Article:

Roux, Alet orcid.org/0000-0001-9176-4468 and Xu, Zhikang (2022) Optimal investment and contingent claim valuation with exponential disutility under proportional transaction costs. International Journal of Theoretical and Applied Finance. ISSN 0219-0249

https://doi.org/10.1142/S0219024922500170

International Journal of Theoretical and Applied Finance

OPTIMAL INVESTMENT AND CONTINGENT CLAIM VALUATION WITH EXPONENTIAL DISUTILITY UNDER PROPORTIONAL TRANSACTION COSTS

ALET ROUX
Department of Mathematics, University of York, Heslington, YO10 5DD, United Kingdom
alet.roux@york.ac.uk

ZHIKANG XU
*Willis Towers Watson, 51 Lime St, London, EC3M 7DQ, United Kingdom**
zk.xu@outlook.com

We consider indifference pricing of contingent claims consisting of payment flows in a discrete time model with proportional transaction costs and under exponential disutility. This setting covers utility maximisation of terminal wealth as a special case. A dual representation is obtained for the associated disutility minimisation problem, together with a dynamic procedure for solving it. This leads to efficient and convergent numerical procedures for indifference pricing, optimal trading strategies and shadow prices that apply to a wide range of payoffs, a large range of time steps and all magnitudes of transaction costs.

Keywords: transaction costs, option pricing, utility maximisation, entropy, indifference pricing, generalised convex hull, dynamic programming

1. Introduction

The price of a contingent claim in a complete market is uniquely determined as the cost of replicating its payoff, equally, the discounted expected payoff under the unique martingale measure. However, the presence of transaction costs can lead to the curious contradiction that superreplicating a claim may involve less trading (and lower transaction costs) than exact replication, and therefore be less expensive, so that the replication price can in fact lead to arbitrage. Furthermore, financial markets with proportional transaction costs and liquid cash generally admit many different martingale measures, leading to intervals of no-arbitrage claim prices. This means that subjective factors, such as an investor's risk appetite, come into play when determining the price of a claim. The indifference principle offers a compelling alternative to replication and arbitrage pricing: it states that the seller of a claim

*Most of the research presented in this paper was conducted while this author was a PhD student in the Department of Mathematics, University of York.

will charge (at least) a price that will allow him to sell the claim without increasing the risk of his existing financial position. This is called the *indifference price*.

Indifference pricing based on utility maximisation has been well studied in the literature on proportional transaction costs. Work in continuous time has mostly focused on adapting stochastic optimal control and other techniques from friction-free models (such as the Black-Scholes model), and in recent years have led to numerical approximation and asymptotics for small transaction costs; see the works by Bichuch (2014), Davis (1997), Davis et al. (1993), Hodges and Neuberger (1989), Kallsen and Muhle-Karbe (2015), Monoyios (2003, 2004), Whalley and Wilmott (1997), for example. Results obtained in continuous time models typically assume continuous trading, which limits their applicability in realistic settings (Dorfleitner and Gerer 2016), hence motivating the need for continued theoretical and numerical work in the discrete time setting.

The present paper is motivated by a recent strand of work by Pennanen (2014*a*,*b*), Pennanen and Perkkiö (2018) and others, which studies optimal investment and indifference pricing in a very general discrete time setting, including proportional transaction costs. In view of the fact that financial liabilities in banking and insurance often consist of sequences of payment streams, such as swaps, coupon paying bonds, insurance premia, etc, the classical utility maximisation framework, which focuses on the expected disutility of hedging shortfall at the expiration date of the liability faced by an investor (and insists on self-financing trading at other times), can be extended to a more flexible framework in which which hedging is allowed to fall short at intermediate steps, the expected total disutility of hedging shortfall at all steps is taken into account, and theoretical results can be derived for contingent claims consisting of cash and physical payment streams and a very general class of disutility functions.

Allowing hedging to fall short at intermediate time steps means that there is also a connection between the current work and another important strand in the transaction cost literature, namely maximising utility from consumption. An important notion in the study of these problems is the *shadow price*, which is a price process taking values in the bid-ask spread of the model with proportional transaction costs, with the property that maximising expected utility from consumption in the friction-free model with this price process, leads to the same maximal utility as in the original market with transaction costs. Kallsen and Muhle-Karbe (2011) and Rogala and Stettner (2015) showed that shadow prices exist in discrete time in a similar (though incompatible) technical setting to the current paper. Working in general discrete time models, Czichowsky et al. (2014) demonstrated that there is a link between the solution to the dual problem, and the existence of a shadow price. The existence of shadow processes in more general models is by no means guaranteed. Additionally, shadow prices may not be tractable, leading to the use of asymptotic expansions and/or restrictions in the magnitude of transaction costs. In the context of continuous-time models, see the earlier paper of Cvitanić and

Karatzas (1996), as well as more recent contributions by Kallsen and Muhle-Karbe (2010), Gerhold et al. (2013), Gerhold et al. (2014), Herczegh and Prokaj (2015), Czichowsky et al. (2017), Czichowsky and Schachermayer (2016, 2017), Lin and Yang (2016) and Gu et al. (2017).

The present paper specialises the model of Pennanen (2014a,b) and Pennanen and Perkkiö (2018) to exponential utility and proportional transaction costs (allowing the use of powerful dual methods) and finite state space (motivated by the need for numerical results). Our results apply to contingent claims with physical delivery (in other words, streams of portfolios rather than just cash). We propose a backward recursive procedure that can be used to solve the utility maximisation problem and compute indifference prices, together with an efficient and convergent numerical approximation method (with error bounds). This procedure has polynomial running time in recombinant models and for path-independent claims, and does not require the construction of a shadow price process, which is in general path-dependent (a known difficulty in models with proportional transaction costs). Nevertheless, the outputs from this procedure can be used to construct a shadow price process and accompanying martingale measure, together with an optimal hedging strategy. This latter construction is performed by (forward) induction, which makes it practical for studying individual scenarios, despite the path-dependence of the objects that are being studied. Our results apply to all magnitudes of transaction costs, and our numerical methods work for a large range of time steps; Xu (2018) reported a number of more demanding numerical results that have not been included in this paper for lack of space.

The results reveal interesting features of disutility minimisation problems and indifference prices. In particular, because asset holdings in the model can be carried over between different time periods and there are no portfolio constraints, the value of the disutility minimisation problem of an investor faced with delivering a portfolio stream depends only on the total payment involved in the stream (suitably discounted), which implies that indifference prices also depend only on the total payment due. Nevertheless, the additional flexibility offered by allowing hedging to fall short at time periods other than the final time leads to smaller spreads in indifference prices, when compared to utility indifference pricing spreads. Our numerical results further suggest that there is a complex relationship between disutility indifference prices and the real-world measure.

The numerical methods and examples work reported in this paper extend and complement the limited work in the literature for discrete time models with proportional transaction costs. The results on disutility minimisation generalise the results of Castañeda-Leyva and Hernández-Hernández (2011) in a one-step binomial model with proportional transaction costs. To put the power of the numerical methods into context, previously reported numerical results are limited to European put options in a 3-step Cox-Ross-Rubinstein binomial model with convex transaction costs and exponential utility (Çetin and Rogers 2007), utility indifference prices of a European call option under exponential utility in a binomial tree model with 6 steps and proportional transaction costs (Quek 2012), and numerical solution of utility maximisation problems under power utility with multiple assets and proportional transaction costs (Cai et al. 2013).

Whilst we restrict our attention to indifference prices (payable at time 0 in cash) rather than indifference swap rates (used by Pennanen 2014*b*) for brevity, we believe that the extension is straightforward (preliminary work reported by Xu 2018). We believe that our work can be generalised to include measuring hedging shortfall in terms of portfolios rather than just cash; this is the subject of ongoing research, as is application of these methods to other classes of utility functions and multi-asset models.

The paper is arranged as follows. Background information on arbitrage and superhedging in discrete time models with proportional transaction costs is collected in Section 2. The disutility minimisation problem that forms the basis of the indifference pricing framework is introduced in Section 3; this includes utility maximisation of terminal wealth as a special case. A dynamic procedure for solving the disutility minimisation problem and computing indifference prices is presented in Section 4, together with a procedure for constructing the shadow price. A procedure for constructing optimal hedging and injection strategies is presented in Section 5. Indifference prices are introduced in Section 6, together with arbitrage pricing bounds. Section 7 proposes a numerical approximation and considers issues such as computational costs, efficiency and convergence rates. Section 8 contains a number of illustrative numerical examples. Appendix A is a self-contained account of a generalisation of the convex hull of convex functions that appears in the dynamic procedure of Section 4 and the numerical approximation of Section 7. Proofs of all results in the main part of the paper appear in Appendix B.

2. Preliminaries

2.1. *Discrete-time model with proportional transaction costs*

In this paper we consider a discrete-time financial market model with a finite time horizon $T \in \mathbb{N}$ and trading dates $t = 0, \dots, T$ on a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t=0}^T$. We assume without loss of generality that $\mathcal{F}_0 = \{\Omega, \emptyset\}$, $\mathcal{F}_T = \mathcal{F} = 2^\Omega$ and $\mathbb{P}(\omega) > 0$ for all $\omega \in \Omega$. For each t, the collection of atoms of $\mathcal{F}_t$ is denoted by Ω_t. The elements of Ω_t are called the *nodes* of the model at time t, and they form a partition of Ω. For each ω and t, denote by ω_t the unique node $\nu \in \Omega_t$ such that $\omega \in \nu$. A node $\nu \in \Omega_{t+1}$ is said to be a *successor* of a node $\mu \in \Omega_t$ if $\nu \subseteq \mu$. Denote the collection of successors of any given node $\mu \in \Omega_t$ by μ^+, and define the transition probability from μ to any successor node $\nu \in \mu^+$ by $p^\nu_{t+1} := \frac{\mathbb{P}(\nu)}{\mathbb{P}(\mu)}$.

For every t and $d = 1, 2$, let $\mathcal{L}^d_t$ be the space of $\mathbb{R}^d$-valued $\mathcal{F}_t$-measurable random variables. Every random variable $x \in \mathcal{L}^d_t$ satisfies $x(\omega) = x(\omega')$ for all $\omega, \omega' \in \nu$ on every node $\nu \in \Omega_t$, and this common value is denoted x^ν. A similar convention applies to $\mathcal{F}_t$-measurable random functions $f : \Omega \times \mathbb{R}^d \to \mathbb{R}$. Let $\mathcal{N}^d$ be the space of adapted $\mathbb{R}^d$-valued processes. We write $\mathcal{L}_t = \mathcal{L}_t^1$ and $\mathcal{N} = \mathcal{N}^1$ for convenience. For $d = 2$ we will adopt the convention that the first and second components of any random variable $c \in \mathcal{L}_t^2$ or process $c \in \mathcal{N}^2$ are denoted c^b and c^s, respectively.

The financial market model consists of a risky and risk-free asset. The price of the risk-free asset, *cash*, is constant and equal to 1 at all times. This is equivalent to assuming that interest rates are zero, or that asset prices are discounted. Trading in the risky asset, the stock, is subject to proportional transaction costs. At any time step t, a share of the stock can be bought for the ask price S_t^a and sold for the bid price S_t^b, where $S_t^a \geq S_t^b > 0$. We assume that $S^a = (S_t^a)_{t=0}^T \in \mathcal{N}$ and $S^b = (S_t^b)_{t=0}^T \in \mathcal{N}$.

The cost of creating a portfolio $x = (x^b, x^s) \in \mathcal{L}_t^2$ at any time t is

$$\phi_t(x) \coloneqq x^b + x_+^s S_t^a - x_-^s S_t^b, \tag{2.1}$$

where $z_+ \coloneqq \max\{z, 0\}$ and $z_- \coloneqq -\min\{z, 0\}$ for all $z \in \mathbb{R}$. The liquidation value of the portfolio x is $x^b + x_+^s S_t^b - x_-^s S_t^a = -\phi_t(-x)$. Define the *solvency cone* $\mathcal{K}_t$ at any time t as the collection of portfolios that can be liquidated into a nonnegative cash amount, in other words,

$$\mathcal{K}_t \coloneqq \left\{x \in \mathcal{L}_t^2 : -\phi_t(-x) \geq 0\right\} = \left\{(x^b, x^s) \in \mathcal{L}_t^2 : x^b + x^s S_t^b \geq 0, x^b + x^s S_t^a \geq 0\right\}.$$

A trading strategy $y = (y_t)_{t=-1}^T$ is an adapted sequence of portfolios, where $y_{-1} \in \mathcal{L}_0^2$ denotes the initial endowment at time 0, the portfolio $y_t \in \mathcal{L}_t^2$ is held between time steps t and $t+1$ for every $t = 0, \ldots, T-1$, and $y_T \in \mathcal{L}_T^2$ is the terminal portfolio created at time T. Denote the collection of trading strategies by $\mathcal{N}^{2\prime}$, and define

$$\Delta y_t \coloneqq y_t - y_{t-1} \text{ for all } t \geq 0.$$

A trading strategy $y \in \mathcal{N}^{2\prime}$ is called *self-financing* if $-\Delta y_t \in \mathcal{K}_t$ for all t. The collection of self-financing trading strategies is defined as

$$\Phi \coloneqq \left\{y \in \mathcal{N}^{2\prime} : -\Delta y_t \in \mathcal{K}_t \ \forall t\right\}.$$

We will also frequently consider the class of trading strategies that start and end with zero holdings (and are not necessarily self-financing). This class of trading strategies is denoted by

$$\Psi \coloneqq \left\{y \in \mathcal{N}^{2\prime} : y_{-1} = 0, y_T = 0\right\}.$$

2.2. *Arbitrage and duality*

There is a connection between the absence of arbitrage and the existence of classes of objects that appear in the study of disutility minimisation problems. To this end, define

$$\begin{aligned} \bar{\mathcal{P}} &\coloneqq \left\{(\mathbb{Q}, S) : \mathbb{Q} \ll \mathbb{P},\, S \text{ a } \mathbb{Q}\text{-martingale},\, S_t^b \leq S_t \leq S_t^a \ \forall t\right\}, \\ \mathcal{P} &\coloneqq \left\{(\mathbb{Q}, S) : \mathbb{Q} \sim \mathbb{P},\, S \text{ a } \mathbb{Q}\text{-martingale},\, S_t^b \leq S_t \leq S_t^a \ \forall t\right\}. \end{aligned} \tag{2.2}$$

We shall refer to the elements of $\bar{\mathcal{P}}$ ($\mathcal{P}$) as (*equivalent*) *martingale pairs*. Observe that $\mathcal{P} \subseteq \bar{\mathcal{P}}$.

The following result characterises the existence of equivalent martingale pairs.

Proposition 2.1 (Kabanov and Stricker (2001 Theorem 1)). *The no-arbitrage condition*

$$\{y_T : y \in \Phi, y_{-1} = 0\} \cap \left\{z \in \mathcal{L}^2_T : z \geq 0\right\} = \{0\} \tag{2.3}$$

holds if and only if $\mathcal{P} \neq \emptyset$.

The definition (2.3) of the no-arbitrage condition is consistent with that of Schachermayer (2004 Def. 1.6) and equivalent, though formally different, to the notion of weak no-arbitrage introduced by Kabanov and Stricker (2001). We will often require a stronger condition in this paper, namely *robust no-arbitrage* (Schachermayer 2004 Def. 1.9), in order to ensure the existence of a solution to the disutility minimisation problem. It is characterised as follows.

Proposition 2.2 (Schachermayer (2004 Theorem 1.7)). *The robust no-arbitrage condition holds if and only if there exists an equivalent martingale pair* $(\mathbb{Q}, S) \in \mathcal{P}$ *such that, for all* t,

$$S_t \in \mathrm{ri}[S^b_t, S^a_t] = \begin{cases} \{S^b_t\} & \textit{on } \{S^b_t = S^a_t\}, \\ (S^b_t, S^a_t) & \textit{on } \{S^b_t < S^a_t\}. \end{cases}$$

The following notation will be useful when working with martingale pairs. For every $\mathbb{Q} \ll \mathbb{P}$, we write

$$\Lambda^{\mathbb{Q}}_t := \mathbb{E}\left[\tfrac{d\mathbb{Q}}{d\mathbb{P}} \middle| \mathcal{F}_t\right] \text{ for all } t = 0, \dots, T, \tag{2.4}$$

where $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon-Nikodym density of $\mathbb{Q}$ with respect to $\mathbb{P}$. As Ω is finite it follows that

$$\Lambda^{\mathbb{Q}\nu}_t = \tfrac{\mathbb{Q}(\nu)}{\mathbb{P}(\nu)} \text{ for all } t \text{ and } \nu \in \Omega_t. \tag{2.5}$$

Define also for all t

$$\Omega^{\mathbb{Q}}_t := \{\nu \in \Omega_t : \mathbb{Q}(\nu) > 0\}$$

as the collection of nodes in Ω_t with positive probability under $\mathbb{Q}$. Moreover, for every $t < T$ and $\mu \in \Omega^{\mathbb{Q}}_t$, denote the transition probability from μ to any successor node $\nu \in \mu^+$ by $q^\nu_{t+1} := \frac{\mathbb{Q}(\nu)}{\mathbb{Q}(\mu)}$. Simple rearrangement of (2.5) then gives

$$\Lambda^{\mathbb{Q}\nu}_{t+1} = \tfrac{\mathbb{Q}(\mu)q^\nu_{t+1}}{\mathbb{P}(\mu)p^\nu_{t+1}} = \Lambda^{\mathbb{Q}\mu}_t \tfrac{q^\nu_{t+1}}{p^\nu_{t+1}} \text{ for all } t < T,\ \mu \in \Omega_t \text{ and } \nu \in \mu^+. \tag{2.6}$$

2.3. *Superhedging*

If the seller of a claim is completely risk-averse, then he would charge (at least) the *superhedging price*, which is the lowest amount that the seller of a claim can charge that will allow him to sell the claim without taking any risk. Such prices are usually lower than the cost of replication (see, for example Bensaid et al. 1992), and have been well studied for European options offering a payoff at a single expiration date; for a selection of contributions at a similar technical level to the current paper, see work by Delbaen, Kabanov and Valkeila (2002), Dempster et al. (2006), Edirisinghe et al. (1993), Jouini and Kallal (1995), Kabanov and Stricker (2001), Löhne and Rudloff (2014), Perrakis and Lefoll (1997), Roux et al. (2008), Roux and Zastawniak (2016).

In this subsection we generalise the theory slightly to the case of payment streams of the form $c \in \mathcal{N}^2$, consisting of sequences of (portfolio) payments $c_t = (c^b_t, c^s_t)$ to be made at all trading dates t. A trading strategy $y \in \mathcal{N}^{2\prime}$ is said to *superhedge* such a payment stream c if it allows a trader to deliver c without risk, in other words, $y_T = 0$ and $-\Delta y_t - c_t \in \mathcal{K}_t$ for all t.

The *seller's superhedging price* of the payment stream c is defined as the smallest cash endowment that is sufficient to superhedge c, in other words,

$$\pi^a(c) := \inf\left\{x \in \mathbb{R} : \exists y \in \mathcal{N}^{2\prime} \text{ superhedging } c \text{ with } y_0 = (x,0)\right\}.$$

The *buyer's superhedging price* of c is defined as

$$\begin{aligned}\pi^b(c) &:= \sup\left\{x \in \mathbb{R} : \exists y \in \mathcal{N}^{2\prime} \text{ superhedging } -c \text{ with } y_0 = (-x,0)\right\} \\ &= -\pi^a(-c). \end{aligned} \tag{2.7}$$

It is the largest cash amount that can be raised without risk by using the payoff of c as collateral. The superhedging prices admit the following dual representation, the proof of which can be found in Appendix B.

Proposition 2.3. *Assume no-arbitrage. For every $c \in \mathcal{N}^2$,*

$$\pi^a(c) = \sup_{(\mathbb{Q},S)\in\mathcal{P}} \sum_{t=0}^{T} \mathbb{E}_{\mathbb{Q}}\left[c^b_t + c^s_t S_T\right] = \max_{(\mathbb{Q},S)\in\bar{\mathcal{P}}} \sum_{t=0}^{T} \mathbb{E}_{\mathbb{Q}}\left[c^b_t + c^s_t S_T\right], \tag{2.8}$$

$$\pi^b(c) = \inf_{(\mathbb{Q},S)\in\mathcal{P}} \sum_{t=0}^{T} \mathbb{E}_{\mathbb{Q}}\left[c^b_t + c^s_t S_T\right] = \min_{(\mathbb{Q},S)\in\bar{\mathcal{P}}} \sum_{t=0}^{T} \mathbb{E}_{\mathbb{Q}}\left[c^b_t + c^s_t S_T\right]. \tag{2.9}$$

The collection of payment streams that can be superhedged from zero will play an important role in the next section. Proposition 2.3 gives that

$$\begin{aligned}\mathcal{Z} &:= \left\{c \in \mathcal{N}^2 : \exists y \in \Psi \text{ superhedging } c\right\} \qquad (2.10)\\ &= \left\{c \in \mathcal{N}^2 : \pi^a(c) \le 0\right\} \\ &= \left\{c \in \mathcal{N}^2 : \textstyle\sum_{t=0}^{T} \mathbb{E}_{\mathbb{Q}}\left[c^b_t + c^s_t S_T\right] \le 0 \ \forall (\mathbb{Q},S) \in \bar{\mathcal{P}}\right\}, \qquad (2.11)\end{aligned}$$

provided that the no-arbitrage condition (2.3) holds. It is self-evident from the representation (2.11) that $\mathcal{Z}$ is a convex cone.

3. Disutility minimisation problem

The ability to manage investments in such a way that their proceeds cover an investor's liabilities as well as possible, is of fundamental importance in financial economics, and has therefore been well studied in the literature; see, for example, the work of Davis (1997), Delbaen, Grandits, Rheinländer, Samperi, Schweizer and Stricker (2002), Guasoni (2002), Hugonnier et al. (2005), Rásonyi and Stettner (2005), Pennanen (2014*a*,*b*), Pennanen and Perkkiö (2018). The purpose of this section is to formulate an optimal investment problem in the model with proportional transaction costs, which will form the basis of the indifference prices that will be studied in Section 6.

Consider an investor who faces the liability of a (given) payment stream $u \in \mathcal{N}^2$. The investor can create a trading strategy $y \in \Psi$ in cash and stock, and is additionally allowed to inject (invest) cash on every trading date in a given set $\mathcal{I} \subseteq \{0, \dots, T\}$. At each trading date $t \in \mathcal{I}$, in order to manage his position, the investor needs to inject $\phi_t(\Delta y_t + u_t)$ in cash in order to manage his position. At trading dates $t \notin \mathcal{I}$, the investor is required to manage his position in a self-financing manner, in other words, $\phi_t(\Delta y_t + u_t) \leq 0$. Denote the number of elements of $\mathcal{I}$ by $|\mathcal{I}|$ and assume that $|\mathcal{I}| > 0$, in other words, injection is allowed at least once. It is *not* assumed that $T \in \mathcal{I}$.

The objective of the investor is to choose y in such a way as to minimise the sum of expected disutility of the cash injections over all the trading dates in $\mathcal{I}$, using for each time step $t \in \mathcal{I}$ the risk-averse exponential disutility (regret) function

$$v_t(x) := e^{\alpha_t x} - 1 \text{ for all } x \in \mathbb{R}$$

with deterministic risk aversion parameter $\alpha_t \in (0, \infty)$. Define for every $t \notin \mathcal{I}$

$$v_t(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ \infty & \text{if } x > 0. \end{cases}$$

The investor's objective can then be written as the unconstrained optimisation problem

$$\text{minimise} \sum_{t=0}^{T} \mathbb{E}[v_t(\phi_t(\Delta y_t + u_t))] \text{ over } y \in \Psi. \tag{3.1}$$

The value function V of (3.1) is defined as

$$V(u) := \inf_{y \in \Psi} \sum_{t=0}^{T} \mathbb{E}[v_t(\phi_t(\Delta y_t + u_t))]. \tag{3.2}$$

We have $-\infty < V(u) \leq 0$ because $v_t(0) = 0$ and v_t is bounded from below for all t.

Remark 3.1. In the special case where $\mathcal{I} = \{T\}$ and $u_t = 0$ for all $t < T$, the problem (3.1) becomes

$$\text{maximise } \mathbb{E}\big[1 - e^{-\alpha_T(-\phi_T(-y_{T-1}+u_T))}\big] \text{ over } y \in \Psi, -\Delta y_t \in \mathcal{K}_t \ \forall t < T. \tag{3.3}$$

Noting that $-\phi_T(-y_{T-1}+u_T)$ is the liquidation value of the portfolio $y_{T-1}-u_T$, this is the classical utility maximisation problem of an investor facing a liability of u_T at time T.

It is possible to rewrite (3.1) directly in terms of the cash injections. The resulting optimal cash injections will be used in Section 5 to construct optimal trading strategies for (3.1). Combining the fact that v_t is nondecreasing for all t with (2.10), we obtain

$$\begin{aligned} V(u) &= \inf\left\{\textstyle\sum_{t=0}^T \mathbb{E}[v_t(x_t)] : (x,y)\in\mathcal{N}\times\Psi, x_t \geq \phi_t(\Delta y_t + u_t)\ \forall t\right\} \\ &= \inf\left\{\textstyle\sum_{t=0}^T \mathbb{E}[v_t(x_t)] : (x,y)\in\mathcal{N}\times\Psi, -\Delta y_t - u_t + (x_t,0)\in\mathcal{K}_t\ \forall t\right\} \\ &= \inf\left\{\textstyle\sum_{t=0}^T \mathbb{E}[v_t(x_t)] : (x,y)\in\mathcal{N}\times\Psi, y \text{ superhedges } u-(x,0)\right\} \\ &= \inf\left\{\textstyle\sum_{t=0}^T \mathbb{E}[v_t(x_t)] : x\in\mathcal{N}, u-(x,0)\in\mathcal{Z}\right\} \\ &= \inf_{x\in\mathcal{A}_u} \sum_{t=0}^T \mathbb{E}[v_t(x_t)], \end{aligned} \tag{3.4}$$

where

$$\mathcal{A}_u := \{x\in\mathcal{N} : u-(x,0)\in\mathcal{Z}\}. \tag{3.5}$$

In conclusion, the problem (3.1) has the same value function as the optimisation problem

$$\text{minimise} \sum_{t=0}^T \mathbb{E}[v_t(x_t)] \text{ over } x\in\mathcal{A}_u. \tag{3.6}$$

Theorem 3.1 below confirms that the optimisation problems (3.1) and (3.6) can be solved for any $u\in\mathcal{N}^2$. It will be shown as part of the construction (in Theorem 5.1) that the optimal cash injection strategy in (3.6) is unique under robust no-arbitrage. The optimal trading strategy in (3.1) is not necessarily unique.

Theorem 3.1 also establishes a dual representation for $V(u)$ that will be key to the construction procedure proposed in Section 4. In order to formulate the result and prepare for the construction, we fix the notation

$$a_t := \sum_{k\in\mathcal{I},k\geq t} \tfrac{1}{\alpha_k} \text{ for all } t \tag{3.7}$$

for brevity, and define for any $X\in\mathcal{L}_T^2$ the functions

$$\begin{aligned} H((\mathbb{Q},S),X) &:= \sum_{t\in\mathcal{I}} \tfrac{1}{\alpha_t}\mathbb{E}_{\mathbb{Q}}\big[\ln\Lambda_t^{\mathbb{Q}}\big] - \mathbb{E}_{\mathbb{Q}}\big[X^b + X^s S_T\big] \text{ for all } (\mathbb{Q},S)\in\bar{\mathcal{P}}, \qquad &(3.8)\\ K(X) &:= \inf_{(\mathbb{Q},S)\in\bar{\mathcal{P}}} H((\mathbb{Q},S),X). &(3.9) \end{aligned}$$

Notice that $K(X)$ is finite because the values of the mapping $x\mapsto x\ln x$ are finite and bounded from below on $[0,\infty)$. In this paper we adopt the convention $0\ln 0 = 0$.

Theorem 3.1. *Under robust no-arbitrage, the infima in* (3.2) *and* (3.4) *are attained for every* $u \in \mathcal{N}^2$ *and*

$$V(u) = \sup_{\lambda>0} \sup_{(\mathbb{Q},S)\in\bar{\mathcal{P}}} \lambda \left(\sum_{t=0}^{T} \mathbb{E}_{\mathbb{Q}}[u_t^b + u_t^s S_t] - \sum_{t\in\mathcal{I}} \tfrac{1}{\alpha_t} \left(\mathbb{E}_{\mathbb{Q}}\left[\ln \Lambda_t^{\mathbb{Q}}\right] + \ln \tfrac{\lambda}{\alpha_t} \right) + a_0 \right) - |\mathcal{I}| \tag{3.10}$$

$$= a_0 \hat{\lambda}_u - |\mathcal{I}|, \tag{3.11}$$

where

$$\hat{\lambda}_u := \exp\left\{ \tfrac{1}{a_0} \left(\textstyle\sum_{t\in\mathcal{I}} \tfrac{\ln \alpha_t}{\alpha_t} - K\left(\textstyle\sum_{t=0}^{T} u_t\right)\right)\right\} > 0. \tag{3.12}$$

The proof of this result relies on theoretical results obtained by Pennanen and Perkkiö (2018) in a general setting. It appears in Appendix B.

Note that Theorem 3.1 implies that $\hat{\lambda}_u$, and hence $V(u)$, depend on the liability u only through the total liability $\sum_{t=0}^{T} u_t$. This is perhaps surprising in view of (3.2), but it is due to the nature of the dual objects in models with proportional transaction costs without trading and portfolio constraints: for example, it can be seen in (2.11) that whether a payment stream can be superhedged from zero depends only on its total payoff.

4. Minimal disutility

It was shown in Section 3 that, under robust no-arbitrage, solving the disutility minimisation problem (3.1) amounts to computing the value of $K(X)$, defined in (3.9), for a suitably chosen random variable X. It will also be shown in Section 6 (see Theorem 6.1) that the same holds true for determining the buyer's and seller's indifference prices.

In this section, we propose a dynamic procedure for determining $K(X)$ for any $X \in \mathcal{L}_T^2$ under the more relaxed no-arbitrage condition. We also present a dynamic procedure for constructing a pair $(\hat{\mathbb{Q}}, \hat{S}) \in \mathcal{P}$ such that

$$K(X) = H((\hat{\mathbb{Q}}, \hat{S}), X) = \sum_{t\in\mathcal{I}} \tfrac{1}{\alpha_t} \mathbb{E}_{\hat{\mathbb{Q}}}\left[\ln \Lambda_t^{\hat{\mathbb{Q}}}\right] - \mathbb{E}_{\hat{\mathbb{Q}}}\left[X^b + X^s \hat{S}_T\right]. \tag{4.1}$$

Remark 4.1. The dynamic procedure can also be used to find the *minimal entropy martingale measure* (Frittelli 2000*a*,*b*). This is the measure $\hat{\mathbb{Q}}$ satisfying

$$K(0) = \mathbb{E}_{\hat{\mathbb{Q}}}\left[\ln \Lambda_T^{\hat{\mathbb{Q}}}\right] = \mathbb{E}\left[\tfrac{d\hat{\mathbb{Q}}}{d\mathbb{P}} \ln \tfrac{d\hat{\mathbb{Q}}}{d\mathbb{P}}\right],$$

in the special case when $\mathcal{I} = \{T\}$ and there are no transaction costs (in other words, $\hat{S} = S^b = S^a$).

The following representation for H in terms of transition probabilities is key to constructing a solution to (4.1) by dynamic programming.

Proposition 4.1. *For all $X \in \mathcal{L}^2_T$ and $(\mathbb{Q}, S) \in \bar{\mathcal{P}}$, we have*

$$H((\mathbb{Q},S),X) = \sum_{t=0}^{T-1} a_{t+1} \sum_{\mu\in\Omega_t^{\mathbb{Q}}} \mathbb{Q}(\mu) \sum_{\nu\in\mu^+} q^\nu_{t+1} \ln \frac{q^\nu_{t+1}}{p^\nu_{t+1}} - \sum_{\mu\in\Omega^{\mathbb{Q}}_{T-1}} \mathbb{Q}(\mu) \sum_{\nu\in\mu^+} q^\nu_T \big(X^{b\nu} + X^{s\nu} S^\nu_T\big). \quad (4.2)$$

The proof of Proposition 4.1 appears in Appendix B. The representation (4.2) suggests that it is possible to construct a sequence $(\hat{q}_t)_{t=1}^T$ of transition probabilities, from which then to assemble the probability measure $\hat{\mathbb{Q}}$. The following construction provides a sequence of auxiliary functions to achieve this aim.

Construction 4.1. For given $X \in \mathcal{L}^2_T$, construct two adapted sequences of random functions $(f_t)_{t=0}^{T-1}$ and $(J_t)_{t=0}^T$ by backward induction. Define $J_T : \Omega\times\mathbb{R} \to \mathbb{R}\cup\{\infty\}$ as

$$J^\nu_T(x) := \begin{cases} -X^{b\nu} - xX^{s\nu} & \text{if } x \in \big[S^{b\nu}_T, S^{a\nu}_T\big], \\ \infty & \text{otherwise.} \end{cases} \quad (4.3)$$

for all $\nu \in \Omega_T$. For every $t < T$, assume that J_{t+1} has already been constructed, and define

$$f^\mu_t(x) := \inf\Big\{\textstyle\sum_{\nu\in\mu^+} q^\nu \Big(J^\nu_{t+1}(x^\nu) + a_{t+1}\ln\frac{q^\nu}{p^\nu_{t+1}}\Big) : q^\nu\in[0,1], \\ x^\nu \in \operatorname{dom} J^\nu_{t+1}\ \forall\nu\in\mu^+, \textstyle\sum_{\nu\in\mu^+} q^\nu = 1, \sum_{\nu\in\mu^+} q^\nu x^\nu = x\Big\}, \quad (4.4)$$

$$J^\mu_t(x) := \begin{cases} f^\mu_t(x) & \text{if } x\in\big[S^{b\nu}_t, S^{a\nu}_t\big], \\ \infty & \text{otherwise.} \end{cases} \quad (4.5)$$

for all $\mu \in \Omega_t$ and $x \in \mathbb{R}$.

The definition (4.4) of f^μ_t is reminiscent of that of the convex hull of the collection $\{J^\nu_{t+1}\}_{\nu\in\mu^+}$ of convex functions, if the term involving the logarithm is disregarded (cf. Rockafellar 1996 Theorem 5.6). The following result summarises the main properties of $(J_t)_{t=0}^T$, with some of the technical arguments involving the generalised convex hull deferred to Appendix A. Recall that the σ-field $\mathcal{F}_0$ is trivial, and therefore J_0 is a deterministic function.

Proposition 4.2. *Assume no-arbitrage. For any $X \in \mathcal{L}^2_T$, let $(J_t)_{t=0}^T$ be the sequence of functions from Construction 4.1. Then for each t and $\nu \in \Omega_t$, the function J^ν_t is convex, bounded from below, continuous on its closed effective domain* $\operatorname{dom} J^\nu_t \subseteq [S^{b\nu}_t, S^{a\nu}_t]$ *and the infimum in* (4.4) *is attained whenever it is finite. Moreover,*

$$J_0(S_0) = \inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}},\bar{S}_0=S_0} H((\bar{\mathbb{Q}},\bar{S}),X) \text{ for all } (\mathbb{Q},S)\in\bar{\mathcal{P}}. \quad (4.6)$$

The proof appears in Appendix B. The following one-step toy model demonstrates the application of Construction 4.1 in a simple setting.

Example 4.1. Let $T = 1$ and $\Omega = \{\mathrm{u}, \mathrm{d}\}$, and take any probability measure $\mathbb{P}$ with $p := \mathbb{P}(\mathrm{u}) \in (0, 1)$. Suppose furthermore that the bid and ask prices in this model satisfy

$$S_1^{b\mathrm{d}} \leq S_1^{a\mathrm{d}} < S_0^b = \bar{S}_0 = S_0^a < S_1^{b\mathrm{u}} \leq S_1^{a\mathrm{u}}. \tag{4.7}$$

Every probability measure $\mathbb{Q}$ in this model can be characterised uniquely by the value of $\mathbb{Q}(\mathrm{u})$. It follows from (4.7) that

$$\mathcal{Q} := \left\{\mathbb{Q}(\mathrm{u}) : (\mathbb{Q}, S) \in \bar{\mathcal{P}}\right\} = \left[\tfrac{\bar{S}_0 - S_1^{a\mathrm{d}}}{S_1^{a\mathrm{u}} - S_1^{a\mathrm{d}}}, \tfrac{\bar{S}_0 - S_1^{b\mathrm{d}}}{S_1^{b\mathrm{u}} - S_1^{b\mathrm{d}}}\right] =: [q_{\min}, q_{\max}].$$

The mid-price process $\bar{S} = (\bar{S}_0, \bar{S}_1) \in \mathcal{N}$ with $\bar{S}_1 := \frac{1}{2}(S_1^a + S_1^b)$ together with the unique probability measure $\bar{\mathbb{Q}}$ with $\bar{\mathbb{Q}}(\mathrm{u}) = \frac{\bar{S}_0 - \bar{S}_1^{\mathrm{d}}}{\bar{S}_1^{\mathrm{u}} - \bar{S}_1^{\mathrm{d}}}$ satisfies the robust no-arbitrage condition of Proposition 2.2.

Let $\mathcal{I} := \{0, 1\}$ and take as given a random variable $X = (X^b, X^s) \in \mathcal{L}_1^2$ with $X^s = 0$. We have

$$J_1^{\mathrm{u}}(x) = \begin{cases} -X^{b\mathrm{u}} & \text{if } x \in \left[S_1^{b\mathrm{u}}, S_1^{a\mathrm{u}}\right], \\ \infty & \text{otherwise,} \end{cases} \qquad J_1^{\mathrm{d}}(x) = \begin{cases} -X^{b\mathrm{d}} & \text{if } x \in \left[S_1^{b\mathrm{d}}, S_1^{a\mathrm{d}}\right], \\ \infty & \text{otherwise,} \end{cases}$$

and then $J_0(\bar{S}_0) = \inf_{[q_{\min}, q_{\max}]} g_X(q)$, where

$$g_X(q) := \tfrac{1}{\alpha}\left(q \ln \tfrac{q}{p} + (1-q)\ln\tfrac{1-q}{1-p}\right) - qX^{b\mathrm{u}} - (1-q)X^{b\mathrm{d}} \text{ for all } q \in [0, 1].$$

The function g_X is continuous and convex on $[0, 1]$, and reaches its minimum at

$$\hat{q}_X := pe^{\alpha X^{b\mathrm{u}}} \Big/ \left(pe^{\alpha X^{b\mathrm{u}}} + (1-p)e^{\alpha X^{b\mathrm{d}}}\right) \in (0, 1).$$

Taking $\bar{q}_X := \min\{\max\{\hat{q}_X, q_{\min}\}, q_{\max}\}$, it follows from Proposition 4.2 that $K(X) = J_0(\bar{S}_0) = g_X(\bar{q}_X)$.

The following construction uses the sequence $(J_t)_{t=0}^T$ of Construction 4.1 to produce a pair $(\hat{\mathbb{Q}}, \hat{S})$ satisfying (4.1). It will be shown in Theorem 4.1 below that this does indeed produce a solution to (3.9).

Construction 4.2. Assume no-arbitrage. For given $X \in \mathcal{L}_T^2$ and associated sequence $(J_t)_{t=0}^T$ from Construction 4.1, construct a process $\hat{S} \in \mathcal{N}$ and a predictable process $(\hat{q}_t)_{t=1}^T$ by induction, as follows. First, choose any $\hat{S}_0$ satisfying

$$J_0(\hat{S}_0) = \min_{x \in [S_0^b, S_0^a]} J_0(x). \tag{4.8}$$

For each $t < T$ and $\mu \in \Omega_t$, assume that $\hat{S}_t^\mu \in [S_t^{b\mu}, S_t^{a\mu}]$ has already been defined, and choose $\hat{q}_{t+1}^\nu \in [0, 1]$, $\hat{S}_{t+1}^\nu \in \left[S_{t+1}^{b\nu}, S_{t+1}^{a\nu}\right]$ for all $\nu \in \mu^+$ such that

$$J_t^\mu(\hat{S}_t^\mu) = \sum_{\nu \in \mu^+} \hat{q}_{t+1}^\nu \left(J_{t+1}^\nu(\hat{S}_{t+1}^\nu) + a_{t+1} \ln \tfrac{\hat{q}_{t+1}^\nu}{p_{t+1}^\nu}\right), \tag{4.9}$$

$$\hat{S}_t^\mu = \sum_{\nu \in \mu^+} \hat{q}_{t+1}^\nu \hat{S}_{t+1}^\nu, \tag{4.10}$$

$$1 = \sum_{\nu \in \mu^+} \hat{q}_{t+1}^\nu. \tag{4.11}$$

Finally, define $\hat{\mathbb{Q}}:\mathcal{F}\to\mathbb{R}$ as $\hat{\mathbb{Q}}(A):=\sum_{\omega\in A}\prod_{t=1}^{T}\hat{q}_t^{\omega_t}$ for all $A\in\mathcal{F}$, where the value of the sum over an empty set is taken to be 0.

Construction 4.2 produces a well-defined pair $(\hat{\mathbb{Q}},\hat{S})$ as long as the model is free of arbitrage. This is because the existence of $\hat{S}_0$ is assured by the continuity of J_0, and the infimum in (4.4) is attained whenever finite. It also produces a solution to the optimization problem (3.9), as claimed at the start of the section.

Theorem 4.1. *Assume no-arbitrage. For $X\in\mathcal{L}_T^2$ given, let $(J_t)_{t=0}^T$ and $(\hat{\mathbb{Q}},\hat{S})=(\hat{\mathbb{Q}},\hat{S})$ be given by Constructions 4.1 and 4.2. Then $(\hat{\mathbb{Q}},\hat{S})\in\mathcal{P}$ is a minimiser in* (3.9) *and*

$$\begin{aligned}K(X)=J_0(\hat{S}_0)&=\min_{x\in[S_0^b,S_0^a]}J_0(x)\\&=H((\hat{\mathbb{Q}},\hat{S}),X)=\min_{(\mathbb{Q},S)\in\bar{\mathcal{P}}}H((\mathbb{Q},S),X)\\&=\min_{(\mathbb{Q},S)\in\bar{\mathcal{P}}}\Big(\textstyle\sum_{t\in\mathcal{I}}\frac{1}{\alpha_t}\mathbb{E}_{\mathbb{Q}}\big[\ln\Lambda_t^{\mathbb{Q}}\big]-\mathbb{E}_{\mathbb{Q}}\big[X^b+X^sS_T\big]\Big).\end{aligned}$$

Moreover, the probability measure $\hat{\mathbb{Q}}$ is unique on nodes at times in $\mathcal{I}$, in the sense that if $(\mathbb{Q},S)\in\mathcal{P}$ is any other pair produced by Construction 4.2, then

$$\hat{\mathbb{Q}}(\nu)=\mathbb{Q}(\nu)\text{ for all }t\in\mathcal{I}\text{ and }\nu\in\Omega_t. \tag{4.12}$$

The proof of this result appears in Appendix B. The property (4.12) ensures that $\hat{\mathbb{Q}}$ is unique as long as the σ-field generated by $\{\nu\in\Omega_t:t\in\mathcal{I}\}$ is 2^Ω. However, the pair $(\hat{\mathbb{Q}},\hat{S})$ is not unique in general, because the solutions to (4.8) and (4.9)–(4.11) might not be unique. Nevertheless, the property (4.12) is sufficient to ensure the uniqueness of the optimal injection strategy, which will be considered in the next section.

5. Optimal injection and investment

In this section we turn to the problem of constructing optimal hedging strategies and optimal cash injection strategies for (3.1) and (3.6), using the objects constructed in the previous section. The optimal cash injection is constructed first, and then used in the construction of the optimal trading strategy.

The primal-dual optimality conditions associated with (3.2) and (3.10) provide a good starting point. When combined with Theorem 4.1, the conditions in the following result can be used to derive some of the properties of the optimal trading and injection strategies. This result, and the others in this section, hold under robust no-arbitrage.

Proposition 5.1. *Assume robust no-arbitrage. For any $u\in\mathcal{N}^2$, let $\hat{\lambda}_u$ be given by* (3.12)*. A trading strategy $\hat{y}\in\Psi$ attains the infimum in* (3.2) *and a martingale pair $(\hat{\mathbb{Q}},\hat{S})\in\bar{\mathcal{P}}$ satisfies*

$$K\big(\textstyle\sum_{t=0}^{T}u_t\big)=H\big((\hat{\mathbb{Q}},\hat{S}),\sum_{t=0}^{T}u_t\big) \tag{5.1}$$

if and only if the following conditions hold true:

(1) *For all $t \in \mathcal{I}$ we have $\hat{\lambda}_u \Lambda_t^{\hat{\mathbb{Q}}} = \alpha_t e^{\alpha_t \phi_t(\Delta \hat{y}_t + u_t)}$.*
(2) *For all $t \notin \mathcal{I}$ we have $\phi_t(\Delta \hat{y}_t + u_t) \leq 0$ and $\{\phi_t(\Delta \hat{y}_t + u_t) < 0\} \subseteq \{\Lambda_t^{\hat{\mathbb{Q}}} = 0\}$.*
(3) *For all t,*

$$\left\{\Delta \hat{y}_t^s + u_t^s > 0\right\} \subseteq \left\{\hat{S}_t = S_t^a\right\} \text{ and } \left\{\Delta \hat{y}_t^s + u_t^s < 0\right\} \subseteq \left\{\hat{S}_t = S_t^b\right\},$$

equivalently,

$$(\Delta \hat{y}_t^s + u_t^s)_+ S_t^a - (\Delta \hat{y}_t^s + u_t^s)_- S_t^b = (\Delta \hat{y}_t^s + u_t^s)\hat{S}_t. \tag{5.2}$$

Just like Theorem 3.1 the proof of this result relies on theoretical results obtained by Pennanen and Perkkiö (2018) in a general setting. It appears in Appendix B.

Proposition 5.1 provides a link with shadow prices, a concept that has been considered in utility optimisation under transaction costs (Kallsen and Muhle-Karbe 2011, Czichowsky et al. 2014 and others). An adapted price process $\hat{S} \in \mathcal{N}$ is called a shadow price process for a given liability $u \in \mathcal{N}^2$ if $S_t^b \leq \hat{S}_t \leq S_t^a$ for all t, and the optimal disutility in the model with bid-ask spread $[S^b, S^a]$ and in the friction-free model with price process $\hat{S}$ coincide. By Proposition 5.1, any martingale pair $(\hat{\mathbb{Q}}, \hat{S})$ given by Theorem 4.1 for $X = \sum_{t=0}^T u_t$ satisfies (5.2), and hence

$$V(u) = \inf_{y \in \Psi} \sum_{t=0}^T \mathbb{E}\left[v_t\left(\Delta y_t^b + u_t^b + (\Delta y_t^s + u_t^s)\hat{S}_t\right)\right], \tag{5.3}$$

in other words, $\hat{S}$ is a shadow price process for u.

The following result gives an explicit formula for the optimal injection strategy. Its proof appears in Appendix B. It is consistent with Corollary 3.4 of Kallsen and Muhle-Karbe (2011) (obtained in a slightly different setting).

Theorem 5.1. *Assume robust no-arbitrage. For any $u \in \mathcal{N}^2$, let $(\hat{\mathbb{Q}}, \hat{S})$ be as in Theorem 4.1 for $X = \sum_{t=0}^T u_t$. Then the process $\hat{x} \in \mathcal{N}$ defined by*

$$\hat{x}_t = \begin{cases} \frac{1}{\alpha_t} \ln \frac{\hat{\lambda}_u \Lambda_t^{\hat{\mathbb{Q}}}}{\alpha_t} & \text{if } t \in \mathcal{I}, \\ 0 & \text{if } t \notin \mathcal{I}, \end{cases} \tag{5.4}$$

where $\hat{\lambda}_u$ is given by (3.12)*, is the unique minimiser in* (3.6).

This result leads to the following important observation about optimal injection and trading strategies.

Remark 5.1. Substituting (5.4) into (3.6) and (5.1) into (3.11), the optimal P&L (cash gain, negative injection) is

$$-\sum_{t \in \mathcal{I}} \hat{x}_t = \sum_{t \in \mathcal{I}} \frac{1}{\alpha_t} \left(\mathbb{E}_{\hat{\mathbb{Q}}}\left[\ln \Lambda_t^{\hat{\mathbb{Q}}}\right] - \ln \Lambda_t^{\hat{\mathbb{Q}}}\right) - \sum_{t=0}^T \mathbb{E}_{\hat{\mathbb{Q}}}[u_t^b + u_t^s \hat{S}_T]. \tag{5.5}$$

The second term on the right hand side arises naturally in the no-arbitrage pricing of the liability u; see Section 2.3. The first term in this expression is effectively a profit that can be achieved from following this particular injection strategy (rather than any other). Taking the expected value of this term under the real-world probability $\mathbb{P}$ gives that

$$\sum_{t\in\mathcal{I}} \frac{1}{\alpha_t} \left(\mathbb{E}_{\hat{\mathbb{Q}}}\big[\ln \Lambda_t^{\hat{\mathbb{Q}}}\big] - \mathbb{E}\big[\ln \Lambda_t^{\hat{\mathbb{Q}}}\big]\right) = \sum_{t\in\mathcal{I}} \frac{1}{\alpha_t} \sum_{\omega\in\Omega} \left(\hat{\mathbb{Q}}(\omega) - \mathbb{P}(\omega)\right) \ln \frac{\hat{\mathbb{Q}}(\omega)}{\mathbb{P}(\omega)} \geq 0.$$

When $\hat{\mathbb{Q}} = \mathbb{P}$, then this term is zero, but whenever $\hat{\mathbb{Q}}$ is distinct from $\mathbb{P}$, there is some room for profit. The numerical results in Example 8.4 support this finding.

Remark 5.2. The optimal injection strategy can be constructed inductively by decomposing (5.4) into transition probabilities and using Theorem 4.1. For given $u \in \mathcal{N}^2$, take the sequence $(J_t)_{t=0}^T$ from Construction 4.1 with $X = \sum_{t=0}^T u_t$ and pair $(\hat{\mathbb{Q}}, \hat{S})$ from Construction 4.2. Then

$$\hat{\lambda}_u = \exp\left\{\tfrac{1}{a_0}\left(\textstyle\sum_{t\in\mathcal{I}} \frac{\ln \alpha_s}{\alpha_s} - J_0(\hat{S}_0)\right)\right\}$$

and

$$\hat{x}_t = \begin{cases} \frac{1}{\alpha_t} \ln \frac{\hat{\lambda}_u}{\alpha_t} & \text{if } t \in \mathcal{I} \cap \{0\}, \\ \frac{1}{\alpha_t} \ln \frac{\hat{\lambda}_u}{\alpha_t} + \frac{1}{\alpha_t} \sum_{s=0}^{t-1} \ln \frac{\hat{q}_{s+1}}{p_{s+1}} & \text{if } t \in \mathcal{I} \setminus \{0\}, \\ 0 & \text{if } t \notin \mathcal{I}. \end{cases}$$

We now turn our attention to the construction of the set of optimal trading strategies, the key idea being that it is sufficient to construct trading strategies that satisfy the conditions in Proposition 5.1. To the point, the uniqueness of the optimal injection strategy and (5.2) means that it is sufficient to construct trading strategies in a friction-model with stock price process $\hat{S}$ for any martingale pair $(\hat{\mathbb{Q}}, \hat{S})$ satisfying the conditions of Theorem 4.1 for $X = \sum_{t=0}^T u_t$. The set of optimal trading strategies is constructed as follows.

Construction 5.1. Assume robust no-arbitrage and take $u \in \mathcal{N}^2$ as given. For the sequence $(J_t)_{t=0}^T$ from Construction 4.1 with $X = \sum_{t=0}^T u_t$ and a pair $(\hat{\mathbb{Q}}, \hat{S})$ from Construction 4.2, construct a sequence of auxiliary sets $(\mathcal{W}_t)_{t=-1}^T$ by induction, where

$$\mathcal{W}_t \subset \mathcal{N}_t^{2\prime} := \left\{(w_k)_{k=-1}^t : w \in \mathcal{N}^{2\prime}\right\} \text{ for all } t,$$

and a set $\mathcal{Y} \subset \mathcal{N}^{2\prime}$.

Define $\mathcal{W}_{-1} := \{0\}$. For each $t = 0, \dots, T-1$, let $\mathcal{W}_t$ be the collection of all processes $(w_k)_{k=-1}^t \in \mathcal{N}_t^{2\prime}$ such that $(w_k)_{k=-1}^{t-1} \in \mathcal{W}_{t-1}$ and the random variable $w_t \in \mathcal{L}_t^2$ solves on each node $\mu \in \Omega_t$ the system of equations

$$\Delta w_t^{s\mu} \hat{S}_t^{\mu} = (\Delta w_t^s)_+ S_t^{a\mu} - (\Delta w_t^s)_- S_t^{b\mu}, \tag{5.6}$$

$$w_t^{b\mu} + w_t^{s\mu} \hat{S}_{t+1}^{\nu} = -J_{t+1}^{\nu}(\hat{S}_{t+1}^{\nu}) - a_{t+1} \ln \frac{\hat{q}_{t+1}^{\nu}}{p_{t+1}^{\nu}} \text{ for all } \nu \in \mu^+, \tag{5.7}$$

where a_{t+1} is given by (3.7). Finally, let $\mathcal{W}_T$ be the collection of all processes $w = (w_t)_{t=-1}^{T} \in \mathcal{N}_T^{2\prime} = \mathcal{N}^{2\prime}$ such that $(w_t)_{t=-1}^{T-1} \in \mathcal{W}_{T-1}$ and the random variable $w_T \in \mathcal{L}_T^2$ satisfies

$$w_T = \sum_{t=0}^{T} u_t, \qquad \Delta w_T^s \hat{S}_T = (\Delta w_T^s)_+ S_T^a - (\Delta w_T^s)_- S_T^b. \tag{5.8}$$

Define $\mathcal{Y}$ to be the collection of all trading strategies $\hat{y} \in \mathcal{N}^{2\prime}$ constructed by induction from some $w \in \mathcal{W}_T$ as $\hat{y}_{-1} := 0$ and

$$\hat{y}_t^b := \begin{cases} \Delta w_0^b + \hat{x}_0 - u_0^b + J_0(\hat{S}_0) & \text{if } t = 0, \\ \hat{y}_{t-1}^b + \Delta w_t^b + \hat{x}_t - u_t^b - a_t \ln \frac{\hat{q}_t}{p_t} & \text{if } t > 0, \end{cases} \tag{5.9}$$

$$\hat{y}_t^s := \hat{y}_{t-1}^s + \Delta w_t^s - u_t^s \text{ for all } t \geq 0. \tag{5.10}$$

Here $\hat{x} \in \mathcal{N}$ is calculated as in Remark 5.2.

Construction 5.1 requires the system of equations (5.6)–(5.7) to be solved at every non-terminal node, and (5.8) at each terminal node, in each case for two variables. Despite being the stock price process of an arbitrage-free model, the shadow price process $\hat{S}$ can be degenerate (for example, under large proportional transaction costs it could be constant), which can lead to these systems of equations being underdetermined, and hence having many solutions. This is the reason why the construction produces a collection of processes, rather than a single strategy. In most practical applications (involving models with two or more successors at each non-terminal node and small to moderate transaction costs), the systems involve two or more equations, and hence the collections produced by this construction are very small. That they are not empty (and hence that the systems are well-determined) comes from the following result. Its proof appears in Appendix B.

Theorem 5.2. *Assume robust no-arbitrage. For given $u \in \mathcal{N}^2$, let $\mathcal{Y}$ be the collection of trading strategies from Construction 5.1. Then $\mathcal{Y} \neq \emptyset$ and every $\hat{y} \in \mathcal{Y}$ is a minimiser in* (3.1).

In practice, the computational cost of constructing an optimal trading strategy $\hat{y}$ grows exponentially in the number of time steps, even in recombinant binary trees. The reason for this is that neither $\hat{x}$ nor $\hat{S}$ are generally recombinant processes, even when $\sum_{t=0}^{T} u_t$ is path-independent and the bid-ask spread $[S^b, S^a]$ is a recombinant process. However, it is very efficient for determining the trading strategy in particular scenarios of interest.

6. Indifference pricing

In this section we consider an investor trading in cash and shares and who is entitled to receive a given portfolio $w_t \in \mathcal{L}_t^2$ at each time step t. We refer to the payment stream $w \in \mathcal{N}^2$ as the *endowment* of the investor (though it may in fact represent a liability if negative). The minimal disutility of the investor in this situation is $V(-w)$.

Indifference pricing provides a way for such an investor to determine the value of derivatives, or payment streams. We will introduce disutility indifference prices for the seller and buyer of a payment stream $c \in \mathcal{N}^2$. Consider the situation where the investor is selling the payment stream c. He receives a single payment of $\delta \in \mathbb{R}$ in cash at time 0, and then delivers the portfolio c_t at each time step t. After selling c, the investor's minimum disutility becomes $V(c - \delta\mathbb{1} - w)$, where the process $\mathbb{1} = (\mathbb{1}_t)_{t=0}^T$ is defined as

$$\mathbb{1}_t := \begin{cases} (1,0) & \text{if } t = 0, \\ (0,0) & \text{if } t > 0. \end{cases}$$

The *seller's disutility indifference price* $\pi^{ai}(c; w)$ of c is defined as the lowest price for which he could sell c without increasing his minimal disutility, in other words,

$$\pi^{ai}(c; w) := \inf\{\delta \in \mathbb{R} : V(c - \delta\mathbb{1} - w) \leq V(-w)\}. \tag{6.1}$$

The *buyer's disutility indifference price* $\pi^{bi}(c; w)$ is similarly defined as the highest price at which the investor could buy the payment stream (and receive c_t at each time step t) without increasing his minimal disutility, in other words,

$$\begin{aligned} \pi^{bi}(c; w) &:= \sup\{\delta \in \mathbb{R} : V(-c + \delta\mathbb{1} - w) \leq V(-w)\} \\ &= -\inf\{\delta \in \mathbb{R} : V(-c - \delta\mathbb{1} - w) \leq V(-w)\} = -\pi^{ai}(-c; w). \end{aligned} \tag{6.2}$$

The following theorem gives formulae for computing the buyer's and seller's indifference prices. These pricing formulae resemble existing formulae for utility indifference prices in friction-free models under exponential utility, in particular those obtained by Delbaen, Grandits, Rheinländer, Samperi, Schweizer and Stricker (2002) and Rouge and El Karoui (2000) in general continuous-time market models without transaction costs, and Musiela and Zariphopoulou (2004) in a discrete time friction-free model with a non-traded asset.

Theorem 6.1. *Assume robust no-arbitrage. We have for any $c, w \in \mathcal{N}^2$ that*

$$\pi^{ai}(c; w) = K\big(- \textstyle\sum_{t=0}^T w_t\big) - K\big(\sum_{t=0}^T (c_t - w_t)\big), \tag{6.3}$$

$$\pi^{bi}(c; w) = K\big(- \textstyle\sum_{t=0}^T (c_t + w_t)\big) - K\big(- \sum_{t=0}^T w_t\big). \tag{6.4}$$

Notice that, in order to determine the buyer's and seller's indifference prices of a payment stream, it is sufficient to be able to determine the value of K for three different random variables. The proof of this result appears in Appendix B.

Remark 6.1. Similar results can be obtained for a related method of valuation, namely *reservation pricing*, which is often encountered in financial reporting and supervision of financial institutions; see work by Davis et al. (1993), Jaschke and Küchler (2001), Pennanen (2014*b*), and many others. The reservation value of a liability $c \in \mathcal{N}^2$ is defined as

$$\pi^r(c) := \inf\{\delta \in \mathbb{R} : V(c - \delta\mathbb{1}) \leq 0\}.$$

Similar arguments as in the proof of Theorem 6.1 give that

$$\pi^r(c) = a_0 \ln \tfrac{a_0}{|\mathcal{I}|} + \sum_{t\in\mathcal{I}} \tfrac{\ln \alpha_t}{\alpha_t} - K(c)$$

under robust no-arbitrage.

We conclude this section by verifying that indifference prices do indeed produce smaller bid-ask intervals than superhedging prices.

Theorem 6.2. *Assume robust no-arbitrage. We have for any $c, w \in \mathcal{N}^2$ that*

$$\pi^b(c) \leq \pi^{bi}(c; w) \leq \pi^{ai}(c; w) \leq \pi^a(c).$$

Moreover, the mapping $u \mapsto \pi^{ai}(u; w)$ is convex, and $u \mapsto \pi^{bi}(u; w)$ is concave.

The proof can be found in Appendix B.

7. Numerical approximation

The adapted process of functions $(J_t)_{t=0}^T$ of Construction 4.1 is the key to determining minimal disutility, indifference prices as well as the construction of optimal trading strategies. It is possible to derive these functions analytically in simple cases (see Example 4.1), but numerical approximation is required, in general.

In this section we propose a numerical method for approximating these functions. For simplicity of exposition and availability of error bounds we assume throughout that the model satisfies the relatively mild condition

$$\min_{\nu\in\mu^+} S_{t+1}^{b\nu} < S_t^{b\mu} \leq S_t^{a\mu} < \max_{\nu\in\mu^+} S_{t+1}^{a\nu} \text{ for all } t < T, \mu \in \Omega_t, \tag{7.1}$$

which implies robust no-arbitrage and $\operatorname{dom} J_t = [S_t^b, S_t^a]$ for all t. Similarly, the assumption that bid-ask intervals are subdivided into the same number $n \in \mathbb{N}$ intervals of equal length is made for simplicity, and can be relaxed.

The following construction of the *upper approximation* is based on the discussion in Appendix A.2. It is applied backwards in time from J_T at time T, and produces adapted sequences of piecewise linear functions by dividing the bid-ask interval at each node into a finite number of subintervals.

Construction 7.1. Assume (7.1) and take $X \in \mathcal{L}_T^2$ and $n \in \mathbb{N}$ as given. Construct an adapted sequence of piecewise linear random functions $(\hat{J}_t)_{t=0}^T$ by backward induction as follows. Define $\hat{J}_T : \Omega \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ as

$$\hat{J}_T^\nu(x) := \begin{cases} -X^{b\nu} - xX^{s\nu} & \text{if } x \in \left[S_T^{b\nu}, S_T^{a\nu}\right], \\ \infty & \text{otherwise.} \end{cases} \tag{7.2}$$

for all $\nu \in \Omega_T$. For every $t < T$, assume that $\hat{J}_{t+1}$ has already been constructed, and define for all $\mu \in \Omega_t$ and $l = 0, \dots, n$

$$\begin{aligned}\hat{x}_l^\mu &:= \tfrac{n-l}{n} S_t^{b\mu} + \tfrac{l}{n} S_t^{a\mu} \text{ for all } l = 0, \dots, n,\\ \hat{g}_l^\mu &:= \min\Big\{ \textstyle\sum_{\nu\in\mu^+} q^\nu \left(\hat{J}_{t+1}^\nu(x^\nu) + a_{t+1} \ln \frac{q^\nu}{p_{t+1}^\nu} \right) : q^\nu \in [0,1],\\ &\qquad x^\nu \in [S_{t+1}^{b\nu}, S_{t+1}^{a\nu}]\, \forall \nu \in \mu^+, \textstyle\sum_{\nu\in\mu^+} q^\nu = 1, \sum_{\nu\in\mu^+} q^\nu x^\nu = \hat{x}_l^\mu \Big\}, \qquad (7.3)\end{aligned}$$

and finally for all $x \in \mathbb{R}$

$$\hat{J}_t^\mu(x) := \begin{cases} \hat{g}_l^\mu & \text{if } x = \hat{x}_l^\mu \text{ for some } l,\\ \frac{\hat{x}_l^\mu - x}{\hat{x}_l^\mu - \hat{x}_{l-1}^\mu} \hat{g}_{l-1}^\mu + \frac{x - \hat{x}_{l-1}^\mu}{\hat{x}_l^\mu - \hat{x}_{l-1}^\mu} \hat{g}_l^\mu & \text{if } x \in (\hat{x}_{l-1}^\mu, \hat{x}_l^\mu) \text{ for any } l,\\ \infty & \text{if } x \in \mathbb{R} \setminus [S_t^{b\mu}, S_t^{a\mu}]. \end{cases}$$

It follows from the discussion in Appendix A.2 that the upper approximation is indeed true to its name in that $\hat{J}_t \geq J_t$ for all t. Moreover, repeated application of Proposition A.5 gives the following error bound for the upper approximation.

Theorem 7.1. *Assume* (7.1) *and take* $X \in \mathcal{L}_T^2$ *and* $n \in \mathbb{N}$ *as given. Let* $(J_t)_{t=0}^T$ *be the process given by Construction 4.1, let* $(\hat{J}_t)_{t=0}^T$ *be the upper approximation of Construction 7.1, and define*

$$\Delta_n := \tfrac{1}{n} \max \left\{ S_t^{a\mu} - S_t^{b\mu} : t = 0, \dots, T, \mu \in \Omega_t \right\}.$$

Then there exists a constant $c > 0$ *such that*

$$\left| J_0(x) - \hat{J}_0(x) \right| \leq c\Delta_n \textit{ for all } x \in [S_0^b, S_0^a].$$

The constant c in this result depends on the (unknown) Lipschitz coefficients of the functions $(J_t)_{t=0}^T$. In practice, one could instead use the lower approximation described in Appendix A.2, which produces a sequence of piecewise linear functions $(\check{J}_t)_{t=0}^T$ with $\check{J}_T = J_T$ and which satisfies $\check{J}_t \leq J_t$ for all t. Whilst the lower approximation has proved to be less efficient in numerical experiments, and no error bound is known, it provides a lower bound that can be used to assess the accuracy of the upper approximation. This is demonstrated in Example 8.1.

The upper approximation requires determination of the value of the generalised convex hull of piecewise linear functions (in (7.3)) at each of the $n+1$ endpoints of the subintervals. Exact solutions exist for this (Xu 2018 Section 4.3), and involves inspecting $O(n^{|\mu^+|})$ different combinations of the pieces of the piecewise linear functions at each of the successors of a node μ. One could alternatively use a non-linear optimiser, with fixed computational cost for each endpoint; however this risks introducing numerical errors and might break the convexity of the approximating piecewise linear functions. As there are $n+1$ endpoints, the computational cost of determining the approximation for J at each node μ is $O(n^{|\mu^+|+1})$. The overall computational complexity depends on the number of nodes (which in turn depends on the structure and size of the model tree). The computational costs and the required memory are both proportional to the number of nodes in the tree.

Consider the example of a binary tree model satisfying (7.1). In order to achieve an accuracy of $O(\epsilon)$ when approximating J_0, the computational cost is $O(\epsilon^{-3})$. However the number of nodes play an important role; if the model is recombinant and the payoff X is path-independent, then one need only consider the $\frac{1}{2}T(T+1)$ "identifiable" nodes; in general one may need to consider up to $2^{T+1}-1$ nodes.

Once the upper approximation $(\hat{J}_t)_{t=0}^T$ has been constructed, then it is natural to approximate K via Theorem 4.1, and π^{ai} and π^{bi} via Theorem 6.1. The process $(\hat{J}_t)_{t=0}^T$ can also be used to construct optimal injection and trading strategies along any given path by means of Constructions 4.2 and 5.1. As these constructions require repeated calculations along a stock price path, and generally lead to path-dependent objects, the computational cost is proportional to the number of scenarios in the model.

8. Numerical examples

Consider a friction-free binomial tree model with $T = 52$ steps representing one year in real time with weekly rehedging, where the stock price $S = (S_t)_{t=0}^{52}$ satisfies $S_0 = 100$ and

$$S_{t+1} = \begin{cases} e^{\sigma\sqrt{1/52}}S_t & \text{with probability } p, \\ e^{-\sigma\sqrt{1/52}}S_t & \text{with probability } 1-p \end{cases}$$

for all $t < 52$. Here $\sigma = 0.2$ is the annual volatility of the return on stock, and the model is assumed to have an annual effective interest rate of $r_e = 0.02$. Define the bid and ask prices of the stock as

$$S_t^a := (1+k)S_t, \qquad S_t^b := (1-k)S_t$$

for all $t > 0$, where k is the proportional transaction cost parameter. We assume that there are no transaction costs at time 0, in other words $S_0^a := S_0^b := S_0 = 100$.

The investor's endowment is $w = 0$, and that the risk aversion coefficient is constant, in other words, $\alpha_t = \alpha$ for all $t \in \mathcal{I}$. Consider a call option with expiry one year, strike 100 and physical delivery (based on the underlying). This corresponds to the payment stream $C = (C_t)_{t=0}^{52}$ where $C_t = 0$ for all $t < 52$ and

$$C_{52} = (-100, 1)\mathbb{1}_{\{S_{52}>100\}}.$$

The numerical results in this section were obtained by applying the upper and lower approximation methods introduced described in Section 7. Superhedging bid and ask prices are also provided for the purposes of comparison, calculated using methods previously reported by Roux et al. (2008).

Example 8.1. Table 1 contains approximate indifference prices for the seller and buyer of the call option in the case where $p = 0.5$, $k = 0.005$, $\mathcal{I} = \{0, \dots, 52\}$ and $\alpha = 0.1$, as computed by both the upper and lower approximation methods. In each case, the approximation is obtained by dividing each (discounted) bid-ask interval into n subintervals of equal length.

Table 1. Indifference prices by approximation method (Example 8.1)

n	20	50	100	150	200	300
Upper approximation method						
$\pi^{bi}(C;0)$	8.5759	8.5673	8.5658	8.5655	8.5654	8.5654
$\pi^{ai}(C;0)$	9.1596	9.1672	9.1684	9.1687	9.1687	9.1688
Lower approximation method						
$\pi^{bi}(C;0)$	8.4974	8.5533	8.5633	8.5647	8.5652	8.5653
$\pi^{ai}(C;0)$	9.2357	9.1797	9.171	9.1692	9.1690	9.1690

The results from the two approximation methods are consistent in that they converge to the same limit, but the upper approximation converges much faster than the lower approximation. The results suggest that taking $n = 150$ results in accuracy up to 3 decimal places, which is perfectly adequate for graphical representation.

The indifference pricing spread (between the seller's and buyer's indifference prices) is considerably smaller than the (superhedging) bid-ask spread; note that the ask and bid prices in this case are $\pi^a(C) = 10.4788$ and $\pi^b(C) = 6.9694$.

Different possibilities for the set $\mathcal{I}$ of dates on which injection is allowed will be considered below. The case $\mathcal{I} = \{52\}$, in particular, corresponds to the classical utility indifference pricing framework, where the cash injection at time 52 reflects the hedging shortfall at the expiration date of the option under exponential utility.

Example 8.2. Figure 1 illustrates seller's and buyer's indifference prices for a range of values of the risk aversion coefficient α and transaction cost parameter k in the case where $p = 0.5$. Observe that the indifference pricing spread (between the seller's and buyer's indifference prices) is smaller for $\mathcal{I} = \{0, \dots, 52\}$ than $\mathcal{I} = \{52\}$. This is because being able to inject cash at different time steps introduces considerable flexibility, which in turn results in decreased hedging costs.

As seen in part (a), indifference pricing spreads increase as α increases. The indifference pricing spread remains well within the superhedging bid-ask spread for a large range of values of α.

Indifference pricing spreads increase with k, the intuitive reason being that increased transaction costs results in increased trading costs. This is illustrated in part (b). Observe finally that the indifference pricing spreads remain well within the superhedging bid-ask spread for all values of k, and also expand slower as k increases.

Example 8.3. Buyer's and seller's indifference prices for a range of values of the market probability parameter p in the case where $k = 0.005$ and $\alpha = 0.1$, are illustrated in Figure 2. It can be seen in part (a) that indifference pricing spreads tend to be at their largest when p is close to the value of the friction-free risk-neutral

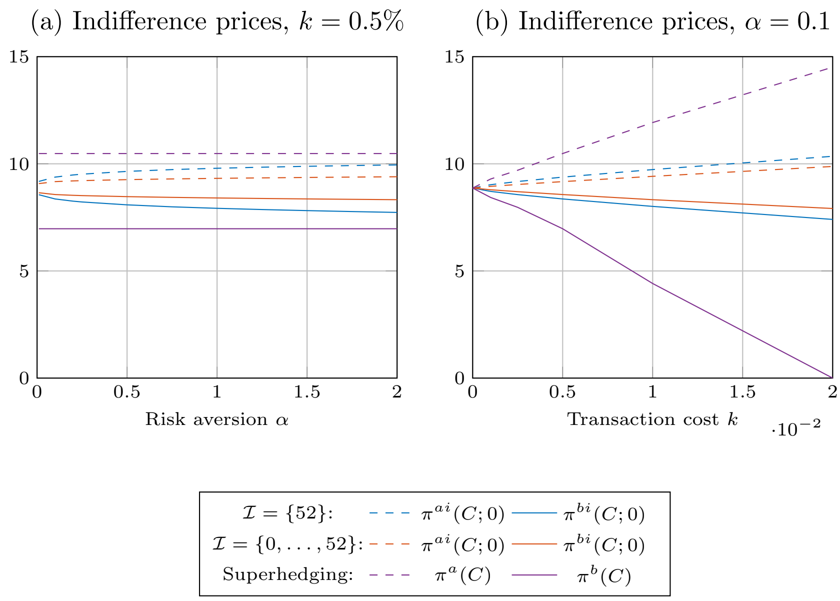

Fig. 1. Indifference prices, transaction costs and risk aversion (Example 8.2)

probability in this model, which is

$$q = \left((1+r_e)^{1/52} - e^{-\sigma\sqrt{1/52}}\right) \Big/ \left(e^{\sigma\sqrt{1/52}} - e^{-\sigma\sqrt{1/52}}\right) \approx 0.4999.$$

The effect is more pronounced when injection is allowed at more trading dates. It can be explained by examining the behaviour of $K(0)$, $K(-C_{52})$ and $K(C_{52})$ for different values of p, illustrated in part (b). Whilst the dependence of these values on p appear to be convex, they vary in steepness, both within groups associated with the same choice and $\mathcal{I}$, and between groups associated with different choices of $\mathcal{I}$. This then has consequences for the vertical differences $\pi^{bi}(C;0) = K(-C_{52}) - K(0)$ and $\pi^{ai}(C;0) = K(0) - K(C_{52})$.

Example 8.4. Figure 3 illustrates a number of numerical results related to optimal injection and hedging strategies for $\mathcal{I} = \{52\}$ and $\mathcal{I} = \{0, 13, \dots, 52\}$ and for different values of the probability p. The risk-aversion parameter is $\alpha = 0.2$ throughout.

Parts (a) and (b) contain histograms of the optimal P&L $-\sum_{t\in\mathcal{I}} \hat{x}_t$ for 100000 randomly generated scenarios in the case where $k = 0.005$. It is clear that the P&L tends to be larger if the real-world probability is further away from the risk-neutral probability (calculated in Example 8.3), thus confirming the analysis in Remark 5.1. The distribution of P&L depends on $\mathcal{I}$, too, with distributions being much wider in the case where $\mathcal{I} = \{0, 13, \dots, 52\}$. Making injections quarterly, instead of at the terminal time step, allows an investor to reduce their regret by taking advantage of

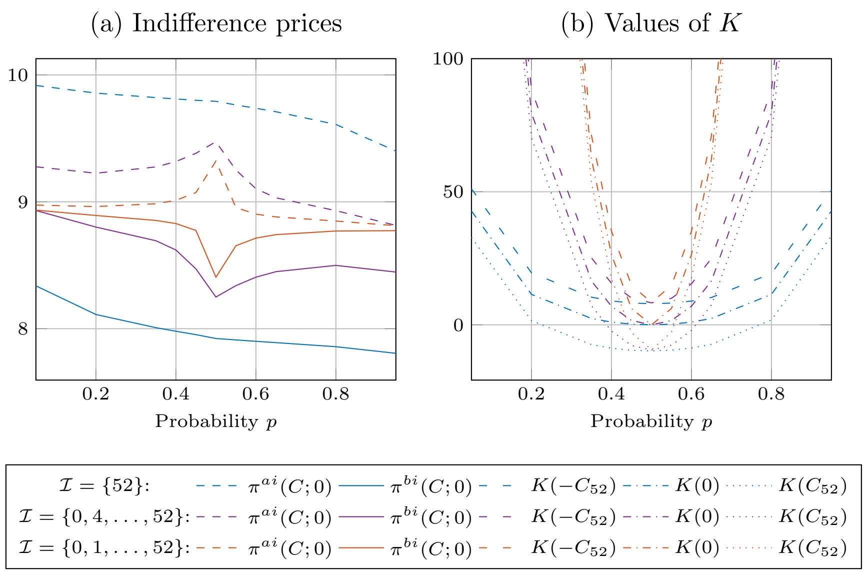

Fig. 2. Indifference prices and market probability (Example 8.3)

the convexity of the disutility function.

Due to the smallness of the transaction costs, Construction 5.1 produces a unique optimal trading strategy $\hat{y} = (\hat{y}_t)_{t=-1}^{52}$ in this model. Parts (c)–(f) illustrate the optimal stock positions $(\hat{y}_t^s)_{t=0}^{52}$ associated with this strategy in two scenarios. The stock positions should be compared to the stock positions associated with the replicating strategy in the binary model without transaction costs (pictured).

Parts (c) and (e) focus on the stock positions when $\mathcal{I} = \{52\}$ in the case of no transaction costs ($k = 0$) and $k = 0.005$. The presence of transaction costs lead to smoother stock positions due to a reduction in trading. Stock positions tend to be higher for higher values of p; this indicates that the investor is taking advantage of market information.

The corresponding results for the case $\mathcal{I} = \{0, 13, \ldots, 52\}$ are provided in (d) and (f). In this case the tendency is for stock holdings to be larger (in absolute value) initially, but with larger adjustments each quarter, and tending to similar values in the final quarter as in the case $\mathcal{I} = \{52\}$.

Xu (2018 Section 5.5) reported a large number of numerical examples illustrating the methods of this paper, for a selection of options with cash and physical delivery, and for a range of values of r_e and T.

Appendix A. Generalised convex hull

The constructions in Section 4 involve a generalisation of the convex hull of convex functions. This section outlines the main properties used in this paper. For $k =$

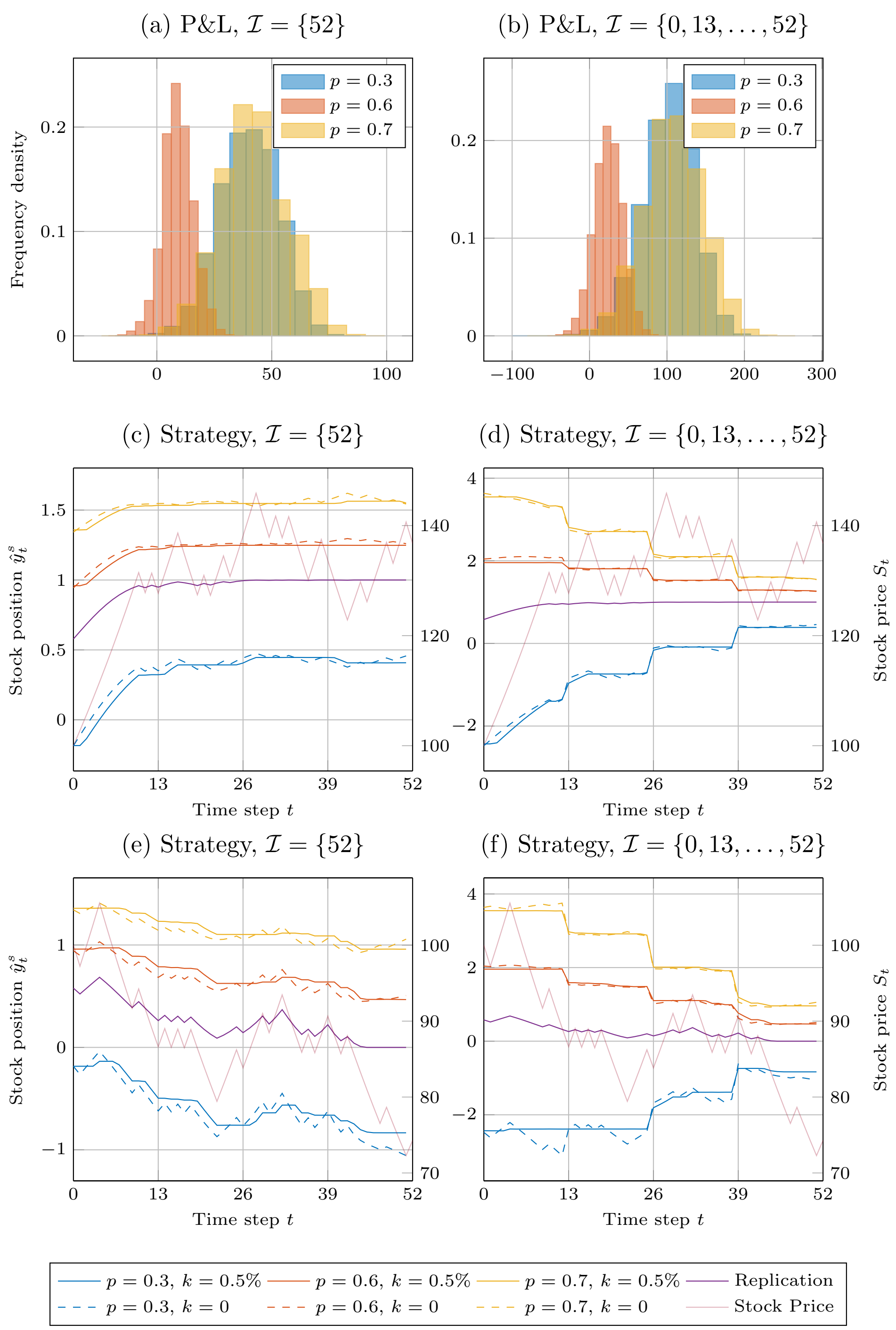

Fig. 3. Optimal injection and trading strategies

$1,\dots,m$, let $f_k, g_k : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be proper convex functions that are continuous on their effective domains $\operatorname{dom} f_k = [b_k, a_k]$ for some $b_k, a_k \in \mathbb{R}$ and $\operatorname{dom} g_k = [0,1]$, and

$$g_k(0) = 0. \tag{A.1}$$

Define the *generalised convex hull* $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ of $f_1, \dots, f_m$ and $g_1, \dots g_m$ as

$$\begin{aligned} f(x) \coloneqq \inf \{ \textstyle\sum_{k=1}^m (q_k f_k(x_k) + g_k(q_k)) : q_k \in [0,1], x_k \in [b_k, a_k] \text{ for all } k, \\ \textstyle\sum_{k=1}^m q_k = 1, \sum_{k=1}^m q_k x_k = x \} \, . \end{aligned} \tag{A.2}$$

A.1. *General properties*

The main aim of this section is to establish the key properties needed in Section 4. Further detail on the arguments below, in a slightly more general setting, were presented by Xu (2018 Chapter 4).

The first result establishes the convexity and boundedness of f, as well as the compactness of its effective domain.

Proposition A.1. *The function f in* (A.2) *is proper, convex, and*

$$\operatorname{dom} f = \operatorname{conv} \bigcup_{k=1}^m [b_k, a_k] = \Big[\min_k b_k, \max_k a_k \Big]. \tag{A.3}$$

Proof. The effective domain $\operatorname{dom} f$ is compact (Rockafellar 1996 Corollary 9.8.2). The properness of f follows from the fact that continuous proper convex functions with compact domains are bounded from below. The convexity of f comes from the convexity of the f_k's and g_k's. □

The remainder of this section is devoted to establishing the closedness of the epigraph of f. This then allows us to establish the desired properties; see Proposition A.4 at the end of the appendix. Define

$$A_k^g \coloneqq \{(q, qx, qy + g_k(q)) : q \in [0,1], (x,y) \in \operatorname{epi} f_k\} \text{ for all } k. \tag{A.4}$$

If $q = 0$, then $(q,a,b) \in A_k^g$ if and only if $a = b = 0$. This also implies that $A_k^g \neq \emptyset$. Moreover, if $(q,a,b) \in A_k^g$ satisfies $q > 0$, then $(q,a,b) + U \subset A_k^g$, where

$$U \coloneqq \{(0,0,b) \in \mathbb{R}^3 : b \geq 0\}.$$

The properties of A_k^g in the next result will be used in Proposition A.3. The recession cone of a set $C \subseteq \mathbb{R}^n$ is defined as $0^+C \coloneqq \{y \in \mathbb{R}^n : C + y \subseteq C\}$ (Rockafellar 1996 Th. 8.1).

Proposition A.2. *The following holds true for the set A_k^g in* (A.4) *for any k:*

(1) The set A_k^g is convex.
(2) The closure of A_k^g is $\operatorname{cl} A_k^g = U \cup A_k^g$.
(3) The recession cone of $\operatorname{cl} A_k^g$ *is* $0^+(\operatorname{cl} A_k^g) = U$.

Proof. Item (1): Fix any $\lambda \in (0,1)$, $q_1, q_2 \in [0,1]$ and $(x_1,y_1),(x_2,y_2) \in \operatorname{epi} f_k$ and define $q := \lambda q_1 + (1-\lambda)q_2$ and

$$z := \lambda(q_1, q_1x_1, q_1y_1 + g_k(q_1)) + (1-\lambda)(q_2, q_2x_2, q_2y_2 + g_k(q_2)).$$

If $q = 0$, then $q_1 = q_2 = 0$, after which $x_1 = y_1 = x_2 = y_2 = 0$ by the observation above, so that $z = 0 \in A_k^g$. If $q > 0$, then define $\varepsilon := \lambda g_k(q_1) + (1-\lambda)g_k(q_2) - g_k(q)$ and $(x,y) := \frac{1}{q}(\lambda q_1(x_1,y_1) + (1-\lambda)q_2(x_2,y_2) + (0,\varepsilon))$. Then $\varepsilon \geq 0$ because g_k is convex and $(x,y) \in \operatorname{epi} f_k$ because $\operatorname{epi} f_k$ is convex and unbounded from above. Thus $z = (q, qx, qy + g_k(q)) \in A_k^g$, so that A_k^g is convex.

Item (2): Define $A_k := \operatorname{cone}(\{1\} \times \operatorname{epi} f_k) = \{\lambda(1,z) : \lambda \geq 0, z \in \operatorname{epi} f_k\}$; then $\operatorname{cl} A_k = U \cup A_k$ due to the compactness of $\operatorname{dom} f_k$ (Rockafellar 1996 Theorem 8.2). For every $(0,0,b) \in U \subset \operatorname{cl} A_k$ there exist $(q_n)_{n\geq 1}$ in $[0,1]$ and $(x_n,y_n)_{n\geq 1}$ in $\operatorname{epi} f_k$ such that

$$(0,0,b) = \lim_{n\to\infty} q_n(1,x_n,y_n) = \lim_{n\to\infty} q_n(1,x_n,y_n + g_k(q_n)),$$

with the last equality due to (A.1) and the continuity of g_n. Thus $(0,0,b) \in \operatorname{cl} A_k^g$. Combining this with $A_k^g \subseteq \operatorname{cl} A_k^g$ gives that $U \cup A_k^g \subseteq \operatorname{cl} A_k^g$.

To establish the opposite inclusion, suppose that $(q,a,b) \in \operatorname{cl} A_k^g$. Then there exist $(q_n)_{n\geq 1}$ in $[0,1]$ and $(x_n,y_n)_{n\geq 1}$ in $\operatorname{epi} f_k$ such that

$$(q,a,b) = \lim_{n\to\infty}(q_n, q_nx_n, q_ny_n + g_k(q_n)).$$

Observe that $\lim_{n\to\infty} g_k(q_n) = g_k(q)$ by the continuity of g_k, so that

$$b - g_k(q) = \lim_{n\to\infty} q_ny_n.$$

Moreover, since $q_n(1,x_n,y_n) \in A_k$ for all $n \in \mathbb{N}$ it follows that

$$(q,a,b-g_k(q)) = \lim_{n\to\infty} q_n(1,x_n,y_n) \in \operatorname{cl} A_k = U \cup A_k.$$

There are now two possibilities. If $(q,a,b-g_k(q)) \in U$, then $q = 0$ and therefore $(q,a,b) \in U$ by (A.1). If $(q,a,b-g_k(q)) \in A_k$ then there exist $(x,y) \in \operatorname{epi} f_k$ such that $(q,a,b-g_k(q)) = q(1,x,y)$, in other words, $(q,a,b) = (q,qx,qy+g_k(q)) \in A_k^g$.

Item (3): The comments just before this proposition together with item (2) gives that $U \subseteq 0^+(\operatorname{cl} A_k^g)$. For the opposite inclusion, take any $(q,a,b) \in 0^+(\operatorname{cl} A_k^g)$. Since $0 \in \operatorname{cl} A_k^g$, this implies that

$$\lambda(q,a,b) = 0 + \lambda(q,a,b) \in \operatorname{cl} A_k^g = U \cup A_k^g \text{ for all } \lambda > 0.$$

It then follows from (A.4) and the comments following it that $q = a = 0$, whence $(q,a,b) \in U$. □

Proposition A.3. *The set*

$$\begin{aligned} E_f &:= \{(a,b) : (1,a,b) \in \textstyle\sum_{k=1}^m A_k^g\} && \text{(A.5)} \\ &= \{\textstyle\sum_{k=1}^m (q_kx_k, q_ky_k + g_k(q_k)) : q_k \in [0,1], (x_k,y_k) \in \operatorname{epi} f_k \ \forall k, \sum_{k=1}^m q_k = 1\} && \text{(A.6)} \end{aligned}$$

is closed.

Proof. We first show that

$$\{1\} \times E_f = M \cap \textstyle\sum_{k=1}^m \operatorname{cl} A_k^g, \tag{A.7}$$

where $M := \{1\} \times \mathbb{R}^2$. Equation (A.5) gives $\{1\} \times E_f \subseteq M \cap \sum_{k=1}^m \operatorname{cl} A_k^g$. To establish the opposite inclusion, fix any $(q,a,b) \in M \cap \sum_{k=1}^m \operatorname{cl} A_k^g$; then $q = 1$ and by Proposition A.2(2) there exist $(q_k, a_k, b_k) \in U \cup A_k^g$ for every k such that

$$(1,a,b) = \textstyle\sum_{k=1}^m (q_k, a_k, b_k).$$

Define $B := \{k : (q_k,a_k,b_k) \in U\}$ and $C := \{k : (q_k,a_k,b_k) \in A_k^g \setminus U\}$. For each $k \in B$, we have $q_k = a_k = 0$ and $b_k \geq 0$; select any $(x_k, y_k) \in \operatorname{epi} f_k$ and observe that $(q_k, q_k x_k, q_k y_k + g_k(q_k)) = 0 = (q_k, a_k, b_k - b_k)$. Noting that $C \neq \emptyset$ (because $q_k > 0$ for at least one k), define $c := \frac{1}{|C|}\sum_{k\in B} b_k \geq 0$. For each $k \in C$ there exists some $(x_k, y_k') \in \operatorname{epi} f_k$ such that $(q_k, a_k, b_k) = (q_k, q_k x_k, q_k y_k' + g_k(q_k))$. Define $y_k := y_k' + \frac{c}{q_k} \geq y_k'$; then $(x_k, y_k) \in \operatorname{epi} f_k$ and

$$(q_k, q_k x_k, q_k y_k + g_k(q_k)) = (q_k, a_k, b_k + c).$$

Finally, rearrangement gives that

$$(1,a,b) = \textstyle\sum_{k\in C}(q_k, a_k, b_k + c) = \sum_{k=1}^m (q_k, q_k x_k, q_k y_k + g_k(q_k)) \in M \cap \sum_{k=1}^m \operatorname{cl} A_k^g,$$

which establishes (A.7).

Note that $\sum_{k=1}^m A_k^g$ is convex (Rockafellar 1996 Theorem 3.1). Furthermore, if $z_k \in 0^+(\operatorname{cl} A_k^g) = U$ for all k satisfies $\sum_{k=1}^m z_k = 0$, then $z_1 = \cdots = z_m = 0 \in U \cap (-U)$; this means that

$$\operatorname{cl} \textstyle\sum_{k=1}^m A_k^g = \sum_{k=1}^m \operatorname{cl} A_k^g \tag{A.8}$$

(Rockafellar 1996 Corollary 9.1.1). It remains to show that

$$M \cap \operatorname{ri} \textstyle\sum_{k=1}^m A_k^g \neq \emptyset, \tag{A.9}$$

because then the closedness E_f follows from (A.8), (A.7) and

$$M \cap \operatorname{cl} \textstyle\sum_{k=1}^m A_k^g = \operatorname{cl}\,(M \cap \sum_{k=1}^m A_k^g)$$

(Rockafellar 1996 Corollary 6.5.1).

To establish (A.9), observe that $\operatorname{ri} \sum_{k=1}^m A_k^g \neq \emptyset$ because $\sum_{k=1}^m A_k^g \neq \emptyset$. Thus there exist $q_k \in [0,1]$ and $(x_k, y_k) \in \operatorname{epi} f_k$ for all k such that

$$(q,a,b) := \textstyle\sum_{k=1}^m (q_k, q_k x_k, q_k y_k + g_k(q_k)) \in \operatorname{ri} \sum_{k=1}^m A_k^g.$$

This can now be used to construct a point $z \in M \cap \operatorname{ri} \sum_{k=1}^m A_k^g$. There are two possibilities, depending on the value of q. If $q \geq 1$, define $z := \frac{1}{q}(q,a,b)$. Then clearly $z \in M$ and moreover z can be written as the convex combination

$$z = \tfrac{1}{q}(q,a,b) + \big(1 - \tfrac{1}{q}\big)(0,0,0) \in \operatorname{ri} \textstyle\sum_{k=1}^m A_k^g$$

(Rockafellar 1996 Theorem 6.1). If $q \in [0,1]$, define $q'_k := \frac{1}{m}(2-q) > 0$ for all k and

$$z' := \textstyle\sum_{k=1}^m (q'_k, q'_k x_k, q'_k y_k + g_k(q'_k)) \in \sum_{k=1}^m A_k^g.$$

Then $z := \frac{1}{2}(q,a,b) + \frac{1}{2}z' \in \operatorname{ri} \sum_{k=1}^m A_k^g$ (Rockafellar 1996 Theorem 6.1) and $z \in M$ because $\frac{1}{2}q + \frac{1}{2}\sum_{k=1}^m q'_k = 1$. □

The following result concludes this section.

Proposition A.4. *The function f in* (A.2) *is continuous on* $\operatorname{dom} f$*, and the infimum in* (A.2) *is attained for all $x \in \operatorname{dom} f$.*

Proof. It is sufficient to show that $\operatorname{epi} f = E_f$, for then f is lower semicontinuous by Proposition A.3, hence continuous on $\operatorname{dom} f$ because it is a closed bounded interval (Rockafellar 1996 Theorems 10.2, 20.5). The fact that the infimum in (A.2) is attained for all $x \in \operatorname{dom} f$ follows from the properties of E_f.

Suppose that $(x,y) \in E_f$. Thus there exist $q_k \in [0,1]$ and $(x_k, y_k) \in \operatorname{epi} f_k$ for all k such that
$\sum_{k=1}^m q_k = 1$, $\sum_{k=1}^m q_k x_k = x$ and $\sum_{k=1}^m (q_k y_k + g_k(q_k)) = y$. Then

$$y = \textstyle\sum_{k=1}^m (q_k y_k + g_k(q_k)) \geq \sum_{k=1}^m (q_k f_k(x_k) + g_k(q_k)) \geq f(x),$$

and so $(x,y) \in \operatorname{epi} f$.

Conversely, suppose that $(x,y) \in \operatorname{epi} f$. Then $f(x) < \infty$ and so by (A.2) there exists a sequence $(q_{1n}, \ldots, x_{mn}, x_{1n}, \ldots, x_{mn})_{n \geq 1}$ such that

$$f(x) = \lim_{n\to\infty} \textstyle\sum_{k=1}^m (q_{kn} f_k(x_{kn}) + g_k(q_{kn}))$$

and for all $n \in \mathbb{N}$ we have $q_{kn} \in [0,1]$ and $x_{kn} \in [b_k, a_k]$ for all k, and $\sum_{k=1}^m q_{kn} = 1$ and $\sum_{k=1}^m q_{kn} x_{kn} = 1$. For each $n \in \mathbb{N}$ and $k = 1, \ldots, m$ define

$$y_{kn} := f_k(x_{kn}) + y - f(x) \geq f_k(x_{kn});$$

then $(x_{kn}, y_{kn}) \in \operatorname{epi} f_k$. Define moreover for all $n \in \mathbb{N}$

$$y_n := \textstyle\sum_{k=1}^m (q_{kn} y_{kn} + g_k(q_{kn})) = \sum_{k=1}^m (q_{kn} f_k(x_{kn}) + g_k(q_{kn})) + y - f(x);$$

then $(x, y_n) \in E_f$ and $\lim_{n\to\infty} y_n = y$. This implies that $(x,y) \in \operatorname{cl} E_f = E_f$ by Proposition A.3, which concludes the proof that $\operatorname{epi} f = E_f$. □

A.2. *Numerical approximation*

Computer implementation of the generalised convex hull necessitates a numerical approximation in all but a few special cases. In this section we propose such a numerical approximation, together with error bounds, that will be suitable for use in the dynamic procedure proposed in Section 4. It is based on approximation of $f_1, \ldots, f_m$ and f by piecewise linear functions. We will refer to this as the *upper approximation* as it approximates the generalised convex hull f from above.

For every k, divide $\operatorname{dom} f_k = [b_k, a_k]$ into n_k subintervals. If $b_k = a_k$, then define $\hat{x}_{k0} \coloneqq \hat{x}_{k1} \coloneqq \cdots \coloneqq \hat{x}_{kn_k} \coloneqq a_k$, and if $b_k < a_k$, choose any $(\hat{x}_{kl})_{l=0}^{n_k}$ such that $b_k \eqqcolon \hat{x}_{k0} < \cdots < \hat{x}_{kn_k} \coloneqq a_k$. Define $\hat{f}_k : \mathbb{R} \to \{\infty\}$ as

$$\hat{f}_k(x) \coloneqq \begin{cases} f(\hat{x}_{kl}) & \text{if } x = \hat{x}_{kl} \text{ for some } l, \\ \frac{\hat{x}_{kl}-x}{\hat{x}_{kl}-\hat{x}_{k[l-1]}}\hat{f}_k(\hat{x}_{k[l-1]}) + \frac{x-\hat{x}_{k[l-1]}}{\hat{x}_{kl}-\hat{x}_{k[l-1]}}\hat{f}_k(\hat{x}_{kl}) & \text{if } x \in (\hat{x}_{k[l-1]}, \hat{x}_{kl}) \text{ for any } l, \\ \infty & \text{if } x \in \mathbb{R} \setminus \operatorname{dom} f_k. \end{cases} \tag{A.10}$$

Observe that $\hat{f}_k \geq f_k$ by virtue of the convexity of f_k.

Let $\hat{g}$ be the generalised convex hull of $\hat{f}_1, \ldots, \hat{f}_m$ and $g_1, \ldots g_m$, in other words,

$$\hat{g}(x) \coloneqq \inf\Big\{\textstyle\sum_{k=1}^m (q_k \hat{f}_k(x_k) + g_k(q_k)) : q_k \in [0,1], x_k \in [b_k, a_k]\,\forall k, \\ \textstyle\sum_{k=1}^m q_k = 1, \sum_{k=1}^m q_k x_k = x\Big\}. \tag{A.11}$$

Then $\hat{g} \geq f$ by definition, and it follows from the arguments in the previous subsection that $\hat{g}$ is convex and continuous on its effective domain $\operatorname{dom}\hat{g} = \operatorname{dom} f$, and that the infimum in (A.11) is attained for all $x \in \operatorname{dom}\hat{g} = \operatorname{dom} f$.

In practical applications, one often needs to approximate f on some subinterval $[b, a] \subset \operatorname{dom} f$. Divide this interval into n subintervals, as follows: if $b = a$, then define $\hat{x}_0 \coloneqq \hat{x}_1 \coloneqq \cdots \coloneqq \hat{x}_n \coloneqq a_k$, and if $b < a$, choose $(\hat{x}_l)_{l=0}^n$ such that $b \eqqcolon \hat{x}_0 < \cdots < \hat{x}_n \coloneqq a$. Finally, define

$$\hat{f}(x) \coloneqq \begin{cases} \hat{g}(\hat{x}_l) & \text{if } x = \hat{x}_l \text{ for some } l, \\ \frac{\hat{x}_l - x}{\hat{x}_l - \hat{x}_{l-1}}\hat{g}(\hat{x}_{l-1}) + \frac{x-\hat{x}_{l-1}}{\hat{x}_l-\hat{x}_{l-1}}\hat{g}(\hat{x}_l) & \text{if } x \in (\hat{x}_{l-1}, \hat{x}_l) \text{ for any } l, \\ \infty & \text{if } x \in \mathbb{R} \setminus [b, a]. \end{cases} \tag{A.12}$$

Then $\hat{f}$ is piecewise linear on its effective domain, and moreover $\hat{f} \geq \hat{g} \geq f$.

Define the mesh size of the approximation as

$$\Delta \coloneqq \max\Big\{\max_{k,l}(\hat{x}_{kl} - \hat{x}_{k[l-1]}), \max_l(\hat{x}_l - \hat{x}_{l-1})\Big\}.$$

We now have the following result.

Proposition A.5. *Let f be defined by* (A.2)*, the function $\hat{f}_k$ by* (A.10) *for all k, and $\hat{f}$ by* (A.12)*. If $[b, a] \subseteq \operatorname{ri}\operatorname{dom} f$ and there exists $c_k \geq 0$ for each k such that $\big|\hat{f}_k(x) - f_k(x)\big| \leq c_k\Delta$ for all $x \in \operatorname{dom} f_k$, then there exists $c \geq 0$ such that $\big|\hat{f}(x) - f(x)\big| \leq c\Delta$ for all $x \in [a, b]$.*

Proof. For any $l = 0, \ldots, n$ we have

$$\begin{aligned} 0 \leq \hat{f}(\hat{x}_l) - f(\hat{x}_l) &\leq \sup\Big\{\textstyle\sum_{k=1}^m q_k\big(\hat{f}_k(x_k) - f_k(x_k)\big) : q_k \in [0,1], x_k \in [b_k, a_k]\,\forall k, \\ &\qquad \textstyle\sum_{k=1}^m q_k = 1, \sum_{k=1}^m q_k x_k = \hat{x}_l\Big\} \\ &\leq \Delta \sup\{\textstyle\sum_{k=1}^m q_k c_k : q_k \in [0,1]\,\forall k, \sum_{k=1}^m q_k = 1\} \\ &= \Delta \max_k c_k. \end{aligned} \tag{A.13}$$

The function f is Lipschitz on $[b, a]$ (Rockafellar 1996 Theorem 10.4), and so there exists some $d \geq 0$ such that

$$|f(x) - f(y)| \leq d|x - y| \text{ for all } x, y \in [a, b]. \tag{A.14}$$

For any $x \in [b, a]$ such that $\hat{x}_{l-1} < x < \hat{x}_l$ for some $l > 0$, choose $l^* \in \{l-1, l\}$ such that $\hat{f}(\hat{x}_{l^*}) = \max\left\{\hat{f}(\hat{x}_{l-1}), \hat{f}(\hat{x}_l)\right\}$. Then

$$|\hat{f}(x) - f(x)| \leq |\hat{f}(\hat{x}_{l^*}) - f(x)| \leq |\hat{f}(\hat{x}_{l^*}) - f(\hat{x}_{l^*})| + |f(\hat{x}_{l^*}) - f(x)|$$

by (A.12) and the triangle inequality. Combining this with (A.13) and (A.14) then gives the desired result after taking $c := d + \max_k c_k$. □

The upper approximation $\hat{f}$ depends on $\hat{g}$ only via the values $\hat{g}(\hat{x}_0), \ldots, \hat{g}(\hat{x}_n)$. It is possible to calculate these values explicitly in the case where $g_k(q) = q \ln \frac{q}{p_k}$ by using standard techniques from calculus (Xu 2018 Section 4.3).

The theoretical error bound in Proposition A.5 ensures that the upper approximation $\hat{f}$ will converge uniformly to f on $[b, a]$ if the mesh size converges to zero. However, it relies on the Lipschitz coefficient of f, which is typically unknown in situations that require approximation (and could well be large). We now present a *lower approximation*, which, while slightly less computationally efficient than the upper approximation, can be used in practical applications to estimate the error of the upper approximation.

For each k, let $\check{f}_k$ be any convex piecewise linear function with $\operatorname{dom} \check{f}_k = [b_k, a_k]$ and such that $\check{f}_k \leq f_k$. Then let $\check{g}$ be the generalised convex hull of $\check{f}_1, \ldots, \check{f}_m$ and $g_1, \ldots g_m$, in other words,

$$\begin{aligned}\check{g}(x) := \inf\left\{\textstyle\sum_{k=1}^m (q_k \check{f}_k(x_k) + g_k(q_k)) : q_k \in [0,1], x_k \in [b_k, a_k]\,\forall k,\right.\\ \left.\textstyle\sum_{k=1}^m q_k = 1, \sum_{k=1}^m q_k x_k = x\right\}.\end{aligned} \tag{A.15}$$

Then $\check{g}$ is clearly convex and continuous on $\operatorname{dom} \check{g} = \operatorname{dom} f$, and the infimum in (A.15) is attained for all $x \in \operatorname{dom} \check{g}$. Furthermore, $\check{g} \leq f \leq \hat{g}$.

If $b = a$, then define

$$\check{f}(x) := \begin{cases} \check{g}(x) & \text{if } x = a, \\ \infty & \text{otherwise;} \end{cases}$$

then clearly $\check{f}(a) \leq f(a) \leq \hat{f}(a)$. Assume for the remainder that $b < a$; this implies that $[b, a] \subset \operatorname{int} \operatorname{dom} f$. Similar to the upper approximation, divide $[b, a]$ into $n - 1$ subintervals by choosing $(\check{x}_l)_{l=1}^n$ such that $b =: \check{x}_1 < \cdots < \check{x}_n := a$. Also choose any $\check{x}_0 \in (\min \operatorname{dom} f, b)$ and $\check{x}_{n+1} \in (\max \operatorname{dom} f, a)$, and consider the function $\check{f}$ defined by

$$\check{f}(x) := \begin{cases} \check{g}(\check{x}_l) & \text{if } x = \check{x}_l \text{ for some } l, \\ \frac{\check{x}_l - x}{\check{x}_l - \check{x}_{l-1}} \check{g}(\check{x}_{l-1}) + \frac{x - \check{x}_{l-1}}{\check{x}_l - \check{x}_{l-1}} \check{g}(\check{x}_l) & \text{if } x \in (\check{x}_{l-1}, \check{x}_l) \text{ for any } l > 0, \\ \infty & \text{if } x \in \mathbb{R} \setminus [\check{x}_0, \check{x}_{n+1}]. \end{cases} \tag{A.16}$$

It is convex, piecewise linear and $\breve{g}(x) \leq \breve{f}(x)$ for all $x \in [\breve{x}_0, \breve{x}_{n+1}]$. The graph of $\breve{f}$ consists of $n+1$ line pieces; the l^{th} line piece (where $l = 0, \ldots, n$) connects the points $(\breve{x}_l, \breve{g}(\breve{x}_l))$ and $(\breve{x}_{l+1}, \breve{g}(\breve{x}_{l+1}))$, and has slope $m_l := \frac{\breve{g}(\breve{x}_{l+1}) - \breve{g}(\breve{x}_l)}{\breve{x}_{l+1} - \breve{x}_l}$. These line pieces are now used to determine the lower approximation $\check{f}$ on $[a, b]$. For $l = 1, \ldots, n-1$, determine the point $(\check{x}_l, \check{y}_l)$ by extending the $(l-1)^{\text{th}}$ and $(l+1)^{\text{th}}$ line pieces and finding their intersection, in other words,

$$\check{x}_l := \begin{cases} \frac{m_{l+1}\breve{x}_{l+1} - m_{l-1}\breve{x}_l + \breve{g}(\breve{x}_l) - \breve{g}(\breve{x}_{l+1})}{m_{l+1} - m_{l-1}} & \text{if } m_{l-1} < m_{l+1}, \\ \frac{1}{2}(\breve{x}_l + \breve{x}_{l+1}) & \text{if } m_{l-1} = m_{l+1}, \end{cases}$$

$$\check{y}_l := m_{l-1}(\check{x}_l - \breve{x}_l) + \breve{g}(\breve{x}_l).$$

Finally define $\check{x}_0 := \breve{x}_1 = b$, $\check{y}_0 := \breve{g}(b)$, $\check{x}_n := \breve{x}_n = a$ and $\check{y}_n := \breve{g}(a)$, after which the lower approximation is defined as

$$\check{f}(x) := \begin{cases} \check{y}_l & \text{if } x = \check{x}_l \text{ for some } l, \\ \frac{\check{x}_l - x}{\check{x}_l - \check{x}_{l-1}}\check{y}_{l-1} + \frac{x - \check{x}_{l-1}}{\check{x}_l - \check{x}_{l-1}}\check{y}_l & \text{if } x \in (\check{x}_{l-1}, \check{x}_l) \text{ for any } l > 0, \\ \infty & \text{if } x \in \mathbb{R} \setminus [b, a]. \end{cases} \tag{A.17}$$

The lower approximation $\check{f}$ is piecewise linear. It is also convex due to the convexity of $\breve{f}$. The fact that $\check{f} \leq \breve{g}$ (whence $\check{f} \leq f$) follows from a simple geometric observation: on every interval $[\breve{x}_l, \breve{x}_{l+1}]$, the graph of $\check{f}$ falls below the extensions of both the $(l-1)^{\text{th}}$ and $(l+1)^{\text{th}}$ line pieces of $\breve{f}$, and these extended line pieces in turn fall below the graph of $\breve{g}$, due to the convexity of $\breve{g}$. Xu (2018 Section 5.4) provides full details.

B. Proofs

Proof of Proposition 2.3. A trading strategy $y \in \mathcal{N}^{2\prime}$ superhedges c if and only if $y_T = 0$ and the trading strategy $w \in \mathcal{N}^{2\prime}$ defined as $w_{-1} := y_{-1}$ and $w_t := y_t + \sum_{s=0}^{t} c_s$ for all $t \geq 0$ satisfies $-\Delta w_t \in \mathcal{K}_t$ for all t. The result then follows from Theorem 4.4 of Roux and Zastawniak (2016) and (2.7). □

Proof of Theorem 3.1. The results of Pennanen and Perkkiö (2018) applies to non-adapted claims; however, in view of their Theorem 9.4 the results are applied directly to adapted claims without further comment. Furthermore, as the cash-settled claim in the paper of Pennanen and Perkkiö (2018) is redundant in the setting of this paper (where cash is perfectly liquid), and so the dimensionality of the dual space is reduced in the exposition below, again without comment.

We first obtain the conjugate of V with respect to the bilinear form $(u, y) \mapsto \sum_{t=0}^{T} \mathbb{E}[u_t \cdot y_t]$. Define for all t the positive polar of the solvency cone $\mathcal{K}_t$ as

$$\mathcal{K}_t^+ := \left\{ y \in \mathcal{L}_t^2 : y \cdot x \geq 0 \text{ for all } x \in \mathcal{K}_t \right\}.$$

Define furthermore the collection of consistent pricing processes as

$$\begin{aligned}\bar{\mathcal{C}} &:= \left\{z \in \mathcal{N}^2 : z \text{ a martingale}, z_t \in \mathcal{K}_t^+ \ \forall t\right\} \\ &= \left\{\left(\lambda(1, S_t)\Lambda_t^{\mathbb{Q}}\right)_{t=0}^T : \lambda \geq 0, (\mathbb{Q}, S) \in \bar{\mathcal{P}}\right\}.\end{aligned} \tag{B.1}$$

Then

$$V^*(z) = \sum_{t=0}^T \mathbb{E}\left[v_t^*(z_t^b)\right] = \sum_{t\in\mathcal{I}} \tfrac{1}{\alpha_t}\mathbb{E}\left[z_t^b\left(\ln z_t^b - \ln \alpha_t - 1\right)\right] + |\mathcal{I}|$$

if $z = (z^b, z^s) \in \bar{\mathcal{C}}$, otherwise $V^*(z) = \infty$ (Pennanen and Perkkiö 2018 Th. 9.1).

Robust no-arbitrage and the fact that v_t has a lower bound for all t gives that the infima in (3.2) and (3.4) are attained for all $u \in \mathcal{N}^2$ (Pennanen and Perkkiö 2018 Th. 9.2). Hence

$$\begin{aligned}V(u) &= \sup_{z\in\mathcal{N}^2}\left(\sum_{t=0}^T \mathbb{E}[u_t \cdot z_t] - V^*(z)\right) \\ &= \sup_{z\in\bar{\mathcal{C}}}\left(\sum_{t=0}^T \mathbb{E}[u_t \cdot z_t] - \sum_{t\in\mathcal{I}}\tfrac{1}{\alpha_t}\mathbb{E}\left[z_t^b\left(\ln z_t^b - \ln\alpha_t - 1\right)\right]\right) - |\mathcal{I}| \\ &= \sup_{\lambda>0}\sup_{(\mathbb{Q},S)\in\bar{\mathcal{P}}} \lambda\left(\sum_{t=0}^T \mathbb{E}_{\mathbb{Q}}[u_t^b + u_t^s S_t] - \sum_{t\in\mathcal{I}}\tfrac{1}{\alpha_t}\mathbb{E}_{\mathbb{Q}}\left[\ln\Lambda_t^{\mathbb{Q}} + \ln\tfrac{\lambda}{\alpha_t} - 1\right]\right) - |\mathcal{I}|,\end{aligned} \tag{B.2}$$

which is (3.10). Note that there is a typo (a missing minus sign) in the dual problem on p. 759 of the paper of Pennanen and Perkkiö (2018); the result above should be compared with Example 6.2 in the same paper. Introducing the notation (3.8)–(3.9) leads to

$$V(u) = -\inf_{\lambda>0}\lambda\left(K\left(\textstyle\sum_{t=0}^T u_t\right) + \sum_{t\in\mathcal{I}}\tfrac{1}{\alpha_t}\ln\tfrac{\lambda}{\alpha_t} - a_0\right) - |\mathcal{I}|. \tag{B.3}$$

The unique minimum is attained at the value $\hat{\lambda}_u$ given in (3.12). Substituting (3.12) into (B.3) leads to the formula (3.11). □

Proof of Proposition 4.1. Observe from (2.6) that

$$\textstyle\sum_{\nu\in\mu^+} q_t^\nu \ln\Lambda_t^{\mathbb{Q}\nu} = \ln\Lambda_{t-1}^{\mathbb{Q}\mu} + \sum_{\nu\in\mu^+} q_t^\nu \ln\frac{q_t^\nu}{p_t^\nu} \text{ for all } t>0, \mu\in\Omega_{t-1}^{\mathbb{Q}}, \nu\in\mu^+.$$

Using the nodes in Ω_{t-1} to partition Ω, and noting that $\mathbb{Q}$ and $\Lambda_t^{\mathbb{Q}}$ are nonzero only on the nodes in $\Omega_{t-1}^{\mathbb{Q}}$, leads to

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\ln\Lambda_t^{\mathbb{Q}}\right] &= \textstyle\sum_{\mu\in\Omega_{t-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+} q_t^\nu \ln\Lambda_t^{\mathbb{Q}\nu} \\ &= \textstyle\sum_{\mu\in\Omega_{t-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\ln\Lambda_{t-1}^{\mathbb{Q}\mu} + \sum_{\mu\in\Omega_{t-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+} q_t^\nu\ln\frac{q_t^\nu}{p_t^\nu} \\ &= \mathbb{E}_{\mathbb{Q}}\left[\ln\Lambda_{t-1}^{\mathbb{Q}}\right] + \textstyle\sum_{\mu\in\Omega_{t-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+} q_t^\nu\ln\frac{q_t^\nu}{p_t^\nu}.\end{aligned}$$

Observing that $\mathbb{E}_{\mathbb{Q}}\left[\ln\Lambda_0^{\mathbb{Q}}\right] = 0$, and introducing a telescoping sum, leads to

$$\mathbb{E}_{\mathbb{Q}}\left[\ln\Lambda_t^{\mathbb{Q}}\right] = \textstyle\sum_{k=1}^t\sum_{\mu\in\Omega_{k-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+} q_k^\nu\ln\frac{q_k^\nu}{p_k^\nu}.$$

Then, after collecting like terms, it follows that

$$\begin{aligned}\textstyle\sum_{t\in\mathcal{I}}\frac{1}{\alpha_t}\mathbb{E}_{\mathbb{Q}}\big[\ln\Lambda_t^{\mathbb{Q}}\big] &= \textstyle\sum_{t\in\mathcal{I}\setminus\{0\}}\frac{1}{\alpha_t}\mathbb{E}_{\mathbb{Q}}\big[\ln\Lambda_t^{\mathbb{Q}}\big]\\ &= \textstyle\sum_{t=0}^{T-1}a_{t+1}\sum_{\mu\in\Omega_t^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{t+1}^{\nu}\ln\frac{q_{t+1}^{\nu}}{p_{t+1}^{\nu}}.\end{aligned}$$

The result follows from (3.8) after using the nodes in Ω_{T-1} to partition Ω and observing that

$$\mathbb{E}_{\mathbb{Q}}\big[X^b+X^sS_T\big]=\textstyle\sum_{\mu\in\Omega_{T-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_T^{\nu}\big(X^{b\nu}+X^{s\nu}S_T^{\nu}\big).\qquad\square$$

Proof of Proposition 4.2. The properties of the J_t's are proved by backward induction. The convexity, continuity and boundedness properties of J_T^{ν} is self-evident from (4.3). For every $t<T$, suppose that J_t^{ν} is convex, bounded from below and continuous on its effective domain $\operatorname{dom}J_t^{\nu}\subseteq[S_t^{b\nu},S_t^{a\nu}]$ for all $\nu\in\Omega_{t+1}$. Define

$$g^{\nu}(q):=\begin{cases}a_{t+1}q\ln\frac{q}{p_{t+1}^{\nu}} & \text{if } q\in[0,1],\\ \infty & \text{otherwise}\end{cases}$$

for all $\nu\in\Omega_{t+1}$; then g^{ν} is convex, bounded from below and continuous on its effective domain $\operatorname{dom}g^{\nu}=[0,1]$. Propositions A.1 and A.4 then give that f_t^{μ} is convex, bounded from below and continuous on its effective domain for every $\mu\in\Omega_t$, and that the infimum in (4.4) is attained for all $x\in\operatorname{dom}f_t^{\mu}$. It is then clear from (4.5) that J_t^{μ} has the properties claimed. This concludes the inductive step.

To establish (4.6), fix any $(\mathbb{Q},S)\in\bar{\mathcal{P}}$. We show first by backward induction that

$$\begin{aligned}\inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}}_{t+1}(\mathbb{Q},S)}H((\bar{\mathbb{Q}},\bar{S}),X) &= \textstyle\sum_{k=0}^{t}a_{k+1}\sum_{\mu\in\Omega_k^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{k+1}^{\nu}\ln\frac{q_{k+1}^{\nu}}{p_{k+1}^{\nu}}\\ &\quad+\textstyle\sum_{\mu\in\Omega_t^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{t+1}^{\nu}J_{t+1}^{\nu}(S_{t+1}^{\nu})\end{aligned}\qquad\text{(B.4)}$$

for all $t<T$, where

$$\bar{\mathcal{P}}_t(\mathbb{Q},S):=\{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}}:\bar{\mathbb{Q}}=\mathbb{Q}\text{ on }\mathcal{F}_t,\bar{S}_k=S_k\ \forall k\le t\}\qquad\text{(B.5)}$$

is the collection of martingale pairs that coincide with $(\mathbb{Q},S)$ up to time t. When $t=T-1$, we have $\bar{\mathcal{P}}_T(\mathbb{Q},S)=\{(\mathbb{Q},S)\}$, so that (B.4) follows from (4.2) and (4.3). Assume now that (B.4) holds for some $t=1,\dots,T-1$. Rearrangement gives

$$\begin{aligned}\inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}}_{t+1}(\mathbb{Q},S)}H((\bar{\mathbb{Q}},\bar{S}),X) &= \textstyle\sum_{k=0}^{t-1}a_{k+1}\sum_{\mu\in\Omega_k^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{k+1}^{\nu}\ln\frac{q_{k+1}^{\nu}}{p_{k+1}^{\nu}}\\ &\quad+\textstyle\sum_{\mu\in\Omega_t^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{t+1}^{\nu}\left(a_{t+1}\ln\frac{q_{t+1}^{\nu}}{p_{t+1}^{\nu}}+J_{t+1}^{\nu}(S_{t+1}^{\nu})\right),\end{aligned}$$

after which we obtain from (2.2), (B.5) and (4.5) that

$$\begin{aligned}&\inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}}_{t}(\mathbb{Q},S)}H((\bar{\mathbb{Q}},\bar{S}),X)\\ &= \textstyle\sum_{k=0}^{t-1}a_{k+1}\sum_{\mu\in\Omega_k^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{k+1}^{\nu}\ln\frac{q_{k+1}^{\nu}}{p_{k+1}^{\nu}}+\sum_{\mu\in\Omega_t^{\mathbb{Q}}}\mathbb{Q}(\mu)J_t^{\mu}(S_t^{\mu})\\ &= \textstyle\sum_{k=0}^{t-1}a_{k+1}\sum_{\mu\in\Omega_k^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_{k+1}^{\nu}\ln\frac{q_{k+1}^{\nu}}{p_{k+1}^{\nu}}+\sum_{\mu\in\Omega_{t-1}^{\mathbb{Q}}}\mathbb{Q}(\mu)\sum_{\nu\in\mu^+}q_t^{\nu}J_t^{\mu}(S_t^{\mu}).\end{aligned}$$

This concludes the inductive step.

Finally, when $t=0$, the equation (B.4) reduces to

$$\inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}}_1(\mathbb{Q},S)} H((\bar{\mathbb{Q}},\bar{S}),X) = a_1\sum_{\nu\in\Omega_1} q_1^\nu \ln\frac{q_1^\nu}{p_1^\nu} + \sum_{\nu\in\Omega_1} q_1^\nu J_1^\nu(S_1^\nu),$$

and again combining (2.2), (B.5) and (4.5) yields

$$\inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}},\bar{S}_0=S_0} H((\bar{\mathbb{Q}},\bar{S}),X) = \inf_{(\bar{\mathbb{Q}},\bar{S})\in\bar{\mathcal{P}}_0(\mathbb{Q},S)} H((\bar{\mathbb{Q}},\bar{S}),X) = J_0(S_0).$$

This completes the proof. □

Proof of Theorem 4.1. Standard arguments (Cutland and Roux 2012 Theorem 5.25) can be used to show that $\hat{\mathbb{Q}}$ is a probability measure. The process $\hat{S}$ is a martingale under $\hat{\mathbb{Q}}$ by (4.10), whence $(\hat{\mathbb{Q}},\hat{S})\in\bar{\mathcal{P}}$. Furthermore, recursive expansion of (4.9) gives

$$\begin{aligned} J_0(\hat{S}_0) = \sum_{t=0}^{T-1} a_{t+1}\sum_{\mu\in\Omega_t^{\hat{\mathbb{Q}}}}\hat{\mathbb{Q}}(\mu)\sum_{\nu\in\mu^+}\hat{q}_{t+1}^\nu \ln\frac{\hat{q}_{t+1}^\nu}{p_{t+1}^\nu} \\ + \sum_{\mu\in\Omega_{T-1}^{\hat{\mathbb{Q}}}}\hat{\mathbb{Q}}(\mu)\sum_{\nu\in\mu^+}\hat{q}_T^\nu J_T^\nu(\hat{S}_T^\nu) = H((\hat{\mathbb{Q}},\hat{S}),X) \end{aligned}$$

from (3.8) and (4.3). Then (4.8), Proposition 4.2 and (3.9) combine to give

$$J_0(\hat{S}_0) = \min_{(\mathbb{Q},S)\in\bar{\mathcal{P}}} H((\mathbb{Q},S),X) = K(X).$$

We now show that $(\hat{\mathbb{Q}},\hat{S})\in\mathcal{P}$. Suppose by contradiction that $(\hat{\mathbb{Q}},\hat{S})\in\bar{\mathcal{P}}\backslash\mathcal{P}$, in other words, $\Lambda_t^{\hat{\mathbb{Q}}}(\omega)=0$ for some $t=0,\dots,T$ and $\omega\in\Omega$. Fix any $(\mathbb{Q},S)\in\mathcal{P}$, and define

$$\epsilon := \tfrac{1}{2}\exp\left\{\big(H((\hat{\mathbb{Q}},\hat{S}),X) - H((\mathbb{Q},S),X)\big)\Big/\sum_{t\in\mathcal{I}}\tfrac{1}{\alpha_t}\mathbb{Q}\big(\Lambda_t^{\hat{\mathbb{Q}}}=0\big)\right\}.$$

Observe that $\epsilon\in[0,1)$ because $H((\hat{\mathbb{Q}},\hat{S}),X) = J_0(\hat{S}_0) \le J_0(S_0) \le H((\mathbb{Q},S),X)$. Define a new probability measure $\bar{\mathbb{Q}}:\mathcal{F}\to[0,1]$ and stochastic process $\bar{S}\in\mathcal{N}$ as

$$\bar{\mathbb{Q}} := \epsilon\mathbb{Q} + (1-\epsilon)\hat{\mathbb{Q}}, \tag{B.6}$$

$$\bar{S}_t := \epsilon S_t\mathbb{E}\left[\frac{d\mathbb{Q}}{d\bar{\mathbb{Q}}}\middle|\mathcal{F}_t\right] + (1-\epsilon)\hat{S}_t\mathbb{E}\left[\frac{d\hat{\mathbb{Q}}}{d\bar{\mathbb{Q}}}\middle|\mathcal{F}_t\right] \text{ for all } t. \tag{B.7}$$

Then $(\bar{\mathbb{Q}},\bar{S})\in\mathcal{P}$ (Roux et al. 2008 Lemma 7.2), after which (3.8) gives

$$\begin{aligned} H((\bar{\mathbb{Q}},\bar{S}),X) - H((\hat{\mathbb{Q}},\hat{S}),X) = \sum_{t\in\mathcal{I}}\tfrac{1}{\alpha_t}\mathbb{E}\big[\Lambda_t^{\bar{\mathbb{Q}}}\ln\Lambda_t^{\bar{\mathbb{Q}}} - \Lambda_t^{\hat{\mathbb{Q}}}\ln\Lambda_t^{\hat{\mathbb{Q}}}\big] \\ + \epsilon\big(\mathbb{E}_{\hat{\mathbb{Q}}}\big[X^b + X^s\hat{S}_T\big] - \mathbb{E}_{\mathbb{Q}}\big[X^b + X^sS_T\big]\big). \end{aligned} \tag{B.8}$$

The mapping $x\mapsto x\ln x$ is convex on $[0,\infty)$, and so

$$\Lambda_t^{\bar{\mathbb{Q}}}\ln\Lambda_t^{\bar{\mathbb{Q}}} - \Lambda_t^{\hat{\mathbb{Q}}}\ln\Lambda_t^{\hat{\mathbb{Q}}} \le \epsilon\big(\Lambda_t^{\mathbb{Q}}\ln\Lambda_t^{\mathbb{Q}} - \Lambda_t^{\hat{\mathbb{Q}}}\ln\Lambda_t^{\hat{\mathbb{Q}}}\big) \text{ for all } t. \tag{B.9}$$

Furthermore, on the set $\big\{\Lambda_t^{\hat{\mathbb{Q}}}=0\big\}$, and recalling the convention $0\ln 0=0$, we have

$$\Lambda_t^{\bar{\mathbb{Q}}}\ln\Lambda_t^{\bar{\mathbb{Q}}} - \Lambda_t^{\hat{\mathbb{Q}}}\ln\Lambda_t^{\hat{\mathbb{Q}}} = \epsilon\Lambda_t^{\mathbb{Q}}\ln\epsilon\Lambda_t^{\mathbb{Q}} = \epsilon\big(\Lambda_t^{\mathbb{Q}}\ln\Lambda_t^{\mathbb{Q}} - \Lambda_t^{\hat{\mathbb{Q}}}\ln\Lambda_t^{\hat{\mathbb{Q}}}\big) + \epsilon\Lambda_t^{\mathbb{Q}}\ln\epsilon.$$

Substituting this into (B.8) gives

$$H((\bar{\mathbb{Q}},\bar{S}),X) - H((\hat{\mathbb{Q}},\hat{S}),X) \\ \leq \epsilon\left(H((\mathbb{Q},S),X) - H((\hat{\mathbb{Q}},\hat{S}),X) + \ln\epsilon \textstyle\sum_{t\in\mathcal{I}}\frac{1}{\alpha_t}\mathbb{Q}\big(\Lambda_t^{\hat{\mathbb{Q}}}=0\big)\right).$$

The choice of ϵ implies that $H((\bar{\mathbb{Q}},\bar{S}),X) < H((\hat{\mathbb{Q}},\hat{S}),X)$, which is a contradiction. Hence $\hat{\mathbb{Q}}(\omega) > 0$ for all $\omega\in\Omega$, so that $(\hat{\mathbb{Q}},\hat{S})\in\mathcal{P}$.

The proof is complete upon establishing the uniqueness of $\hat{\mathbb{Q}}$ on the nodes in $\mathcal{I}$. To this end, suppose by contradiction that there exists another pair $(\mathbb{Q},S)\in\mathcal{P}$ such that $H((\hat{\mathbb{Q}},\hat{S}),X) = H((\mathbb{Q},S),X)$ and $\hat{\mathbb{Q}}(\nu')\neq\mathbb{Q}(\nu')$ for some $t'\in\mathcal{I}$ and $\nu'\in\Omega_{t'}$. The argument now proceeds along similar lines as above: take any $\epsilon\in(0,1)$, and use (B.6)–(B.7) to define a new pair $(\bar{\mathbb{Q}},\bar{S})\in\mathcal{P}$. This immediately leads to (B.8) and (B.9), noting in (B.9) that $\Lambda_t^{\hat{\mathbb{Q}}}(\nu')\neq\Lambda_t^{\mathbb{Q}}(\nu')$ gives

$$\Lambda_{t'}^{\bar{\mathbb{Q}}}\ln\Lambda_{t'}^{\bar{\mathbb{Q}}} - \Lambda_{t'}^{\hat{\mathbb{Q}}}\ln\Lambda_{t'}^{\hat{\mathbb{Q}}} < \epsilon\big(\Lambda_{t'}^{\mathbb{Q}}\ln\Lambda_{t'}^{\mathbb{Q}} - \Lambda_{t'}^{\hat{\mathbb{Q}}}\ln\Lambda_{t'}^{\hat{\mathbb{Q}}}\big) \text{ on } \nu'.$$

Substituting into (3.8), it follows that

$$\begin{aligned} &H((\bar{\mathbb{Q}},\bar{S}),X) - H((\hat{\mathbb{Q}},\hat{S}),X) \\ &< \epsilon\textstyle\sum_{t\in\mathcal{I}}\frac{1}{\alpha_t}\mathbb{E}\big[\Lambda_t^{\mathbb{Q}}\ln\Lambda_t^{\mathbb{Q}} - \Lambda_t^{\hat{\mathbb{Q}}}\ln\Lambda_t^{\hat{\mathbb{Q}}}\big] + \epsilon\left(\mathbb{E}_{\mathbb{Q}}\big[X^b + X^s S_T\big] - \mathbb{E}_{\hat{\mathbb{Q}}}\big[X^b + X^s\hat{S}_T\big]\right) \\ &= \epsilon(H((\mathbb{Q},S),X) - H((\hat{\mathbb{Q}},\hat{S}),X)) = 0, \end{aligned}$$

in other words, $H((\bar{\mathbb{Q}},\bar{S}),X) < H((\hat{\mathbb{Q}},\hat{S}),X)$. This contradicts the assumption that $(\hat{\mathbb{Q}},\hat{S})$ is a solution to the optimization problem (3.9). □

Proof of Proposition 5.1. We continue with the notation and conventions of the proof of Theorem 3.1. By Theorem 9.3 of Pennanen and Perkkiö (2018), a pair $\hat{y}\in\Psi$ and $\hat{z}\in\bar{\mathcal{C}}$ attains the infimum and supremum in (3.2) and (B.2), respectively, if and only if

$$(\hat{z}_0^b,\ldots,\hat{z}_T^b)\in\partial v(\phi_0(\Delta\hat{y}_0+u_0),\ldots,\phi_T(\Delta\hat{y}_0+u_T)), \tag{B.10}$$

$$\hat{z}_t\in\hat{z}_t^b\partial\phi_t(\Delta\hat{y}_t+u_t) \text{ for all } t, \tag{B.11}$$

where

$$v(x_0,\ldots,x_T) = \sum_{t=0}^{T} v_t(x_t) \text{ for all } (x_0,\ldots,x_T)\in\mathbb{R}^{T+1}$$

and the subdifferential ∂f of any function $f:\mathbb{R}^d\to\mathbb{R}\cup\{\infty\}$ is defined as

$$\partial f(x) = \big\{y\in\mathbb{R}^d : f(z)\geq f(x) + y\cdot(z-x)\forall z\in\mathbb{R}^d\big\} \text{ for all } x\in\mathbb{R}^d.$$

Standard results from convex analysis (Rockafellar 1996 Ths. 23.8, 25.1, 25.6) give that (B.10)–(B.11) is equivalent to the following:

(1) For all $t\in\mathcal{I}$ we have $\hat{z}_t^b = v_t'(\phi_t(\Delta\hat{y}_t+u_t)) = \alpha_t e^{\alpha_t\phi_t(\Delta\hat{y}_t+u_t)}$.
(2) For all $t\notin\mathcal{I}$ we need $\phi_t(\Delta\hat{y}_t+u_t)\leq 0$ in order to have $\partial\phi_t(\Delta\hat{y}_t+u_t)\neq\emptyset$, and moreover, $\{\phi_t(\Delta\hat{y}_t+u_t)<0\}\subseteq\big\{\hat{z}_t^b=0\big\}$.

(3) For all t,

$$\left\{\Delta\hat{y}^s_t + u^s_t > 0\right\} \subseteq \left\{\hat{z}^s_t = \hat{z}^b_t S^a_t\right\} \text{ and } \left\{\Delta\hat{y}^s_t + u^s_t < 0\right\} \subseteq \left\{\hat{z}^s_t = \hat{z}^b_t S^b_t\right\}.$$

The proof is complete upon observing that $(\hat{\mathbb{Q}}, \hat{S}) \in \bar{\mathcal{P}}$ satisfies (5.1) if and only if $\hat{z} = (\hat{\lambda}_u(1, \hat{S}_t)\Lambda^{\hat{\mathbb{Q}}}_t)^T_{t=0} \in \bar{\mathcal{C}}$ attains the supremum in (B.2). □

Proof of Theorem 5.1. Theorem 4.1 gives that $\hat{\mathbb{Q}} \sim \mathbb{P}$ and that $(\hat{\mathbb{Q}}, \hat{S})$ satisfies (5.1), and therefore by Theorem 3.1 and Proposition 5.1 there is a trading strategy $\hat{y} \in \Psi$ attaining the infimum in (3.2) and such that

$$\phi_t(\Delta\hat{y}_t + u_t) = \hat{x}_t \text{ for all } t.$$

Thus $\hat{x}$ attains the infimum in (3.6).

Finally, notice that the partial uniqueness property of $\hat{\mathbb{Q}}$ in Theorem 4.1 ensures that the process $\hat{x}$ in (5.4) is well defined, in that it does not depend on the choice of the minimiser $(\hat{\mathbb{Q}}, \hat{S})$. For this reason it is also unique. □

Proof of Theorem 5.2. Let $(J_t)^T_{t=0}$ be the sequence of functions from Construction 4.1 with $X = \sum_{t=0}^T u_t$, and let $(\hat{\mathbb{Q}}, \hat{S})$ be the pair from Construction 4.2. Recursive expansion of (4.9) gives

$$J_t(\hat{S}_t) = \mathbb{E}_{\hat{\mathbb{Q}}}\left[-\textstyle\sum_{s=0}^T(u^b_s + u^s_s\hat{S}_T) + \sum_{s=t}^{T-1} a_{s+1}\ln\frac{\hat{q}_{s+1}}{p_{s+1}}\middle|\mathcal{F}_t\right] \text{ for all } t < T. \tag{B.12}$$

It then follows from (5.5) that

$$\textstyle\sum_{t=0}^T \hat{x}_t = \sum_{t=0}^{T-1} a_{t+1}\ln\frac{\hat{q}_{t+1}}{p_{t+1}} + \sum_{t\in\mathcal{I}}\frac{1}{\alpha_t}\ln\frac{\hat{\lambda}_u}{\alpha_t} = \sum_{t=0}^{T-1} a_{t+1}\ln\frac{\hat{q}_{t+1}}{p_{t+1}} - J_0(\hat{S}_0). \tag{B.13}$$

The first step in the proof is to show that the collection $\mathcal{W}_T$ in Construction 5.1 is non-empty. Theorem 3.1 guarantees the existence of a minimiser $\hat{y} \in \Psi$ for (3.1), and by Proposition 5.1 and Theorem 5.1 it follows that $\hat{y}$ satisfies (5.2) and

$$y_{-1} = y_T = 0, \qquad \Delta y^b_t + u^b_t + (\Delta y^s_t + u^s_t)\hat{S}_t = \hat{x}_t \text{ for all } t \geq 0. \tag{B.14}$$

The trading strategy $w \in \mathcal{N}^{2\prime}$ defined by

$$w_{-1} = 0, \qquad w_t := y_t + \textstyle\sum_{s=0}^t (u^b_s - \hat{x}_s, u^s_s) \text{ for all } t = 0, \ldots, T \tag{B.15}$$

satisfies

$$(\Delta w^s_t)_+ S^a_t - (\Delta w^s_t)_- S^b_t = \Delta w^s_t \hat{S}_t \text{ for all } t, \qquad w^s_T = \textstyle\sum_{t=0}^T u^s_t \tag{B.16}$$

by definition and by (B.13)

$$w^b_T = \textstyle\sum_{t=0}^T u^b_t - \sum_{t=0}^{T-1} a_{t+1}\ln\frac{\hat{q}_{t+1}}{p_{t+1}} + J_0(\hat{S}_0). \tag{B.17}$$

Moreover (B.14) gives the self-financing condition

$$\Delta w^b_t + \Delta w^s_t \hat{S}_t = 0 \text{ for all } t \geq 0. \tag{B.18}$$

Combining (B.18) with the fact that $\hat{S}$ is a martingale under $\hat{\mathbb{Q}}$, it follows from standard arguments (cf. Cutland and Roux 2012 Th. 5.40) that

$$w_t^b + w_t^s \hat{S}_{t+1} = \mathbb{E}_{\hat{\mathbb{Q}}}\left[w_T^b + w_T^s \hat{S}_T \middle| \mathcal{F}_{t+1}\right] \text{ for all } t < T. \tag{B.19}$$

For every $t < T$, substituting (B.16), (B.17) and (B.12) leads to

$$\begin{aligned} w_t^b + w_t^s \hat{S}_{t+1} &= \mathbb{E}_{\hat{\mathbb{Q}}}\left[\textstyle\sum_{s=0}^{T}(u_s^b + u_s^s \hat{S}_T) - \sum_{s=0}^{T-1} a_{s+1} \ln \frac{\hat{q}_{t+1}}{p_{s+1}} + J_0(\hat{S}_0) \middle| \mathcal{F}_{t+1}\right] \\ &= -J_{t+1}(\hat{S}_{t+1}) - \textstyle\sum_{s=0}^{t} a_{s+1} \ln \frac{\hat{q}_{s+1}}{p_{s+1}} + J_0(\hat{S}_0). \end{aligned}$$

After defining the stochastic process $(x_t^b)_{t=-1}^T$ as

$$z_t^b := \begin{cases} 0 & \text{if } t = -1, \\ w_0^b - J_0(\hat{S}_0) & \text{if } t = 0, \\ w_t^b + \sum_{s=0}^{t-1} a_{s+1} \ln \frac{\hat{q}_{s+1}}{p_{s+1}} - J_0(\hat{S}_0), & \text{if } t > 0, \end{cases}$$

this can be rewritten as

$$z_t^b + w_t^s \hat{S}_{t+1} = -J_{t+1}(\hat{S}_{t+1}) - a_{t+1} \ln \tfrac{\hat{q}_{t+1}}{p_{t+1}}.$$

When combined with (B.16)–(B.17), this means that $(z_t^b, w_t^s)_{t=-1}^T \in \mathcal{W}_T$ and hence $\mathcal{W}_T \neq \emptyset$.

Now let $\mathcal{W}_T$ and $\mathcal{Y}$ be the collections of processes from Construction 5.1. By Proposition 5.1 and Theorem 5.1 it suffices to show that every $\hat{y} \in \mathcal{Y}$ satisfies (5.2) and (B.14). As $\hat{y} \in \mathcal{Y}$, there exists some $w \in \mathcal{W}_T$ satisfying (5.9)–(5.10). Taking the sum over all t in (5.9)–(5.10) and substituting (B.13) gives that $\hat{y}_T = 0$. Turning to the properties of w, it satisfies (B.16) by construction, which immediately gives (5.2). Moreover,

$$w_t^b + w_t^s \hat{S}_t = -J_t(\hat{S}_t) \text{ for all } t. \tag{B.20}$$

For $t = T$ this comes from (4.3) and (5.8). For $t < T$ it is obtained by taking conditional expectation in (5.7) with respect to $\hat{\mathbb{Q}}$ and $\mathcal{F}_t$, and substituting (4.9). Combining (B.20) with (5.7) furthermore gives

$$\Delta w_t^b + \Delta w_t^s \hat{S}_t = a_t \ln \tfrac{\hat{q}_t}{p_t} \text{ for all } t > 0. \tag{B.21}$$

The equalities (B.20) for $t = 0$ (recall $w_{-1} = 0$) and (B.21) for $t > 0$ now combine with (5.9)–(5.10) to give (B.14), as required. $\square$

Proof of Theorem 6.1. Observe first that (6.4) follows directly from (6.2) and (6.3). For any $\delta \in \mathbb{R}$ we have

$$K\left(\textstyle\sum_{t=0}^T (c_t - w_t) - (\delta, 0)\right) = \delta + K\left(\textstyle\sum_{t=0}^T (c_t - w_t)\right),$$

whence $\hat{\lambda}_{c-\delta\mathbb{1}-w} = e^{-\delta/a_0} \hat{\lambda}_{c-w}$ by (3.12), so that $\delta \mapsto V(c - \delta\mathbb{1} - w)$ is strictly decreasing and continuous. Therefore $\pi^{ai}(c; w)$ is the unique solution $\pi^{ai}(c; w)$ of the equation $V(c - \pi^{ai}(c; w)\mathbb{1} - w) = V(w)$, which by (3.11) is (6.3). $\square$

Proof of Theorem 6.2. Note from (2.8) and (2.11) that $c-\pi^a(c)\mathbb{1}\in\mathcal{Z}$. Furthermore, for any $x\in\mathcal{A}_{-w}$, we have $-w-(x_t,0)_{t=0}^T\in\mathcal{Z}$, and since $\mathcal{Z}$ is a convex cone, it follows that $c-\pi^a(c)\mathbb{1}-w-(x_t,0)_{t=0}^T\in\mathcal{Z}$, and finally $x\in\mathcal{A}_{c-\pi^a(c)\mathbb{1}-w}$. Thus $\mathcal{A}_{-w}\subseteq\mathcal{A}_{c-\pi^a(c)\mathbb{1}-w}$, so $V(c-\pi^a(c)\mathbb{1}-w)\leq V(-w)$ by (3.4). With (6.1) this gives $\pi^{ai}(c;w)\leq\pi^a(c)$. In combination with (6.2) and (2.7), this immediately leads to $\pi^{bi}(c;w)=-\pi^{ai}(-c;w)\geq-\pi^a(-c)=\pi^b(c)$.

The rest of the proof is devoted to showing the convexity of $u\mapsto\pi^{ai}(u;w)$. Once established, it immediately gives that $u\mapsto\pi^{bi}(u;w)$ is concave by (6.2). Moreover, combining the convexity with (6.3) gives for all $c,w\in\mathcal{N}^2$ that

$$0=\pi^{ai}(0;w)\leq\tfrac{1}{2}\pi^{ai}(c;w)+\tfrac{1}{2}\pi^{ai}(-c;w),$$

whence $\pi^{bi}(c;w)=-\pi^{ai}(-c;w)\leq\pi^{ai}(c;w)$.

To establish the convexity, note that $C\coloneqq\{x\in\mathcal{N}^2: V(x-w)\leq V(-w)\}$ is convex because, for all $x,y\in C$ and $\lambda\in[0,1]$ we have

$$V(\lambda x+(1-\lambda)y-w)\leq\lambda V(x-w)+(1-\lambda)V(y-w)\leq V(-w)$$

as V is convex (due to the convexity of the v_t's). For any $c,d\in\mathcal{N}^2$ and $\lambda\in[0,1]$ we have

$$\begin{aligned}\lambda\pi^{ai}(c;w)+(1-\lambda)\pi^{ai}(d;w)&=\lambda\inf\{\gamma: c-\gamma\mathbb{1}\in C\}+(1-\lambda)\inf\{\delta: d-\delta\mathbb{1}\in C\}\\&=\inf\{\lambda\gamma+(1-\lambda)\delta: c-\gamma\mathbb{1}\in C, d-\delta\mathbb{1}\in C\}.\end{aligned}$$

By the convexity of C, the conditions $c-\gamma\mathbb{1}\in C, d-\delta\mathbb{1}\in C$ imply that

$$\lambda c+(1-\lambda)d-(\lambda\gamma+(1-\lambda)\delta)\mathbb{1}=\lambda(c-\gamma\mathbb{1})+(1-\lambda)(d-\delta\mathbb{1})\in C,$$

whence

$$\begin{aligned}\lambda\pi^{ai}(c;w)+(1-\lambda)\pi^{ai}(d;w)&\geq\inf\{\varepsilon:\lambda c+(1-\lambda)d-\varepsilon\mathbb{1}\in C\}\\&=\pi^{ai}(\lambda c+(1-\lambda)d;w).\end{aligned}$$

Thus $u\mapsto\pi^{ai}(u;w)$ is convex and the proof is complete. □

References

Bensaid, B., Lesne, J.-P., Pagès, H. and Scheinkman, J. (1992), 'Derivative asset pricing with transaction costs', *Mathematical Finance* **2**, 63–86.

Bichuch, M. (2014), 'Pricing a contingent claim liability with transaction costs using asymptotic analysis for optimal investment', *Finance and Stochastics* **18**(3), 651–694.

Cai, Y., Judd, K. L. and Xu, R. (2013), Numerical solution of dynamic portfolio optimization with transaction costs, Working paper 18709, National Bureau of Economic Research.

Castañeda-Leyva, N. and Hernández-Hernández, D. (2011), 'Utility maximization in markets with bid–ask spreads', *Stochastics: An International Journal of Probability and Stochastics Processes* **83**(01), 17–43.

Çetin, U. and Rogers, L. C. G. (2007), 'Modelling liquidity effects in discrete time', *Mathematical Finance* **17**(1), 15–29.

Cutland, N. J. and Roux, A. (2012), *Derivative Pricing in Discrete Time*, Springer Undergraduate Mathematics Series, Springer.

Cvitanić, J. and Karatzas, I. (1996), 'Hedging and portfolio optimization under transaction costs: A martingale approach', *Mathematical Finance* **6**, 133–166.

Czichowsky, C., Muhle-Karbe, J. and Schachermayer, W. (2014), 'Transaction costs, shadow prices, and duality in discrete time', *SIAM Journal on Financial Mathematics* **5**(1), 258–277.

Czichowsky, C. and Schachermayer, W. (2016), 'Duality theory for portfolio optimisation under transaction costs', *The Annals of Applied Probability* **26**(3), 1888–1941.

Czichowsky, C. and Schachermayer, W. (2017), 'Portfolio optimisation beyond semimartingales: Shadow prices and fractional Brownian motion', *The Annals of Applied Probability* **27**(3), 1414–1451.

Czichowsky, C., Schachermayer, W. and Yang, J. (2017), 'Shadow prices for continuous processes', *Mathematical Finance* **27**(3), 623–658.

Davis, M. H. A. (1997), Option pricing in incomplete markets, *in* M. A. H. Dempster and S. R. Pliska, eds, 'Mathematics of Derivative Securities', Cambridge University Press, pp. 216–226.

Davis, M. H. A., Panas, V. G. and Zariphopoulou, T. (1993), 'European option pricing with transaction costs', *SIAM Journal on Control and Optimization* **31**(2), 470–493.

Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M. and Stricker, C. (2002), 'Exponential hedging and entropic penalties', *Mathematical Finance* **12**(2), 99–123.

Delbaen, F., Kabanov, Y. M. and Valkeila, E. (2002), 'Hedging under transaction costs in currency markets: A discrete-time model', *Mathematical Finance* **12**, 45–61.

Dempster, M. A. H., Evstigneev, I. V. and Taksar, M. I. (2006), 'Asset pricing and hedging in financial markets with transaction costs: An approach based on the von Neumann-Gale model', *Annals of Finance* **2**, 327–355.

Dorfleitner, G. and Gerer, J. (2016), 'A note on utility indifference pricing', *International Journal of Theoretical and Applied Finance* **19**(6), 1650037.

Edirisinghe, C., Naik, V. and Uppal, R. (1993), 'Optimal replication of options with transactions costs and trading restrictions', *The Journal of Financial and Quantitative Analysis* **28**(1), 117–138.

Frittelli, M. (2000*a*), 'Introduction to a theory of value coherent with the no-arbitrage principle', *Finance and Stochastics* **4**, 275–297. Referenced by janiszewski2001.

Frittelli, M. (2000*b*), 'The minimal entropy martingale measure and the valuation problem in incomplete markets', *Mathematical Finance* **10**(1), 39–52.

Gerhold, S., Guasoni, P., Muhle-Karbe, J. and Schachermayer, W. (2014), 'Transaction costs, trading volume, and the liquidity premium', *Finance and Stochastics* **18**(1), 1–37.

Gerhold, S., Muhle-Karbe, J. and Schachermayer, W. (2013), 'The dual optimizer for the growth-optimal portfolio under transaction costs', *Finance and Stochastics* **17**(2), 325–354.

Gu, L., Lin, Y. and Yang, J. (2017), 'On the existence of shadow prices for optimal investment with random endowment', *Stochastics* **89**(6-7), 1082–1103.

Guasoni, P. (2002), 'Optimal investment with transaction costs and without semimartingales', *The Annals of Applied Probability* **12**(4), 1227–1246.

Herczegh, A. and Prokaj, V. (2015), 'Shadow price in the power utility case', *The Annals of Applied Probability* **25**(5), 2671–2707.

Hodges, S. D. and Neuberger, A. (1989), 'Optimal replication of contingent claims under transaction costs', *Review of Futures Markets* **8**, 222–239.

Hugonnier, J., Kramkov, D. and Schachermayer, W. (2005), 'On utility based pricing of contingent claims in incomplete markets', *Mathematical Finance* **15**(2).

Jaschke, S. and Küchler, U. (2001), 'Coherent risk measures and good-deal bounds', *Finance and Stochastics* **5**(2), 181–200.

Jouini, E. and Kallal, H. (1995), 'Martingales and arbitrage in securities markets with transaction costs', *Journal of Economic Theory* **66**, 178–197.

Kabanov, Y. M. and Stricker, C. (2001), 'The Harrison-Pliska arbitrage pricing theorem under transaction costs', *Journal of Mathematical Economics* **35**, 185–196.

Kallsen, J. and Muhle-Karbe, J. (2010), 'On using shadow prices in portfolio optimization with transaction costs', *The Annals of Applied Probability* **20**(4), 1341–1358.

Kallsen, J. and Muhle-Karbe, J. (2011), 'Existence of shadow prices in finite probability spaces', *Math Meth Oper Res* **73**(2), 251–262.

Kallsen, J. and Muhle-Karbe, J. (2015), 'Option pricing and hedging with small transaction costs', *Mathematical Finance* **25**(4), 702–723. Early view.

Lin, Y. and Yang, J. (2016), 'Utility maximization problem with random endowment and transaction costs: when wealth may become negative', *Stochastic Analysis and Applications* **35**(2), 257–278.

Löhne, A. and Rudloff, B. (2014), 'An algorithm for calculating the set of superhedging portfolios in markets with transaction costs', *International Journal of Theoretical and Applied Finance* **17**(2). Article ID 1450012.

Monoyios, M. (2003), 'Efficient option pricing with transaction costs', *Journal of Computational Finance* **7**(1), 1–22.

Monoyios, M. (2004), 'Option pricing with transaction costs using a Markov chain approximation', *Journal of Economic Dynamics and Control* **28**, 889–913.

Musiela, M. and Zariphopoulou, T. (2004), 'A valuation algorithm for indifference prices in incomplete markets', *Finance and Stochastics* **8**, 399–414.

Pennanen, T. (2014*a*), 'Convex duality in stochastic programming and mathematical finance', *Mathematical Programming Series B* **148**(1-2), 279–295.

Pennanen, T. (2014*b*), 'Optimal investment and contingent claim valuation in illiquid markets', *Finance and Stochastics* **18**(4), 733–754.

Pennanen, T. and Perkkiö, A.-P. (2018), 'Convex duality in optimal investment and contingent claim valuation in illiquid markets', *Finance and Stochastics* **28**, 733–771.

Perrakis, S. and Lefoll, J. (1997), 'Derivative asset pricing with transaction costs: An extension', *Computational Economics* **10**, 359–376.

Quek, G. S. H. (2012), Portfolio optimisation and option pricing in discrete time with transaction costs, PhD thesis, Imperial College London.

Rásonyi, M. and Stettner, L. (2005), 'On utility maximization in discrete-time financial market models', *The Annals of Applied Probability* **15**(2), 1367–1395.

Rockafellar, R. T. (1996), *Convex Analysis*, Princeton Landmarks in Mathematics and Physics, Princeton University Press.

Rogala, T. and Stettner, L. (2015), 'Construction of discrete time shadow price', *Applied Mathematics & Optimization* **72**(3), 391–433.

Rouge, R. and El Karoui, N. (2000), 'Pricing via utility maximization and entropy', *Mathematical Finance* **10**(2), 259–276.

Roux, A., Tokarz, K. and Zastawniak, T. (2008), 'Options under proportional transaction costs: An algorithmic approach to pricing and hedging', *Acta Applicandae Mathematicae* **103**(2), 201–219.

Roux, A. and Zastawniak, T. (2016), 'American and Bermudan options in currency markets under proportional transaction costs', *Acta Applicandae Mathematicae* **141**, 187–225.

Schachermayer, W. (2004), 'The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time', *Mathematical Finance* **14**(1), 19–48.

Whalley, A. E. and Wilmott, P. (1997), 'An asymptotic analysis of an optimal hedging model for option pricing with transaction costs', *Mathematical Finance* **7**(3), 307–324.

Xu, Z. (2018), Option Pricing and Hedging with Regret Optimisation, PhD thesis, University of York.