WILEY MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics

ORIGINAL ARTICLE

The robust pricing–hedging duality for American options in discrete time financial markets

Anna Aksamit[1] | Shuoqing Deng[3] | Jan Obłój[2] | Xiaolu Tan[3]

[1] School of Mathematics and Statistics, University of Sydney, Sydney, NSW, Australia

[2] University of Oxford, Oxford, United Kingdom

[3] University of Paris-Dauphine, PSL University, CNRS, UMR [7534], CEREMADE, Paris, France

Correspondence
Xiaolu Tan, University of Paris-Dauphine, PSL University, CNRS, UMR [7534], CEREMADE, Place du Maréchal De Lattre De Tassigny 75775 Paris Cedex 16, France.
Email: tan@ceremade.dauphine.fr

Abstract

We investigate the pricing–hedging duality for American options in discrete time financial models where some assets are traded dynamically and others, for example, a family of European options, only statically. In the first part of the paper, we consider an abstract setting, which includes the classical case with a fixed reference probability measure as well as the robust framework with a nondominated family of probability measures. Our first insight is that, by considering an enlargement of the space, we can see American options as European options and recover the pricing–hedging duality, which may fail in the original formulation. This can be seen as a weak formulation of the original problem. Our second insight is that a duality gap arises from the lack of dynamic consistency, and hence that a different enlargement, which reintroduces dynamic consistency is sufficient to recover the pricing–hedging duality: It is enough to consider fictitious extensions of the market in which all the assets are traded dynamically. In the second part of the paper, we study two important examples of the robust framework: the setup of Bouchard and Nutz and the martingale optimal transport setup of Beiglböck, Henry-Labordère, and Penkner, and show that our general results apply in both cases and enable us to obtain the pricing–hedging duality for American options.

1 | INTRODUCTION

The robust approach to pricing and hedging has been an active field of research in mathematical finance over recent years. It aims to address one of the key shortcomings of the classical approach, namely, its inability to account for model misspecification risk. Accordingly, the capacity to account for model uncertainty is at the core of the robust approach. In the classical approach, one postulates a fixed probability measure $\mathbb{P}$ to describe the future evolution of prices of risky assets. In contrast, in the robust approach, one considers the pricing and hedging problem simultaneously under a family of probability measures, or pathwise on a set of feasible trajectories. The challenge lies in extending the arbitrage pricing theory, which is well understood in the classical setup, to the robust setting.

In the classical approach, when the reference measure $\mathbb{P}$ is fixed, the absence of arbitrage is equivalent to the existence of a martingale measure equivalent to $\mathbb{P}$, a result known as the first fundamental theorem of asset pricing, see, for example, Delbaen and Schachermayer (2006) or Föllmer and Schied (2004). When the market is complete—that is, when every contingent claim can be perfectly replicated using a self-financing trading strategy—the equivalent martingale measure $\mathbb{Q}$ is unique, and the fair price for a contingent claim is equal to the replication cost of its payoff, and may be computed as the expected value of the discounted payoff under $\mathbb{Q}$. In an incomplete market, where a perfect replication strategy does not always exist, a conservative way of pricing is to use the minimum superreplication cost of the option. Employing duality techniques, this superreplication price can be expressed as the supremum of expectations of the discounted payoff over all martingale measures equivalent to $\mathbb{P}$.

In the robust approach, in the absence of a dominating probability measure, this elegant story often becomes more involved and technical. In continuous time models under *volatility uncertainty*, analogous pricing–hedging duality results have been obtained by, among many others, Denis and Martini (2006), Soner, Touzi, and Zhang (2013), Neufeld and Nutz (2013), and Possamaï, Royer, and Touzi (2013). In discrete time, a general pricing–hedging duality was shown in, for example, Bouchard and Nutz (2015) and Burzoni, Frittelli, and Maggis (2017). Importantly, in a robust setting, one often wants to include additional market instruments, which may be available for trading. In a setup that goes back to the seminal work of Hobson (1998), one often considers dynamic trading in the underlying asset and static trading, that is, buy and hold strategies at time zero, in some European options, often call or put options with a fixed maturity. Naturally, such additional assets constrain the set of martingale measures, which may be used for pricing. General pricing–hedging duality results, in different variations of this setting, both in continuous and in discrete time, can be found in, for example, Acciaio, Beiglböck, Penkner, and Schachermayer (2016), Burzoni, Frittelli, Hou, Maggis, and Obłój (2018), Beiglböck et al. (2013), Dolinsky and Soner (2014), Hou and Obłój (2018), Guo, Tan, and Touzi (2017), Tan and Touzi (2013), and we refer to the survey papers Hobson (2011) and Obłój (2004) for more details.

The main focus in the literature so far has been on the duality for (possibly exotic) European payoffs. However, more recently, some papers have also investigated American options. Cox and Hoeggerl (2016) studied the necessary (and, in some cases, sufficient) conditions for absence of arbitrage among

American put option prices. Dolinsky (2014) studied game options (including American options) in a nondominated discrete time market, but in the absence of any statically traded options. Neuberger (2007) considered a discrete time, discrete space market in the presence of statically traded European vanilla options. He observed that the superhedging price for an American option may be strictly larger than the supremum of its expected (discounted) payoff over all stopping times and all (relevant) martingale measures. We refer to such a situation as a *duality gap*. In Neuberger (2007), the pricing–hedging duality was then restored by using a *weak* dual formulation. This approach was further exploited, with more general results, in Hobson and Neuberger (2017). Bayraktar, Huang, and Zhou (2015) studied the same superhedging problem as in the setup of Bouchard and Nutz (2015), but only considered *strong* stopping times in their dual formulation, which again, in general, leads to a duality gap. More recently, and in parallel to an earlier version of this paper, imposing suitable regularity and integrability conditions on the payoff functions, Bayraktar and Zhou (2017) were able to prove a duality result by considering *randomized* models.

Motivated by the above works, we endeavor here to understand the fundamental reasons why pricing–hedging duality for American options holds or fails, and offer a systematic approach to mend it in the latter case. We derive two main general results, which we then apply to various specific contexts, both classical and robust. Our first insight is that by considering an appropriate enlargement of the space, namely, the time-space product structure, we can see an American option as a European option and recover the pricing–hedging duality, which may fail in the original formulation. This may be seen as a weak formulation of the dual (pricing) problem and leads to considering a large family of stopping times. This formulation of the dual problem is similar in spirit to Neuberger (2007), Hobson and Neuberger (2017), and Bayraktar and Zhou (2017), but our approach leads to duality results in more general settings, and/or under more general conditions, see Remark 2.7 and Section 3.3 and also Hobson and Neuberger (2016). Our second main insight is that the duality gap is caused by the failure of the dynamic programming principle. To recover the duality, under the formulation with *strong* stopping times, it is necessary and sufficient to consider an enlargement, which restores dynamic consistency: it is enough to consider (fictitious) extensions of the market in which all the assets are traded dynamically. As a by-product, we find that the strategies which trade dynamically in options and the semistatic strategies described above lead to the same superhedging cost in various settings.

The first part of the paper, Section 2, presents the above two main insights in a very general discrete time framework, which covers both classical (dominated) and robust (nondominated) settings. In the second part of the paper, we apply our general results in the context of two important examples of the robust framework: the setup of Bouchard and Nutz (2015) in Section 3, and the martingale optimal transport setup of Beiglböck et al. (2013) in Section 4. We obtain suitable pricing–hedging duality for American options in both setups. In the latter case of martingale optimal transport, there is an infinity of assets to consider and we use measure valued martingales to elegantly describe this setting. To allow for a suitable flow of narrative of our main results, technical proofs of the results in Sections 2, 3, and 4 are postponed and presented, respectively, in Sections 5,6 and 7.

Example 1.1. We conclude this introduction with a motivating example showing that the pricing–hedging duality may fail in the presence of statically traded instruments and how it may be recovered when exercise times are allowed to depend on the dynamic price processes of these instruments. This example is summarized in Figure 1. We consider a two period model with stock price process S given by $S_0 = S_1 = 0$ and $S_2 \in \{-2, -1, 1, 2\}$. The American option process Φ is defined as $\Phi_1 \equiv 1$, $\Phi_2(\{S_2 \in \{-2, 2\}\}) = 0$ and $\Phi_2(\{S_2 \in \{-1, 1\}\}) = 2$. The (pathwise) superhedging price of Φ, that is, the minimal initial wealth, which allows superhedging against all possible states and times by trading in the stock, can be easily computed and equals 2 (keeping 2 in cash and not trading in stock). A
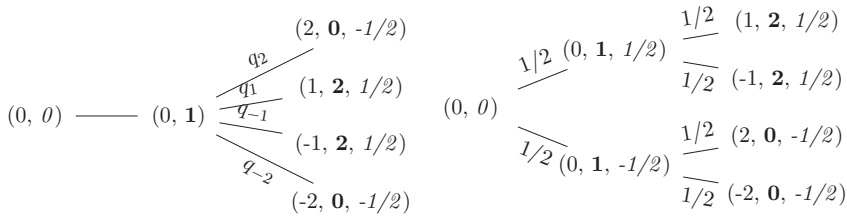
FIGURE 1 The stock prices are written in a regular font, the payoffs of the American option are in bold, and the prices of the European option g are in italic; the model without dynamic trading in g is on the left; the model with dynamic trading in g, in which the duality is recovered, is on the right

probability measure $\mathbb{Q}$ on the space of four possible paths is uniquely described through a choice of $q_i = \mathbb{Q}(S_2 = i) \geq 0$ for $i \in \{-2, -1, 1, 2\}$ satisfying $q_2 + q_1 + q_{-1} + q_{-2} = 1$. The martingale condition is equivalent to $2q_2 + q_1 - q_{-1} - 2q_{-2} = 0$. Note that as there are only two stopping times greater than 0, namely, $\tau_1 = 1$ and $\tau_2 = 2$, the market model price given as the double supremum over all stopping times τ and all martingale measures $\mathbb{Q}$ of $\mathbb{E}^{\mathbb{Q}}[\Phi_\tau]$ also equals 2 (as $\mathbb{E}^{\widetilde{\mathbb{Q}}}[\Phi_{\tau_2}] = 2$ for $\widetilde{\mathbb{Q}}$ given by $\widetilde{q}_1 = \widetilde{q}_{-1} = 1/2$ and $\widetilde{q}_2 = \widetilde{q}_{-2} = 0$) and the two prices agree.

Suppose now that we add a European option g with a payoff $g = \mathbb{1}_{\{|S_2|=1\}} - 1/2$ and an initial price 0, which may be used as a static hedging instrument. With g and S, the superhedging price of Φ drops to $3/2$ (e.g., keep $3/2$ in cash and buy one option g). The presence of g also imposes a calibration constraint on the martingale measures: $q_1 + q_{-1} = 1/2$. Thus, any calibrated martingale measure can be expressed by $(q_2, q_1, q_{-1}, q_{-2}) = (q, 3/4 - 2q, 2q - 1/4, 1/2 - q)$ with $q \in (1/8, 3/8)$, and the market model price equals 1 (as $\mathbb{E}^{\mathbb{Q}}[\Phi_{\tau_1}] = \mathbb{E}^{\mathbb{Q}}[\Phi_{\tau_2}] = 1$ under any calibrated martingale measure). We therefore see that adding a statically traded option breaks the pricing–hedging duality.

Let us now show that the duality is recovered when we consider a fictitious market where the option g is traded dynamically and the exercise policy can depend on its current price. We model this through a process $Y = (Y_t : t = 0, 1, 2)$ given by $Y_2 = g$, $Y_1 = 1/2$ on $\{|S_2| = 1\}$, $Y_1 = -1/2$ on $\{|S_2| = 2\}$, and $Y_0 = 0$. Note that there exists a (unique) measure $\widehat{\mathbb{Q}}$, presented in Figure 1, such that both S and Y are martingales with respect to their joint natural filtration $\widehat{\mathbb{F}}$ and in particular $\widehat{\mathbb{Q}}$ is calibrated: $\mathbb{E}^{\widehat{\mathbb{Q}}}[g] = 0$. The filtration $\widehat{\mathbb{F}}$ is richer than the natural filtration of S alone and allows for an additional stopping time $\tau^* = \mathbb{1}_{\{Y_1 = -1/2\}} + 2\mathbb{1}_{\{Y_1 = 1/2\}}$, and the duality is recovered as $\mathbb{E}^{\widehat{\mathbb{Q}}}[\Phi_{\tau^*}] = 3/2$.

2 | PRICING–HEDGING DUALITY FOR AMERICAN OPTIONS

We present in this section general results that explain when and why the pricing–hedging duality for American options holds. We work in a general discrete time setup, which we now introduce. Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathbb{F} := (\mathcal{F}_k)_{k=0,1,\ldots,N}$ be a filtration, where $\mathcal{F}_0$ is trivial and $N \in \mathbb{N}$ is the time horizon. We denote by $\mathfrak{P}(\Omega)$ the set of all probability measures on $(\Omega, \mathcal{F})$ and consider a subset $\mathcal{P} \subset \mathfrak{P}(\Omega)$. We say that a given property holds $\mathcal{P}$-quasi surely ($\mathcal{P}$-q.s.) if it holds $\mathbb{P}$-almost surely for every $\mathbb{P} \in \mathcal{P}$, and say that a set from $\mathcal{F}$ is $\mathcal{P}$-polar if it is a null set with respect to every $\mathbb{P} \in \mathcal{P}$. We write $\mathbb{Q} \ll \mathcal{P}$ if there exists a $\mathbb{P} \in \mathcal{P}$ such that $\mathbb{Q} \ll \mathbb{P}$. Given a random variable ξ and a sub-σ-field $\mathcal{G} \subset \mathcal{F}$, we define the conditional expectation $\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{G}] := \mathbb{E}^{\mathbb{P}}[\xi^+|\mathcal{G}] - \mathbb{E}^{\mathbb{P}}[\xi^-|\mathcal{G}]$ with the convention $\infty - \infty = -\infty$, where $\xi^+ := \xi \vee 0$ and $\xi^- := -(\xi \wedge 0)$. We consider a market with no transaction costs and with financial assets, some of which are dynamically traded and some of which are only statically traded. The former are modeled by an adapted $\mathbb{R}^d$-valued process S with $d \in \mathbb{N}$. We think of the latter

as European options, which are traded at time $t = 0$ and not at future times. We let $g = (g^\lambda)_{\lambda \in \Lambda}$, where Λ is a set of arbitrary cardinality, be the vector of their payoffs, which are assumed to be $\mathbb{R}$-valued and $\mathcal{F}$-measurable. Up to a constant shift of the payoffs, we may assume, without loss of generality, that all options g^λ have zero initial price. All prices are expressed in units of some numeraire S^0, such as a bank account, whose price is thus normalized, so that $S_t^0 \equiv 1$. We denote by $\mathcal{H}$ the set of all $\mathbb{F}$-predictable $\mathbb{R}^d$-valued processes, and by $\mathfrak{h} = \{h \in \mathbb{R}^\Lambda : \exists \text{ finite subset } \beta \subset \Lambda \text{ s.t. } h^\lambda = 0 \forall \lambda \notin \beta\}$. A self-financing strategy trades dynamically in S and statically in finitely many of g^λ, $\lambda \in \Lambda$ and hence corresponds to a choice of $H \in \mathcal{H}$ and $h \in \mathfrak{h}$. Its associated final payoff is given by

$$(H \circ S)_N + hg = \sum_{j=1}^d \sum_{k=1}^N H_k^j \Delta S_k^j + \sum_{\lambda \in \Lambda} h^\lambda g^\lambda, \tag{1}$$

where $\Delta S_k^j = S_k^j - S_{k-1}^j$. Having defined the trading strategy, we can consider the superhedging price of an option with payoff ξ at time N, given by

$$\pi_g^E(\xi) := \inf\{x : \exists (H, h) \in \mathcal{H} \times \mathfrak{h} \text{ s.t. } x + (H \circ S)_N + hg \geq \xi \ \mathcal{P}\text{-q.s.}\}. \tag{2}$$

In particular, if $\mathcal{P} = \mathfrak{P}(\Omega)$ is the set of all probability measures on $\mathcal{F}$ and $\{\omega\} \in \mathcal{F}$ for all $\omega \in \Omega$, then the superreplication in (2) is pathwise on Ω.

To formulate a duality relationship, we need the dual elements given by rational pricing rules, or martingale measures,

$$\mathcal{M} = \left\{ \mathbb{Q} \in \mathfrak{P}(\Omega) : \mathbb{Q} \ll \mathcal{P} \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}[\Delta S_k | \mathcal{F}_{k-1}] = 0, \quad \forall k = 1, \ldots, N \right\},$$

$$\mathcal{M}_g = \left\{ \mathbb{Q} \in \mathcal{M} : \mathbb{E}^{\mathbb{Q}}[g^\lambda] = 0, \ \forall \lambda \in \Lambda \right\}. \tag{3}$$

Definition 2.1. Let Υ be a given class of real-valued functions defined on Ω. We say that the *(European) pricing–hedging duality holds for the class* Υ if $\mathcal{M}_g \neq \emptyset$ and

$$\pi_g^E(\xi) = \sup_{\mathbb{Q} \in \mathcal{M}_g} \mathbb{E}^{\mathbb{Q}}[\xi], \quad \xi \in \Upsilon. \tag{4}$$

Remark 2.2. Note that the inequality "$\geq$" in (4), called weak pricing–hedging duality, holds automatically from the definition of $\mathcal{M}_g$ in (3).

A number of papers, including Bouchard and Nutz (2015) and Burzoni et al. (2018), proved that the above pricing–hedging duality (4) holds under various further specifications and restrictions on Ω, $\mathbb{F}$, $\mathcal{P}$ and Υ, including in particular an appropriate no-arbitrage condition. We take the above duality for granted here and our aim is to study an analogous duality for American options. We work first in the general setup described above without specifying $\mathbb{F}$ or Υ, as our results will apply to any such further specification. Further, many abstract results in this section also extend to other setups, for example, to trading in continuous time.

2.1 | Superhedging of American options

An American option may be exercised at any time $k \in \mathbb{T} := \{1, \ldots, N\}$ (without loss of generality we exclude exercise at time 0). It is described by its payoff function $\Phi = (\Phi_k)_{1 \leq k \leq N}$, where $\Phi_k : \Omega \to \mathbb{R}$

belongs to Υ and is the payoff, delivered at time N, if the option is exercised at time k. Usually Φ_k is taken to be $\mathcal{F}_k$-measurable, but here we only assume Φ_k to be $\mathcal{F}$-measurable for greater generality, which includes, for example, the case of a portfolio containing a mixture of American and European options. We note that when hedging our exposure to an American option, we should be allowed to adjust our strategy in response to an early exercise. As a consequence, the superhedging cost of the American option Φ using semistatic strategies is given by

$$\pi_g^A(\Phi) := \inf \left\{ x \ : \ \exists (^1H, \dots, {}^NH) \in \mathcal{H}^N \ \text{s.t.} {}^jH_i = {}^kH_i \ \forall 1 \le i \le j \le k \le N \quad \text{and} \quad h \in \mathfrak{h} \right.$$

$$\left. \text{satisfying} \ \ x + (^kH \circ S)_N + hg \ge \Phi_k \ \forall k = 1, \dots, N \ \mathcal{P}\text{-q.s.} \right\}.$$

Remark 2.3. We formulate the problem above with payoff delivered by the seller at maturity N irrespectively of the actual exercise time. In full generality, this is necessary because the payoff is not assumed to be known at the exercise time. However, given that we work in discounted units, if the payoff is known at the exercise time, our convention is equivalent to the one in which the payoff is delivered at its exercise time, via taking a loan, and then the seller has to be able to continue trading in such a way that her final payoff is nonnegative. For this equivalence to hold it is important to allow the seller to adjust the strategy at the time of the exercise. Note that in the more classical setting when Φ_k is $\mathcal{F}_k$-measurable and there are no statically traded options, that is, $\Lambda = \emptyset$, no-arbitrage ensures that to have a final nonnegative payoff the seller has to have a nonnegative wealth after delivering the payoff at the exercise time. She can then just stop trading altogether—the vector of strategies $(^1H, \dots, {}^NH) \in \mathcal{H}^N$ above then reduces to a single trading strategy which is unwound at the exercise time.

Classically, the pricing of an American option is recast as an optimal stopping problem and, extending (4), it would be natural to ask whether

$$\pi_g^A(\Phi) \stackrel{?}{=} \sup_{\mathbb{Q} \in \mathcal{M}_g} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau] \tag{5}$$

holds, where $\mathcal{T}(\mathbb{F})$ denotes the set of $\mathbb{F}$-stopping times. However, as illustrated by the simple example in the introduction, this duality may fail. The "numerical" reason is that the right-hand side in (4) may be too small because the set $\mathcal{M}_g \times \mathcal{T}(\mathbb{F})$ is too small. Our aim here is to understand the fundamental reasons why the duality fails and hence discuss how and why the right-hand side should be modified to obtain the equality in (5).

2.2 | An American option is a European option on an enlarged space

The first key idea of this paper offers a generic enlargement of the underlying probability space which turns all American options into European options. Depending on the particular setup, it may take more or less effort to establish (4) for the enlarged space, but this shifts the difficulty back to the better understood and well studied case of European options. Our reformulation technique—from an American to European option—can be easily extended to other contexts, such as the continuous time case. The enlargement of space is based on construction of random times, previously used, for example, in Jeanblanc and Song (2011a, 2011b) to study the existence of random times with a given survival probability, in El Karoui and Tan (2013) to study a general optimal control/stopping problem, and in Guo, Tan, and Touzi (2016) and Källblad, Tan, and Touzi (2017) to study the optimal Skorokhod embedding problem.

Recalling the notation $\mathbb{T} := \{1, \dots, N\}$, we introduce the space $\overline{\Omega} := \Omega \times \mathbb{T}$ with the canonical time $T : \overline{\Omega} \to \mathbb{T}$ given by $T(\overline{\omega}) := \theta$, where $\overline{\omega} := (\omega, \theta)$, the filtration $\overline{\mathbb{F}} := (\overline{\mathcal{F}}_k)_{k=0,1,\dots,N}$ with $\overline{\mathcal{F}}_k =$

$\mathcal{F}_k \otimes \vartheta_k$ and $\vartheta_k = \sigma(T \wedge (k+1))$, and the σ-field $\overline{\mathcal{F}} = \mathcal{F} \otimes \vartheta_N$. By definition, T is an $\overline{\mathbb{F}}$-stopping time. We denote by $\overline{\mathcal{H}}$ the class of $\overline{\mathbb{F}}$-predictable processes and extend naturally the definitions of S and g^λ from Ω to $\overline{\Omega}$ via $S(\overline{\omega}) = S(\omega)$ and $g^\lambda(\overline{\omega}) = g^\lambda(\omega)$ for $\overline{\omega} = (\omega, \theta) \in \overline{\Omega}$. We let $\overline{\Upsilon}$ be the class of random variables $\overline{\xi} : \overline{\Omega} \to \mathbb{R}$ such that $\overline{\xi}(\cdot, k) \in \Upsilon$ for all $k \in \mathbb{T}$ and we let $\overline{\pi}_g^E(\overline{\xi})$ denote the superreplication cost of $\overline{\xi}$. We may, and will, identify $\overline{\Upsilon}$ with Υ^N via $\overline{\xi}(\overline{\omega}) = \Phi_\theta(\omega)$. Finally, we introduce

$$\overline{\mathcal{P}} = \{\overline{\mathbb{P}} \in \mathfrak{P}(\overline{\Omega}) : \overline{\mathbb{P}}_{|\Omega} \in \mathcal{P}\},$$

$$\overline{\mathcal{M}} = \{\overline{\mathbb{Q}} \in \mathfrak{P}(\overline{\Omega}) : \overline{\mathbb{Q}} \ll \overline{\mathcal{P}} \quad \text{and} \quad \mathbb{E}^{\overline{\mathbb{Q}}}[\Delta S_k | \overline{\mathcal{F}}_{k-1}] = 0 \; \forall k \in \mathbb{T}\},$$

$$\overline{\mathcal{M}}_g = \{\overline{\mathbb{Q}} \in \overline{\mathcal{M}} : \mathbb{E}^{\overline{\mathbb{Q}}}[g^\lambda] = 0 \; \forall \lambda \in \Lambda\}. \tag{6}$$

Theorem 2.4. *For any $\Phi \in \Upsilon^N = \overline{\Upsilon}$, we have*

$$\pi_g^A(\Phi) = \overline{\pi}_g^E(\Phi) := \inf\{x : \exists (\overline{H}, h) \in \overline{\mathcal{H}} \times \mathfrak{h} \; s.t. \; x + (\overline{H} \circ S)_N + hg \geq \xi \; \overline{\mathcal{P}}\text{-}q.s.\}. \tag{7}$$

In particular, if the European pricing–hedging duality on $\overline{\Omega}$ holds for Φ, then

$$\pi_g^A(\Phi) = \overline{\pi}_g^E(\Phi) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]. \tag{8}$$

Proof. First note that

$$\overline{\mathcal{H}} = \{\overline{H} = (\overline{H}(\cdot, 1), \dots, \overline{H}(\cdot, N)) \in \mathcal{H}^N \; : \; \overline{H}_i(\cdot, j) = \overline{H}_i(\cdot, k) \; \forall 1 \leq i \leq j \leq k \leq N\},$$

and hence that the dynamic strategies used for superhedging in π_g^A and in $\overline{\pi}_g^E$ are the same. The equality now follows by observing that a set $\Gamma \in \overline{\mathcal{F}}_N$ is $\overline{\mathcal{P}}$-polar if and only if its k-sections $\Gamma_k = \{\omega : (\omega, k) \in \Gamma\}$ are $\mathcal{P}$-polar for all $k \in \mathbb{T}$. Indeed, for one implication assume that $\overline{\mathbb{P}}(\Gamma) = 0$ for each $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$. For arbitrary $\mathbb{P} \in \mathcal{P}$ and $k \in \mathbb{T}$ we can define $\overline{\mathbb{P}} = \mathbb{P} \otimes \delta_k$, which belongs to $\overline{\mathcal{P}}$, and hence $\mathbb{P}(\Gamma_k) = 0$ follows.

To show the reverse implication, assume that $\mathbb{P}(\Gamma_k) = 0$ for each $\mathbb{P} \in \mathcal{P}$ and $k \in \mathbb{T}$. Observe that, for any $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$,

$$\overline{\mathbb{P}}(\Gamma) = \sum_{k \in \mathbb{T}} \overline{\mathbb{P}}(\Gamma_k \times \{k\}) \leq \sum_{k \in \mathbb{T}} \overline{\mathbb{P}}_{|\Omega}(\Gamma_k) = 0$$

as $\overline{\mathbb{P}}_{|\Omega} \in \mathcal{P}$. This completes the proof. □

Remark 2.5. If the pricing–hedging duality holds with respect to the filtration $\mathbb{F}$, then it also holds for any filtration $\mathbb{H} \supset \mathbb{F}$ such that $\mathbb{H}$ and $\mathbb{F}$ only differ up to $\mathcal{M}_g$-polar sets. Indeed, this follows from Remark 2.2, observing that such a change does not affect $\mathcal{M}_g$ and can only decrease the superhedging cost as one has more trading strategies available.

Remark 2.6. We note that the set $\overline{\mathcal{M}}_g$ in (8) is potentially much larger than the set of all pushforward measures induced by $\omega \mapsto (\omega, \tau(\omega))$ and $\mathbb{Q} \in \mathcal{M}_g$ for all $\tau \in \mathcal{T}(\mathbb{F})$. Indeed, instead of stopping times relative to $\mathbb{F}$, it allows us to consider any *random* time, which can be made into a stopping time under some calibrated martingale measure. We can rephrase this as saying that $\overline{\mathcal{M}}_g$ is equivalent to a *weak*

formulation of the initial problem on the right-hand side of (5). To make this precise, let us define a weak stopping rule α as a collection

$$\alpha = \left(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{Q}^\alpha, \mathbb{F}^\alpha = \left(\mathcal{F}_k^\alpha \right)_{0 \leq k \leq N}, \left(S_k^\alpha \right)_{0 \leq k \leq N}, (g^{\lambda,\alpha})_{\lambda \in \Lambda}, \left(\Phi_k^\alpha \right)_{k \in \mathbb{T}}, \tau^\alpha \right)$$

with $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{Q}^\alpha, \mathbb{F}^\alpha)$ a filtered probability space, τ^α a $\mathbb{T}$-valued $\mathbb{F}^\alpha$-stopping time, an $\mathbb{R}^d$-valued $(\mathbb{Q}^\alpha, \mathbb{F}^\alpha)$-martingale S^α, and a collection of random variables $g^{\lambda,\alpha}, \Phi_k^\alpha$, and such that there is a measurable surjective mapping $\mathtt{i}_\alpha : \Omega^\alpha \to \Omega$ with $\mathbb{Q} = \mathbb{Q}^\alpha \circ \mathtt{i}_\alpha^{-1} \in \mathcal{M}$ and $\mathtt{i}_\alpha^{-1}(\mathcal{F}_k) \subset \mathcal{F}_k^\alpha$, $\mathtt{i}_\alpha^{-1}(\mathcal{F}) \subset \mathcal{F}^\alpha$, and finally $\mathcal{L}_{\mathbb{Q}^\alpha}(S^\alpha, g^\alpha, \Phi^\alpha) = \mathcal{L}_{\mathbb{Q}}(S, g, \Phi)$. Denote by $\mathcal{A}_g$ the collection of all weak stopping rules α such that $\mathbb{E}^{\mathbb{Q}^\alpha}[g^{\lambda,\alpha}] = 0$ for each $\lambda \in \Lambda$. It follows that any $\alpha \in \mathcal{A}_g$ induces a probability measure $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g$ and $\mathbb{E}^{\mathbb{Q}^\alpha}[\Phi_{\tau^\alpha}^\alpha] = \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]$. Reciprocally, any $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g$, together with the space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}})$ and (S, g, Φ), provides a weak stopping rule in $\mathcal{A}_g$. As a consequence,

$$\sup_{\alpha \in \mathcal{A}_g} \mathbb{E}^{\mathbb{Q}^\alpha} \left[\Phi_{\tau^\alpha} \right] = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi].$$

In summary, and similarly to a number of other contexts, see the introduction in Pham and Zhang (2014), the weak formulation (and not the strong one) offers the right framework to compute the value of the problem. In fact, the set $\overline{\mathcal{M}}_g$ is large enough to make the problem static, or European, again. However, although it offers a solution and a corrected version of (5), it does not offer a fundamental insight into why (5) may fail and if there is a canonical "smaller" way of enlarging the objects on the right-hand side thereof to preserve the equality. These questions are addressed in the subsequent section.

Remark 2.7. Neuberger (2007) and Hobson and Neuberger (2017) studied the same superhedging problem in a Markovian setting, where the underlying process S takes values in a discrete lattice $\mathcal{X}$. By considering the weak formulation (which is equivalent to our formulation, as shown in Remark 2.6 above), they obtain similar duality results. However, they only consider $\Phi_k = \phi(S_k)$, where $\phi : \mathbb{R}^d \to \mathbb{R}$. Then, the authors show that in the optimization problem $\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]$ given in (7) one may restrict to only Markovian martingale measures. The primal and the dual problem then turn out to be linear programming problems under linear constraints, which can be solved numerically. Their arguments have also been extended to a more general context, where S takes values in $\mathbb{R}_+$. Comparing to Neuberger (2007) and Hobson and Neuberger (2017), our weak formulation is very similar to theirs. However, our setting is much more general and, when considering the specific setups in Sections 3 and 4, we rely on entirely different arguments to prove the duality.

2.3 | The loss and recovery of the dynamic programming principle and the natural duality for American options

The classical pricing of American options, on which the duality in (5) was modeled, relies on optimal stopping techniques, which subsume a certain dynamic consistency, or a dynamic programming principle, as explained below. Our second key observation in this paper is that if the pricing–hedging duality (5) for American options fails it is because the introduction of static trading of European options g at time $t = 0$ destroys the dynamic programming principle. Indeed, $\pi_g^E(\xi)$ will typically be lower than the superhedging price at time $t = 0$ of the capital needed at time $t = 1$ to superhedge from thereon. To reinstate such dynamic consistency, we need to enlarge the model and consider dynamic trading in options in g. This will generate a richer filtration than $\mathbb{F}$ and one which will carry enough stopping times

to obtain the correct natural duality in the spirit of (5). In particular, if $g = 0$ (or equivalently $\Lambda = \emptyset$), then (5) should hold. We now first prove this statement and then present the necessary extension when g is nontrivial.

Let Υ be a class of $\mathcal{F}$-measurable random variables such that $-\infty \in \Upsilon$, we denote $\mathcal{E}(\xi) := \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\xi]$, and suppose that there is a family of operators $\mathcal{E}_k : \Upsilon \to \Upsilon$ for $k \in \{0, \dots, N-1\}$ such that $\mathcal{E}_k(\xi)$ is $\mathcal{F}_k$-measurable for all $\xi \in \Upsilon$ and $\mathcal{E}_k(-\infty) = -\infty$. Notice that $\mathcal{F}_0$ is assumed to be trivial so that $\mathcal{E}_0(\xi)$ is deterministic. We say that the family $(\mathcal{E}_k)$ provides a dynamic programming representation of $\mathcal{E}$ if

$$\mathcal{E}(\xi) = \mathcal{E}^0(\xi) \ \forall \xi \in \Upsilon, \quad \text{where} \quad \mathcal{E}^k(\xi) := \mathcal{E}_k \circ \cdots \circ \mathcal{E}_{N-1}(\xi), \quad 0 \le k \le N-1. \tag{9}$$

The family $(\mathcal{E}_k)$ naturally extends to $(\overline{\mathcal{E}}_k)$, $0 \le k \le N-1$, defined for any $\Phi \in \overline{\Upsilon} = \Upsilon^N$ by

$$\overline{\mathcal{E}}_0(\Phi) := \mathcal{E}_0(\Phi(\cdot, 1)),$$

$$\overline{\mathcal{E}}_k(\Phi)(\overline{\omega}) := \begin{cases} \mathcal{E}_k(\Phi(\cdot, \theta))(\omega) & \text{if } \theta < k, \\ \mathcal{E}_k(\Phi(\cdot, k))(\omega) \vee \mathcal{E}_k(\Phi(\cdot, k+1))(\omega) & \text{if } \theta \ge k, \end{cases} \quad \text{for } 1 \le k \le N-1. \tag{10}$$

Assume that $f \vee f' \in \Upsilon$ for $f, f' \in \Upsilon$ so that $(\overline{\mathcal{E}}_k)$ maps functionals from $\overline{\Upsilon}$ to $\overline{\Upsilon}$. We introduce the following process:

$$\overline{\mathcal{E}}^k(\Phi) := \overline{\mathcal{E}}_k \circ \cdots \circ \overline{\mathcal{E}}_{N-1}(\Phi), \quad 0 \le k \le N-1, \tag{11}$$

which, under suitable assumptions (see Proposition 2.9 below) represents the $\mathcal{M}$-Snell envelope process of an American option $\Phi \in \overline{\Upsilon}$. To illustrate how the operator $\overline{\mathcal{E}}^0$ works, we develop it for the case $\mathbb{T} = \{1, 2, 3\}$,

$$\overline{\mathcal{E}}^0(\Phi) = \mathcal{E}_0 \left(\mathcal{E}_1 \circ \mathcal{E}_2(\Phi(\cdot, 1)) \bigvee \mathcal{E}_1(\mathcal{E}_2(\Phi(\cdot, 2)) \vee \mathcal{E}_2(\Phi(\cdot, 3))) \right).$$

We say that the family $(\overline{\mathcal{E}}_k)$ provides a dynamic programming representation of $\overline{\mathcal{E}}(\Phi) := \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]$ if

$$\overline{\mathcal{E}}(\Phi) = \overline{\mathcal{E}}^0(\Phi), \quad \forall \Phi \in \overline{\Upsilon}. \tag{12}$$

Typically we will consider $\mathcal{E}_k$ to be a supremum over conditional expectations with respect to $\mathcal{F}_k$ (see Examples 2.11 and 2.12 below), and in such setups we automatically obtain

$$\overline{\mathcal{E}}(\Phi) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \le \overline{\mathcal{E}}^0(\Phi), \quad \Phi \in \overline{\Upsilon}. \tag{13}$$

Theorem 2.8. *Assume that $\Lambda = \emptyset$, $\mathcal{E}_k$ satisfies (9), that (13) holds true, and that $f \vee f' \in \Upsilon$ for all $f, f' \in \Upsilon$. Then, for all $\Phi \in \Upsilon^N = \overline{\Upsilon}$,*

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \sup_{\mathbb{Q} \in \mathcal{M}} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau]. \tag{14}$$

If, further, the European pricing–hedging duality holds on $\overline{\Omega}$ for the class $\overline{\Upsilon}$, then

$$\pi^A(\Phi) = \sup_{\mathbb{Q}\in\mathcal{M}} \sup_{\tau\in\mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau].$$

The second assertion follows instantly from the first one and Theorem 2.4. The first assertion is reformulated and proved in Proposition 2.9 below, which also allows us to identify the optimal stopping time on the right-hand side of (14).

Proposition 2.9. *Assume that $\Lambda = \emptyset$ and $f \vee f' \in \Upsilon$ for all $f, f' \in \Upsilon$. Then the dynamic programming representation (12) holds if and only if (9) and (13) hold true. Moreover, under condition (12), the $\mathbb{F}$-stopping time*

$$\tau^*(\omega) := \min\{k \geq 1 : \mathcal{E}^k(\Phi(\cdot, k))(\omega) = \overline{\mathcal{E}}^k(\Phi)(\omega, k)\} \tag{15}$$

provides the optimal exercise policy for $\Phi \in \overline{\Upsilon}$, in sense that

$$\sup_{\overline{\mathbb{Q}}\in\overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \sup_{\mathbb{Q}\in\mathcal{M}} \sup_{\tau\in\mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau] = \sup_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}\left[\Phi_{\tau^*}\right] = \overline{\mathcal{E}}^0(\Phi). \tag{16}$$

Remark 2.10. The proof of Proposition 2.9 will be provided in Section 5. The results in Theorem 2.8 and Proposition 2.9 are stated on $(\Omega, \mathcal{F})$, where there are only finitely many dynamically traded risky assets. However, their proofs do not rely on the fact that the number of risky assets is finite, and the same results still hold true if there are infinitely many dynamically traded risky assets.

Next, we give two examples of operators $(\mathcal{E}_k)_{k\leq N-1}$ satisfying (9), (13), and therefore, by Proposition 2.9, also (12).

Example 2.11. The model-specific setting is recovered by taking $\mathcal{P} = \{\mathbb{P}\}$, for a fixed probability measure $\mathbb{P}$. Then, taking Υ to be the set of all $\mathcal{F}$-measurable random variables and

$$\mathcal{E}_k(\xi) = \text{ess sup}_{\mathbb{Q}\in\mathcal{M}_g} \mathbb{E}^{\mathbb{Q}}[\xi|\mathcal{F}_k], \tag{17}$$

where the essential supremum is taken with respect to $\mathbb{P}$, leads to a family of operators satisfying (9), (13), and therefore also (12). See the literature on dynamic coherent risk measures for further discussion, for example, Acciaio and Penner (2011) for an overview. In particular, Theorem 2.8 recovers the classical superhedging theorem for American options (see, e.g., Myneni, 1992).

Example 2.12. Let (Ω, d) be a Polish space, $\mathcal{F}$ the universally completed Borel σ-field, $\mathcal{P}$ a given set of probability measures on $(\Omega, \mathcal{F})$, and $\mathcal{M}$ be defined by (3). We are given a filtration $\mathbb{G} := (\mathcal{G}_k)_{k\leq N}$ such that $\mathcal{G}_0 = \{\emptyset, \Omega\}$ and each σ-field $\mathcal{G}_k$ is countably generated. Let $\mathcal{F}_k$ be the universal completion of $\mathcal{G}_k$. Notice that $\mathbb{E}^{\mathbb{P}}[\xi|\mathcal{G}_k] = \mathbb{E}^{\mathbb{P}}[\xi|\mathcal{F}_k]$, $\mathbb{P}$-a.s. for all $\mathcal{F}$-measurable ξ and $\mathbb{P} \in \mathcal{P}$, and the fact that $\mathcal{G}_k$ is countably generated ensures the existence of a regular conditional probability of $\mathbb{P}$ with respect to $\mathcal{G}_k$.

Assume there exists a family $(\mathcal{M}_k(\omega))_{k\leq N-1, \omega\in\Omega}$ of sets of measures satisfying:

- $\mathbb{Q}([\omega]_{\mathcal{G}_k}) = 1$ for all $\mathbb{Q} \in \mathcal{M}_k(\omega)$, where $[\omega]_{\mathcal{G}_k}$ is the atom of $\mathcal{G}_k$ containing ω, that is,

$$[\omega]_{\mathcal{G}_k} = \bigcap_{F\in\mathcal{G}_k : \omega\in F} F. \tag{18}$$

Note that $[\omega]_{\mathcal{G}_k} \in \mathcal{G}_k$ because the latter is countably generated.

- For every $\mathbb{Q} \in \mathcal{M}$ and every family of regular conditional probabilities $(\mathbb{Q}_\omega)_{\omega \in \Omega}$ of $\mathbb{Q}$ with respect to $\mathcal{G}_k$, one has $\mathbb{Q}_\omega \in \mathcal{M}_k(\omega)$, for $\mathbb{Q}$-a.e. ω.

Define the family $(\mathcal{E}_k)_{k \leq N-1}$ by

$$\mathcal{E}_k(\xi)(\omega) = \sup_{\mathbb{Q} \in \mathcal{M}_k(\omega)} \mathbb{E}^{\mathbb{Q}}[\xi].$$

If we furthermore assume that $\mathcal{E}_k(\xi) \in \Upsilon$ for any $\xi \in \Upsilon$, then the family $(\mathcal{E}_k)_{k \leq N-1}$ satisfies (13) (see Proposition 5.1). Moreover, under suitable assumptions on $(\Omega, \mathbb{F}, \mathcal{P})$, $\mathcal{M}_k(\omega)$, and Υ, we shall also prove that (9) holds for this family. This holds in particular in the setup of Bouchard and Nutz (2015) as shown therein; see (4.12) in Bouchard and Nutz (2015).

Let us consider the case with statically traded options: $\Lambda \neq \emptyset$. We saw in Example 1.1 that this can break down dynamic consistency as the universe of traded assets differs at time $t = 0$ and times $t \geq 1$. To remedy this, one has to embed the market into a fictitious larger one where both S and all the options g^λ, $\lambda \in \Lambda$, are traded dynamically.

Definition 2.13. Let $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}})$ be a filtered space satisfying the following properties:

1. There exists a surjective mapping $\mathrm{i} : \widehat{\Omega} \to \Omega$.

2. For each k, $\mathrm{i}^{-1}(\mathcal{F}_k) \subset \widehat{\mathcal{F}}_k$.

3. There exists a family of $\widehat{\mathbb{F}}$-adapted processes $Y = (Y^\lambda)_{\lambda \in \Lambda}$ such that $Y_0^\lambda = 0$ and $Y_N^\lambda(\widehat{\omega}) = g^\lambda(\mathrm{i}(\widehat{\omega}))$.

4. Let $S(\widehat{\omega}) := S(\mathrm{i}(\widehat{\omega}))$, $\widehat{S} := (S, Y)$, and define

$$\widehat{\mathcal{P}} := \{\widehat{\mathbb{P}} \in \mathfrak{P}(\widehat{\Omega}) : \widehat{\mathbb{P}} \circ \mathrm{i}^{-1} \in \mathcal{P}\}, \quad \widehat{\mathcal{M}} := \{\widehat{\mathbb{Q}} \ll \widehat{\mathcal{P}} : \widehat{S} \text{ is an } (\widehat{\mathbb{Q}}, \widehat{\mathbb{F}})\text{-martingale}\}. \tag{19}$$

There exists a mapping $\mathrm{J} : \mathcal{M}_g \to \widehat{\mathcal{M}}$ such that for $\mathbb{Q} \in \mathcal{M}_g$, $\mathrm{J}(\mathbb{Q}) \circ \mathrm{i}^{-1} = \mathbb{Q}$ and

$$\mathcal{L}_{\mathrm{J}(\mathbb{Q})}(\widehat{S}) = \mathcal{L}_{\mathbb{Q}}(S, Y^{\mathbb{Q}}),$$

where $Y^{\mathbb{Q}} = (Y^{\lambda, \mathbb{Q}})_{\lambda \in \Lambda}$ and $Y^{\lambda, \mathbb{Q}} := (\mathbb{E}^{\mathbb{Q}}[g^\lambda | \mathcal{F}_k])_{k \leq N}$.

The collection $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, \widehat{\mathcal{P}}, Y, \mathrm{i}, \mathrm{J})$ satisfying Properties 1–4 above is called a dynamic extension of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$. In short, we shall say that $\widehat{\Omega}$ is a dynamic extension of Ω.

Remark 2.14. A measure $\mathbb{Q} \in \mathcal{M}_g$ is an admissible pricing measure under which the time-k prices for European options g^λ are given by $Y_k^{\lambda, \mathbb{Q}}$. Property 4 in the above definition says that any such $\mathbb{Q} \in \mathcal{M}_g$ can be lifted to a measure $\mathrm{J}(\mathbb{Q}) \in \widehat{\mathcal{M}}$, which preserves the joint distribution of the stock and option prices. In general, we do not expect the reverse to be true and we may have $\mathrm{J}(\mathcal{M}_g) \subsetneq \widehat{\mathcal{M}}$. More precisely, $\widehat{\mathcal{M}}$ may offer scope for a richer description and dynamics so that the mapping $\widehat{\mathcal{M}} \ni \widehat{\mathbb{Q}} \mapsto \widehat{\mathbb{Q}} \circ \mathrm{i}^{-1} \in \mathcal{M}_g$ is surjective but typically not injective.

Example 2.15. In practice, the map J in a dynamic extension is often built from a family of mappings from Ω to $\widehat{\Omega}$. Assume that, for each $\mathbb{Q} \in \mathcal{M}_g$, one has a mapping $\mathrm{j}_{\mathbb{Q}} : \Omega \to \widehat{\Omega}$ such that $\mathrm{i} \circ \mathrm{j}_{\mathbb{Q}} = id$ and $\mathcal{L}_{\mathbb{Q} \circ \mathrm{j}_{\mathbb{Q}}^{-1}}(Y) = \mathcal{L}_{\mathbb{Q}}(Y^{\mathbb{Q}})$. Let $\mathrm{J}(\mathbb{Q}) := \mathbb{Q} \circ \mathrm{j}_{\mathbb{Q}}^{-1}$. Then, $\mathrm{J} : \mathcal{M}_g \to \widehat{\mathcal{M}}$ satisfies Property 4 of Definition 2.13.

Let us illustrate this with an example of a dynamic extension of Ω in the case of finitely many statically traded options, that is, $\Lambda = \{1, \ldots, e\}$ for some $e \in \mathbb{N}$. Consider the space $\widehat{\Omega} = \Omega \times \mathbb{R}^{(N-1) \times e}$. An element $\widehat{\omega}$ of $\widehat{\Omega}$ can be written as $\widehat{\omega} = (\omega, y)$, where $y = (y^1, \ldots, y^e) \in \mathbb{R}^{(N-1) \times e}$ with $y^i =$

$(y_1^j, \dots, y_{N-1}^j)$. Define a mapping $\mathtt{i} : \widehat{\Omega} \to \Omega$ by $\mathtt{i}(\widehat{\omega}) = \omega$, which is clearly surjective. We also introduce the process Y as $Y_k(\widehat{\omega}) = y_k = (y_k^1, \dots, y_k^e)$ for $k \in \{1, \dots, N-1\}$, $Y_0(\widehat{\omega}) = 0$, and $Y_N(\widehat{\omega}) = g(\widehat{\omega}) = g(\omega)$. Let $\mathcal{Y}_k := \sigma(Y_n : n \leq k)$ and $\widehat{\mathcal{F}}_k$ be the universal completion of $\mathcal{F}_k \otimes \mathcal{Y}_k$, so that one obtains the filtration $\widehat{\mathbb{F}} := (\widehat{\mathcal{F}}_k)_{k=0,1,\dots,N}$. In this context, we define $\mathtt{j}_\mathbb{Q}(\omega) := (\omega, (y_k^{\mathbb{Q},i})_{i \leq e, k \leq N-1})$, where $y_k^{\mathbb{Q},i}$ is a version of $\mathbb{E}^\mathbb{Q}[g^i | \mathcal{F}_k](\omega)$ for each $\mathbb{Q} \in \mathcal{M}_g$. Then, it is clear that $\mathtt{i} \circ \mathtt{j}_\mathbb{Q} = id$ and $\mathcal{L}_{\mathbb{Q} \circ \mathtt{j}_\mathbb{Q}^{-1}}(Y) = \mathcal{L}_\mathbb{Q}(Y^\mathbb{Q})$. This implies that $\mathtt{J} : \mathbb{Q} \mapsto \mathbb{Q} \circ \mathtt{j}_\mathbb{Q}^{-1}$ satisfies Property 4 of Definition 2.13. In line with Remark 2.14, the inverse of J, given by $\widehat{\mathcal{M}} \ni \widehat{\mathbb{Q}} \mapsto \widehat{\mathbb{Q}} \circ \mathtt{i}^{-1} = \widehat{\mathbb{Q}}_{|\Omega} \in \mathcal{M}_g$ is surjective.

We consider a class of functions $\widehat{\Upsilon}$ on $\widehat{\Omega}$ and assume that $\Upsilon \subset \widehat{\Upsilon}$ in the sense that for $\xi \in \Upsilon$, $\xi(\widehat{\omega}) := \xi(\mathtt{i}(\widehat{\omega}))$ belongs to $\widehat{\Upsilon}$. The relationship between $\widehat{\mathcal{M}}$ and $\mathcal{M}_g$ observed in Remark 2.14 then yields

$$\sup_{\mathbb{Q} \in \widehat{\mathcal{M}}} \mathbb{E}^\mathbb{Q}[\xi] \geq \sup_{\mathbb{Q} \in \mathcal{M}_g} \mathbb{E}^\mathbb{Q}[\xi] \quad \text{for any } \xi \in \Upsilon.$$

We can apply the enlargement construction introduced in Section 2.2 to the space $\widehat{\Omega}$, which leads to the set of martingale measures $\overline{\widehat{\mathcal{M}}}$ on $\widehat{\Omega}$, and the above inequality extends to

$$\sup_{\mathbb{Q} \in \overline{\widehat{\mathcal{M}}}} \mathbb{E}^\mathbb{Q}[\Phi] \geq \sup_{\mathbb{Q} \in \mathcal{M}_g} \mathbb{E}^\mathbb{Q}[\Phi] \quad \text{for any } \Phi \in \overline{\Upsilon}. \tag{20}$$

We now consider a dynamic extension $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, \widehat{\mathcal{P}}, Y, \mathtt{i}, \mathtt{J})$ of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$ as a fictitious market in which we can trade dynamically in $\widehat{S} = (S, Y)$ using the class of trading strategies $\widehat{H}$, which are the $\widehat{\mathbb{F}}$-predictable $\mathbb{R}^{\widehat{\Lambda}}$-valued processes, which have only finitely many nonzero coordinates, where $\widehat{\Lambda} = \{(i, s) : i \in \{1, \dots, d\}\} \cup \{(\lambda, y) : \lambda \in \Lambda\}$, that is,

$$\widehat{\mathcal{H}} = \left\{ \widehat{H} = \left(\widehat{H}_k^{\widehat{\lambda}} : \widehat{\lambda} \in \widehat{\Lambda} \right)_{k \leq N} : \widehat{\mathbb{F}}\text{-predictable } \mathbb{R}^{\widehat{\Lambda}}\text{-valued process s.t.} \right.$$

$$\left. \exists \text{ finite subset } \widehat{\Lambda}_0 \subset \widehat{\Lambda} \text{ s.t.} \widehat{H}_k^{\widehat{\lambda}} = 0, \; \forall k, \forall \widehat{\lambda} \notin \widehat{\Lambda}_0 \right\}.$$

A self-financing strategy corresponds to a choice of $\widehat{H} \in \widehat{\mathcal{H}}$ and yields a final payoff of

$$(\widehat{H} \circ \widehat{S})_N = \sum_{j=1}^d \sum_{k=1}^N \widehat{H}_k^{(j,s)} \Delta S_k^j + \sum_{\lambda \in \Lambda} \sum_{k=1}^N \widehat{H}_k^{(\lambda,y)} \Delta Y_k^\lambda. \tag{21}$$

Note that our choice of trading strategies ensures that the sums are finite. The supehedging costs of a European option $\widehat{\xi}$ and an American option $\widehat{\Phi} = (\widehat{\Phi}_k)_{k \leq N}$ on $\widehat{\Omega}$ are given by

$$\widehat{\pi}^E(\widehat{\xi}) = \inf\{x : \exists \widehat{H} \in \widehat{\mathcal{H}} \text{ s.t. } x + (\widehat{H} \circ \widehat{S})_N \geq \widehat{\xi}, \; \widehat{\mathcal{P}}\text{-q.s.}\},$$

$$\widehat{\pi}^A(\widehat{\Phi}) = \inf\{x : \exists ({}^1\widehat{H}, \dots, {}^N\widehat{H}) \in \widehat{\mathcal{H}}^N \text{ s.t. } {}^j\widehat{H}_i = {}^k\widehat{H}_i \; \forall \, 1 \leq i \leq j \leq k \leq N \tag{22}$$

$$\text{and s.t. } x + ({}^k\widehat{H} \circ \widehat{S})_N \geq \widehat{\Phi}_k, \; \forall k = 1, \dots, N, \; \widehat{\mathcal{P}}\text{-q.s.}\}.$$

Remark 2.16. Clearly $\widehat{\mathbb{F}}$ is much richer than $\mathbb{F}$ as it captures not only the evolution of prices of S but also of all the vanilla options. The inequality $\widehat{\pi}^A(\Phi) \leq \pi_g^A(\Phi)$ holds trivially as a buy-and-hold strategy is a special case of a dynamic trading strategy and $\mathcal{P} = \widehat{\mathcal{P}} \circ \mathtt{i}^{-1}$.

The following result shows that if the pricing–hedging duality holds then the superhedging prices in the fictitious dynamic extension market are the same as in the original market. This will apply to the setups we consider in Sections 3 and 4 below.

Proposition 2.17. *Let $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, \widehat{\mathcal{P}}, Y, \mathtt{i}, \mathtt{J})$ be a dynamic extension of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$ with its superhedging prices given by (22).*

(a) *Assume that the European pricing–hedging duality holds for the class Υ on Ω. Then*

$$\pi_g^E(\xi) = \widehat{\pi}^E(\xi), \quad \xi \in \Upsilon.$$

(b) *Assume that the European pricing–hedging duality holds for the class Υ^N on $\overline{\Omega}$. Then*

$$\pi_g^A(\Phi) = \widehat{\pi}^A(\Phi), \quad \Phi \in \Upsilon^N.$$

Proof. Note that $\pi_g^A \geq \widehat{\pi}^A$ holds by Remark 2.16. Using (7) twice we obtain

$$\overline{\pi}_g^E(\Phi) = \pi_g^A(\Phi) \geq \widehat{\pi}^A(\Phi) = \overline{\widehat{\pi}}^E(\Phi) \geq \sup_{\widehat{\mathbb{Q}} \in \widehat{\overline{\mathcal{M}}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\Phi] \geq \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi], \tag{23}$$

where the penultimate inequality always holds by Remark 2.2 and the last inequality follows by (20). The assumed pricing–hedging duality on $\overline{\Omega}$ implies that we have equalities throughout. The proof of (a) is analogous but simpler. □

Remark 2.18. The above result may at first seem surprising. The dynamic extension introduces many new dynamically traded assets, yet the superhedging prices remain the same. The intuition behind this is that under pricing–hedging duality, the cheapest superhedge is a perfect hedge (or very nearly so) under some (worst case) model. Our dynamic extensions do not introduce any constraints on the prices of options g and hence do not restrict the set of martingale measures. The worst case model will remain an admissible model and for this model the additional traded assets make no difference. They could however make a difference in many other (specific) models. If we considered a restricted version of dynamic trading in which we make further assumptions about the price dynamics of vanilla options, then this could imply that $\mathtt{J} \circ \mathtt{i}^{-1}$ is not surjective and the superhedging prices might strictly decrease. Such a setup is studied in Nadtochiy and Obłój (2017), where the authors consider restrictions on the levels of implied volatility through time.

Let $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, \widehat{\mathcal{P}}, Y, \mathtt{i}, \mathtt{J})$ be a dynamic extension of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$ and $\widehat{\mathcal{E}}_k$ be a family of operators on the space $\widehat{\Upsilon}$ of functionals on $\widehat{\Omega}$. One can define the corresponding extended operators $\widehat{\mathcal{E}}^k$ as well as $\overline{\widehat{\mathcal{E}}}^k$ as in (9) and (10). We can then apply Theorem 2.8 and Proposition 2.17 to obtain the following result:

Corollary 2.19. *Let $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, \widehat{\mathcal{P}}, Y, \mathtt{i}, \mathtt{J})$ be a dynamic extension of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$ with operators $\widehat{\mathcal{E}}_k : \widehat{\Upsilon} \to \widehat{\Upsilon}$ and the corresponding extended operators $\widehat{\mathcal{E}}^k$ as well as $\overline{\widehat{\mathcal{E}}}^k$ satisfying (9) and (13). Assume that the European pricing–hedging duality holds for the class Υ^N on $\overline{\Omega}$, and $f \vee f' \in \Upsilon$ for all $f, f' \in \Upsilon$. Then, for all $\Phi \in \Upsilon^N$,*

$$\pi_g^A(\Phi) = \widehat{\pi}^A(\Phi) = \sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \sup_{\widehat{\tau} \in \mathcal{T}(\widehat{\mathbb{F}})} \mathbb{E}^{\widehat{\mathbb{Q}}}[\Phi_{\widehat{\tau}}] = \sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\Phi] = \sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\Phi_{\widehat{\tau}^*}], \tag{24}$$

$$(3, \mathbf{0}, \mathit{3/4}) \quad \overset{3|4}{\underset{1/4}{\overline{0}}}$$

$$(4, \mathbf{8}, \mathit{1})$$
$$(2, \mathbf{0}, \mathit{0})$$
$$(0, \mathbf{0}, \mathit{0})$$

$$(2, \mathit{2/5}) \quad \overset{1/2}{\diagup} \quad (3, \mathbf{0}, \mathit{3/4})$$

$$\overset{1/10}{\overline{}} (1, \mathbf{1}, \mathit{1/4}) \quad \overset{1|4}{\underset{3/4}{\overline{0}}}$$

$$(4, \mathbf{8}, \mathit{1})$$
$$(2, \mathbf{0}, \mathit{0})$$
$$(0, \mathbf{0}, \mathit{0})$$

$$\overset{2/5}{\diagdown} \quad (1, \mathbf{1}, \mathit{0}) \quad \overset{1/2}{\underset{1/2}{\overline{0}}}$$

$$(4, \mathbf{8}, \mathit{1})$$
$$(2, \mathbf{0}, \mathit{0})$$
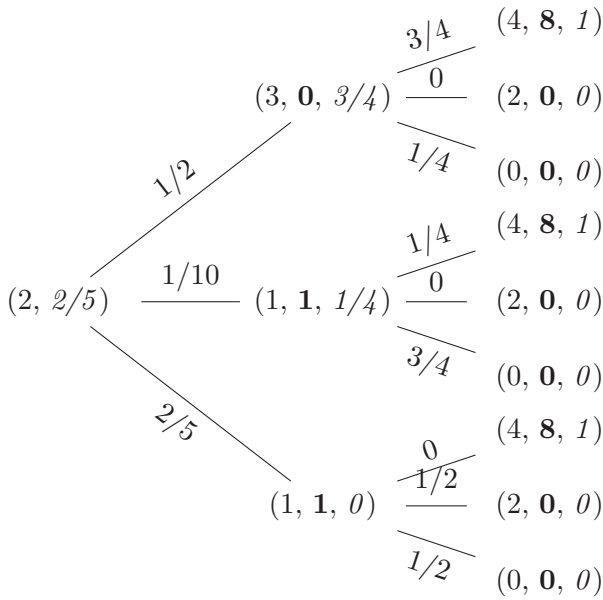$$(0, \mathbf{0}, \mathit{0})$$

FIGURE 2 The model on $\widehat{\Omega}$ which corresponds to the mixture model in Hobson and Neuberger (2016) attaining the superhedging price; prices of the stock are written in regular font, payoffs of the American option in bold and prices of European option in italic

where

$$\hat{\tau}^* := \min\{k \geq 1 : \widehat{\mathcal{E}}^k(\Phi(\cdot, k))(\omega) = \overline{\widehat{\mathcal{E}}}^k(\Phi)(\omega, k)\}.$$

Remark 2.20. In Section 3, in the context of Bouchard and Nutz (2015), we will adopt the dynamic extension introduced in Example 2.15, and show that it admits a family of operators $\widehat{\mathcal{E}}_k$ to which we can apply Corollary 2.19.

Remark 2.21. We believe that Corollary 2.19 describes a canonical, and in some sense *minimal*, solution to the pricing–hedging duality of American option, when compared to addition of all consistent random times, as discussed in Remark 2.6. A dynamic extension $\widehat{\Omega}$ is crucial to establish the Dynamic Programming Principle (DPP), which in turn allows one to define the optimal stopping time $\hat{\tau}^*$.

Remark 2.22. Let us consider the two period ($N = 2$) example of Hobson and Neuberger (2016); see Figure 2. For simplicity, we introduce only one statically traded option g with payoff $\mathbb{1}_{\{S_2=4\}}$ at time $t = 2$ and price $2/5$ at time $t = 0$. This already destroys the pricing–hedging duality for the American option Φ. In Hobson and Neuberger (2016), the duality is recovered by considering a (calibrated) mixture of martingale measures. It is insightful to observe that their mixture model is nothing else but a martingale measure for an augmented setup with dynamic trading in g, which, following Corollary 2.19, restores the dynamic programming principle and the pricing–hedging duality for American options. To show this, let Y denote the price process of the option g, so that $Y_0 = 2/5$ and $Y_2 = g$. Figure 2 illustrates a martingale measure $\mathbb{Q}$ along with the intermediate prices Y_1 such that the processes S and Y are martingales. With $\tau = \mathbb{1}_{\{S_1=1,Y_1=0\}} + 2\mathbb{1}_{\{S_1=1,Y_1=1/4\}\cup\{S_1=3\}}$, we find $\mathbb{E}^{\mathbb{Q}}[\Phi_\tau] = 18/5$, which is the superhedging price, and the duality is recovered.

2.4 | Pseudo-stopping times

In this subsection we study the connection of our problem to pseudo-stopping times in the filtration $\mathbb{F}$ which form a bigger class than $\mathbb{F}$-stopping times. We refer the reader to Williams (2002), Nikeghbali and Yor (2005) and Mansuy and Yor (2006) for an introduction to pseudo-stopping times.

It follows from Theorem 2.4 that in general we expect to see

$$\pi_g^A(\Phi) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \geq \sup_{\mathbb{Q} \in \mathcal{M}_g} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau],$$

where the inequality may be strict. We showed above that this is linked with the necessity to use random times beyond $\tau \in \mathcal{T}(\mathbb{F})$. To conclude our general results, we explore this property from another angle and identify the subset(s) of $\overline{\mathcal{M}}_g$, which lead to equality in the place of the inequality above. We introduce

$$\underline{\mathcal{M}}_g := \{\overline{\mathbb{Q}} \in \mathfrak{P}(\overline{\Omega}) : \overline{\mathbb{Q}} \ll \overline{\mathcal{P}}, \ \mathbb{E}^{\overline{\mathbb{Q}}}[g^\lambda] = 0, \ \lambda \in \Lambda, \ S \text{ is an } (\mathbb{F}, \overline{\mathbb{Q}})\text{-martingale},$$

$$\mathbb{E}^{\overline{\mathbb{Q}}}[M_T] = \mathbb{E}^{\overline{\mathbb{Q}}}[M_0], \quad \text{for all bounded } (\mathbb{F}, \overline{\mathbb{Q}}) \text{-martingales } M\}, \tag{25}$$

as the set of measures that make S an $\mathbb{F}$-martingale and T an $\mathbb{F}$-pseudo-stopping time. These are natural because the martingale part of the Snell envelope can be stopped at the pseudo-stopping time with null expectation.

Proposition 2.23. *Assume that $\mathcal{M}_g \neq \emptyset$. Then*

$$\sup_{\overline{\mathbb{Q}} \in \underline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \sup_{\mathbb{Q} \in \mathcal{M}_g} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau]. \tag{26}$$

Proof. Let $\overline{\mathbb{Q}} \in \underline{\mathcal{M}}_g$ such that $\mathbb{E}^{\overline{\mathbb{Q}}}[|g^\lambda|] < \infty$ and $\mathbb{E}^{\overline{\mathbb{Q}}}[|\Phi_k|] < \infty$ for all $\lambda \in \Lambda$ and $k = 1, \ldots, N$. We next consider the optimal stopping problem $\sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi_\tau]$. Define its Snell envelope $(Z_k)_{0 \leq k \leq N}$ by

$$Z_k := \operatorname*{ess\,sup}_{\tau \in \mathcal{T}(\mathbb{F}), \tau \geq k} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi_\tau | \mathcal{F}_k],$$

which is an $(\mathbb{F}, \overline{\mathbb{Q}})$-supermartingale. Its Doob–Meyer decomposition is given by

$$Z_k = Z_0 + M_k - A_k, \quad \text{where } A = (A_k)_{0 \leq k \leq N} \text{ is an } \mathbb{F}\text{-predictable increasing process,}$$

and $A_0 = M_0 = 0$. It follows that

$$\mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \mathbb{E}^{\overline{\mathbb{Q}}}[Z_T] \leq Z_0 + \mathbb{E}^{\overline{\mathbb{Q}}}[M_T] = Z_0. \tag{27}$$

We deduce that $\sup_{\overline{\mathbb{Q}} \in \underline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \sup_{\mathbb{Q} \in \mathcal{M}_g} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau]$. Then, (26) holds as every stopping time $\tau \in \mathcal{T}(\mathbb{F})$ is a pseudo-stopping time and hence the inverse inequality is trivial. □

Remark 2.24. The above allows us to see that it is not enough to use randomized stopping times to recover the equality in (5). Such a time corresponds to an $\mathbb{F}$-adapted increasing process V with $V_0 = 0$ and $V_N = 1$. It may be seen as a distribution over all possible stopping times, in our setup a distribution

η on $\mathbb{T}$ such that $\eta(\{k\}) := \Delta V_k = V_k - V_{k-1}$ for each $k \in \mathbb{T}$. For any pseudo-stopping time τ, the dual optional projection of the process $\mathbb{1}_{[\![\tau, N]\!]}$ is a randomized stopping time. Conversely, for a given V, if we take a uniformly distributed random variable Θ independent of V, possibly enlarging the probability space, then $\tau := \inf\{t : V_t \geq \Theta\}$ is $\mathbb{F}$-pseudo-stopping time, which generates V. Let $\mathcal{R}$ be the set of such randomized stopping times. Then, from Proposition 2.23 and the definition of the dual optional projection,

$$\sup_{\mathbb{Q} \in \mathcal{M}_g} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau] = \sup_{\mathbb{Q} \in \mathcal{M}_g} \sup_{V \in \mathcal{R}} \mathbb{E}^{\mathbb{Q}} \left[\sum_k \Phi_k \Delta V_k \right].$$

Remark 2.25. Nikeghbali and Yor (2005) showed that under a progressive enlargement with pseudo-stopping time τ, all martingales from the smaller filtration stopped at τ remain martingales in the larger filtration. One can relate this to a more restrictive situation, when all martingales from the smaller filtration remain martingales in the bigger filtration, which is called the immersion property in the context of filtration enlargement. Clearly each random time satisfying the immersion property is a pseudo-stopping time. Thus, keeping the equality (26) true, the pseudo-stopping time property in the definition of $\underline{\mathcal{M}}_g$ above can be replaced by a stronger condition characterizing the immersion property,

$$\overline{\mathbb{Q}}[T > k | \mathcal{F}_n] = \overline{\mathbb{Q}}[T > k | \mathcal{F}_k], \quad \text{for all} \quad 0 \leq k \leq n \leq N. \tag{28}$$

See section 3.1.2 of Blanchet-Scalliet, Jeanblanc, and Romero (2016) for the discrete time context of progressive enlargement of filtration and Aksamit and Li (2016) for connections between pseudo-stopping times, the immersion property, and projections.

3 | A DETAILED STUDY OF THE NONDOMINATED SETUP OF BOUCHARD AND NUTZ (2015)

In this section we work in the nondominated setup introduced in Bouchard and Nutz (2015), which is a special case of Example 2.12. We let $\Omega_0 = \{\omega_0\}$ be a singleton and Ω_1 be a Polish space. For each $k \in \{1, \dots, N\}$, we define $\Omega_k := \{\omega_0\} \times \Omega_1^k$ as the k-fold Cartesian product. For each k, we denote $\mathcal{G}_k := \mathcal{B}(\Omega_k)$ and by $\mathcal{F}_k$ its universal completion. In particular, $\mathcal{G}_0$ and $\mathcal{F}_0$ are trivial, and $\mathbb{E}^{\mathbb{P}}[\xi | \mathcal{G}_k] = \mathbb{E}^{\mathbb{P}}[\xi | \mathcal{F}_k]$ for all $\xi \in \mathcal{F}_N$ and every probability measure $\mathbb{P}$ on $(\Omega_N, \mathcal{F}_N)$. We shall often see $\mathcal{G}_k$ and $\mathcal{F}_k$ as sub-σ-fields of $\mathcal{F}_N$, and hence obtain two filtrations $\mathbb{G} = (\mathcal{G}_k)_{0 \leq k \leq N}$ and $\mathbb{F} = (\mathcal{F}_k)_{0 \leq k \leq N}$ on Ω. Denote

$$\Omega := \Omega_N, \mathcal{G} := \mathcal{G}_N \quad \text{and} \quad \mathcal{F} := \mathcal{F}_N.$$

Recall that a subset of a Polish space Ω is analytic if it is the image of a Borel subset of another Polish space under a Borel measurable mapping. We take Υ to be the class of upper semianalytic functions $f : \Omega \to \overline{\mathbb{R}} := [-\infty, \infty]$, that is, such that $\{\omega \in \Omega : f(\omega) > c\}$ is analytic for all $c \in \mathbb{R}$.

The price process S is a $\mathbb{G}$-adapted $\mathbb{R}^d$-valued process and the collection of options $g = (g^1, \dots, g^e)$ is a $\mathcal{G}$-measurable $\mathbb{R}^e$-valued vector for $e \in \mathbb{N}$ (thus $\Lambda = \{1, \dots, e\}$).

Let $k \in \{0, \dots, N-1\}$ and $\omega \in \Omega_k$. We are given a nonempty convex set $\mathcal{P}_k(\omega) \subseteq \mathfrak{P}(\Omega_1)$ of probability measures, which represents the set of all possible models for the $(k+1)$th period, given state ω at times $0, 1, \dots, k$. We assume that for each k,

$$\text{graph}(\mathcal{P}_k) := \{(\omega, \mathbb{P}) : \omega \in \Omega_k, P \in \mathcal{P}_k(\omega)\} \subseteq \Omega_k \times \mathfrak{P}(\Omega_1) \text{ is analytic.} \tag{29}$$

Given a universally measurable kernel $\mathbb{P}_k : \Omega_k \to \mathfrak{P}(\Omega_1)$ for each $k \in \{0, 1, \ldots, N-1\}$, we define a probability measure $\mathbb{P} = \mathbb{P}_0 \otimes \mathbb{P}_1 \otimes \cdots \otimes \mathbb{P}_{N-1}$ on Ω by

$$\mathbb{P}(A) := \int_{\Omega_1} \cdots \int_{\Omega_1} \mathbf{1}_A(\omega_1, \omega_2, \ldots, \omega_N) \mathbb{P}_{N-1}(\omega_1, \ldots, \omega_{N-1}; d\omega_N) \cdots \mathbb{P}_0(d\omega_1).$$

We can then introduce the set $\mathcal{P} \subseteq \mathfrak{P}(\Omega)$ of possible models for the multiperiod market up to time N by

$$\mathcal{P} := \left\{ \mathbb{P}_0 \otimes \mathbb{P}_1 \otimes \cdots \otimes \mathbb{P}_{N-1} : \mathbb{P}_k(\cdot) \in \mathcal{P}_k(\cdot), k = 0, 1, \ldots, N-1 \right\}. \tag{30}$$

Notice that the condition (29) ensures that $\mathcal{P}_k$ always has a universally measurable selector: $\mathbb{P}_k : \Omega_k \to \mathfrak{P}(\Omega_1)$ such that $\mathbb{P}_k(\omega) \in \mathcal{P}_k(\omega)$ for all $\omega \in \Omega_k$. Then the set $\mathcal{P}$ defined in (30) is nonempty. We also denote

$$\mathcal{M}_{k,k+1}(\omega) := \{\mathbb{Q} \in \mathfrak{P}(\Omega_1) : \mathbb{Q} \ll \mathcal{P}_k(\omega) \quad \text{and} \quad \mathbb{E}^{\delta_\omega \otimes_k \mathbb{Q}}[\Delta S_{k+1}] = 0\}, \tag{31}$$

where $\delta_\omega \otimes_k \mathbb{Q} := \delta_{(\omega_1, \ldots, \omega_k)} \otimes \mathbb{Q}$ is a Borel probability measure on $\Omega_{k+1} := \Omega_k \times \Omega_1$, and

$$\mathcal{M}_k(\omega) := \{\delta_\omega \otimes_k \mathbb{Q}_k \otimes \cdots \otimes \mathbb{Q}_{N-1} : \mathbb{Q}_i(\cdot) \in \mathcal{M}_{i,i+1}(\cdot), i = k, \ldots, N-1\}. \tag{32}$$

The following notion of no-arbitrage NA($\mathcal{P}$) has been introduced in Bouchard and Nutz (2015): NA($\mathcal{P}$) holds if for all $(H, h) \in \mathcal{H} \times \mathbb{R}^e$,

$$(H \circ S)_N + hg \geq 0 \quad \mathcal{P}\text{-q.s.} \quad \Rightarrow \quad (H \circ S)_N + hg = 0 \quad \mathcal{P}\text{-q.s.}$$

Analogously, we will say that NA($\overline{\mathcal{P}}$) holds if for all $(\overline{H}, h) \in \overline{\mathcal{H}} \times \mathbb{R}^e$,

$$(\overline{H} \circ S)_N + hg \geq 0 \quad \overline{\mathcal{P}}\text{-q.s.} \quad \Rightarrow \quad (\overline{H} \circ S)_N + hg = 0 \quad \overline{\mathcal{P}}\text{-q.s.} \tag{33}$$

Recall also that $\mathcal{M}_g$ and $\overline{\mathcal{M}}_g$ have been defined in (3) and (6). As established in Bouchard and Nutz (2015), the condition NA($\mathcal{P}$) is equivalent to the statement that $\mathcal{P}$ and $\mathcal{M}_g$ have the same polar sets. The following lemma extends this result to $\overline{\Omega}$.

Lemma 3.1. NA($\mathcal{P}$) $\Longleftrightarrow$ NA($\overline{\mathcal{P}}$) $\Longleftrightarrow$ [$\overline{\mathcal{P}}$ and $\overline{\mathcal{M}}_g$ have the same polar sets.]

Proof. The two conditions NA($\mathcal{P}$) and NA($\overline{\mathcal{P}}$) are equivalent by the same arguments as in proving (7). It is enough to show that $\overline{\mathcal{P}}$ and $\overline{\mathcal{M}}_g$ have the same polar sets if and only if $\mathcal{P}$ and $\mathcal{M}_g$ have the same polar sets. This boils down to proving that a set $\Gamma \in \overline{\Omega}$ is an $\overline{\mathcal{M}}_g$ polar set if and only if the k-section $\Gamma_k = \{\omega : (\omega, k) \in \Gamma\}$ is an $\mathcal{M}_g$ polar set for each $k \in \mathbb{T}$, which can be shown similarly to the analogous statement involving $\mathcal{P}$ and $\overline{\mathcal{P}}$ established in the proof of Theorem 2.4. $\square$

3.1 | Duality on the enlarged space $\overline{\Omega}$

Our first main result is the following duality under the no-arbitrage condition (33):

Theorem 3.2. *Let* $\mathrm{NA}(\overline{\mathcal{P}})$ *hold. Then the set* $\overline{\mathcal{M}}_g$ *is nonempty, and, for any upper semianalytic* Φ : $\overline{\Omega} \to \mathbb{R}$, *one has*

$$\overline{\pi}_g^E(\Phi) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi], \tag{34}$$

and in particular the pricing–hedging duality (8) holds. Moreover, there exists $(\overline{H}, h) \in \overline{\mathcal{H}} \times \mathbb{R}^e$ *such that*

$$\overline{\pi}_g^E(\Phi) + (\overline{H} \circ S)_N + hg \geq \Phi, \quad \overline{\mathcal{P}}\text{-q.s.}$$

The proof is postponed to Section 6 and uses the following lemma. Let us work with the operators $\mathcal{E}_k$ introduced in Example 2.12 with $\mathcal{M}_k(\omega)$ defined as in (32). Observe that

$$\mathcal{E}_k \circ \cdots \circ \mathcal{E}_{N-1}(\xi)(\omega) = \mathcal{E}_{k,k+1} \circ \cdots \circ \mathcal{E}_{N-1,N}(\xi)(\omega), \quad \xi \in \Upsilon,$$

where $\mathcal{E}_{k,k+1}(\xi)(\omega) = \sup_{\mathcal{M}_{k,k+1}(\omega)} \mathbb{E}^{\mathbb{Q}}[\xi]$. By Proposition 2.9, (4.12) in Bouchard and Nutz (2015) and using that the maximum of upper semianalytic functions is still upper semianalytic, we conclude the following:

Lemma 3.3. *Consider the case* $e = 0$, *that is,* $\Lambda = \emptyset$. *Let* $\Phi \in \overline{\Upsilon}$. *Then* $\overline{\mathcal{E}}_k(\Phi)$ *in (10) is also upper semianalytic and*

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \overline{\mathcal{E}}^0(\Phi) := \overline{\mathcal{E}}_0 \circ \cdots \circ \overline{\mathcal{E}}_{N-1}(\Phi).$$

3.2 | Dynamic programming principle on $\widehat{\Omega}$

We consider a dynamic extension $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, \widehat{\mathcal{P}}, Y, \mathtt{i}, \mathtt{J})$ of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$ as defined in Example 2.15, where $\widehat{\Omega} = \Omega \times \mathbb{R}^{(N-1)\times e}$ has the same product structure as Ω, and $\widehat{\mathcal{P}}$ is defined by (19). Moreover, $\widehat{\mathcal{G}}_k := \mathcal{G}_k \otimes \sigma(Y_i, i \leq k)$ is countably generated and $\widehat{\mathcal{F}}_k$ is the universal completion of $\widehat{\mathcal{G}}_k$. For $k = 0, \ldots, N-1$ and $\widehat{\omega} = (\omega, y) \in \widehat{\Omega}$, we define

$$\widehat{\mathcal{P}}_k(\omega, y) := \{\widehat{\mathbb{P}} : \widehat{\mathbb{P}}|_{\Omega_1} \in \mathcal{P}_k(\omega)\}.$$

By the definition of $\mathcal{P}$ in (30), it is clear that $\widehat{\mathcal{P}}$ has the same product structure, that is,

$$\widehat{\mathcal{P}} = \{\widehat{\mathbb{P}}_0 \otimes \cdots \otimes \widehat{\mathbb{P}}_{N-1} : \widehat{\mathbb{P}}_k(\cdot) \in \widehat{\mathcal{P}}_k(\cdot), \ k = 0, \ldots, N-1\}.$$

Further, recalling the definitions of $\mathcal{M}_{k,k+1}(\omega)$ and $\mathcal{M}_k(\omega)$ in (31) and (32), we define

$$\widehat{\mathcal{M}}_{k,k+1}(\omega, y) := \{\widehat{\mathbb{Q}} \in \mathfrak{P}(\Omega_1 \times \mathbb{R}^e) : \widehat{\mathbb{Q}}|_{\Omega_1} \in \mathcal{M}_{k,k+1}(\omega), \mathbb{E}^{\widehat{\mathbb{Q}}}[(\Delta S_{k+1}, \Delta Y_{k+1})] = 0\},$$

and

$$\widehat{\mathcal{M}}_k(\widehat{\omega}) := \{\delta_{\widehat{\omega}} \otimes_k \widehat{\mathbb{Q}}_k \otimes \cdots \otimes \widehat{\mathbb{Q}}_{N-1} : \widehat{\mathbb{Q}}_i(\cdot) \in \widehat{\mathcal{M}}_{i,i+1}(\cdot), i = k, \ldots, N-1\},$$

with $\Delta S_{k+1} := S_{k+1} - S_k$ and $\Delta Y_{k+1} := Y_{k+1} - Y_k$. Similarly to the definition of $\mathcal{E}_k$ in Example 2.12 with $\mathcal{M}_k(\omega)$, we define a family of operators $(\widehat{\mathcal{E}}_k)$ on functionals on $\widehat{\Omega}$ using $\widehat{\mathcal{M}}_k(\widehat{\omega})$, and show in the

following theorem that it provides a dynamic programming representation of $\sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\,\cdot\,]$, where $\widehat{\mathcal{M}}$ is defined in Property 4 of Definition 2.13. The proof of Theorem 3.4 is left to Section 6.3.

Theorem 3.4. *Let* $\widehat{\xi} : \widehat{\Omega} \to \mathbb{R}$ *be an upper semianalytic functional. Then,* $\widehat{\mathcal{E}}_k(\widehat{\xi})$ *is also upper semianalytic and*

$$\sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}] = \sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_0} \mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}] = \widehat{\mathcal{E}}^0(\widehat{\xi}) \; := \; \widehat{\mathcal{E}}_0 \circ \cdots \circ \widehat{\mathcal{E}}_{N-1}(\widehat{\xi}). \tag{35}$$

Remark 3.5.

(a) The above result is in fact a classical dynamic programming principle result studied in Bertsekas and Shreve (2007) and Dellacherie (1985). The only crucial step is to prove that the graph set $[[\widehat{\mathcal{M}}_{k,k+1}]] := \{(\widehat{\omega}, \mathbb{Q}) \; : \; \mathbb{Q} \in \mathcal{M}_{k,k+1}(\widehat{\omega})\}$ is analytic.

(b) Assume that NA($\mathcal{P}$) holds. It then follows by Lemma 3.1 and Theorem 3.2 that the pricing–hedging duality on $\overline{\Omega}$ in (34) holds. Further, by defining $\widehat{\mathcal{E}}^k$ and $\overline{\widehat{\mathcal{E}}}^k$ with $\widehat{\mathcal{E}}_k$ as in (9) and (11), one has that (9) holds for $\widehat{\mathcal{E}}^k$ from Theorem 3.4, and moreover that (13) holds for $\overline{\widehat{\mathcal{E}}}^k$ as it is a special case of Example 2.12. It then follows by Corollary 2.19 that (24) holds in this framework.

3.3 | Comparison with Bayraktar et al. (2015) and Bayraktar and Zhou (2017)

In Bayraktar et al. (2015) the authors considered the same superhedging problem $\pi_g^A(\Phi)$ with the finite set $\Lambda = \{1, \dots, e\}$, and established the duality

$$\pi_g^A(\Phi) = \inf_{h \in \mathbb{R}^e} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \sup_{\mathbb{Q} \in \mathcal{M}_0} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau - hg], \tag{36}$$

under some regularity conditions (see proposition 3.1 in Bayraktar et al., 2015). Our duality in Theorem 3.2 is more general and more complete, and moreover, together with Lemma 3.3, it induces the above duality (36). In exchange, Bayraktar et al. (2015) also studied another subhedging problem $\sup_{\tau \in \mathcal{T}(\mathbb{F})} \inf_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau]$, which we do not consider here.

More recently, Bayraktar and Zhou (2017) consider the "randomized" stopping times, and obtain a more complete duality for $\pi_g^A(\Phi)$. The dual formulations in Bayraktar and Zhou (2017) and in our results are more or less in the same spirit, as in Neuberger (2007) and Hobson and Neuberger (2017). Nevertheless, the duality in Bayraktar and Zhou (2017) is established under strong integrability conditions and an abstract condition, which is checked under regularity conditions (see their assumption 2.1 and remark 2.1). In particular, when $\mathcal{P}$ is the class of all probability measures on Ω, the integrability condition in their assumption 2.1 is equivalent to saying that Φ_k and g^i are all uniformly bounded. In our paper, we only assume that g^i are Borel measurable, Φ_k are upper semianalytic and all are $\mathbb{R}$-valued.

Technically, Bayraktar and Zhou (2017) use the duality results in Bouchard and Nutz (2015) together with a minimax theorem to prove their results. Our first main result consists of introducing an enlarged canonical space (together with an enlarged canonical filtration) to reformulate the main problem as a superhedging problem for European options. Then, by adapting the arguments in Bouchard and Nutz (2015), we establish our duality under general conditions as in Bouchard and Nutz (2015). Moreover, we do not assume that Φ_k is $\mathcal{F}_k$-measurable, which enables us to study the superhedging problem for a portfolio containing an American option and some European options. Finally, our setting enables us to

apply an approximation argument to study a new class of martingale optimal transport problems and to obtain a Kantorovich duality as in Section 4.

4 | A MARTINGALE (OPTIMAL) TRANSPORT SETUP

In this section, we study the duality for American options in a martingale optimal transport setup, with canonical space $\Omega := \{s_0\} \times \mathbb{R}^{d \times N}$ for some $s_0 \in \mathbb{R}^d$, the canonical process S on Ω, and $\mathcal{P} := \mathfrak{P}(\Omega)$. Then, with $\overline{\Omega} := \Omega \times \mathbb{T}$, we have $\overline{\mathcal{P}} = \mathfrak{P}(\overline{\Omega})$. We assume that the statically traded options on the market are all vanilla options and are arbitrage-free, see Cox and Obłój (2011) and Cox, Hou, and Obłój (2016), and numerous enough such that one can recover the marginal distribution of the underlying process S at some maturity times $\mathbb{T}_0 = \{t_1, \dots, t_M\} \subseteq \mathbb{T}$, where $t_M = N$. More precisely, we are given a vector $\mu = (\mu_1, \dots, \mu_M)$ of marginal distributions. We write $\mu(f) := (\int f(x)\mu_1(dx), \dots, \int f(x)\mu_M(dx))$ and we assume that $\mu(|\cdot|) < \infty$ and

$$\mu_i(f) \leq \mu_j(f) \quad \text{for all } i \leq j, \ i, j \leq M, \quad \text{and any convex function } f : \mathbb{R}^d \to \mathbb{R}. \tag{37}$$

The condition (37) ensures the existence of a calibrated martingale measure, that is, that the following sets are nonempty:

$$\mathcal{M}_\mu := \left\{ \mathbb{Q} \in \mathfrak{P}(\Omega) \ : \ \mathcal{L}_{\mathbb{Q}}(S_{t_i}) = \mu_i, \ i \leq M, \quad \text{and } S \text{ is a } (\mathbb{Q}, \mathbb{F})\text{-martingale} \right\},$$

$$\overline{\mathcal{M}}_\mu := \left\{ \overline{\mathbb{Q}} \in \mathfrak{P}(\overline{\Omega}) \ : \ \mathcal{L}_{\overline{\mathbb{Q}}}(S_{t_i}) = \mu_i, \ i \leq M, \quad \text{and } S \text{ is a } (\overline{\mathbb{Q}}, \overline{\mathbb{F}})\text{-martingale} \right\}.$$

Let Λ_0 be the class of all Lipschitz functions $\lambda : \mathbb{R}^d \to \mathbb{R}$, and denote $\Lambda := \Lambda_0^M$. The statically traded options $g = (g^\lambda)_{\lambda \in \Lambda}$ are given by $g^\lambda(\omega) := \lambda(\omega) - \mu(\lambda)$, where $\lambda(\omega) := \sum_{i=1}^M \lambda_i(\omega_{t_i})$ and $\mu(\lambda) := \sum_{i=1}^M \mu_{t_i}(\lambda_i)$. Recall that $\mathcal{M}_g = \mathcal{M}_\mu$. As Λ is a linear space, the superhedging cost of the American option Φ using semistatic strategies $\pi_g^A(\Phi)$ defined in Section 2.1 can be rewritten as

$$\pi_g^A(\Phi) = \pi_\mu^A(\Phi) := \inf\{\mu(\lambda) \ : \ \exists ({}^1 H, \dots, {}^N H) \in \mathcal{H}^N \text{ s.t. } {}^j H_i = {}^k H_i \ \forall 1 \leq i \leq j \leq k \leq N$$

$$\text{and } \lambda \in \Lambda \text{ satisfying } \lambda(\omega) + ({}^k H \circ S)_N(\omega) \geq \Phi_k(\omega) \quad \text{for all } k \in \mathbb{T}, \ \omega \in \Omega\}.$$

Similarly, we denote by $\overline{\pi}_\mu^E(\Phi)$ the corresponding superhedging cost for a European option with payoff Φ defined on $\overline{\Omega}$, and one has $\pi_\mu^A(\Phi) = \overline{\pi}_\mu^E(\Phi)$ by Theorem 2.4.

Example 4.1. We shall construct an example similar to Example 1.1 to highlight that we may have a strict inequality in (5). Consider the case $N = 2$, $\mathbb{T}_0 = \mathbb{T} = \{1, 2\}$, $\mu_1 = \delta_{\{0\}}$, and $\mu_2 = \frac{1}{4}(\delta_{\{-2\}} + \delta_{\{-1\}} + \delta_{\{1\}} + \delta_{\{2\}})$. Let $\Phi_1(\{S_1 = 0\}) = 1$, $\Phi_2(\{|S_2| = 1\}) = 2$, and $\Phi_2(\{|S_2| = 2\}) = 0$. Then, $\mathcal{M}_\mu$ contains only one probability measure $\mathbb{Q}$, and by direct computation, one has

$$\mathbb{E}^{\mathbb{Q}}[\Phi_\tau] = 1, \quad \text{for all } \tau \in \mathcal{T}(\mathbb{F}).$$

Let us now construct a martingale measure $\overline{\mathbb{Q}}_0$ by

$$\overline{\mathbb{Q}}_0(d\omega, d\theta) := \frac{1}{4}\delta_{\{1\}}(d\theta) \otimes (\delta_{(0,1)} + \delta_{(0,-1)})(d\omega) + \frac{1}{4}\delta_{\{2\}}(d\theta) \otimes (\delta_{(0,2)} + \delta_{(0,-2)})(d\omega).$$

Then, one can check that $\overline{\mathbb{Q}}_0 \in \overline{\mathcal{M}}_\mu$, and it follows that

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_\mu} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \geq \mathbb{E}^{\overline{\mathbb{Q}}_0}[\Phi] = \frac{3}{2} > 1 = \sup_{\mathbb{Q} \in \mathcal{M}_\mu} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau].$$

The superhedging price of Φ is equal to $3/2$, as one can consider a superhedging strategy consisting of holding $3/2$ in cash and one option g from Example 1.1. In a similar way as in Example 1.1, the duality may be recovered by allowing dynamic trading options.

4.1 | Duality on the enlarged space $\overline{\Omega}$

The following theorem shows the duality for $\overline{\Omega}$. Its proof is postponed to Section 7.

Theorem 4.2. *Suppose that* $\Phi : \overline{\Omega} \to \mathbb{R}$ *is bounded from above and upper semicontinuous. Then, there exists an optimal martingale measure* $\overline{\mathbb{Q}}^* \in \overline{\mathcal{M}}_\mu$ *and the pricing–hedging duality holds,*

$$\mathbb{E}^{\overline{\mathbb{Q}}^*}[\Phi] = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_\mu} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \overline{\pi}_\mu^E(\Phi),$$

and in particular (8) holds.

Remark 4.3. Note that in the above formulation each μ_i is an element of $\mathfrak{P}(\mathbb{R}^d)$. Instead one could take μ_i to be an element of $(\mathfrak{P}(\mathbb{R}))^d$, and the same statements with analogous proofs would still hold. This alternative formulation has a more transparent financial interpretation as it corresponds only to marginal laws of terminal values of each stock price as opposed to the full distribution; see also Lim (2016) for a related discussion.

4.2 | Dynamic programming principle on $\widehat{\Omega}$

Eldan (2016) and Cox and Källblad (2017) studied the Skorokhod embedding and martingale optimal transport problems in continuous time using measure-valued martingales. This point of view enables one to obtain the dynamic programming principle with marginal constraint because the terminal constraint is transformed into the initial constraint. We adopt this perspective which proves to be very useful.

As before, we work with the set of marginal times $\mathbb{T}_0 = \{t_1, \ldots, t_M\} \subset \{1, \ldots, N\}$ such that $t_M = N$, and a vector of marginal measures $\mu = (\mu_1, \ldots, \mu_M)$ satisfying (37), where each μ_i is a probability measure on $\mathbb{R}^d$. We let $\mathfrak{P}_1(\mathbb{R}^d) = \{\eta \in \mathfrak{P}(\mathbb{R}^d) : \eta(|\cdot|) < \infty\}$ be the set of probability measures with finite first moment which we equip with the 1-Wasserstein distance, that is, $\eta_n \to \eta_0$ if and only if

$$\int_{\mathbb{R}^d} f(x) \eta_n(dx) \to \int_{\mathbb{R}^d} f(x) \eta_0(dx), \quad \forall f \in C_1,$$

where C_1 denotes the set of all continuous functions on $\mathbb{R}^d$ with at most linear growth, which makes $\mathfrak{P}_1(\mathbb{R}^d)$ a Polish space. Continuing with the construction from Example 2.15, $\widehat{\Omega}$ has to be an infinite dimensional space, and it is convenient to parameterize it as the canonical space for the measure-valued processes

$$\widehat{\Omega} := \{\mu\} \times (\mathfrak{P}_1(\mathbb{R}^d))^{M \times N}$$

and denote by $\widehat{X} = (\widehat{X}_k^1, \dots, \widehat{X}_k^M)_{0 \le k \le N}$ the canonical process on $\widehat{\Omega}$. Let $\widehat{\mathbb{G}} = (\widehat{\mathcal{G}}_k)_{0 \le k \le N}$ be the canonical filtration and $\widehat{\mathbb{F}} = (\widehat{\mathcal{F}}_k)_{0 \le k \le N}$ its universal completion. Denote by $\mathcal{T}(\widehat{\mathbb{F}})$ the collection of all $\widehat{\mathbb{F}}$-stopping times. For $f \in C_1$ we denote the process of its integrals against $\widehat{X}$ as

$$\widehat{X}_k(f) = \left(\widehat{X}_k^1(f), \dots, \widehat{X}_k^M(f) \right), \quad \text{where} \quad \widehat{X}_k^i(f) := \int_{\mathbb{R}^d} f(x)\widehat{X}_k^i(dx) \text{ and}$$

$$\widehat{X}_k(id) = \left(\widehat{X}_k^1(id), \dots, \widehat{X}_k^M(id) \right), \quad \text{where} \quad \widehat{X}_k^i(id) = \int_{\mathbb{R}^d} x\widehat{X}_k^i(dx).$$

Define $\mathbf{i} : \widehat{\Omega} \to \Omega$ by $\mathbf{i}(\widehat{\omega}) = (\widehat{X}_0^M(id)(\widehat{\omega}), \dots, \widehat{X}_N^M(id)(\widehat{\omega}))$, which is surjective and naturally extends processes on Ω to processes on $\widehat{\Omega}$. In particular the price process extends via $S_k(\widehat{\omega}) = S_k(\mathbf{i}(\widehat{\omega})) = \widehat{X}_k^M(id)(\widehat{\omega})$ and the statically traded options via $g^\lambda(\widehat{\omega}) = g^\lambda(\mathbf{i}(\widehat{\omega})) = \lambda(\mathbf{i}(\widehat{\omega})) - \mu(\lambda)$. Define a family of processes $Y = (Y^\lambda)_{\lambda \in \Lambda}$ by $Y^\lambda = \sum_{i=1}^M Y^{\lambda_i}$, where

$$Y_k^{\lambda_i} = \begin{cases} \widehat{X}_k^i(\lambda_i) - \mu_i(\lambda_i) & 0 \le k \le t_i - 1, \\ g^{\lambda_i} = \lambda_i(\widehat{X}_{t_i}^i(id)) - \mu_i(\lambda_i) & t_i \le k \le N. \end{cases}$$

Note that $Y_0^{\lambda_i} = 0$.

For any $\mathbb{Q} \in \mathcal{M}_\mu$ we define a mapping $\mathbf{j}_\mathbb{Q} : \Omega \to \widehat{\Omega}$ by $\mathbf{j}_\mathbb{Q}(\omega) = (\mathcal{L}_\mathbb{Q}(S_{t_i}|\mathcal{F}_k)(\omega))_{i \le M, k \le N}$. As in Example 2.15, the map $\mathbf{J} : \mathcal{M}_\mu \to \widehat{\mathcal{M}}$ defined by $\mathbf{J}(\mathbb{Q}) = \mathbb{Q} \circ \mathbf{j}_\mathbb{Q}^{-1}$ satisfies Property 4 of Definition 2.13.

Definition 4.4. (a) A probability measure $\widehat{\mathbb{Q}}$ on $(\widehat{\Omega}, \widehat{\mathcal{F}})$ is called a measure-valued martingale measure (MVM measure) if the process $(\widehat{X}_k(f))_{0 \le k \le N}$ is a $(\widehat{\mathbb{Q}}, \widehat{\mathbb{F}})$-martingale for all $f \in C_1$.

(b) An MVM measure $\widehat{\mathbb{Q}}$ is terminating if for $i \in \{1, \dots, M\}$,

$$\widehat{X}_{t_i}^i \in \Delta := \{\eta \in \mathfrak{P}(\mathbb{R}^d) : \eta = \delta_x, \ x \in \mathbb{R}^d\} \quad \widehat{\mathbb{Q}}\text{-a.s.}$$

(c) An MVM measure $\widehat{\mathbb{Q}}$ is consistent if $S_k = \widehat{X}_k^i(id)$ for $k \le t_i$ and $i \in \{1, \dots, M\}$, $\widehat{\mathbb{Q}}$-a.s.

Let us denote by

$$\widehat{\mathcal{M}}_\mu = \{\widehat{\mathbb{Q}} \in \mathfrak{P}(\widehat{\Omega}) : \widehat{\mathbb{Q}} \text{ is a terminating, consistent, MVM measure}\}.$$

The following lemma shows that the marginal distribution of S at t_i is equal to μ_i, $\widehat{\mathcal{M}}_\mu$-q.s., and hence $\widehat{\mathbb{Q}} \circ \mathbf{i}^{-1} \in \mathcal{M}_\mu$ for any $\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_\mu$.

Lemma 4.5. *For a measure $\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_\mu$ the following holds:*

(a) $\mathcal{L}_{\widehat{\mathbb{Q}}}(S_{t_i}|\widehat{\mathcal{F}}_k) = \widehat{X}_k^i$ $\widehat{\mathbb{Q}}$-*a.s. for* $k \le t_i$, *and in particular* $\mathcal{L}_{\widehat{\mathbb{Q}}}(S_{t_i}) = \mu_i$.

(b) *For* $k \le t_j \le t_i$, $\widehat{X}_k^j \le \widehat{X}_k^i$ $\widehat{\mathbb{Q}}$-*a.s., that is, for any convex function* f,

$$\int_{\mathbb{R}^d} f(x)\widehat{X}_k^j(dx) \le \int_{\mathbb{R}^d} f(x)\widehat{X}_k^i(dx) \quad \widehat{\mathbb{Q}}\text{-a.s.}$$

Proof.

(a) Let $A \subset \mathbb{R}^d$ and recall that $S_k = \widehat{X}_k^i(id)$ $\widehat{\mathbb{Q}}$-a.s. Then, we have

$$\int_{\mathbb{R}^d} \mathbb{1}_A(x) \mathcal{L}_{\widehat{\mathbb{Q}}}\left(\widehat{X}_{t_i}^i(id) \big| \widehat{\mathcal{F}}_k\right)(dx) = \mathbb{E}^{\widehat{\mathbb{Q}}}\left[\mathbb{1}_{\{\widehat{X}_{t_i}^i(id) \in A\}} \big| \widehat{\mathcal{F}}_k\right] = \mathbb{E}^{\widehat{\mathbb{Q}}}\left[\widehat{X}_{t_i}^i(\mathbb{1}_A) \big| \widehat{\mathcal{F}}_k\right] = \widehat{X}_k^i(\mathbb{1}_A),$$

where the second equality holds as $\widehat{\mathbb{Q}}$ is terminating and the third one follows by the definition of the MVM measure in Definition 4.4 as well as remark 2.2 of Cox and Källblad (2017). Hence, the first assertion is proven.

(b) Let $j \leq i$, $k \leq t_j$ and f be a convex function. Then,

$$\int_{\mathbb{R}^d} f(x)\widehat{X}_k^i(dx) = \mathbb{E}^{\widehat{\mathbb{Q}}}\left[f\left(\widehat{X}_{t_i}^i(id)\right) | \widehat{\mathcal{F}}_k\right]$$

$$\geq \mathbb{E}^{\widehat{\mathbb{Q}}}\left[f\left(\mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{X}_{t_i}^i(id) | \widehat{\mathcal{F}}_{t_j}]\right) | \widehat{\mathcal{F}}_k\right]$$

$$= \mathbb{E}^{\widehat{\mathbb{Q}}}\left[f\left(\widehat{X}_{t_j}^i(id)\right) | \widehat{\mathcal{F}}_k\right] = \mathbb{E}^{\widehat{\mathbb{Q}}}\left[f\left(\widehat{X}_{t_j}^j(id)\right) | \widehat{\mathcal{F}}_k\right]$$

$$= \int_{\mathbb{R}^d} f(x)\widehat{X}_k^j(dx),$$

where the first and the last equalities follow by (a), the penultimate is due to the consistency of $\widehat{\mathbb{Q}}$, and the inequality follows by conditional Jensen's inequality. $\qquad\square$

Recall the set $\widehat{\mathcal{M}}$ of martingale measures in Definition 2.13. The following lemma shows how to build the map J and that $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, Y, \mathtt{i}, J)$ is a dynamic extension of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$.

Lemma 4.6.

(a) *Under any $\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_\mu$, the processes S and Y^λ, for $\lambda \in \Lambda$, are $(\widehat{\mathbb{Q}}, \widehat{\mathbb{F}})$-martingales. In particular, one has $\widehat{\mathcal{M}}_\mu \subset \widehat{\mathcal{M}}$.*

(b) *For $\mathbb{Q} \in \mathcal{M}_\mu$, let $J(\mathbb{Q})$ be the distribution of the measure-valued process $\eta = (\eta_k^1, \ldots, \eta_k^M)_{k \leq N}$, where $\eta_k^i = \mathcal{L}_{\mathbb{Q}}(S_{t_i} | \mathcal{F}_k)$. Then, $J : \mathcal{M}_\mu \to \widehat{\mathcal{M}}$ and $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\mathcal{F}}, Y, \mathtt{i}, J)$ is a dynamic extension of $(\Omega, \mathbb{F}, \mathcal{F}, \mathcal{P}, S, g)$.*

Proof.

(a) The process $S = \widehat{X}^M(id)$ is a $(\widehat{\mathbb{Q}}, \widehat{\mathbb{F}})$-martingale as $\widehat{\mathbb{Q}}$ is an MVM measure. To prove that Y^λ is a $(\widehat{\mathbb{Q}}, \widehat{\mathbb{F}})$-martingale for any $\lambda \in \Lambda$, it is enough to show that for any $i \leq M$ and $\lambda \in \Lambda_0$ one has $\mathbb{E}^{\widehat{\mathbb{Q}}}[\lambda(\widehat{X}_{t_i}^M(id)) | \widehat{\mathcal{F}}_k] = \widehat{X}_k^i(\lambda)$ for any $k < t_i$. The latter holds because

$$\mathbb{E}^{\widehat{\mathbb{Q}}}\left[\lambda\left(\widehat{X}_{t_i}^M(id)\right) | \widehat{\mathcal{F}}_k\right] = \mathbb{E}^{\widehat{\mathbb{Q}}}\left[\lambda\left(\widehat{X}_{t_i}^i(id)\right) | \widehat{\mathcal{F}}_k\right] = \mathbb{E}^{\widehat{\mathbb{Q}}}\left[\widehat{X}_{t_i}^i(\lambda) | \widehat{\mathcal{F}}_k\right] = \widehat{X}_k^i(\lambda),$$

where the first equality follows by consistency of $\widehat{\mathbb{Q}}$, the second holds as $\widehat{\mathbb{Q}}$ is terminating, and the last one holds because $\widehat{\mathbb{Q}}$ is an MVM measure.

(b) We have $\eta_{t_i}^i = \mathcal{L}_{\widehat{\mathbb{Q}}}(S_{t_i} | \mathcal{F}_{t_i}) = \delta_{S_{t_i}} \in \Delta$, and it follows that $J(\mathbb{Q}) \in \widehat{\mathcal{M}}_\mu \subset \widehat{\mathcal{M}}$. $\qquad\square$

For $\hat{\omega} \in \hat{\Omega}$, we define the set $[\hat{\omega}]_{\hat{\mathcal{G}}_k}$ as in (18), and denote by $\widehat{\mathcal{M}}_\mu^k(\hat{\omega})$ the following set of measures:

$$\widehat{\mathcal{M}}_\mu^k(\hat{\omega}) := \{\hat{\mathbb{Q}} \in \mathfrak{P}(\hat{\Omega}) : \hat{\mathbb{Q}} \text{ is terminating and consistent,}$$

$$\hat{\mathbb{Q}}([\hat{\omega}]_{\hat{\mathcal{G}}_k}) = 1, \quad \text{and } (\hat{X}_l)_{k \le l \le N} \text{ is a} (\hat{\mathbb{Q}}, \hat{\mathbb{F}}) \text{ -MVM}\}.$$

Let us define a family of operators $\hat{\mathcal{E}}_k$, and so forth, as in Example 2.12,

$$\hat{\mathcal{E}}_k(\hat{\xi})(\hat{\omega}) = \sup_{\mathbb{Q} \in \widehat{\mathcal{M}}_\mu^k(\hat{\omega})} \mathbb{E}^{\mathbb{Q}}[\hat{\xi}], \quad \hat{\xi} \in \hat{\Upsilon},$$

and then the extension $\overline{\hat{\mathcal{E}}}_k$ as well as $\overline{\hat{\mathcal{E}}}^0$ on the enlarged space as in Section 2.2. We then have the following theorem.

Theorem 4.7. *For all upper semianalytic functionals $\hat{\xi} : \hat{\Omega} \to \mathbb{R}$, $\overline{\hat{\mathcal{E}}}_k(\hat{\xi})$ is also upper semianalytic and*

$$\sup_{\overline{\hat{\mathbb{Q}}} \in \overline{\widehat{\mathcal{M}}}_\mu} \mathbb{E}^{\overline{\hat{\mathbb{Q}}}}[\hat{\xi}] = \overline{\hat{\mathcal{E}}}^0(\hat{\xi}). \tag{38}$$

In particular the pricing–hedging duality (24) holds in this martingale optimal transport context for all functionals $\Phi : \overline{\Omega} \to \mathbb{R}^N$, which are upper semicontinuous and bounded from above.

Proof. Notice that the pricing–hedging duality on $\overline{\Omega}$ holds by Theorem 4.2. Then, by Corollary 2.19, it is enough to establish the dynamic programming principle on $\hat{\Omega}$ to prove the pricing–hedging duality (24). Using exactly the same arguments as in (4.12) of Bouchard and Nutz (2015), to establish the dynamic programming principle on $\hat{\Omega}$, it is enough to argue that $\widehat{\mathcal{M}}_\mu$ is such that

$$\left\{ (\hat{\omega}, \hat{\mathbb{Q}}) : \hat{\mathbb{Q}} \in \widehat{\mathcal{M}}_\mu^k(\hat{\omega}) \right\} \text{ is analytic.}$$

To prove the above analyticity property, we first observe that

$$\hat{\mathcal{E}}_k \circ \cdots \circ \hat{\mathcal{E}}_{N-1}(\hat{\xi})(\hat{\omega}) = \hat{\mathcal{E}}_{k,k+1} \circ \cdots \circ \hat{\mathcal{E}}_{N-1,N}(\hat{\xi})(\hat{\omega}), \quad \hat{\xi} \in \hat{\Upsilon},$$

where $\mathcal{E}_{k,k+1}(\hat{\xi})(\omega) = \sup_{\mathcal{M}_\mu^{k,k+1}(\omega)} \mathbb{E}^{\mathbb{Q}}[\hat{\xi}]$ and

$$\widehat{\mathcal{M}}_\mu^{k,k+1}(\hat{\omega}) := \{\hat{\mathbb{Q}} \in \mathfrak{P}(\hat{\Omega}) : \hat{\mathbb{Q}} \text{ is terminating and consistent,}$$

$$\hat{\mathbb{Q}}([\hat{\omega}]_{\hat{\mathcal{G}}_k}) = 1, \quad \text{and} \quad \hat{\omega}_k(f) = \mathbb{E}^{\hat{\mathbb{Q}}}[\hat{X}_{k+1}(f)], \forall f \in \mathcal{C}_1\}.$$

Next, let $\mathcal{C}_1^0$ denote a countable dense subset of $\mathcal{C}_1$ under the uniform convergence topology. Then, it is clear that for each $k \in \mathbb{T}$,

$$\left\{ (\hat{\omega}, \hat{\mathbb{Q}}) \in \hat{\Omega} \times \mathfrak{P}(\hat{\Omega}) : \hat{\mathbb{Q}} \in \widehat{\mathcal{M}}_\mu^{k,k+1}(\hat{\omega}) \right\} = \{ (\hat{\omega}, \hat{\mathbb{Q}}) \in \hat{\Omega} \times \mathfrak{P}(\hat{\Omega}) : \hat{\mathbb{Q}}([\hat{\omega}]_{\hat{\mathcal{G}}_k}) = 1$$

$$\hat{\mathbb{Q}} \text{ is terminating and consistent,} \quad \text{and} \quad \hat{\omega}_k(f) = \mathbb{E}^{\hat{\mathbb{Q}}}[\hat{X}_{k+1}(f)], \ \forall f \in \mathcal{C}_1^0 \}$$

is a Borel set. $\qquad \square$

5 | PROOFS FOR SECTION 2

We recall that Section 2 stays in a context with an abstract space $(\Omega, \mathcal{F})$ equipped with an underlying process S and a family $\mathcal{P}$ of probability measures. $\mathcal{M}$ denotes the collection of all measures $\mathbb{Q}$ dominated by some $\mathbb{P} \in \mathcal{P}$ and such that S is a $\mathbb{Q}$-martingale, and $\sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\cdot]$ admits a dynamic programming representation by $\mathcal{E}_k$ (see (9)), from which one defines the family of operators $\widehat{\mathcal{E}}_k$ in (10). A first enlarged space $\overline{\Omega} := \Omega \times \{1, \dots, N\}$ is introduced in Section 2.2 to reduce an American option to a European option, and a dynamic extension $\widehat{\Omega}$ of Ω is defined in Definition 2.13 to introduce a fictitious market allowing dynamic trading of options.

Proof of Proposition 2.9. First we prove that (12) implies (9). For a given ξ on Ω let us define Φ on $\overline{\Omega}$ by

$$\Phi(\omega, k) = -\infty \quad \text{if } k \in \{1, \dots, N-1\},$$

$$\Phi(\omega, N) = \xi(\omega).$$

The definition of Φ combined with (12) implies that

$$\overline{\mathcal{E}}_0 \circ \overline{\mathcal{E}}_1 \circ \cdots \circ \overline{\mathcal{E}}_{N-1}(\xi) = \mathcal{E}_0 \circ \mathcal{E}_1 \circ \cdots \circ \mathcal{E}_{N-1}(\xi).$$

Moreover, one has that

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Phi_N] = \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\xi],$$

because for a measure $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}$ such that $\overline{\mathbb{Q}}(\Omega \times \{1, \dots, N-1\}) > 0$, the expected value drops to $-\infty$.

Now let us prove that (9) and (13) imply (12). Define an $\mathbb{F}$-stopping time τ^* by

$$\tau^*(\omega) := \min \left\{ k \geq 1 : \mathcal{E}^k(\Phi(\cdot, k))(\omega) = \overline{\mathcal{E}}^k(\Phi)(\omega, k) \right\} \tag{39}$$

$$= \min \left\{ k \geq 1 : \mathcal{E}^k \left(\mathcal{E}^{k+1}(\Phi(\cdot, k)) \right)(\omega) \geq \mathcal{E}_k \left(\overline{\mathcal{E}}^{k+1}(\Phi)(\cdot, k+1) \right)(\omega) \right\}.$$

Note that on $\{k < \tau^*\}$ one has

$$\mathcal{E}^k(\Phi(\cdot, k))(\omega) < \overline{\mathcal{E}}^k(\Phi)(\omega, k) = \mathcal{E}_k \left(\overline{\mathcal{E}}^{k+1}(\Phi)(\cdot, k+1) \right)(\omega). \tag{40}$$

Then,

$$\overline{\mathcal{E}}_0 \circ \cdots \circ \overline{\mathcal{E}}_{N-1}(\Phi)$$

$$= \mathcal{E}_0 \left(\mathbb{1}_{\{\tau^*=1\}} \mathcal{E}_1 \left(\overline{\mathcal{E}}^2(\Phi)(\cdot, 1) \right) + \mathbb{1}_{\{\tau^*>1\}} \mathcal{E}_1 \left(\overline{\mathcal{E}}^2(\Phi)(\cdot, 2) \right) \right)$$

$$= \dots$$

$$= \mathcal{E}_0 \circ \mathcal{E}_1 \circ \cdots \circ \mathcal{E}_{N-1}(\Phi_{\tau^*})$$

$$= \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Phi_{\tau^*}],$$

where the last equality follows from the Dynamic Programming Principle (DPP) on Ω, (9). Note as well that

$$\overline{\mathcal{E}}_0 \circ \cdots \circ \overline{\mathcal{E}}_{N-1}(\Phi) = \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Phi_{\tau^*}] \leq \sup_{\mathbb{Q} \in \mathcal{M}} \sup_{\tau \in \mathcal{T}(\mathbb{F})} \mathbb{E}^{\mathbb{Q}}[\Phi_\tau] \leq \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]. \tag{41}$$

Combining this with (13), we conclude the proof. $\qquad\square$

Proposition 5.1. *The family $(\overline{\mathcal{E}}_k)$ given in Example 2.12, with $\mathcal{M}_k(\omega)$ defined by (32), satisfies (13).*

Proof. In the context of Example 2.12, the family $(\overline{\mathcal{E}}_k)$ take the following form:

$$\overline{\mathcal{E}}_0(\Phi) := \sup_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Phi(\cdot, 1)], \tag{42}$$

$$\overline{\mathcal{E}}_k(\Phi)(\overline{\omega}) := \begin{cases} \sup_{\mathbb{Q} \in \mathcal{M}_k(\omega)} \mathbb{E}^{\mathbb{Q}}[\Phi(\cdot, \theta)] & \text{if } \theta < k, \\ \sup_{\mathbb{Q} \in \mathcal{M}_k(\omega)} \mathbb{E}^{\mathbb{Q}}[\Phi(\cdot, k)] \vee \sup_{\mathbb{Q} \in \mathcal{M}_k(\omega)} \mathbb{E}^{\mathbb{Q}}[\Phi(\cdot, k+1)] & \text{if } \theta \geq k. \end{cases}$$

To see that (13) holds, it is insightful to rewrite $\overline{\mathcal{E}}^0$ in a slightly different way, as $\widetilde{\mathcal{E}}^0$ below. Recall that $\mathcal{F}_k$ is the universal completion of $\mathcal{G}_k$, where the latter is countably generated. Let

$$\overline{\mathcal{G}}_k^- := \mathcal{G}_k \otimes \sigma(T \wedge k) \subset \overline{\mathcal{G}}_k := \mathcal{G}_k \otimes \sigma(T \wedge (k+1)) \subset \mathcal{F}_k \otimes \sigma(T \wedge (k+1)) =: \overline{\mathcal{F}}_k,$$

$$\overline{\mathcal{M}}_k^-(\overline{\omega}) := \{\overline{\mathbb{Q}} \ll \overline{\mathcal{P}} : \overline{\mathbb{Q}}\left[[\overline{\omega}]_{\overline{\mathcal{G}}_k^-}\right] = 1 \quad and \quad \mathbb{E}^{\overline{\mathbb{Q}}}[\Delta S_n | \overline{\mathcal{F}}_{n-1}] = 0 \, \forall n \in \{k+1, \ldots, N\}\},$$

where $[\overline{\omega}]_{\overline{\mathcal{G}}_k^-}$ is defined as in (18). Next, for $\Phi \in \overline{\Upsilon}$, let us introduce the operators

$$\widetilde{\mathcal{E}}_0(\Phi) := \sup_{\mathbb{Q} \in \overline{\mathcal{M}}} \mathbb{E}^{\mathbb{Q}}[\Phi(\cdot, 1)], \qquad \widetilde{\mathcal{E}}_k(\Phi)(\overline{\omega}) := \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_k^-(\overline{\omega})} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi], \, k \leq N-1.$$

Denote $\overline{\mathcal{E}}^k(\cdot) := \overline{\mathcal{E}}_k \circ \cdots \circ \overline{\mathcal{E}}_{N-1}(\cdot)$ and $\widetilde{\mathcal{E}}^k(\cdot) := \widetilde{\mathcal{E}}_k \circ \cdots \circ \widetilde{\mathcal{E}}_{N-1}(\cdot)$. We claim that

$$\overline{\mathcal{E}}^k(\Phi)(\overline{\omega}) = \widetilde{\mathcal{E}}^k(\Phi)(\overline{\omega}), \, 0 \leq k < N, \, \Phi \in \overline{\Upsilon}. \tag{43}$$

Note that the regular conditional probabilities of any $\mathbb{Q} \in \overline{\mathcal{M}}$ with respect to $\overline{\mathcal{G}}_k^-$, denoted $\mathbb{Q}_{\overline{\omega}}$, satisfy $\mathbb{Q}[\{\overline{\omega} : \mathbb{Q}_{\overline{\omega}} \in \overline{\mathcal{M}}_k^-(\overline{\omega})\}] = 1$ and one has $\mathbb{E}^{\mathbb{Q}}[\Phi | \overline{\mathcal{F}}_k^-] \leq \widetilde{\mathcal{E}}_k(\Phi)$, $\mathbb{Q}$-a.s., which implies (13) by the tower property of conditional expectations and the definition of $\widetilde{\mathcal{E}}^0$.

It is then enough to prove the claim (43). Note that, for $\overline{\omega} = (\omega, \theta)$ with $\theta \leq k-1$, a measure $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_k^-(\overline{\omega})$ satisfies $\overline{\mathbb{Q}}_{|\Omega} \in \mathcal{M}_k(\omega)$ and $\overline{\mathbb{Q}}(\Omega \times \{\theta\}) = 1$; and a measure $\mathbb{Q} \in \mathcal{M}_k(\omega)$ satisfies $\mathbb{Q} \otimes \delta_\theta \in \overline{\mathcal{M}}_k^-(\overline{\omega})$. It is thus clear that, in this case, $\overline{\mathcal{E}}^k(f)(\overline{\omega}) = \widetilde{\mathcal{E}}^k(f)(\overline{\omega})$.

As a second step, for $\overline{\omega} = (\omega, \theta)$ with $\theta \geq k$, we show that $\overline{\mathcal{E}}_k(f)(\overline{\omega}) \leq \widetilde{\mathcal{E}}_k(f)(\overline{\omega})$. Take any $\mathbb{Q} \in \mathcal{M}_k(\omega)$. Then, for $n \in \{k, \ldots, N\}$, $\mathbb{Q} \otimes \delta_n \in \overline{\mathcal{M}}_k^-(\overline{\omega})$ and $\mathbb{Q} \otimes \delta_n(\Omega \times \{n\}) = 1$. Hence, it follows that $\overline{\mathcal{E}}_k(f)(\overline{\omega}) \leq \widetilde{\mathcal{E}}_k(f)(\overline{\omega})$.

In a final step, we show that, for $\overline{\omega} = (\omega, \theta)$ with $\theta \geq k$, $\overline{\mathcal{E}}_k(f)(\overline{\omega}) \geq \widetilde{\mathcal{E}}_k(f)(\overline{\omega})$ holds. Let us start with $k = N-1$. Take any $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{N-1}^-(\overline{\omega})$ and consider its regular conditional probability with respect to $\overline{\mathcal{G}}_{N-1}$ (the atom $\{\omega\} \times \{N-1, N\}$ is divided into atoms $\{\omega\} \times \{N-1\}$ and $\{\omega\} \times \{N\}$) denoted by $\overline{\mathbb{Q}}_N$ and $\overline{\mathbb{Q}}_{N-1}$. Then, clearly, $\overline{\mathbb{Q}}_{N|\Omega}$ and $\overline{\mathbb{Q}}_{N-1|\Omega}$ belong to $\mathcal{M}^{N-1}(\omega)$, and $\overline{\mathbb{Q}}_N(\{\omega\} \times \{N\}) = 1$ and $\overline{\mathbb{Q}}_{N-1}(\{\omega\} \times \{N-1\}) = 1$. Thus, it follows that $\overline{\mathcal{E}}_{N-1}(f)(\overline{\omega}) \geq \widetilde{\mathcal{E}}_{N-1}(f)(\overline{\omega})$.

Finally, to complete the proof, we need to show that $\overline{\mathcal{E}}^{k+1}(f)(\overline{\omega}) = \widetilde{\mathcal{E}}^{k+1}(f)(\overline{\omega})$ implies $\overline{\mathcal{E}}^{k}(f)(\overline{\omega}) \geq \widetilde{\mathcal{E}}^{k}(f)(\overline{\omega})$ for $\overline{\omega} = (\omega, \theta)$ with $\theta \geq k$. First note that $\overline{\mathcal{E}}^{k+1}(f)(\overline{\omega}) = \widetilde{\mathcal{E}}^{k+1}(f)(\overline{\omega})$ is constant on $\theta \in \{k, \dots, N\}$, that is,

$$\overline{\mathcal{E}}^{k+1}(f)(\omega, \theta_1) = \overline{\mathcal{E}}^{k+1}(f)(\omega, \theta_2) \quad \text{for all } \omega \in \Omega \quad \text{and} \quad \theta_1, \theta_2 \in \{k, \dots, N\}. \tag{44}$$

Take any $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_k(\overline{\omega})$ and consider its regular conditional probability with respect to ϑ_N (the atom $\{\omega\} \times \{k, \dots, N\}$ is divided into atoms $\{\omega\} \times \{n\}$ for $n = k, \dots, N$) denoted by $\overline{\mathbb{Q}}_n$ for $n = k, \dots, N$. Then, clearly, $\overline{\mathbb{Q}}_{n|\Omega} \in \mathcal{M}_k(\omega)$ and $\overline{\mathbb{Q}}_n([\omega]_k \times \{n\}) = 1$, where $[\omega]_k$ denotes an atom of $\mathcal{G}_k$, which contains ω. Thus, combining with (44), it follows that $\overline{\mathcal{E}}_k(f)(\overline{\omega}) \geq \widetilde{\mathcal{E}}_k(f)(\overline{\omega})$. $\qquad\square$

6 | PROOFS FOR SECTION 3

We now recall the context of Section 3, where $\Omega_0 := \{\omega_0\}$ is a singleton, Ω_1 is a nonempty Polish space and $\Omega := \Omega_0 \times \Omega_1^N$. For technical reasons, we introduce a Ω_1-valued canonical process $X = (X_k)_{0 \leq k \leq N}$ on the enlarged space $\overline{\Omega}$ by $X_k(\overline{\omega}) := \omega_k$ for all $\overline{\omega} := (\omega, \theta) \in \overline{\Omega}$, and an enlarged filtration $\overline{\mathbb{G}} = (\overline{\mathcal{G}}_k)_{0 \leq k \leq N}$ by

$$\overline{\mathcal{G}}_0 := \{\emptyset, \overline{\Omega}\} \quad \text{and} \quad \overline{\mathcal{G}}_k := \sigma\{X_i, \{T \leq i\}, i = 1, \dots, k\},$$

and the universally completed filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_k)_{0 \leq k \leq N}$ by defining $\overline{\mathcal{F}}_k$ as the universal completion of $\overline{\mathcal{G}}_k$. It follows that the random time $T : \overline{\Omega} \to \mathbb{T}$ is a $\overline{\mathbb{G}}$-stopping time. We also define a restricted space $\overline{\Omega}_k$, for every $k = 1, \dots, N$,

$$\overline{\Omega}_k := \Omega_k \times \{1, \dots, k\} = \Omega_1^k \times \{1, \dots, k\}.$$

Lemma 6.1. *Let $\overline{\mathbb{P}} \in \overline{\mathcal{P}}$ be a probability measure on $(\overline{\Omega}, \overline{\mathcal{G}}_N)$, and $(\overline{\mathbb{P}}_{\overline{\omega}})_{\overline{\omega} \in \overline{\Omega}}$ be a family of regular conditional probability distributions of $\overline{\mathbb{P}}$ with respect to $\overline{\mathcal{G}}_k$. Then, for every $k \in \{0, 1, \dots, N-1\}$, one has $\overline{\mathbb{P}}_{\overline{\omega}} \circ X_{k+1}^{-1} \in \mathcal{P}_k(\omega)$ for $\overline{\mathbb{P}}$-a.e. $\overline{\omega} = (\omega, \theta) \in \overline{\Omega}$.*

Let us introduce the following set of measures

$$\overline{\mathcal{M}}_g^{loc} := \{\overline{\mathbb{Q}} : \overline{\mathbb{Q}} \ll \overline{\mathcal{P}}, \mathbb{E}^{\overline{\mathbb{Q}}}[g^i] = 0, i \in \{1, \dots, e\}$$

$$\text{and } S \text{ is an } (\overline{\mathbb{F}}, \overline{\mathbb{Q}})\text{-local martingale}\}.$$

Lemma 6.2. *Let Φ be upper semianalytic and $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g^{loc}$. Then, for any $x \in \mathbb{R}$ and $(\overline{H}, h) \in \overline{\mathcal{H}} \times \mathbb{R}^e$ such that $x + (\overline{H} \circ S)_N(\overline{\omega}) + hg(\omega) \geq \Phi(\overline{\omega})$, $\overline{\mathbb{Q}}$-a.s., one has $\mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq x$.*

Proof. The proof follows by exactly the same arguments as in Lemma A.2 of Bouchard and Nutz (2015), using the discrete time local martingale characterization in their Lemma A.1.

By Lemma 6.2, we easily obtain the weak duality for all upper semianalytic Φ

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g^{loc}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \overline{\pi}_g^E(\Phi). \tag{45}$$

The following lemma shows that, for a fixed Φ, we can restrict to martingale measures satisfying a further integrability constraint.

Lemma 6.3. *Let Φ be upper semianalytic, $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_0^{loc}$ and $\varphi : \Omega \to [1, \infty)$ be such that $|\Phi(\omega, k)| \leq \varphi(\omega)$ for all $\overline{\omega} = (\omega, k) \in \overline{\Omega}$. Then $\overline{\mathcal{M}}^{\varphi, \overline{\mathbb{Q}}} \neq \emptyset$ and*

$$\mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \sup_{\overline{\mathbb{Q}}' \in \overline{\mathcal{M}}^{\varphi, \overline{\mathbb{Q}}}} \mathbb{E}^{\overline{\mathbb{Q}}'}[\Phi],$$

where

$$\overline{\mathcal{M}}^{\varphi, \overline{\mathbb{Q}}} := \{\overline{\mathbb{Q}}' \sim \overline{\mathbb{Q}} \; : \; \mathbb{E}^{\overline{\mathbb{Q}}'}[\varphi] < \infty, \quad and \quad S \text{ is an } (\overline{\mathbb{F}}, \overline{\mathbb{Q}}')\text{-martingale}\}.$$

Proof. First, by Lemma A.3 of Bouchard and Nutz (2015), there exists a probability measure $\overline{\mathbb{P}}_*$, equivalent to $\overline{\mathbb{Q}}$ on $(\overline{\Omega}, \overline{\mathcal{F}}_N)$, such that $\mathbb{E}^{\overline{\mathbb{P}}_*}[\varphi] < \infty$. On the filtered probability space $(\overline{\Omega}, \overline{\mathcal{F}}_N, \overline{\mathbb{F}}, \overline{\mathbb{P}}_*)$, one defines $\overline{\mathcal{M}}_*^{loc}$ as the collection of all probability measures $\overline{\mathbb{Q}}' \sim \overline{\mathbb{Q}} \sim \overline{\mathbb{P}}_*$ under which S is an $\overline{\mathbb{F}}$-local martingale. Denote

$$\pi_0^{E, \overline{\mathbb{Q}}}(\Phi) := \inf \left\{ x \; : \; \exists \overline{H} \in \overline{\mathcal{H}} \text{ s.t. } x + (\overline{H} \circ S)_N \geq \Phi, \overline{\mathbb{Q}}\text{-a.s.} \right\},$$

then by the classical arguments for the dominated discrete time market, such as Kabanov (2008) and Kabanov and Stricker (2001), see also lemma A.3 of Bouchard and Nutz (2015), one can easily obtain the inequality

$$\mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \sup_{\overline{\mathbb{Q}}' \in \overline{\mathcal{M}}_*^{loc}} \mathbb{E}^{\overline{\mathbb{Q}}'}[\Phi] \leq \overline{\pi}_0^{E, \overline{\mathbb{Q}}}(\Phi) \leq \sup_{\overline{\mathbb{Q}}' \in \overline{\mathcal{M}}^{\varphi, \overline{\mathbb{Q}}}} \mathbb{E}^{\overline{\mathbb{Q}}'}[\Phi],$$

which concludes the proof. □

Using theorem 2.2 of Bouchard and Nutz (2015), which is stated for a general abstract space $(\Omega, \mathcal{F})$, one directly obtains a closeness result for the set of all payoffs, which can be superreplicated from initial capital $x = 0$, in our context. Let us denote by $\overline{\mathcal{L}}_+^0$ the set of all positive random variables on $\overline{\Omega}$, and define

$$\overline{C} := \left\{ (\overline{H} \circ S)_N + hg : \overline{H} \in \overline{\mathcal{H}}, h \in \mathbb{R}^e \right\} - \overline{\mathcal{L}}_+^0.$$

Lemma 6.4 (**Theorem 2.2 of Bouchard and Nutz (2015)**). *Let Φ be upper semianalytic and assume that $\mathrm{NA}(\overline{\mathcal{P}})$ holds. Then the set $\overline{C}$ is closed in the following sense. Whenever $(W^n)_{n \geq 1} \subset \overline{C}$ and W is a random variable such that $W^n \to W$, $\overline{\mathcal{P}}$-q.s., then $W \in \overline{C}$.*

6.1 | Proof of Theorem 3.2: the case $e = 0$, equivalently $\Lambda = \emptyset$

For each $1 \leq i \leq j \leq N$, we introduce a map from Ω_j to Ω_i (resp. $\overline{\Omega}_j$ to $\overline{\Omega}_i$) by

$$[\omega]_i := (\omega_1, \ldots, \omega_i), \quad \text{for all} \quad \omega \in \Omega_j \ (\text{resp.} [\overline{\omega}]_i := ([\omega]_i, \theta \wedge i), \quad \text{for all} \quad \overline{\omega} = (\omega, \theta) \in \overline{\Omega}_j).$$

Note that $\overline{\mathcal{F}}_k$ is the smallest σ-field on $\overline{\Omega}$ generated by $[\,\cdot\,]_k : \overline{\Omega} \to \overline{\Omega}_k$; equivalently, an $\overline{\mathcal{F}}_k$-measurable random variable f defined on $\overline{\Omega}$ can be identified as a Borel measurable function on $\overline{\Omega}_k$. The process S is naturally defined on the restricted spaces Ω_k and $\overline{\Omega}_k$.

We next recall the notion of local no-arbitrage condition $NA(\mathcal{P}_k(\omega))$ introduced at the beginning of section 4.2 in Bouchard and Nutz (2015). Given a fixed $\omega \in \Omega_k$, we can consider $\Delta S_{k+1}(\omega, \cdot) := S_{k+1}(\omega, \cdot) - S_k(\omega)$ as a random variable on Ω_1, which determines a one-period market on $(\Omega_1, \mathcal{B}(\Omega_1))$ endowed with a class $\mathcal{P}_k(\omega)$ of probability measures. Then, $NA(\mathcal{P}_k(\omega))$ denotes the corresponding no-arbitrage condition in this one-period market, that is, $NA(\mathcal{P}_k(\omega))$ holds if for all $H \in \mathbb{R}^d$,

$$H\Delta S_{k+1}(\omega, \cdot) \geq 0 \quad \mathcal{P}_k(\omega)\text{-q.s.} \implies H\Delta S_{k+1}(\omega, \cdot) = 0 \quad \mathcal{P}_k(\omega)\text{-q.s.}$$

Lemma 6.5. *In the context of Section 3, let $f : \overline{\Omega}_{k+1} \to \overline{\mathbb{R}}$ be upper semianalytic. Then, $\overline{\mathcal{E}}_k(f) : \overline{\Omega}_k \to \overline{\mathbb{R}}$ is still upper semianalytic. Moreover, there exist two universally measurable functions $(y_1, y_2) : \overline{\Omega}_k \to \mathbb{R}^d \times \mathbb{R}^d$ such that*

$$\overline{\mathcal{E}}_k(f)(\overline{\omega}) + y_1(\overline{\omega})\Delta S_{k+1}(\omega, \cdot) \geq f(\omega, \cdot, \theta) \quad \mathcal{P}_k(\omega)\text{-q.s.,}$$

$$\overline{\mathcal{E}}_k(f)(\overline{\omega}) + y_2(\overline{\omega})\Delta S_{k+1}(\omega, \cdot) \geq f(\omega, \cdot, k+1) \quad \mathcal{P}_k(\omega)\text{-q.s.,}$$

for all $\overline{\omega} = (\omega, \theta) \in \overline{\Omega}_k$ such that $NA(\mathcal{P}_k(\omega))$ holds and $f(\omega, \cdot, \theta) > -\infty$, $\mathcal{P}_k(\omega)$-q.s. $f(\omega, \cdot, k+1) > -\infty$, $\mathcal{P}_k(\omega)$-q.s.

Proof. Notice that $f_1 \vee f_2$ is upper semianalytic whenever f_1 and f_2 are both upper semianalytic. Using the definition of $\overline{\mathcal{E}}_k$, the above lemma follows by applying lemma 4.10 of Bouchard and Nutz (2015) for every fixed θ.

Proof of Theorem 3.2 (the case $e = 0$). First, one has the weak duality as in (45)

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \leq \overline{\pi}_g^E(\Phi).$$

Next, for the inverse inequality, we can assume, without loss of generality, that Φ is bounded from above. Indeed, by Lemma 6.4, one has $\lim_{n \to \infty} \overline{\pi}_g^E(\Phi \wedge n) = \overline{\pi}_g^E(\Phi)$; see also the proof of theorem 3.4 of Bouchard and Nutz (2015). Besides, the approximation $\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \lim_{n \to \infty} \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi \wedge n]$ is an easy consequence of the monotone convergence theorem.

When Φ is bounded from above, by Lemma 3.3, it is enough to prove that there is some $\overline{H} \in \overline{\mathcal{H}}$ such that

$$\overline{\mathcal{E}}^0[\Phi] + (\overline{H} \circ S)_N \geq \Phi \quad \overline{\mathcal{P}}\text{-q.s.} \tag{46}$$

In view of Lemma 6.3, we know that $\overline{\mathcal{E}}^k(\Phi)(\overline{\omega}) > -\infty$ for all $\overline{\omega} \in \overline{\Omega}_k$. Further, by Lemma 6.5, there exist two universally measurable functions $(y_1^k, y_2^k) : \overline{\Omega}_k \to \mathbb{R}^d \times \mathbb{R}^d$ such that

$$y_1^k(\overline{\omega})\Delta S_{k+1}(\omega, \cdot) \geq \overline{\mathcal{E}}^{k+1}(\Phi)(\omega, \cdot, \theta) - \overline{\mathcal{E}}^k(\Phi)(\overline{\omega}) \quad \mathcal{P}_k(\omega)\text{-q.s.,}$$

$$y_2^k(\overline{\omega})\Delta S_{k+1}(\omega, \cdot) \geq \overline{\mathcal{E}}^{k+1}(\Phi)(\omega, \cdot, k+1) - \overline{\mathcal{E}}^k(\Phi)(\overline{\omega}) \quad \mathcal{P}_k(\omega)\text{-q.s.,}$$

for all $\overline{\omega} = (\omega, \theta) \in \overline{\Omega}_k$ such that $NA(\mathcal{P}_k(\omega))$ holds.

As $N_k := \{\omega_k : \text{NA}(\mathcal{P}_k(\omega)) \text{ fails}\}$ is $\mathcal{P}$-polar by theorem 4.5 of Bouchard and Nutz (2015), it follows that, with $\overline{H}_{k+1}(\overline{\omega}) := y_1^k([\overline{\omega}]_k)\mathbf{1}_{\{\theta \leq k\}} + y_1^k([\overline{\omega}]_k)\mathbf{1}_{\{\theta > k\}}$, one has

$$\sum_{k=0}^{N-1} \overline{H}_{k+1}\Delta S_{k+1} \geq \sum_{k=0}^{N-1} \left(\overline{\mathcal{E}}^{k+1}(\Phi) - \overline{\mathcal{E}}^k(\Phi)\right) = \Phi - \overline{\mathcal{E}}(\Phi), \quad \mathcal{P}\text{-q.s.}$$

To conclude, it is enough to notice that the above $\overline{H}$ is an optimal dual strategy for the case of Φ being bounded from above. The existence of the optimal dual strategy for general Φ is then a consequence of Lemma 6.4.

6.2 | Proof of Theorem 3.2: The case $e \geq 1$, equivalently $\Lambda \neq \emptyset$

We will adapt the arguments in section 5 of Bouchard and Nutz (2015) to prove Theorem 3.2 in the context with finitely many options $e \geq 1$.

For technical reasons, we introduce

$$\varphi(\omega, \theta) := 1 + |g^1(\omega)| + \cdots + |g^e(\omega)| + \max_{1 \leq k \leq N} |\Phi_k(\omega)|,$$

which depends only on ω, and

$$\overline{\mathcal{M}}_g^\varphi := \left\{\overline{\mathbb{Q}} \in \overline{\mathcal{M}} : \mathbb{E}^{\overline{\mathbb{Q}}}[\varphi] < \infty \quad \text{and} \quad \mathbb{E}^{\overline{\mathbb{Q}}}[g^i] = 0 \quad \text{for } i = 1, \ldots, e\right\}. \tag{47}$$

Moreover, in view of Lemma 6.3, one has

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g^\varphi} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi].$$

Proof of Theorem 3.2 (the case $e \geq 1$). The existence of some $\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g$ is an easy consequence of theorem 5.1 of Bouchard and Nutz (2015) under NA($\mathcal{P}$). Moreover, similarly to Bouchard and Nutz (2015), there exists an optimal dual strategy by Lemma 6.4. Let us now focus on the duality results.

First, the duality $\overline{\pi}_g^E(\Phi) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]$ in (34) has already been proved for the case $e = 0$. We will use an inductive argument. Suppose that the duality (34) holds true for the case with $e \geq 0$, We aim to prove the duality with $e + 1$ options

$$\overline{\pi}_{(g,f)}^E(\Phi) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{(g,f)}^\varphi} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi],$$

where the additional option has a Borel measurable payoff function $f \equiv g^{e+1}$ such that $|f| \leq \varphi$, and has an initial price $f_0 = 0$. By the weak duality in (45) and Lemma 6.3, the "$\geq$" side of the inequality holds true, so we will focus on the "$\leq$" side of the inequality, that is,

$$\overline{\pi}_{(g,f)}^E(\Phi) \leq \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{(f,g)}^\varphi} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi]. \tag{48}$$

If f is replicable by some semistatic strategy with underlying S and options $(g^1, \ldots, g^e)$ in the sense that $\exists \overline{H} \in \overline{\mathcal{H}}, h \in \mathbb{R}^e$, such that $f = (\overline{H} \circ S)_N + hg$, $\overline{\mathcal{P}}$-q.s. (or equivalently, $\exists H \in \mathcal{H}, h \in \mathbb{R}^e$, such that $f = (H \circ S)_N + hg$, $\mathcal{P}$-q.s.), then the problem is reduced to the case with e options and the result is

trivial. Let us assume that f is not replicable, and we claim that there exists a sequence $(\overline{\mathbb{Q}}_n)_{n\geq 1} \subset \overline{\mathcal{M}}_g^\varphi$ such that

$$\mathbb{E}^{\overline{\mathbb{Q}}_n}[f] \to f_0 \quad \text{and} \quad \mathbb{E}^{\overline{\mathbb{Q}}_n}[\Phi] \to \overline{\pi}_{(g,f)}^E(\Phi), \quad \text{as } n \to \infty. \tag{49}$$

Next, denote by $\overline{\pi}_g^E(f)$ the minimum superhedging cost of the European option f using S and $(g^1, \ldots, g^e)$, that is,

$$\pi_g^E(f) = \overline{\pi}_g^E(f) = \inf\{x : \exists \overline{H} \in \overline{\mathcal{H}}, h \in \mathbb{R}^e, \text{ s.t. } x + (\overline{H} \circ S)_N + hg \geq f, \ \overline{\mathcal{P}}\text{-q.s.}\}.$$

As f is not replicable, by the second fundamental theorem in Theorem 5.1.(c) of Bouchard and Nutz (2015), we have that $\overline{\mathbb{Q}} \mapsto \mathbb{E}^{\overline{\mathbb{Q}}}[f]$ is not constant on $\overline{\mathcal{M}}_g^\varphi$. Then, under the no-arbitrage condition, one has $0 = f_0 < \overline{\pi}_g^E(f)$. It follows that $0 = f_0 < \overline{\pi}_g^E(f) = \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g^\varphi} \mathbb{E}^{\overline{\mathbb{Q}}}[f]$. Thus, there exists some $\overline{\mathbb{Q}}_+ \in \overline{\mathcal{M}}_g^\varphi$, such that $0 < \mathbb{E}^{\overline{\mathbb{Q}}_+}[f] < \overline{\pi}_g^E(f)$. With the same argument on $-f$, we can find another $\overline{\mathbb{Q}}_- \in \overline{\mathcal{M}}_g^\varphi$ such that

$$-\overline{\pi}_g^E(-f) < \mathbb{E}^{\overline{\mathbb{Q}}_-}[f] < f_0 < \mathbb{E}^{\overline{\mathbb{Q}}_+}[f] < \overline{\pi}_g^E(f).$$

Then, one can choose an appropriate sequence of weights $(\lambda_-^n, \lambda_0^n, \lambda_+^n) \in \mathbb{R}_+^3$, such that $\lambda_-^n + \lambda_0^n + \lambda_+^n = 1$, $\lambda_\pm^n \to 0$,

$$\overline{\mathbb{Q}}_n' := \lambda_-^n \overline{\mathbb{Q}}_- + \lambda_0^n \overline{\mathbb{Q}}_n + \lambda_+^n \overline{\mathbb{Q}}_+ \in \overline{\mathcal{M}}_g, \quad \text{and} \quad \mathbb{E}^{\overline{\mathbb{Q}}_n'}[f] = f_0 = 0,$$

that is, $\overline{\mathbb{Q}}_n' \in \overline{\mathcal{M}}_{(g,f)}^\varphi$. Moreover, as $\lambda_\pm^n \to 0$, it follows that $\mathbb{E}^{\overline{\mathbb{Q}}_n'}[\Phi] \to \overline{\pi}_{(g,f)}^E(\Phi)$, and we hence have the inequality (48).

It is enough to prove the claim (49), for which we suppose without loss of generality that $\overline{\pi}_{(g,f)}^E(\Phi) = 0$. Assuming that (49) fails, one then has

$$0 \notin \overline{\left\{ \mathbb{E}^{\overline{\mathbb{Q}}}[(f, \Phi)] : \overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g^\varphi \right\}} \subseteq \mathbb{R}^2.$$

By the convexity of the above set and the separation argument, there exists $(y, z) \in \mathbb{R}^2$ with $|(y, z)| = 1$, such that

$$0 > \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_g^\varphi} \mathbb{E}^{\overline{\mathbb{Q}}}[yf + z\Phi] = \overline{\pi}_g^E(yf + z\Phi) \geq \overline{\pi}_{(g,f)}^E(z\Phi). \tag{50}$$

The strict inequality $\overline{\pi}_{(g,f)}^E(z\Phi) < 0$ implies that $z \neq 0$. Now, if $z > 0$, then we have $\overline{\pi}_{(g,f)}^E(\Phi) < 0$, which contradicts $\overline{\pi}_{(f,g)}^E(\Phi) = 0$. If $z < 0$, then by (50) one has $0 > \mathbb{E}^{\overline{\mathbb{Q}}'}[yf + z\Phi] = \mathbb{E}^{\overline{\mathbb{Q}}'}[z\Phi]$ for some $\overline{\mathbb{Q}}' \in \overline{\mathcal{M}}_{(g,f)} \subseteq \overline{\mathcal{M}}_g$, as $\overline{\mathcal{M}}_{(f,g)}$ is nonempty under the NA($\overline{\mathcal{P}}$) assumption in the case of $e + 1$ options. Then, in the case $z < 0$, one has $\mathbb{E}^{\overline{\mathbb{Q}}'}[\Phi] > 0 = \overline{\pi}_{(f,g)}^E(\Phi)$, which contradicts the weak duality result (45), and we hence conclude the proof of the duality.

6.3 ⎪ **Proof of Theorem 3.4.**

(a) For the first equality in (35), we first notice that for every $\hat{\mathbb{Q}} \in \widehat{\mathcal{M}}$ and a regular conditional probability measure $(\hat{\mathbb{Q}}_{\hat{\omega}})$ of $\hat{\mathbb{Q}}$ knowing $\hat{\mathcal{G}}_k$, one has $\hat{\mathbb{Q}}_{\hat{\omega}} \in \widehat{\mathcal{M}}_k(\hat{\omega})$. It then follows that $\sup_{\hat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\hat{\mathbb{Q}}}[\hat{\xi}] \leq$

$\sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_0} \mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}]$. Next, let $\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_0$. By its definition, one has $\widehat{\mathbb{Q}} \ll \widehat{\mathcal{P}}$ and (S, Y) is a generalized martingale and hence a local martingale under $\widehat{\mathbb{Q}}$; see, for example, lemma A.1 of Bouchard and Nutz (2015). Using lemma A.3 of Bouchard and Nutz (2015), there is a $\widehat{\mathbb{Q}}' \sim \widehat{\mathbb{Q}}$ such that $\mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}] \leq \mathbb{E}^{\widehat{\mathbb{Q}}'}[\widehat{\xi}]$, and (S, Y) is a $\widehat{\mathbb{Q}}'$-martingale. As $\widehat{\mathbb{Q}}' \sim \widehat{\mathbb{Q}} \ll \widehat{\mathcal{M}}$, then $\widehat{\mathbb{Q}}' \in \widehat{\mathcal{M}}$ and hence $\mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}] \leq \sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}]$. Then, it follows that $\sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_0} \mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}] \leq \sup_{\widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}} \mathbb{E}^{\widehat{\mathbb{Q}}}[\widehat{\xi}]$, and one obtains the first equality of (35).

(b) Next, we claim that the graph set

$$\left[\left[\widehat{\mathcal{M}}_{k,k+1}\right]\right] := \left\{ (\widehat{\omega}, \widehat{\mathbb{Q}}) \ : \ \widehat{\mathbb{Q}} \in \widehat{\mathcal{M}}_{k,k+1}(\widehat{\omega}) \right\} \tag{51}$$

is analytic. Then using the (analytic) measurable projection theorem; see, for example, proposition 4.47 of Bertsekas and Shreve (2007), $\widehat{\mathcal{E}}_k(\widehat{\xi})$ is upper semianalytic whenever $\widehat{\xi}$ is upper semianalytic. Further, the second equality in (35) is just a classical dynamic programming principle result, which follows by the measurable selection arguments as in Dellacherie (1985) or Bertsekas and Shreve (2007).

It is enough to prove the claim (51), for which we notice that

$$\left[\left[\widehat{\mathcal{M}}_{k,k+1}\right]\right] = \left\{ (\omega, y, \widehat{\mathbb{Q}}) \ : \ (\omega, \widehat{\mathbb{Q}}|_{\Omega_1}) \in [[\mathcal{M}_{k,k+1}]], \ \mathbb{E}^{\widehat{\mathbb{Q}}}[(\Delta S_{k+1}, \Delta Y_{k+1})] = 0 \right\},$$

where the graph $[[\mathcal{M}_{k,k+1}]]$ is analytic by lemma 4.8 of Bouchard and Nutz (2015). As $\widehat{\mathbb{Q}} \mapsto \mathbb{E}^{\widehat{\mathbb{Q}}}[(\Delta S_{k+1}, \Delta Y_{k+1})]$ is Borel measurable, $(\widehat{\omega}, \widehat{\mathbb{Q}}) \mapsto (\omega, \widehat{\mathbb{Q}}|_{\Omega_1})$ is continuous and hence is also Borel measurable. Then, by proposition 7.40 of Bertsekas and Shreve (2007), $[[\widehat{\mathcal{M}}_{k,k+1}]]$ is also analytic. □

7 | PROOFS FOR SECTION 4

We finally complete here the proof of Theorem 4.2, which concerns the pricing–hedging duality in the martingale optimal transport problem setup. Recall that, in this setup, $\overline{\Omega} := \Omega \times \{1, \dots, N\}$, with $\Omega := \mathbb{R}^{d \times N}$, and $\mathcal{P}$ the collection of all Borel probability measures on Ω.

A first idea of how to prove Theorem 4.2 could be the following two-step argument as in Guo et al. (2016). First, under the condition that Φ is bounded from above and upper semicontinuous, one could prove that

$$\mathfrak{P}((\mathbb{R}^d)^M) \ni \mu \mapsto \sup_{\overline{\mathbb{Q}} \in \mathcal{M}_\mu} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \in \mathbb{R}$$

is concave and upper semicontinuous, where we equip $\mathfrak{P}((\mathbb{R}^d)^M)$ with a Wasserstein type topology. Second, using the Fenchel–Moreau theorem, it follows that

$$\sup_{\overline{\mathbb{Q}} \in \mathcal{M}_\mu} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \overline{\pi}_{\mu,0}^E(\Phi) := \inf_{\lambda \in \Lambda} \left\{ \mu(\lambda) + \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi - \lambda] \right\}. \tag{52}$$

Solving the maximization problem (52), by using Theorem 3.2, concludes the proof of Theorem 4.2.

However, in the following, we will provide another proof based on an approximation argument. For simplicity, we suppose that $\mathbb{T}_0 = \{N\}$, where the same arguments work for more general $\mathbb{T}_0$. In

preparation, let us provide a technical lemma. In the context of the martingale optimal transport problem, we introduce a sequence of payoff functions $(g^i)_{i\geq 1}$ by

$$g^i(\omega) := f^i(\omega_N) - c^i \quad \text{with} \quad c^i := \int_{\mathbb{R}^d} f^i(x)\mu(dx),$$

where $f^i : \mathbb{R}^d \to \mathbb{R}$ is Lipschitz and $(f^i)_{i\geq 1}$ is dense in the space of all Lipschitz functions on $\mathbb{R}^d$ under the uniform convergence topology, and moreover, it contains all functions of the form $(x_j - K_n)^+$, $(-K_n - x_j)^+$ for $j = 1, \dots, d$ and $n \geq 1$, where $(K_n)_{n\geq 1} \subset \mathbb{R}$ is a sequence such that $K_n \to \infty$. Notice that μ has finite first order and hence the c^i are all finite constants.

Next, let us introduce an approximate dual problem by

$$\pi_{\mu,m}^A(\Phi) := \inf \left\{ x \; : \; \exists (\overline{H}, h) \in \overline{\mathcal{H}} \times \mathbb{R}^m \text{ s.t.} \quad \text{for all } k \in \mathbb{T}, \ \omega \in \Omega, \right.$$

$$\left. x + \sum_{i=1}^m h^i g^i(\omega_N) + (\overline{H}(\cdot, k) \circ S)_N(\omega) \geq \Phi_k(\omega) \right\}.$$

Similarly,

$$\overline{\mathcal{M}}_{\mu,m} := \{ \overline{\mathbb{Q}} \in \overline{\mathcal{M}} \; : \; \mathbb{E}^{\overline{\mathbb{Q}}}[g^i] = 0 \quad \text{for} \quad i = 1, \dots, m \},$$

and

$$P_{\mu,m} := \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{\mu,m}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi].$$

Lemma 7.1. *Let $(\overline{\mathbb{Q}}_m)_{m\geq 1} \subset \overline{\mathcal{M}}$ be a sequence of martingale measures such that $\overline{\mathbb{Q}}_m \in \overline{\mathcal{M}}_{\mu,m}$ for each $m \geq 1$. Then,*

(a) *$(\overline{\mathbb{Q}}_m)_{m\geq 1}$ is relatively compact under the weak convergence topology.*

(b) *The sequence $(S_N, \overline{\mathbb{Q}}_m)_{m\geq 1}$ is uniformly integrable, and any accumulation point of $(\overline{\mathbb{Q}}_m)_{m\geq 1}$ belongs to $\mathcal{M}_\mu$.*

Proof.

(a) Without loss of generality, we assume that $f_1(x) = \sum_{i=1}^d |x_i|$ so that

$$\sup_{m\geq 1} \mathbb{E}^{\overline{\mathbb{Q}}_m} \left[\sum_{i=1}^d |S_N^i| \right] < \int_{\mathbb{R}^d} \sum_{i=1}^d |x_i| \mu(dx) < \infty.$$

Let us first prove the relative compactness of $(\overline{\mathbb{Q}}_m)_{m\geq 1}$. By the Prokhorov theorem, it is enough to find, for every $\varepsilon > 0$, a compact set $D_\varepsilon \subset \mathbb{R}^d$ such that $\overline{\mathbb{Q}}_m[S_k \notin D_\varepsilon] \leq \varepsilon$ for all $k = 1, \dots, N$. It is then enough to find, for every $\varepsilon > 0$, a constant $K_\varepsilon > 0$ such that $\overline{\mathbb{Q}}_m[|S_k^i| \geq K_\varepsilon] \leq \varepsilon$ for all $i = 1, \dots, d$ and $k = 1, \dots, N$. Next, by the martingale property, one has $\mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_k^i|] \leq \mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_N^i|]$. Then, for every $\varepsilon > 0$, one can choose $K_\varepsilon > 0$ such that $\sup_{m\geq 1} \mathbb{E}^{\overline{\mathbb{Q}}_m}[\sum_{i=1}^d |S_N^i|] \leq K_\varepsilon \varepsilon$. It follows that $\overline{\mathbb{Q}}_m[|S_k^i| \geq K_\varepsilon] \leq \dfrac{\mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_k^i|]}{K_\varepsilon} \leq \varepsilon$, and hence that $(\overline{\mathbb{Q}}_m)_{m\geq 1}$ is relatively compact.

(b) To see that the sequence $(S_N, \overline{\mathbb{Q}}_m)_{m \geq 1}$ is uniformly integrable, it is enough to notice that $|x_i| \mathbf{1}_{|x_i| \geq 2K_n} \leq 2(|x_i| - K_n) \mathbf{1}_{|x_i| \geq K_n}$, where the latter is a payoff function contained in the sequence $(f_k)_{k \geq 1}$.

(c) Let $\overline{\mathbb{Q}}_0$ be an accumulation point of $(\overline{\mathbb{Q}}_m)_{m \geq 1}$, then one has $\mathbb{E}^{\overline{\mathbb{Q}}_0}[f^i(S_N)] = \mu(f^i) := \int_{\mathbb{R}^d} f^i(x)\mu(dx)$ for all $i \geq 1$. As $(f^i)_{i \geq 1}$ is supposed to be dense in the space of all Lipschtiz functions on $\mathbb{R}^d$ under the uniformly convergence topology, by dominated convergence, it follows that $\mathbb{E}^{\overline{\mathbb{Q}}_0}[f(S_N)] = \mu(f) := \int_{\mathbb{R}^d} f(x)\mu(dx)$ for all Lipschitz functions f. Therefore, one has $\overline{\mathbb{Q}}_0 \circ S_N^{-1} = \mu$.

(d) To conclude the proof, it is enough to show that the martingale property is preserved for the limiting measure $\overline{\mathbb{Q}}_0$. By extracting a subsequence, we assume that $\overline{\mathbb{Q}}_m \to \overline{\mathbb{Q}}_0$ weakly.

First, for any $k = 1, \dots, N$, $i = 1, \dots, d$ and $K > 0$, one has

$$\mathbb{E}^{\overline{\mathbb{Q}}_0}[|S_k^i| \wedge K] = \lim_{m \to \infty} \mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_k^i| \wedge K] \leq \limsup_{m \to \infty} \mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_k^i|]$$

$$\leq \limsup_{m \to \infty} \mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_N^i|] = \int_{\mathbb{R}^d} |x^i|\mu(dx) < \infty.$$

Let $K \to \infty$, it follows by the monotone convergence theorem that $\mathbb{E}^{\overline{\mathbb{Q}}_0}[|S_k^i|] < \infty$.

Next, we prove that for all $1 \leq k_1 < k_2 \leq N$, and any bounded continuous function $\varphi : (\mathbb{R}^d)^{k_1} \times \mathbb{T} \to \mathbb{R}$, one has

$$\mathbb{E}^{\overline{\mathbb{Q}}_0}[\varphi(S_1, \dots, S_{k_1}, T \wedge (k_1 + 1))(S_{k_2} - S_{k_1})] = 0. \tag{53}$$

Let $K > 0$, and $\chi_K : \mathbb{R}^d \to \mathbb{R}^d$ a continuous function, uniformly bounded by K, satisfying $\chi_K(x) = x$ when $\|x\| \leq K$, and $\chi_K(x) = 0$ when $\|x\| \geq K + 1$. Then, for every $m = 0$ or $m \geq 1$, one has

$$\left| \mathbb{E}^{\overline{\mathbb{Q}}_m} \left[\varphi(S, T)(S_{k_2} - S_{k_1}) \right] \right| \leq \left| \mathbb{E}^{\overline{\mathbb{Q}}_m} \left[\varphi(S, T) \left(\chi_K(S_{k_2}) - \chi_K(S_{k_1}) \right) \right] \right|$$

$$+ |\varphi|_\infty \mathbb{E}^{\overline{\mathbb{Q}}_m} \left[|S_{k_2}| \mathbf{1}_{|S_{k_2}| \geq K} + |S_{k_1}| \mathbf{1}_{|S_{k_1}| \geq K} \right], \tag{54}$$

where we simplify $\varphi(S_1, \dots, S_{k_1}, T \wedge (k_1 + 1))$ to $\varphi(S, T)$.

For every $\varepsilon > 0$, by uniformly integrability of $(S_N, \overline{\mathbb{Q}}_m)_{m \geq 1}$, there is $K_\varepsilon > 0$ such that

$$|\varphi|_\infty \mathbb{E}^{\overline{\mathbb{Q}}_m}[|S_{k_2}| \mathbf{1}_{|S_{k_2}| \geq K_\varepsilon} + |S_{k_1}| \mathbf{1}_{|S_{k_1}| \geq K_\varepsilon}] \leq \varepsilon, \quad \text{for all } m = 0, 1, \dots \tag{55}$$

Moreover, for $m \geq 1$, $\overline{\mathbb{Q}}_m$ is a martingale measure, so that $\mathbb{E}^{\overline{\mathbb{Q}}_m}[\varphi(S, T)(S_{k_2} - S_{k_1})] = 0$ and hence $|\mathbb{E}^{\overline{\mathbb{Q}}_m}[\varphi(S, T)(\chi_K(S_{k_2}) - \chi_K(S_{k_1}))]| \leq \varepsilon$. Then, by taking the limit as $m \to \infty$, it follows that

$$\left| \mathbb{E}^{\overline{\mathbb{Q}}_0} \left[\varphi(S, T) \left(\chi_K(S_{k_2}) - \chi_K(S_{k_1}) \right) \right] \right| \leq \varepsilon. \tag{56}$$

Combining (54), (55), and (56), and by the arbitrariness of $\varepsilon > 0$, it follows that (53) holds true for all bounded continuous functions φ.

Recall that $\overline{\mathcal{F}}_{k_1} = \sigma(S_1, \dots, S_{k_1}, T \wedge (k_1 + 1))$, as defined at the beginning of Section 2.2, and observe that $\overline{\mathcal{F}}_{k_1} = \sigma(\mathcal{K})$, where

$$\mathcal{K} := \{\xi := \varphi(S_1, \dots, S_{k_1}, T \wedge (k_1 + 1)) : \varphi \text{ is bounded and continuous}\}.$$

It follows that $\mathcal{K}$ is included in the vector space of all bounded $\overline{\mathcal{F}}_{k_1}$-measurable random variables ξ for which $\mathbb{E}^{\overline{\mathbb{Q}}_0}[\xi(S_{k_2} - S_{k_1})] = 0$. An application of the monotone class theorem (see, e.g., theorem I.8 of Protter, 2005) yields

$$\mathbb{E}^{\overline{\mathbb{Q}}_0}[\xi(S_{k_2} - S_{k_1})] = 0, \quad \text{for all bounded } \overline{\mathcal{F}}_{k_1}\text{-measurable random variables } \xi,$$

which is equivalent to S being a $\overline{\mathbb{Q}}_0$-martingale, and concludes the proof.

Proof of Theorem 4.2. We notice that by Theorem 3.2,

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{\mu,m}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \pi^A_{\mu,m}(\Phi) \geq \pi^A_\mu(\Phi).$$

Let $(\overline{\mathbb{Q}}_m)_{m \geq 1}$ be a sequence of probability measures such that $\overline{\mathbb{Q}}_m \in \overline{\mathcal{M}}_{\mu,m}$ for each $m \geq 1$ and

$$\limsup_{m \to \infty} \mathbb{E}^{\overline{\mathbb{Q}}_m}[\Phi_T(S)] = \limsup_{m \to \infty} \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{\mu,m}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi].$$

It follows by Lemma 7.1 that there exists some $\overline{\mathbb{Q}}_0 \in \overline{\mathcal{M}}_\mu$ and a subsequence such that $\overline{\mathbb{Q}}_{m_k} \to \overline{\mathbb{Q}}_0$ under the weak convergence topology. Using upper semicontinuity of Φ and by Fatou's lemma, it follows that $\mathbb{E}^{\overline{\mathbb{Q}}_0}[\Phi_T(S)] \geq \limsup_{m \to \infty} \mathbb{E}^{\overline{\mathbb{Q}}_m}[\Phi_T(S)]$. This leads to the inequality

$$\sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_\mu} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] \geq \mathbb{E}^{\overline{\mathbb{Q}}_0}[\Phi] \geq \limsup_{m \to \infty} \sup_{\overline{\mathbb{Q}} \in \overline{\mathcal{M}}_{\mu,m}} \mathbb{E}^{\overline{\mathbb{Q}}}[\Phi] = \limsup_{m \to \infty} \pi^A_{\mu,m}(\Phi) \geq \pi^A_\mu(\Phi),$$

and we hence conclude the proof by the weak duality (45).

CONFLICT OF INTEREST

ACKNOWLEDGMENTS

REFERENCES

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

Acciaio, B., & Penner, I. (2011). Dynamic risk measures. In D. Nunno & Øksendal (Eds.), *Advanced mathematical methods for finance* (pp. 1–34). Berlin Heidelberg: Springer.

Aksamit, A., & Li, L. (2016). Projections, pseudo-stopping times and the immersion property. In C. Donati-Martin & A. R. A. Lejay (Eds.), *Séminaire de Probabilités XLVIII, number 2168 in lecture notes in math* (pp. 459–467). Berlin and New York: Springer.

Bayraktar, E., Huang, Y.-J., & Zhou, Z. (2015). On hedging American options under model uncertainty. *SIAM Journal on Financial Mathematics*, *6*(1), 425–447.

Bayraktar, E., & Zhou, Z. (2017). Super-hedging American options with semi-static trading strategies under model uncertainty. *International Journal of Theoretical and Applied Finance*, *20*(6), 1–10.

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

Bertsekas, D. P., & Shreve, S. E. (2007). *Stochastic optimal control: The discrete-time case*. Belmont MA: Athena Scientific.

Blanchet-Scalliet, C., Jeanblanc, M., & Romero, R. R. (2016). Enlargement of filtration in discrete time. preprint, https://hal.archives-ouvertes.fr/hal-01253214.

Bouchard, B., & Nutz, M. (2015). Arbitrage and duality in nondominated discrete-time models. *Annals of Applied Probability*, *25*(2), 823–859.

Burzoni, M., Frittelli, M., Hou, Z., Maggis, M., & Obłój, J. (2018). Pointwise arbitrage pricing theory in discrete time. *Mathematics of Operations Research*. Preprint, available at https://doi.org/10.1287/moor.2018.0956

Burzoni, M., Frittelli, M., & Maggis, M. (2017). Model-free superhedging duality. *Annals of Applied Probability*, *27*(3), 1452–1477.

Cox, A. M. G., & Hoeggerl, C. (2016). Model-independent no-arbitrage conditions on American put options. *Mathematical Finance*, *26*(2), 431–458.

Cox, A. M. G., Hou, Z., & Obłój, J. (2016). Robust pricing and hedging under trading restrictions and the emergence of local martingale models. *Finance and Stochastics*, *20*(3), 669–704.

Cox, A. M. G., & Källblad, S. (2017). Model-independent bounds for asian options: A dynamic programming approach. *SIAM Journal on Control and Optimization*, *55*(6), 3409–3436.

Cox, A. M. G., & Obłój, J. (2011). Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, *15*(3), 573–605.

Delbaen, F., & Schachermayer, W. (2006). *The mathematics of arbitrage*. Berlin: Springer Finance, Springer-Verlag.

Dellacherie, C. (1985). Quelques résultats sur les maisons de jeu analytiques. In Azéma, J. and Yor, M. (Eds.), *Séminaire de Probabilités, XIX*, *volume 1123 of lecture notes in math* (pp. 222–229). Berlin and New York: Springer.

Denis, L., & Martini, C. (2006). A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Annals of Applied Probability*, *16*(2), 827–852.

Dolinsky, Y. (2014). Hedging of game options under model uncertainty in discrete time. *Electronic Communications in Probability*, *19*(19), 1–11.

Dolinsky, Y., & Soner, H. M. (2014). Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields*, *160*(1-2), 391–427.

El Karoui, N., & Tan, X. (2013). Capacities, measurable selection and dynamic programming part ii: Application in stochastic control problems. arXiv preprint arXiv:1310.3364.

Eldan, R. (2016). Skorokhod embeddings via stochastic flows on the space of Gaussian measures. *Annales de l'Institut Henri Poincaré B: Probability and Statistics*, *52*(3), 1259–1280.

Föllmer, H., & Schied, A. (2004). *Stochastic finance, an introduction in discrete time* (2nd ed.). Berlin, New York: Walter de Gruyter.

Guo, G., Tan, X., & Touzi, N. (2016). Optimal Skorokhod embedding under finitely many marginal constraints. *SIAM Journal on Control and Optimization*, *54*(4), 2174–2201.

Guo, G., Tan, X., & Touzi, N. (2017). Tightness and duality of martingale transport on the Skorokhod space. *Stochastic Processes and Their Applications*, *127*(3), 927–956.

Hobson, D. (1998). Robust hedging of the lookback option. *Finance and Stochastics*, *2*(4), 329–347.

Hobson, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton lectures on mathematical finance 2010*, *volume 2003 of lecture notes in math* (pp. 267–318). Berlin: Springer.

Hobson, D., & Neuberger, A. (2016). More on hedging American options under model uncertainty. arXiv preprint arXiv:1604.02274.

Hobson, D., & Neuberger, A. (2017). Model uncertainty and the pricing of American options. *Finance and Stochastics*, *21*(1), 285–329.

Hou, Z., & Obłój, J. (2018). Robust pricing–hedging dualities in continuous time. *Finance and Stochastics*, *22*(3), 511–567.

Jeanblanc, M., & Song, S. (2011a). An explicit model of default time with given survival probability. *Stochastic Processes and Their Applications*, *121*(8), 1678–1704.

Jeanblanc, M., & Song, S. (2011b). Random times with given survival probability and their f-martingale decomposition formula. *Stochastic Processes and Their Applications*, *121*(6), 1389–1410.

Kabanov, Y. (2008). In discrete time a local martingale is a martingale under an equivalent probability measure. *Finance and Stochastics*, *12*(3), 293–297.

Kabanov, Y., & Stricker, C. (2001). A teacher's note on no-arbitrage criteria. In Azéma, J., Emery, M., Ledoux, M., and Yor, M. (Eds.), *Séminaire de Probabilités, XXXV*, *volume 1755 of lecture notes in math* (pp. 149–152). Berlin: Springer.

Källblad, S., Tan, X., & Touzi, N. (2017). Optimal Skorokhod embedding given full marginals and Azéma-Yor peacocks. *Annals of Applied Probability*, *27*(2), 686–719.

Lim, T. (2016). Multi-martingale optimal transport.

Mansuy, R., & Yor, M. (2006). *Random times and enlargements of filtrations in a Brownian setting*, *volume 1873 of lecture notes in math*. Berlin: Springer.

Myneni, R. (1992). The pricing of the American option. *Annals of Applied Probability*, *2*(1), 1–23.

Nadtochiy, S., & Obłój, J. (2017). Robust trading of implied skew. *International Journal of Theoretical and Applied Finance*, *20*(02), 1750008.

Neuberger, A. (2007). Bounds on the American option. Preprint, http://ssrn.com/abstract=966333.

Neufeld, A., & Nutz, M. (2013). Superreplication under volatility uncertainty for measurable claims. *Electronic Journal of Probability*, *18*(48), 1–14.

Nikeghbali, A., & Yor, M. (2005). A definition and some characteristic properties of pseudo-stopping times. *Annals of Applied Probability*, *33*(5), 1804–1824.

Obłój, J. (2004). The Skorokhod embedding problem and its offspring. *Probability Surveys*, *1*, 321–392.

Pham, T., & Zhang, J. (2014). Two person zero-sum game in weak formulation and path dependent Bellman–Isaacs equation. *SIAM Journal on Control and Optimization*, *52*(4), 2090–2121.

Possamaï, D., Royer, G., & Touzi, N. (2013). On the robust superhedging of measurable claims. *Electronic Communications in Probability*, *18*, 1–13.

Protter, P. (2005). *Stochastic integration and differential equations*. Berlin Heidelberg: Springer-Verlag.

Soner, H. M., Touzi, N., & Zhang, J. (2013). Dual formulation of second order target problems. *Annals of Applied Probability*, *23*(1), 308–347.

Tan, X., & Touzi, N. (2013). Optimal transportation under controlled stochastic dynamics. *Annals of Probability*, *41*(5), 3201–3240.

Williams, D. (2002). A "non-stopping" time with the optional-stopping property. *Bulletin of the London Mathematical Society*, *34*(5), 610–612.