA) and is nondecreasing; thus,

$$\sup_{\Delta\in\mathcal{H}} \sup_{\psi\in\Phi_\Delta(c)} S^{\widehat{U}}\left(\sum_{t=0}^T \psi_t \right) = \sup_{\Delta\in\mathcal{H}} \sup_{\psi\in\Phi_\Delta(c)} S^{\widehat{U}}\left(\sum_{t=0}^T \psi_t + I^\Delta \right) = S^{\widehat{U}}(c)$$

by definition of $\Phi_\Delta(c)$ and since $c \in \mathcal{C}_b(\Omega)$. □

6 Conclusion and future research

This work provides a detailed insight on entropy martingale optimal transport theory, introduced in Doldi and Frittelli (2023) as an extension and application of the EOT theory by Liero et al. (2018). We describe how EMOT can be naturally embedded in the current literature in Mathematical Finance, building a bridge between robust pricing-hedging dualities and the theory of convex risk measures. We introduce the concept of stock additivity, which allows for mimicking, in a typically nonlinear setup, some of the key motivating arguments in the formulation of the subhedging problem. Considering a compact underlying canonical space, we obtain two main results. The first one extends to general penalty functions the divergence-based EOT duality in Liero et al. (2018). We then deduce the EMOT duality from the EOT duality, replicating the core minimax argument yielding the MOT duality in Beiglböck et al. (2013) from the classical Kantorovich duality, thus giving an alternative and simpler proof of a similar result obtained in Doldi and Frittelli (2023). Two main applications are discussed: First, we consider the valuation induced by expected utility, modifying the well-known optimized certainty equivalent to account for stock additivity. In this case, divergence functionals, as those in Liero et al. (2018), appear in the EMOT dual problem. Secondly, we provide a dual representation for the generalized optimized certainty equivalent associated with the indirect utility function.

The EMOT theory we have presented opens up a wide range of possible generalizations and applications. From a theoretical perspective, the challenge remains open regarding extending the results to a continuous-time setup. Issues of stability and convergence in EMOT can also be further investigated, starting from a preliminary study in Doldi and Frittelli (2023). Among other potential developments, we include the analysis of the impact of penalty functions on numerical approximation, following in spirit Cuturi (2013), for example, and considering the convergence results in Doldi and Frittelli (2023). Finally, it is worth noting that the problem of actually developing algorithms for computing primal/dual values in EMOT is challenging, and is currently under investigation.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors have no competing interest to declare that are relevant to the content of this article.

A Appendix

A.1 Setting

A.1.1 Measures

We start fixing our setup and some notation. Let Ω be a Polish space and endow it with the Borel sigma algebra $\mathcal{B}(\Omega)$ generated by its open sets. A set function $\mu : \mathcal{B}(\Omega) \to \mathbb{R}$ is a **finite signed measure** if $\mu(\emptyset) = 0$ and μ is σ-additive. A **finite measure** μ is a finite signed measure such that $\mu(B) \geq 0$ for all $B \in \mathcal{B}(\Omega)$. A finite measure μ such that $\mu(\Omega) = 1$ will be called a **probability measure**. Recall from Sect. 4.1 the notations for $\mathrm{ca}(\Omega)$, $\mathrm{Meas}(\Omega)$, $\mathrm{Prob}(\Omega)$. The following result is well known, see, e.g., Billingsley (1999) Theorem 1.1 and 1.3.

Proposition A.1 *Every finite measure μ on $\mathcal{B}(\Omega)$ is a Radon measure, that is for every $B \in \mathcal{B}(\Omega)$ and every $\varepsilon > 0$ there exists a compact $K_\varepsilon \subseteq B$ such that $\mu(B \setminus K_\varepsilon) \leq \varepsilon$.*

We also introduce for $M \in \mathbb{N}$, $M \geq 1$ the sets

$$\mathcal{C}_b(\Omega) := \mathcal{C}_b(\Omega, \mathbb{R}) = \{\varphi : \Omega \to \mathbb{R} \mid \varphi \text{ is bounded and continuous on } \Omega\},$$
$$\mathcal{C}_b(\Omega, \mathbb{R}^M) := (\mathcal{C}_b(\Omega))^M = \{\varphi : \Omega \to \mathbb{R}^M \mid \varphi \text{ is bounded and continuous on } \Omega\}.$$

Given a vector subspace $\mathcal{E} \subseteq \mathcal{C}_b(\Omega, \mathbb{R}^{M+1})$, we will consider the dual pair $(\mathrm{ca}(\Omega), \mathcal{C}_b(\Omega, \mathbb{R}^{M+1}))$ with pairing given by the bilinear functional $(\gamma, \varphi) \mapsto$

$\int_\Omega \left(\sum_{m=0}^M \varphi_m\right) \mathrm{d}\gamma$. We will induce on $\mathrm{ca}(\Omega)$ the topology $\sigma(\mathrm{ca}(\Omega), \mathcal{E})$, which is the coarsest topology on $\mathrm{ca}(\Omega)$ making the functional $\gamma \mapsto \int_\Omega \left(\sum_{m=0}^M \varphi_m\right) \mathrm{d}\gamma$ continuous for each $\varphi \in \mathcal{E}$. Similarly, we will induce on $\mathcal{E}$ the topology $\sigma(\mathcal{E}, \mathrm{ca}(\Omega))$ which is the coarsest topology on $\mathcal{E}$ making the functional $\gamma \mapsto \int_\Omega \left(\sum_{m=0}^M \varphi_m\right) \mathrm{d}\gamma$ continuous for each $\gamma \in \mathrm{ca}(\Omega)$. The **Weak Topology** on $\mathrm{Meas}(\Omega)$ is the coarsest (Hausdorff) topology for which all maps $\mu \mapsto \int_\Omega \varphi \, \mathrm{d}\mu$ are continuous, for all $\varphi \in \mathcal{C}_b(\Omega)$.

Remark A.2 The weak topology on $\mathrm{Meas}(\Omega)$ is the topology $\sigma(\mathrm{Meas}(\Omega), \mathcal{C}_b(\Omega, \mathbb{R})$, which is the relative topology $\sigma(\mathrm{ca}(\Omega), \mathcal{C}_b(\Omega, \mathbb{R})|_{\mathrm{Meas}(\Omega)}$ induced by $\sigma(\mathrm{ca}(\Omega), \mathcal{C}_b(\Omega, \mathbb{R})$ on $\mathrm{Meas}(\Omega) \subseteq \mathrm{ca}(\Omega)$ (see Aliprantis and Border (2006) Lemma 2.53).

A.2 Auxiliary results and proofs

Lemma A.3 *Take compact $K_1, \dots, K_T \subseteq \mathbb{R}$, suppose that $K_0 = \{x_0\}$ and that if K_t is a singleton for some $0 \le t \le T$, so is K_s, $0 \le s \le t$. Take $\mathcal{E} = \mathcal{E}_0 \times \cdots \times \mathcal{E}_T$ for vector subspaces $\mathcal{E}_t \subseteq \mathcal{C}_b(K_t)$ with $\mathbb{R} \subseteq \mathcal{E}_t$ and $\mathrm{Id}_t \in \mathcal{E}_t$, $t = 0, \dots, T$. Suppose there exist $\varphi, \psi \in \mathcal{E}$ and $\Delta \in \mathcal{H}$, where $\mathcal{H}$ is defined in (7), such that $\sum_{t=0}^T \varphi_t = \sum_{t=0}^T \psi_t + I^\Delta$. Then there exist constants $k_0, \dots, k_T, h_0, \dots, h_T \in \mathbb{R}$ such that for each $t = 0, \dots, T$ $\psi_t(x_t) = \varphi_t(x_t) + k_t x_t + h_t$, $\forall x_t \in K_t$. In particular for $S_t : \mathcal{E}_t \to \mathbb{R}$, $t = 0, \dots, T$ Stock Additive functionals we have*

$$\sum_{t=0}^T S_t(\varphi_t) = \sum_{t=0}^T S_t(\psi_t) \,.$$

and for $\mathcal{V} := \sum_{t=0}^T \mathcal{E}_t + \mathcal{I}$ (see (8)) the map

$$v = \sum_{t=0}^T \varphi_t + I^\Delta \mapsto S(v) := \sum_{t=0}^T S_t(\varphi_t)$$

is well defined on $\mathcal{V}$, (CA) and (IA).

Proof *Step 1* we prove that if $\sum_{t=0}^T \varphi_t = \sum_{t=0}^T \psi_t + I^\Delta$ then $\Delta = [\Delta_0, \dots, \Delta_{T-1}] \in \mathcal{H}$ is a deterministic vector $\Delta \in \mathbb{R}^T$. If K_T is a singleton then the claim is trivial. Otherwise,

$$\begin{aligned}
&\varphi_T(x_T) - \psi_T(x_T) \\
&\quad = \sum_{t=0}^{T-1} (\psi(x_t) - \varphi_t(x_t)) + \sum_{t=0}^{T-2} \Delta_t(x_0, \dots, x_t)(x_{t+1} - x_t) \\
&\qquad + \Delta_{T-1}(x_0, \dots, x_{T-1})(x_T - x_{T-1}) \\
&\quad = f(x_0, \dots, x_{T-1}) + \Delta_{T-1}(x_0, \dots, x_{T-1}) x_T
\end{aligned}$$

for some function f. If Δ_{T-1} were not constant, on two points it would assume values $a \neq b$, with corresponding values of f that we call f_a, f_b. Then $f_a + a x_T = f_b + b x_T$

has a unique solution, contradicting the fact that all the equalities need to hold on the whole $K_0, \dots, K_T$ and in particular for two different values of x_T. We proceed one step backward. If K_{T-1} is a singleton the claim trivially follows, given our previous step. Otherwise, similarly to the previous computation,

$$\begin{aligned}\varphi_{T-1}(x_{T-1}) - \psi_{T-1}(x_{T-1}) &= \sum_{s\neq T-1}(\psi_s(x_s) - \varphi_s(x_s)) \\ &+ \sum_{t=0}^{T-3}\Delta_t(x_0,\dots,x_t)(x_{t+1}-x_t) + \\ &+\Delta_{T-2}(x_0,\dots,x_{T-2})(x_{T-1}-x_{T-2}) + \Delta_{T-1}(x_T - x_{T-1}) \\ &= f(x_s, s\neq T-1) + (\Delta_{T-2}(x_0,\dots,x_{T-2}) - \Delta_{T-1})x_{T-1}.\end{aligned}$$

An argument similar to the one we used in the previous time step shows that $\Delta_{T-2}(x_0,\dots,x_{T-2}) - \Delta_{T-1}$ is constant, hence so is Δ_{T-2}. Our argument can be clearly be iterated up to Δ_0.

Step 2 we prove existence of the vectors $k, h \in \mathbb{R}^{T+1}$, as stated in the lemma. From Step 1, it is clear that there exist constants $k_0,\dots,k_T$ such that $I^\Delta(x) = \sum_{t=0}^T k_t x_t$. Hence, $\sum_{t=0}^T \varphi_t(x_t) = \sum_{t=0}^T(\psi_t(x_t) + k_t x_t)$ for all $x\in\Omega$, which yields for each $t=0,\dots,T$ that $\varphi_t(x_t) - (\psi_t(x_t) + k_t x_t)$ does not depend on x_t, hence is constant, call it $-h_t$. Then $k_0,\dots,k_T, h_0,\dots,h_T \in\mathbb{R}$ satisfy our requirements. The last claim $\sum_{t=0}^T S_t(\varphi_t) = \sum_{t=0}^T S_t(\psi_t)$ is then an easy consequence of stock additivity.

Step 3 well-posedness and properties of S. Observe that whenever $\varphi,\psi\in\mathcal{E}$, $\Delta, H\in\mathcal{H}$ are given with $\sum_{t=0}^T\varphi_t + I^\Delta = \sum_{t=0}^T\psi_t + I^H$ we have by Steps 1-2 that $\sum_{t=0}^T S_t(\varphi_t) = \sum_{t=0}^T S_t(\psi_t)$. As a consequence, S is well defined. Cash additivity is inherited from $S_0,\dots,S_T$ while integral additivity is trivial from the definition. □

Lemma A.4 *Fix a probability space $(\Omega,\mathcal{F},P)$ and take $X, Z\in L^\infty(\Omega,\mathcal{F},P)$. Then*

$$\lim_{\alpha\to+\infty}\frac{E[X\exp(Z+\alpha X)]}{E[\exp(Z+\alpha X)]} = \operatorname{ess\,sup}_P(X) \quad \lim_{\alpha\to-\infty}\frac{E[X\exp(Z+\alpha X)]}{E[\exp(Z+\alpha X)]} = \operatorname{ess\,inf}_P(X) \tag{70}$$

Proof We only show the former equality, as the latter follows by replacing X with $-X$. Observe that $\alpha\mapsto\Theta(\alpha) := \frac{E[X\exp(Z+\alpha X)]}{E[\exp(Z+\alpha X)]}$ is nondecreasing: By taking derivatives, we get

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\Theta(\alpha) = E\left[X^2\frac{\exp(Z+\alpha X)}{E[\exp(Z+\alpha X)]}\right] - \left(E\left[X\frac{\exp(Z+\alpha X)}{E[\exp(Z+\alpha X)]}\right]\right)^2 = \operatorname{Var}_{P_\alpha}(X)\geq 0,$$

where the change of measure is given by $\frac{rmdP_\alpha}{\mathrm{d}P} = \frac{\exp(Z+\alpha X)}{E[\exp(Z+\alpha X)]}$. Thus, $\lim_{\alpha\to+\infty}\Theta(\alpha) =: L\in(-\infty,+\infty]$ exists. Since we may rewrite $\Theta(\alpha)$ as

$$\Theta(\alpha) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\log\left(E[\exp(Z+\alpha X)]\right) = \frac{\frac{\mathrm{d}}{\mathrm{d}\alpha}\log\left(E[\exp(Z+\alpha X)]\right)}{\frac{\mathrm{d}}{\mathrm{d}\alpha}\alpha},$$

we deduce, by a version of l'Hôpital's rule (see (Rudin 1976), Theorem 5.13), that

$$\lim_{\alpha\to+\infty}\frac{\log\left(E[\exp(Z+\alpha X)]\right)}{\alpha} = \frac{\frac{\mathrm{d}}{\mathrm{d}\alpha}\log\left(E[\exp(Z+\alpha X)]\right)}{\frac{\mathrm{d}}{\mathrm{d}\alpha}\alpha} = \lim_{\alpha\to+\infty}\Theta(\alpha).$$

From

$$-\|Z\|_\infty + \log\left(E[\exp(\alpha X)]\right) \le \log\left(E[\exp(Z+\alpha X)]\right) \le \|Z\|_\infty + \log\left(E[\exp(\alpha X)]\right),$$

we conclude:

$$\begin{aligned}\lim_{\alpha\to+\infty} \Theta(\alpha) &= \lim_{\alpha\to+\infty} \frac{\log\left(E[\exp(Z+\alpha X)]\right)}{\alpha} \\ &= \lim_{\alpha\to+\infty} \frac{1}{\alpha}\log\left(E[\exp(\alpha X)]\right) = \operatorname{ess\,sup}_P(X),\end{aligned}$$

where the last equality is well known, see, e.g., Carmona (2009) Sect. 3.2.3. □

A.3 On minimax duality theorem

The following theorem is stated, without the proof, in Liero et al. (2018), Th. 2.4. For the sake of completeness and without claiming any originality, we here provide the short proof.

Theorem A.5 *(Minimax Duality Theorem) Let A, B be nonempty convex subsets of some vector spaces and suppose A is endowed with a Hausdorff topology. Let $L : A \times B \to \mathbb{R}$ be a function such that*

1. *$a \mapsto L(a, b)$ is convex and lower semicontinuous in A for every $b \in B$.*
2. *$b \mapsto L(a, b)$ is concave in B for every $a \in A$*

When $\alpha := \sup_{b\in B} \inf_{a\in A} L(a, b) < +\infty$, suppose that there exist $C > \alpha$ and $b^\star \in B$ such that $\{a \in A \mid L(a, b^\star) \le C\}$ is compact in A. Then

$$\inf_{a\in A} \sup_{b\in B} L(a, b) = \sup_{b\in B} \inf_{a\in A} L(a, b)\,. \tag{71}$$

Proof We start observing that in general $\inf_{a\in A} \sup_{b\in B} L(a, b) \ge \sup_{b\in B} \inf_{a\in A} L(a, b)$; hence, if $\alpha = +\infty$ then (71) trivially holds. We then assume $\alpha < +\infty$ and modify the proof of Simons (1998) Theorem 3.1. Let $b_1, \dots, b_N \in B$ be given and set $b_0 = b^\star$. By Simons (1998) Lemma 2.1.(a), using $f_i(\cdot) := L(\cdot, b_i)$ we get constants $\lambda_0, \dots, \lambda_N \ge 0$ with $\sum_{i=0}^N \lambda_i = 1$ such that

$$\begin{aligned}\inf_{a\in A}\left(\max_{i=0,\dots,N} L(a, b_i)\right) &= \inf_{a\in A}\left(\sum_{i=0}^N \lambda_i L(a, b_i)\right) \\ &\le \inf_{a\in A} L\left(a, \sum_{i=0}^N \lambda_i b_i\right) \le \sup_{b\in B}\inf_{a\in A} L(a, b) = \alpha,\end{aligned}$$

where we used the concavity in B to obtain the first inequality. We now observe that for all $\varepsilon > 0$ there exists an $a \in A$ such that

$$a \in \left\{ \max_{i=0,\dots,N} L(a, b_i) \le \alpha + \varepsilon \right\} = \bigcap_{i=0}^{N} \{L(a, b_i) \le \alpha + \varepsilon\} = \bigcap_{i=1}^{N} \{L(a, b_i) \le \alpha + \varepsilon\} \cap \left\{ L(a, b^\star) \le \alpha + \varepsilon \right\}.$$

Hence, for $A^\star = \{L(a, b^\star) \le \alpha + \varepsilon\}$ the family $A_b^\varepsilon := \{a \in A^\star \mid L(a, b) \le \alpha + \varepsilon\}$ is a collection of closed subsets of $A^\star$ having the finite intersection property. Now take $\varepsilon > 0$ such that $\alpha + \varepsilon < C$. Then $A^\star$ is Hausdorff and compact, being a closed subset of the compact set $\{a \in A \mid L(a, b^\star) \le C\}$. As a consequence, $\bigcap_{b\in B} A_b^\varepsilon \neq \emptyset$. This yields the existence of an $a^\star$ such that $a^\star \in A^\star$ and $L(a^\star, b) \le \alpha + \varepsilon \; \forall\, b \in B$. Hence,

$$\inf_{a\in A} \sup_{b\in B} L(a, b) \le \sup_{b\in B} L(a^\star, b) \le \varepsilon + \alpha,$$

and letting $\varepsilon \downarrow 0$, we get

$$\inf_{a\in A} \sup_{b\in B} L(a, b) \le \sup_{b\in B} \inf_{a\in A} L(a, b) \le \inf_{a\in A} \sup_{b\in B} L(a, b).$$

□

A.4 Summary of symbols

To ease readability, we provide the following table that summarizes the relevant symbols and notations introduced in the paper.

$\mathcal{P}(P)$: set of all probability measures on Ω that are absolutely continuous wrt P

$\mathrm{Mart}(\Omega)$: set of all probability measures on Ω under which X is a martingale (pg. 5)

$\mathcal{M}(P) = \mathcal{P}(P) \cap \mathrm{Mart}(\Omega)$

$K_0, \dots, K_T$: subsets of $\mathbb{R}$ with $\Omega := K_0 \times \cdots \times K_T$ (pg. 5)

$\mathcal{M}(\widehat{Q}_0, \widehat{Q}_1, \dots \widehat{Q}_T) = \left\{ Q \in \mathrm{Mart}(\Omega) \mid X_t \sim_Q \widehat{Q}_t \text{ for each } t = 0, \dots, T \right\}$ (pg. 5)

$\mathcal{H}$: set of admissible trading strategies (see (7))

$\mathcal{I}$: set of stochastic integrals of trading strategies in $\mathcal{H}$ (see Eq. (8))

I^Δ, $\Delta \in \mathcal{H}$: the stochastic integral corresponding to Δ (see Eq. (8))

$\mathcal{D}_{F_t, \widehat{Q}_t}(\mu_t) = \int_{K_t} F_t \left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\widehat{Q}_t} \right) \mathrm{d}\widehat{Q}_t$: divergence functional (see Eq. (11))

$\mathfrak{D}(c) = \inf_{Q\in\mathrm{Mart}(\Omega)} (E_Q[c] + \sum_{t=0}^{T} \mathcal{D}_t(Q_t))$ (see Eq. (13))

$\mathfrak{P}(c) = \sup\{\sum_{t=0}^{T} S_t(\varphi_t) \mid \exists \Delta \in \mathcal{H} \text{ s.t. } \sum_{t=0}^{T} \varphi_t + I^\Delta \le c \text{ on } \Omega\}$ (see Eq. (18))

$S^U(\varphi) = \sup_{\xi\in\mathbb{R}^{T+1}} \left(U(\varphi + \xi) - \sum_{t=0}^{T} \xi_t \right)$: generalized OCE (see Eq. (27))

$\mathcal{D}(\gamma) = \sup_{\substack{\varphi\in\mathcal{E} \\ \varphi=(\varphi_1,\dots,\varphi_M)}} \left(U(\varphi) - \sum_{m=0}^{M} \int_\Omega \varphi_m \mathrm{d}\gamma \right)$: penalty function, conjugate of U (see Eq. (31))

$\Phi_\Delta(c)$: see Eq. (44)

$\mathcal{S}_{\text{sub}}(c)$, $\mathcal{S}_{\text{sup}}(c)$: static parts of sub/super replicating semistatic strategies (see Eqs. (54) and (55))

u_t, v_t, v_t^*: utility functions on $\mathbb{R}$ and corresponding convex functions and conjugates (see Assumption 4.9)

$\mathcal{D}_{v_t^*,\widehat{\mu}_t}(\mu)$: divergence functional associated with v_t^* and $\widehat{\mu}_t$ (see Eq. (59))

$U_{\widehat{Q}_t}(\varphi_t)$, $V_{\widehat{Q}_t}(\varphi_t)$: stock additive functionals induced by utility functions (see Eq. (61)).

References

Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W.: A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. Math. Finance **26**(2), 233–251 (2016)

Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis. Springer, Berlin, third edition (2006)

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Finance **9**(3), 203–228 (1999)

Backhoff-Veraguas, J., Pammer, G.: Stability of martingale optimal transport and weak optimal transport. Ann. Appl. Probab. **32**(1), 721–752 (2022). https://doi.org/10.1214/21-AAP1694

Bartl, D., Kupper, M., Prömel, D.J., Tangpi, L.: Duality for pathwise superhedging in continuous time. Finance Stoch. **23**(3), 697–728 (2019)

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

Bayraktar, E., Zhou, Z.: Super-hedging American options with semi-static trading strategies under model uncertainty. Int. J. Theor. Appl. Finance **20**(6), 1750036 (2017)

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

Bellini, F., Frittelli, M.: On the existence of minimax martingale measures. Math. Finance **12**(1), 1–21 (2002)

Ben-Tal, A., Teboulle, M.: Expected utility, penalty functions, and duality in stochastic nonlinear programming. Manag. Sci. **32**(11), 1445–1466 (1986)

Ben-Tal, A., Teboulle, M.: An old-new concept of convex risk measures: the optimized certainty equivalent. Math. Finance **17**(3), 449–476 (2007)

Bernton, E., Ghosal, P., Nutz, M.: Entropic optimal transport: Geometry and large deviations. Duke Math. J. **171**(16), 3363–3400 (2022). https://doi.org/10.1215/00127094-2022-0035

Biagini, S., Frittelli, M.: On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. In Optimality and risk—Modern Trends in Mathematical Finance, pp. 1–28. Springer, Berlin, (2009)

Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1999)

Blanchet, J., Jambulapati, A., Kent, C., Sidford, A.: Towards optimal running times for optimal transport. Oper. Res. Lett. **52**, 107054 (2024). https://doi.org/10.1016/j.orl.2023.11.007

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

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

Brown, H., Hobson, D., Rogers, L.C.G.: Robust hedging of barrier options. Math. Finance **11**(3), 285–314 (2001)

Burzoni, M., Frittelli, M., Hou, Z., Maggis, M., Obłój, J.: Pointwise arbitrage pricing theory in discrete time. Math. Oper. Res. **44**(3), 1034–1057 (2019)

Burzoni, M., Frittelli, M., Maggis, M.: Universal arbitrage aggregator in discrete-time markets under uncertainty. Finance Stoch. **20**(1), 1–50 (2016)

Burzoni, M., Frittelli, M., Maggis, M.: Model-free superhedging duality. Ann. Appl. Probab. **27**(3), 1452–1477 (2017)

Carmona, R. (ed.): Indifference Pricing. Theory and Applications. Princeton University Press, Princeton (2009)

Cheridito, P., Kiiski, M., Prömel, D.J., Soner, H.M.: Martingale optimal transport duality. Math. Annal. **379**, 1685–1712 (2020)
Cheridito, P., Kupper, M., Tangpi, L.: Duality formulas for robust pricing and hedging in discrete time. SIAM J. Financ. Math. **8**(1), 738–765 (2017)
Cohen, S.N.: Quasi-sure analysis, aggregation and dual representations of sublinear expectations in general spaces. Electron. J. Probab. **17**, 62 (2012)
Cox, A.M.G., Obłój, J.: Robust hedging of double touch barrier options. SIAM J. Financ. Math. **2**(1), 141–182 (2011)
Cox, A.M.G., Obłój, J.: Robust pricing and hedging of double no-touch options. Finance Stoch. **15**(3), 573–605 (2011)
Cox, A.M.G., Wang, J.: Root's barrier: construction, optimality and applications to variance options. Ann. Appl. Probab. **23**(3), 859–894 (2013)
Cuturi, Marco: Sinkhorn distances: lightspeed computation of optimal transport. Adv. Neural Inf. Process. Syst. **26**, 2292–2300 (2013)
Davis, M., Obłój, J., Raval, V.: Arbitrage bounds for prices of weighted variance swaps. Math. Finance **24**(4), 821–854 (2014)
De March, H., Henry-Labordère, P.: Building arbitrage-free implied volatility: Sinkhorn's algorithm and variants. Preprint: arXiv:1902.04456v2, (2020)
Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M., Stricker, C.: Exponential hedging and entropic penalties. Math. Finance **12**(2), 99–123 (2002)
Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**(3), 463–520 (1994)
Denis, L., Martini, C.: A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. Ann. Appl. Probab. **16**(2), 827–852 (2006)
Doldi, A., Frittelli, M.: Entropy martingale optimal transport and nonlinear pricing-hedging duality. Finance Stoch. **27**(2), 255–304 (2023)
Dolinsky, Y., Soner, H.M.: Martingale optimal transport and robust hedging in continuous time. Probab. Theory Related Fields **160**(1–2), 391–427 (2014)
Dolinsky, Y., Soner, H.M.: Martingale optimal transport in the Skorokhod space. Stochastic Process. Appl. **125**(10), 3893–3931 (2015)
El Karoui, N., Quenez, M.-C.: Dynamic programming and pricing of contingent claims in an incomplete market. SIAM J. Control Optim. **33**(1), 29–66 (1995)
Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Finance Stoch. **6**(4), 429–447 (2002)
Föllmer, H., Schied, A.: Stochastic Finance, fourth revised and extended. An introduction in discrete time De Gruyter Graduate. De Gruyter, Berlin (2016)
Frittelli, M.: Introduction to a theory of value coherent with the no-arbitrage principle. Finance Stoch. **4**(3), 275–297 (2000)
Frittelli, M., Rosazza Gianin, E.: Putting order in risk measures. J. Bank. Finance **26**(7), 1473–1486 (2002)
Frittelli, M., Scandolo, G.: Risk measures and capital requirements for processes. Math. Finance **16**(4), 589–612 (2006)
Galichon, A., Henry-Labordère, P., Touzi, N.: A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. Ann. Appl. Probab. **24**(1), 312–336 (2014)
Ghosal, P., Nutz, M., Bernton, E.: Stability of entropic optimal transport and Schrödinger bridges. J. Funct. Anal. **283**(9), 109622 (2022). https://doi.org/10.1016/j.jfa.2022.109622
Guo, G., Obłój, J.: Computational methods for martingale optimal transport problems. Ann. Appl. Probab. **29**(6), 3311–3347 (2019)
Henry-Labordère, P.: Automated option pricing: numerical methods. Int. J. Theor. Appl. Finance **16**(8), 1350042 (2013)
Henry-Labordère, P.: From (Martingale) Schrodinger bridges to a new class of Stochastic Volatility Models. Preprint: arXiv:1904.04554, (2019)
Henry-Labordère, P., Obłój, J., Spoida, P., Touzi, N.: The maximum maximum of a martingale with given n marginals. Ann. Appl. Probab. **26**(1), 1–44 (2016)
Hobson, D.: The Skorokhod embedding problem and model-independent bounds for option prices. In Paris-Princeton Lectures on Mathematical Finance 2010, volume 2003 of Lecture Notes in Math., pp. 267–318. Springer, Berlin (2011)

Hobson, D., Klimmek, M.: Maximizing functionals of the maximum in the Skorokhod embedding problem and an application to variance swaps. Ann. Appl. Probab. **23**(5), 2020–2052 (2013)
Hobson, D.G.: Robust hedging of the lookback option. Finance Stochast. **2**(4), 329–347 (1998)
Hodges, S.D., Neuberger, A.: Optimal replication of contingent claims under transaction costs. Rev. Fut. Market **8**, 222–239 (1989)
Hou, Z., Obłój, J.: Robust pricing-hedging dualities in continuous time. Finance Stoch. **22**(3), 511–567 (2018)
Karatzas, I.: Lectures on the mathematics of finance. CRM Monograph Series, vol. 8. American Mathematical Society, Providence, RI (1997)
Liero, M., Mielke, A., Savaré, G.: Optimal entropy-transport problems and a new Hellinger-Kantorovich distance between positive measures. Invent. Math. **211**(3), 969–1117 (2018)
Neufeld, A., Sester, J.: On the stability of the martingale optimal transport problem: a set-valued map approach. Statist. Probab. Lett. **176**, 109131 (2021)
Nutz, M., Wiesel, J.: Entropic optimal transport: Convergence of potentials. Probab. Theory Relat. Fields **184**, 401–424 (2022). https://doi.org/10.1007/s00440-021-01096-8
Peng, S.: Nonlinear expectations and stochastic calculus under uncertainty. Probability Theory and Stochastic Modelling, vol. 95. Springer, Berlin (2019)
Pennanen, T., Perkkiö, A.-P.: Convex duality in nonlinear optimal transport. J. Funct. Anal. **277**(4), 1029–1060 (2019)
Peyré, G., Cuturi, M.: Computational optimal transport: with applications to data science. Found. Trends® Mach. Learn. **11**, 355–607 (2019)
Riedel, F.: Financial economics without probabilistic prior assumptions. Decis. Econ. Finance **38**(1), 75–91 (2015)
Rouge, R., El Karoui, N.: Pricing via utility maximization and entropy. Math. Finance **10**(2), 259–276 (2000)
Rudin, W.: Principles of mathematical analysis. In: International Series in Pure and Applied Mathematics, 3rd edn. McGraw-Hill Book Co., New York (1976)
Sester, J.: A multi-marginal c-convex duality theorem for martingale optimal transport. Preprint: arXiv:2305.03344v2 (2023)
Sester, J.: On intermediate marginals in martingale optimal transportation. Math. Finan. Econ. **17**, 615–654 (2023). https://doi.org/10.1007/s11579-023-00345-9
Simons, S.: Minimax and monotonicity. Lecture Notes in Mathematics. Springer, Berlin (1998)
Soner, H.M., Touzi, N., Zhang, J.: Quasi-sure stochastic analysis through aggregation. Electron. J. Probab. **16**(67), 1844–1879 (2011)
Tan, X., Touzi, N.: Optimal transportation under controlled stochastic dynamics. Ann. Probab. **41**(5), 3201–3240 (2013)