Finance Stoch
DOI 10.1007/s00780-012-0192-1

Duality and convergence for binomial markets with friction

Yan Dolinsky · Halil Mete Soner

Received: 15 June 2011 / Accepted: 27 February 2012

Abstract We prove limit theorems for the super-replication cost of European options in a binomial model with friction. Examples covered are markets with proportional transaction costs and illiquid markets. A dual representation for the super-replication cost in these models is obtained and used to prove the limit theorems. In particular, the existence of a liquidity premium for the continuous-time limit of the model proposed in Çetin et al. (Finance Stoch. 8:311–341, 2004) is proved. Hence, this paper extends the previous convergence result of Gökay and Soner (Math Finance 22:250–276, 2012) to the general non-Markovian case. Moreover, the special case of small transaction costs yields, in the continuous limit, the G-expectation of Peng as earlier proved by Kusuoka (Ann. Appl. Probab. 5:198–221, 1995).

Keywords Super-replication · Liquidity · Binomial model · Limit theorems · G-Expectation

Mathematics Subject Classification (2010) 91G10 · 60F05 · 60H30

JEL Classification G11 · G13 · D52

1 Introduction

We consider a one-dimensional binomial model in which the size of a trade has an immediate but temporary effect on the price of the asset. Indeed, let $g(t, \nu)$ be the

Y. Dolinsky (✉) · H.M. Soner
Dept. of Mathematics, ETH Zurich, Zurich, Switzerland
e-mail: yan.dolinsky@math.ethz.ch

H.M. Soner
Swiss Finance Institute, Zurich, Switzerland
e-mail: mete.soner@math.ethz.ch

Published online: 10 July 2012

cost of trading ν shares at time t. We simply assume that g is adapted to the natural filtration and convex in ν with $g(t,0)=0$. In this generality, this model corresponds to the classical transaction cost model when $g(t,\nu)=\lambda|\nu|$ with given constant $\lambda>0$. However, it also covers the illiquidity model considered in [4] and [13], which is the binomial version of the model introduced by Çetin et al. in [5] for continuous time.

In continuous time, the super-replication cost of a European option behaves quite differently depending on the structure of g. In the case of proportional transaction costs (i.e., when g is nondifferentiable at the origin), the super-replication cost is very high as proved in [8, 14, 17, 22]. In several papers [2, 16], asymptotic problems with vanishing transaction costs are considered to obtain nontrivial pricing equations. On the other hand, when g is differentiable, then any continuous trading strategy which has finite variation has no liquidity cost. Thus, one may avoid the liquidity cost entirely as shown in [5] and also in [1]. However, in [6] it is shown that mild constraints on the admissible strategies render these approximation inadmissible and one has a liquidity premium. This result is further verified in [13], which derives the same premium as the continuous-time limit of binomial models. The equation satisfied by this limit is a nonlinear Black–Scholes equation

$$-u_t(t,s)+\frac{\sigma^2 s^2}{2}H\big(u_{ss}(t,s)\big)=0 \quad \forall t<T,\ s>0, \tag{1.1}$$

where t is time, s is the current stock price, and H is a convex nonlinear function of the second derivative derived explicitly in [6, 13]. Since H is convex, the above equation is the dynamic programming equation of a stochastic optimal control problem. Then this problem may be considered as the dual of the original super-replication problem.

The proof given in [13] depends on homogenization techniques for viscosity solutions. Thus, it is limited to Markovian claims. Moreover, the mentioned duality result is obtained only through the partial differential equation and not by a direct argument.

In this paper, we extend the study of [13] to non-Markovian claims and to more general liquidity functions g. The model is again a simple one-dimensional model with trading cost g. In this formulation, the super-replication problem is a convex optimization problem, and its dual can be derived by classical theory. This derivation is another advantage of the discrete model as a derivation of the dual in continuous time is essentially an open problem. However, a new approach is now developed in [24]. The dual is an optimal control problem in which the controller is allowed to choose different probability measures. We then use this dual representation to formally identify the limit optimal control problem. The dynamic programming equation of this optimal control problem is given by (1.1) in the Markov case. This representation also allows us to prove the continuous-time limit.

Our approach is purely probabilistic and allows us to deal with path-dependent payoffs and path-dependent penalty functions g. One of the key steps is a construction of Kusuoka given in the context of transaction costs. Indeed, given a martingale M on the Brownian probability space whose volatility satisfies some regularity conditions, Kusuoka in [16] constructs on the discrete probability space $\{-1,1\}^{\mathbb{N}}$ a sequence of martingales of a specific form which converge in law to M. Moreover, the quadratic variation of M is also approximated through this powerful procedure of Kusuoka.

This construction is our main tool in proving the lower bound (i.e., the existence of a liquidity premium) for the continuous-time limit of the super-replication costs. The upper bound follows from compactness and two general lemmas (Lemmas 7.1 and 7.2).

As remarked before, the super-replication cost can be quite expensive in markets with transaction costs. Therefore, if $g(t,\nu)=\lambda|\nu|$ and $\lambda>0$ is a given constant, one obtains a trivial result in the continuous-time limit. So we need to scale the proportionality constant λ as the time discretization gets finer. Indeed, if we take $\lambda_n=\Lambda/\sqrt{n}$ in an n-step model, then the limit problem is the uncertain volatility model or equivalently G-expectation of Peng [18]. This is exactly the main result of Kusuoka in [16]. In fact, relatedly, the authors in joint work with Nutz [9] provide a different discretization of the G-expectation.

The paper is organized as follows. In the next section, we introduce the setup. In Sect. 3, we formulate the main results. In Sect. 4, we prove Theorem 3.1, which is a duality result for the super-replication prices in the binomial models. The main tool that is used in this section is the Kuhn–Tucker theory for convex optimization. Theorem 3.5 which describes the asymptotic behavior of the super-replication prices is proved in Sect. 5. In Sect. 6, we state the main results from [16] which are used in this paper. In particular, we give a short formulation of the main properties of the Kusuoka construction, which is the main tool in proving the lower bound (liquidity premium) of Theorem 3.5. In Sect. 7, we derive auxiliary lemmas that are used in the proof of Theorem 3.5.

2 Preliminaries and the model

Let $\Omega=\{-1,1\}^{\mathbb{N}}$ be the space of infinite sequences $\omega=(\omega_1,\omega_2,\ldots)$, where $\omega_i\in\{-1,1\}$, with the product probability $\mathbb{Q}=\{\frac{1}{2},\frac{1}{2}\}^{\mathbb{N}}$. Define the canonical sequence of independent and identically distributed (i.i.d.) random variables $\xi_1,\xi_2,\ldots$ by

$$\xi_i(\omega)=\omega_i,\quad i\in\mathbb{N},$$

and consider the natural filtration $\mathcal{F}_k=\sigma\{\xi_1,\ldots,\xi_k\}$, $k\geq 1$, and let $\mathcal{F}_0$ be trivial.

For any $T>0$, denote by $\mathcal{C}[0,T]$ the space of all continuous functions on $[0,T]$ with the uniform topology induced by the norm $\|y\|_\infty=\sup_{0\leq t\leq T}|y(t)|$. Let $F:\mathcal{C}[0,T]\to\mathbb{R}_+$ be a continuous map such that there exist constants $C,p>0$ for which

$$F(y)\leq C\big(1+\|y\|_\infty^p\big)\quad\forall y\in\mathcal{C}[0,T].\tag{2.1}$$

Without loss of generality, we take $T=1$.

Next, we introduce a sequence of binomial models for which the volatility of the stock price is a constant $\sigma>0$ (which is independent of n). For any n, consider the n-step binomial model of a financial market which is active at times $0,1/n,2/n,\ldots,1$. We assume that the market consists of a savings account and a stock. Without loss of generality (by discounting), we assume that the savings account price is a constant which equals 1 and the stock price at time k/n is given by

$$S^{(n)}(k) = s_0 \exp\left(\sigma\sqrt{\frac{1}{n}}\sum_{i=1}^{k}\xi_i\right), \quad k = 0, 1, \dots, n, \tag{2.2}$$

where $s_0 > 0$ is the initial stock price. For any $n \in \mathbb{N}$, let $\mathcal{W}_n : \mathbb{R}^{n+1} \to \mathcal{C}[0, 1]$ be the linear interpolation operator given by

$$\mathcal{W}_n(y)(t) := \big([nt] + 1 - nt\big)y\big([nt]\big) + \big(nt - [nt]\big)y\big([nt] + 1\big) \quad \forall t \in [0, 1],$$

where $y = \{y(k)\}_{k=0}^{n} \in \mathbb{R}^{n+1}$, and $[z]$ denotes the integer part of z. We consider a (possibly path-dependent) European contingent claim with maturity $T = 1$ and payoff

$$F_n := F\big(\mathcal{W}_n\big(S^{(n)}\big)\big), \tag{2.3}$$

where, by definition, we consider $\mathcal{W}_n(S^{(n)})$ as a random element in $\mathcal{C}[0, 1]$.

For future reference, let $\mathcal{C}^+[0, 1]$ be the set of all strictly positive continuous functions on $[0, 1]$ with the uniform topology. Then clearly $\mathcal{W}_n(S^{(n)})$ is an element of $\mathcal{C}^+[0, 1]$.

2.1 Wealth dynamics and super-replication

Next, we define the notion of a self-financing portfolio in these models. Fix $n \in \mathbb{N}$ and consider an n-step binomial model with a penalty function g. This function represents the cost of trading in this market, namely $g(t, s, \nu)$ is the transaction cost which the investor pays at time t given that the stock price process and the trading volume at this moment are equal to s and ν, respectively. We assume the following.

Assumption 2.1 The *trading cost function*

$$g : [0, 1] \times \mathcal{C}^+[0, 1] \times \mathbb{R} \to [0, \infty)$$

is assumed to be nonnegative, adapted, and such that $g(t, S, 0) = 0$. Moreover, we assume that $g(t, S, \cdot)$ is convex for every $(t, S) \in [0, 1] \times \mathcal{C}^+[0, 1]$.

In this simple setting, the adaptedness of g simply means that $g(t, S, \nu)$ depends only on the restriction of S to the interval $[0, t]$, namely

$$g(t, S, \nu) = g(t, \hat{S}, \nu) \quad \text{whenever } S(s) = \hat{S}(s)\ \forall s \le t.$$

A self-financing portfolio π with initial capital x is a pair $\pi = (x, \{\gamma(k)\}_{k=0}^{n})$ where $\gamma(0) = 0$ and for any $k \ge 1$, $\gamma(k)$ is an $\mathcal{F}_{k-1}$-measurable random variable. Here $\gamma(k)$ represents the number of stocks that the investor holds at the moment k/n *before* a transfer is made at this time. The portfolio value $Y^\pi(k) := Y^\pi(k : g)$ of a trading strategy π is given by the difference equation

$$\begin{aligned} Y^\pi(k+1) = {} & Y^\pi(k) + \gamma(k+1)\big(S^{(n)}(k+1) - S^{(n)}(k)\big) \\ & - g\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), \gamma(k+1) - \gamma(k)\right) \end{aligned} \tag{2.4}$$

for $k = 0, \dots, n-1$ and with initial data $Y^\pi(0) = x$.

The term $Y^\pi(k)$ is the portfolio value at time k/n *before* a transfer is made at this time, and the last term in (2.4) represents the cost of trading and is the only source of friction in the model. Note that the random variable $g(\frac{k}{n}, \mathcal{W}_n(S^{(n)}), \gamma(k+1) - \gamma(k))$ is $\mathcal{F}_k$-measurable. We mostly use the notation $Y^\pi(k)$ when the dependence on the penalty function is clear.

Let $\mathcal{A}_n(x)$ be the set of all portfolios with initial capital x. The problem we consider is the super-replication of a European claim whose payoff is given in (2.3). The formal definition of the super-replication price is given by

$$V_n := V_n(g, F_n) = \inf\bigl\{x \bigm| \exists \pi \in \mathcal{A}_n(x) \text{ with } Y^\pi(n:g) \geq F_n \ \mathbb{Q}\text{-a.s.}\bigr\}. \tag{2.5}$$

2.2 Trading cost

In this subsection, we state the main assumption on g in addition to Assumption 2.1. We also provide several examples and make the connection to models with proportional transaction costs and models with price impact.

Let $G : [0,1] \times \mathcal{C}^+[0,1] \times \mathbb{R} \to [0,\infty]$ be the Legendre transform (or convex conjugate) of g,

$$G(t,S,y) = \sup_{\nu\in\mathbb{R}} \bigl(\nu y - g(t,S,\nu)\bigr), \quad \forall (t,S,y) \in [0,1] \times \mathcal{C}^+[0,1] \times \mathbb{R}. \tag{2.6}$$

Observe that G may become infinite. It is well known that we have the dual relation

$$g(t,S,\nu) = \sup_{y\in\mathbb{R}} \bigl(\nu y - G(t,S,y)\bigr), \quad \forall (t,S,\nu) \in [0,1] \times \mathcal{C}^+[0,1] \times \mathbb{R}.$$

Example 2.2 The following three cases provide the essential examples of the theory developed in this paper.

(a) For a given constant $\Lambda > 0$, let

$$g(t,S,\nu) = \Lambda \nu^2.$$

In this example, we directly calculate that

$$G(t,S,y) = y^2/4\Lambda.$$

This penalty function is the binomial version of the linear liquidity model of Çetin et al. [5] that was studied in [13] (see Remark 2.4 below).

In [23], it is proved that the optimal trading strategies in continuous time with a smooth g do not have jumps. Hence, one expects that in a binomial model with large n, the optimal portfolio changes are also small. Thus, any trading cost g which is twice differentiable essentially behaves like this example with $\Lambda = g_{\nu\nu}(t,S,0)/2$.

(b) This example corresponds to the example of proportional transaction costs. For fixed n, there have recently been interesting results in relation to arbitrage. We refer to the papers of Schachermayer [21], Pennanen and Penner [19], and the references therein. But as remarked earlier, a fixed transaction cost forces the super-replication to be very costly as n tends to infinity. Hence, we take a sequence of problems with vanishing transaction costs,

$$g_n^c(t,S,\nu) = \frac{c}{\sqrt{n}} S(t)|\nu|,$$

where $c>0$ is a constant. This discrete financial market with vanishing transaction costs is the model studied in [16] by Kusuoka. In this case, the dual function is

$$G_n^c(t,S,y)=\begin{cases}0 & \text{if } |y|\le c\,S(t)/\sqrt{n},\\ +\infty & \text{else.}\end{cases}$$

(c) This example is a mixture of the previous two. It is obtained by appropriately modifying the liquidity example. In our analysis, this modification will be used in several places. For a given constant c, let

$$G_n^c(t,S,y)=\begin{cases}y^2/4\Lambda & \text{if } |y|\le cS(t)/\sqrt{n},\\ +\infty & \text{else.}\end{cases}$$

We directly calculate that

$$g_n^c(t,S,\nu)=\begin{cases}\Lambda\nu^2 & \text{if } |\nu|\le \frac{cS(t)}{2\sqrt{n}\Lambda},\\ \frac{c}{\sqrt{n}}S(t)|\nu|-\frac{c^2S^2(t)}{4n\Lambda} & \text{else.}\end{cases}$$

In the above, the third example is obtained from the first one through an appropriate truncation of the dual cost function G. One may perform the same modification to all given penalty functions g. The following definition formalizes this.

Definition 2.3 Let $g:[0,1]\times\mathcal{C}^+[0,1]\times\mathbb{R}\to[0,\infty]$ be a convex function (in the third argument) with $g(t,S,0)=0$. Then the *truncation of g at level c* is given by

$$g_n^c(t,S,\nu):=g_n^c(t,S,\nu:g)=\sup\left\{\nu y-G(t,s,y)\;\middle|\;|y|\le cS(t)/\sqrt{n}\right\},$$

where G is the convex conjugate of the original g.

An important but simple observation is the structure of the dual function of g_n^c. Indeed, it is clear that the Legendre transform G_n^c of g_n^c is simply given by

$$G_n^c(t,S,y)=\begin{cases}G(t,S,y) & \text{if } |y|\le cS(t)/\sqrt{n},\\ +\infty & \text{else,}\end{cases}\tag{2.7}$$

where G is the Legendre transform of g.

Note that for any $n\in\mathbb{N}$, g_n^c converges monotonically to g as c tends to infinity. Also, observe that Example 2.2(b) is the truncation of the function

$$g(t,S,\nu)=\begin{cases}0 & \text{if } \nu=0,\\ +\infty & \text{else.}\end{cases}$$

Example 2.2(c), however, corresponds to the truncation of $g(t,S,\nu)=\Lambda\nu^2$.

We close this subsection by connecting the above model to the discrete liquidity models.

Remark 2.4 Following the liquidity model which was introduced in [5], we introduce a path-dependent supply curve

$$\mathbf{S}:[0,1]\times\mathcal{C}^+[0,1]\times\mathbb{R}\to\mathbb{R}.$$

We assume that $\mathbf{S}(t, S, \cdot)$ is adapted, i.e., it depends only on the restriction of S to the interval $[0, t]$, namely

$$\mathbf{S}(t, S, \nu) = \mathbf{S}(t, \hat{S}, \nu) \quad \text{whenever } S(s) = \hat{S}(s)\ \forall s \le t.$$

In the n-step binomial model, the price per share at time t is $\mathbf{S}(t, \mathcal{W}_n(S^{(n)}), \nu)$, where ν is the size of the transaction of the investor. The penalty which represents the liquidity effect of the model is then given by

$$g(t, S, \nu) = \big(\mathbf{S}(t, S, \nu) - S(t)\big)\nu \quad \forall (t, S, \nu) \in [0, 1] \times \mathcal{C}^+[0, 1] \times \mathbb{R}.$$

3 Main results

Our first result is a characterization of the dual problem. We believe that this simple result is of independent interest as well. Also, it will be the essential tool to study the asymptotic behavior of the super-replication cost.

Recall that F_n and V_n are given, respectively, in (2.3) and (2.5). Moreover, g is the trading cost function, and G is its Legendre transform.

Theorem 3.1 (Duality) *Let $\mathcal{Q}_n$ be the set of all probability measures on $(\Omega, \mathcal{F}_n)$. Then*

$$V_n = \sup_{\mathbb{P} \in \mathcal{Q}_n} \mathbb{E}^{\mathbb{P}} \left(F_n - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), \mathbb{E}^{\mathbb{P}}\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k)\right)\right),$$

where $\mathbb{E}^{\mathbb{P}}$ denotes the expectation with respect to a probability measure $\mathbb{P}$.

The above duality is proved in the next section.

We continue by discussing the main assumption on the trading costs. We assume that the Legendre transform G of the convex penalty function g satisfies the following.

Assumption 3.2 We assume that G satisfies the following growth and scaling conditions:

(a) There are constants $C, p > 0$, and $\beta \ge 2$ such that

$$G(t, S, y) \le C y^{\beta} \big(1 + \|S\|_\infty\big)^p \quad \forall (t, S, y) \in [0, 1] \times \mathcal{C}^+[0, 1] \times \mathbb{R}. \tag{3.1}$$

(b) There exists a continuous function

$$\widehat{G} : [0, 1] \times \mathcal{C}^+[0, 1] \times \mathbb{R} \to [0, \infty]$$

such that for any bounded sequence $\{\alpha_n\}$, discrete-valued sequence $\{\xi_n\}$ in $\{-1, 1\}$, and convergent sequences $t_n \to t$, $S^{(n)} \to S$ (in the $\|\cdot\|_\infty$-norm), we have

$$\lim_{n\to\infty} \left| nG\left(t_n, S^{(n)}, \frac{\xi_n \alpha_n}{\sqrt{n}} S^{(n)}(t_n)\right) - \widehat{G}\big(t, S, \alpha_n S(t)\big)\right| = 0. \tag{3.2}$$

It is straightforward to show that $\widehat{G}$ is quadratic in the y-variable. Moreover, the above Assumption 3.2 is essentially equivalent to assuming that G is twice differentiable at the origin. Indeed, if the latter holds, a Taylor approximation implies that

$$\widehat{G}(t,S,y)=\frac{1}{2}y^2G_{yy}(t,S,0).$$

We give the following example to clarify the above assumption.

Example 3.3 For $\gamma \geq 1$, let

$$g_\gamma(\nu)=\frac{1}{\gamma}|\nu|^\gamma.$$

Then, for $\gamma>1$,

$$G_\gamma(y)=\frac{1}{\gamma^*}\,|y|^{\gamma^*},\qquad \gamma^*=\frac{\gamma}{\gamma-1}.$$

For $\gamma=1$, $G_1(y)=0$ for $|y|\leq 1$ and is equal to infinity otherwise. Moreover, we directly calculate that $\widehat{G}_\gamma(0)=0$ and for $y\neq 0$,

$$\widehat{G}_\gamma(y):=\lim_{n\to\infty} nG_\gamma\left(\frac{y}{\sqrt{n}}\right)=\begin{cases} G_2(t,y) & \text{if } \gamma=2,\\ 0 & \text{if } \gamma\in[1,2),\\ \infty & \text{if } \gamma>2.\end{cases}$$

Notice that G_γ is twice differentiable at the origin only for $\gamma\in[1,2]$.

To describe the continuous-time limit, we need to introduce some further notation. Let $(\Omega_W,\mathcal{F}^W,\mathbb{P}^W)$ be a complete probability space together with a standard one-dimensional Brownian motion W and the right-continuous filtration given by $\mathcal{F}^W_t=\sigma(\sigma\{W(s)|s\leq t\}\cup\mathcal{N})$, where $\mathcal{N}$ is the collection of all $\mathbb{P}^W$-null sets. For any progressively measurable, bounded, nonnegative α, let S_α be the continuous martingale given by

$$S_\alpha(t)=s_0\exp\left(\int_0^t\alpha(u)\,dW(u)-\frac{1}{2}\int_0^t\alpha^2(u)\,du\right),\qquad t\in[0,1]. \tag{3.3}$$

We also introduce the following notation which is related to the quadratic variation density of $\ln S_\alpha$. Recall that the constant σ is the volatility that was already introduced in the dynamics of the discrete stock price process in (2.2). Set

$$a(t:S_\alpha):=\frac{\frac{d\langle\ln S_\alpha\rangle(t)}{dt}-\sigma^2}{2\sigma}=\frac{\alpha^2(t)-\sigma^2}{2\sigma}. \tag{3.4}$$

The continuous limit is given through an optimal control problem in which α is the control and S_α is the controlled state process. To complete the description of this control problem, we need to specify the set of admissible controls.

Definition 3.4 For any constant $c>0$, an *admissible control at the level c* is a progressively measurable, nonnegative process $\alpha(\cdot)$ satisfying

$$\bigl|a(\cdot:S_\alpha)\bigr|\leq c\quad \mathcal{L}\otimes\mathbb{P}^W\text{-a.s.},$$

where $\mathcal{L}$ is the Lebesgue measure on $[0,1]$. The set of all admissible controls at the level c is denoted by $\mathcal{A}^c$.

As before, g is the penalty function, and g_n^c is the truncation of g at the level c as in Definition 2.3. Let F_n be a given claim, and $V_n = V_n(g, F_n)$ the super-replication cost defined in (2.5). For any level c, let $V_n^c = V_n(g_n^c, F_n)$.

The following theorem, which will be proved in Sect. 5, is the main result of the paper. It provides the asymptotic behavior of the truncated super-replication costs V_n^c. Since $V_n^c \leq V_n$ for every c, the result below can be used to show the existence of a liquidity premium as it was done for a Markovian example in [13]; see Corollary 3.6 and Remark 3.7 below.

Theorem 3.5 (Convergence) *Let G be a dual function satisfying Assumption 3.2, and let $\widehat{G}$ be as in (3.2). Then, for every $c > 0$,*

$$\lim_{n\to\infty} V_n^c = \sup_{\alpha\in\mathcal{A}^c} J(S_\alpha),$$

where

$$J(S_\alpha) := \mathbb{E}^W\left(F(S_\alpha) - \int_0^1 \widehat{G}\big(t, S_\alpha, a(t:S_\alpha)S_\alpha(t)\big)\,dt\right), \tag{3.5}$$

and $\mathbb{E}^W$ denotes the expectation with respect to $\mathbb{P}^W$.

Since $V_n^c \leq V_n$ for every $c > 0$, we have the following immediate corollary.

Corollary 3.6

$$\liminf_{n\to\infty} V_n \geq \sup_{\alpha\in\mathcal{A}} \mathbb{E}^W\left(F(S_\alpha) - \int_0^1 \widehat{G}\big(t, S_\alpha, a(t:S_\alpha)S_\alpha(t)\big)\,dt\right), \tag{3.6}$$

where $\mathcal{A}$ is the set of all bounded, nonnegative, progressively measurable processes.

A natural question which for now remains open is under which assumptions the above inequality is in fact an equality. For the specific quadratic penalty and Markovian payoffs, [13] proves the equality.

Remark 3.7 (Liquidity premium) A related question is whether the limiting super-replication cost contains a liquidity premium, i.e., whether the right-hand side of (3.6) is strictly bigger than the Black–Scholes price $V_{BS}(F)$. For Markovian non-affine payoffs, this was proved in [6]. Notice that the standard Black–Scholes price is given by $V_{BS}(F) := \mathbb{E}^W(F(S_\sigma))$, and this can be achieved by simply choosing the control $\alpha \equiv \sigma$ in the right-hand side of (3.6).

In the generality considered in this paper, the following argument might be utilized to establish a liquidity premium. Fix $\epsilon > 0$. From (3.1) one can prove the estimate

$$\sup_{\alpha\in\mathcal{A}^\epsilon} \mathbb{E}^W\left(\int_0^1 G\big(t, S_\alpha, a(t:S_\alpha)S_\alpha(t)\big)\,dt\right) = O\big(\epsilon^2\big).$$

Thus, in order to prove the strict inequality, it remains to show that there exists a constant $C > 0$ such that

$$\sup_{\alpha\in\mathcal{A}^\epsilon} \mathbb{E}^W\big(F(S_\alpha)\big) \geq \mathbb{E}^W\big(F(S_\sigma)\big) + C\epsilon.$$

Notice that $\sup_{\alpha\in\mathcal{A}^\epsilon}\mathbb{E}^W(F(S_\alpha))$ is exactly the G-expectation of Peng. For many classes of payoffs, this methodology can be used to prove the existence of a liquidity premium. Indeed, for convex types of payoffs such as put options, call options, and Asian (put or call) options, this can be verified directly by observing that the maximum in the above expression is achieved for $\alpha\equiv\sqrt{\sigma(\sigma+2\epsilon)}$.

We close this section by revisiting Example 3.3.

Example 3.8 Let g_γ be the power penalty function given in Example 3.3. In the case of $\gamma=2$, $\widehat{G}$ is also a quadratic function. Hence, the limit stochastic optimal control problem is exactly the one derived and studied in [6, 13]. The case $\gamma>2$ is not covered by our hypothesis, but formally the limit value function is equal to the Black–Scholes price as $\widehat{G}$ is finite and zero only when $\alpha\equiv\sigma$. This result can be proved from our results by appropriate approximation arguments. The case $\gamma\in[1,2)$ is included in our hypothesis, and the limit of the truncated problem is the G-expectation. Namely, only volatility processes α that are in a certain interval are admissible.

Since in these markets the investors make only small transactions, a larger γ means less trading cost. Hence, when γ is sufficiently large (i.e., $\gamma>2$), the trading penalty is completely avoided in the limit. Hence, for these values of γ, the limiting super-replication cost is simply the usual replication price in a complete market.

4 Duality

In this section, we prove the duality result in Theorem 3.1. Fix $n\in\mathbb{N}$ and consider the n-step binomial model with penalty function g. We first motivate the result and prove one of the inequalities. Then the proof is completed by casting the super-replication problem as a convex program and using the standard duality. Indeed, for any $k=0,\dots,n-1$,

$$Y^\pi(k+1)=Y^\pi(k)+\gamma(k+1)\big(S^{(n)}(k+1)-S^{(n)}(k)\big)-g\left(\frac{k}{n},\gamma(k+1)-\gamma(k)\right).$$

Since $\gamma(0)=0$ and $Y^\pi(0)=x$, we sum over k to arrive at

$$\begin{aligned}Y^\pi(n)-x&=\sum_{k=0}^{n-1}\left(\gamma(k+1)\big(S^{(n)}(k+1)-S^{(n)}(k)\big)-g\left(\frac{k}{n},\gamma(k+1)-\gamma(k)\right)\right)\\&=\sum_{k=0}^{n-1}\Bigg(\big(\gamma(k+1)-\gamma(k)\big)\big(S^{(n)}(n)-S^{(n)}(k)\big)\\&\qquad-g\left(\frac{k}{n},\gamma(k+1)-\gamma(k)\right)\Bigg).\end{aligned}$$

Let $\mathbb{P}$ be a probability measure in $\mathcal{Q}_n$. We take the conditional expectations and use the definition of the dual function G to obtain

$$\mathbb{E}^{\mathbb{P}}\big(Y^{\pi}(n)\big) = x + \mathbb{E}^{\mathbb{P}}\left(\sum_{k=0}^{n-1}\big(\gamma(k+1)-\gamma(k)\big)\,\big(\mathbb{E}^{\mathbb{P}}\big(S^{(n)}(n)\,\big|\,\mathcal{F}_k\big) - S^{(n)}(k)\big)\right.$$
$$\left. - g\left(\frac{k}{n}, \gamma(k+1)-\gamma(k)\right)\right)$$
$$\le x + \mathbb{E}^{\mathbb{P}}\left(\sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathbb{E}^{\mathbb{P}}\big(S^{(n)}(n)\,\big|\,\mathcal{F}_k\big) - S^{(n)}(k)\right)\right).$$

If π is a super-replicating strategy with initial wealth x, then $Y^{\pi}(n) \ge F_n$ and

$$x \ge \mathbb{E}^{\mathbb{P}}\left(F_n - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathbb{E}^{\mathbb{P}}\big(S^{(n)}(n)\,\big|\,\mathcal{F}_k\big) - S^{(n)}(k)\right)\right).$$

Since $\mathbb{P} \in \mathcal{Q}_n$ is arbitrary, the above calculation proves that

$$V_n \ge \sup_{\mathbb{P}\in\mathcal{Q}_n} \mathbb{E}^{\mathbb{P}}\left(F_n - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathbb{E}^{\mathbb{P}}\big(S^{(n)}(n)\,\big|\,\mathcal{F}_k\big) - S^{(n)}(k)\right)\right).$$

The opposite inequality is proved using the standard duality. Indeed, the proof that follows does not use the above calculations.

Proof of Theorem 3.1 We model the n-step binomial model as in [7]. Consider a tree whose paths are sequences of the form $(a_1,\dots,a_k) \in \{-1,1\}^k,\ 0 \le k \le n$. The set of all paths is denoted by $\mathbb{V}$. The empty path (corresponding to $k=0$) is the root of the tree and denoted by $\emptyset$. In our model, each path of the form $u=(u_1,\dots,u_k)\in\{-1,1\}^k$, $k<n$, has two immediate successor paths $(u_1,\dots,u_k,1)$ and $(u_1,\dots,u_k,-1)$. Let $\mathbb{T} := \{-1,1\}^n$ be the set of all paths with length n. For $u \in \mathbb{V}\setminus\mathbb{T}$, denote by u^+ the set of all paths which consist of the immediate successors of u. The unique immediate predecessor of a path $u=(u_1,\dots,u_k)\in\mathbb{V}\setminus\{\emptyset\}$ is denoted by $u^- := (u_1,\dots,u_{k-1})$. For $u=(u_1,\dots,u_k)\in\mathbb{V}\setminus\mathbb{T}$, let

$$\mathbb{T}(u) := \{v \in \mathbb{T} \mid v_i = u_i,\ 1 \le i \le k\}$$

with $\mathbb{T}(\{\emptyset\}) = \mathbb{T}$. For $u \in \mathbb{V}$, $\ell(u)$ is the number of elements in the sequence u, where we set $\ell(\emptyset)=0$. Finally, we define the functions $S:\mathbb{V}\to\mathbb{R}$, $\hat{S}:\mathbb{V}\to\mathcal{C}^+[0,1]$, and $\hat{F}:\mathbb{T}\to\mathbb{R}_+$ by

$$S(u) = s_0 \exp\left(\frac{\sigma}{\sqrt{n}}\sum_{i=1}^{\ell(u)} u_i\right),$$
$$\hat{S}(u) = \mathcal{W}_n\big(\big\{S(u_1,\dots,u_{k\wedge\ell(u)})\big\}_{k=0}^{n}\big),$$
$$\hat{F}(v) = F\big(\hat{S}(v)\big).$$

In this notation, the super-replication cost V_n is the solution of the convex minimization problem to

$$\text{minimize } Y(\emptyset) \tag{4.1}$$

over all β, γ, Y subject to the constraints

$$\gamma(\emptyset)=0, \tag{4.2}$$

$$\gamma(u)-\gamma(u^-)-\beta(u^-)=0 \quad \forall u\in\mathbb{V}\setminus\{\emptyset\}, \tag{4.3}$$

$$\begin{aligned} Y(u)+g\left(\frac{\ell(u^-)}{n},\hat{S}(u^-),\beta(u^-)\right)& \\ -\gamma(u)\big(S(u)-S(u^-)\big)-Y(u^-)&\le 0 \quad \forall u\in\mathbb{V}\setminus\{\emptyset\}, \end{aligned} \tag{4.4}$$

$$Y(u)\ge \hat{F}(u) \quad \forall u\in\mathbb{T}. \tag{4.5}$$

Notice that (2.4) implies that constraint (4.4) should in fact be an equality. However, this modification of the constraint does not alter the value of the optimization problem. The optimization problem given by (4.1)–(4.5) is an ordinary convex program on the space $\mathbb{R}^{\mathbb{V}\setminus\mathbb{T}}\times\mathbb{R}^{\mathbb{V}}\times\mathbb{R}^{\mathbb{V}}$. Following the Kuhn–Tucker theory (see [20]), we define the Lagrangian $L:\mathbb{R}^{\mathbb{V}}\times\mathbb{R}_+^{\mathbb{V}\setminus\{\emptyset\}}\times\mathbb{R}_+^{\mathbb{T}}\times\mathbb{R}^{\mathbb{V}\setminus\mathbb{T}}\times\mathbb{R}^{\mathbb{V}}\times\mathbb{R}^{\mathbb{V}}\to\mathbb{R}$ by

$$\begin{aligned} L(\Upsilon,\Phi,\Theta,\beta,\gamma,Y)&=Y(\emptyset)+\Upsilon(\emptyset)\gamma(\emptyset)+\sum_{u\in\mathbb{V}\setminus\{\emptyset\}}\Upsilon(u)\big(\gamma(u)-\gamma(u^-)-\beta(u^-)\big)\\ &\quad+\sum_{u\in\mathbb{V}\setminus\{\emptyset\}}\Phi(u)\bigg(Y(u)+g\left(\frac{\ell(u^-)}{n},\hat{S}(u^-),\beta(u^-)\right)\\ &\quad-\gamma(u)\big(S(u)-S(u^-)\big)-Y(u^-)\bigg)\\ &\quad+\sum_{u\in\mathbb{T}}\Theta(u)\big(\hat{F}(u)-Y(u)\big). \end{aligned}$$

We rearrange the above expressions to arrive at

$$\begin{aligned} &L(\Upsilon,\Phi,\Theta,\beta,\gamma,Y)\\ &\quad=Y(\emptyset)\bigg(1-\sum_{u\in\emptyset^+}\Phi(u)\bigg)+\sum_{u\in\mathbb{V}\setminus(\{\emptyset\}\cup\mathbb{T})}Y(u)\bigg(\Phi(u)-\sum_{\tilde{u}\in u^+}\Phi(\tilde{u})\bigg)\\ &\qquad+\sum_{u\in\mathbb{T}}Y(u)\big(\Phi(u)-\Theta(u)\big)+\gamma(\emptyset)\bigg(\Upsilon(\emptyset)-\sum_{u\in\emptyset^+}\Upsilon(u)\bigg)\\ &\qquad+\sum_{u\in\mathbb{V}\setminus\{\emptyset\}}\gamma(u)\bigg(\Upsilon(u)-\sum_{\tilde{u}\in u^+}\Upsilon(\tilde{u})\Phi(u)\big(S(u)-S(u^-)\big)\bigg)+\sum_{u\in\mathbb{T}}\Theta(u)\hat{F}(u)\\ &\qquad+\sum_{u\in\mathbb{V}\setminus\mathbb{T}}\bigg(\sum_{\tilde{u}\in u^+}\Phi(\tilde{u})g\left(\frac{\ell(u)}{n},\hat{S}(u),\beta(u)\right)-\beta(u)\sum_{\tilde{u}\in u^+}\Upsilon(\tilde{u})\bigg). \end{aligned} \tag{4.6}$$

By Theorem 28.2 in [20], we conclude that the value of the optimization problem (4.1)–(4.5) is also equal to

$$V_n=\sup_{(\Upsilon,\Phi,\Theta)\in\mathbb{R}^{\mathbb{V}}\times\mathbb{R}_+^{\mathbb{V}\setminus\{\emptyset\}}\times\mathbb{R}_+^{\mathbb{T}}}\ \inf_{(\beta,\gamma,Y)\in\mathbb{R}^{\mathbb{V}\setminus\mathbb{T}}\times\mathbb{R}^{\mathbb{V}}\times\mathbb{R}^{\mathbb{V}}} L(\Upsilon,\Phi,\Theta,\beta,\gamma,Y). \tag{4.7}$$

Using (4.6) and (4.7), we conclude that

$$V_n = \sup_{(\Upsilon,\Phi,\Theta)\in D} \inf_{\beta\in\mathbb{R}^{\mathbb{V}\setminus\mathbb{T}}} \Bigg(\sum_{u\in\mathbb{T}} \Theta(u)\hat{F}(u) \\ + \sum_{u\in\mathbb{V}\setminus\mathbb{T}} \bigg(\sum_{\tilde{u}\in\, u^+} \Phi(\tilde{u}) g\Big(\frac{\ell(u)}{n}, \hat{S}(u), \beta(u)\Big) - \beta(u) \sum_{\tilde{u}\in\, u^+} \Upsilon(\tilde{u}) \bigg)\Bigg), \quad (4.8)$$

where $D \subset \mathbb{R}^{\mathbb{V}} \times \mathbb{R}^{\mathbb{V}\setminus\{\emptyset\}} \times \mathbb{R}_+^{\mathbb{T}}$ is the subset of all (Υ, Φ, Θ) satisfying the constraints

$$\sum_{u\in\emptyset^+} \Phi(u) = 1,$$

$$\sum_{\tilde{u}\in\, u^+} \Phi(\tilde{u}) = \Phi(u) \quad \forall u \in \mathbb{V} \setminus \big(\mathbb{T} \cup \{\emptyset\}\big), \quad (4.9)$$

$$\Upsilon(u) = \Phi(u)\big(S(u) - S\big(u^-\big)\big) + \sum_{\tilde{u}\in\, u^+} \Upsilon(\tilde{u}) \quad \forall u \in \mathbb{V} \setminus \{\emptyset\}, \quad (4.10)$$

$$\Phi(u) = \Theta(u) \quad \forall u \in \mathbb{T}. \quad (4.11)$$

By (4.9) and (4.10), we obtain that for any $(\Upsilon, \Phi, \Psi) \in D$,

$$\frac{\sum_{\tilde{u}\in\, u^+} \Upsilon(\tilde{u})}{\sum_{\tilde{u}\in\, u^+} \Phi(\tilde{u})} = \frac{\sum_{\tilde{u}\in\mathbb{T}(u)} \Phi(\tilde{u}) S(\tilde{u})}{\Phi(u)} - S(u) \quad \forall u \in \mathbb{V} \setminus \mathbb{T}, \quad (4.12)$$

with the convention $0/0 = 0$ (observe that $\sum_{\tilde{u}\in\mathbb{T}(u)} \Phi(\tilde{u}) S(\tilde{u}) = 0$ if $\Phi(u) = 0$). Let $\mathbb{D} \subset \mathbb{R}_+^{\mathbb{V}\setminus\{\emptyset\}}$ be the set of all functions $\Phi : \mathbb{V} \setminus \{\emptyset\} \to \mathbb{R}_+$ which satisfy (4.9). In view of (2.6), (4.8), (4.9) and (4.11), (4.12),

$$V_n = \sup_{\Phi\in\mathbb{D}} \sum_{u\in\mathbb{T}} \Phi(u) \bigg(\hat{F}(u) - G\bigg(\frac{\ell(u)}{n}, \hat{S}(u), \frac{\sum_{\tilde{u}\in\mathbb{T}(u)} \Phi(\tilde{u}) S(\tilde{u})}{\Phi(u)} - S(u) \bigg)\bigg). \quad (4.13)$$

Clearly, there is a natural bijection $\pi : \mathbb{D} \to \mathcal{Q}_n$ such that for any $\Phi \in \mathbb{D}$, the probability measure $P := \pi(\Phi)$ is given by

$$\mathbb{P}(\xi_1 = u_1, \xi_2 = u_2, \ldots, \xi_n = u_n) = \Phi(u) \quad \forall u = (u_1, \ldots, u_n) \in \mathbb{T}. \quad (4.14)$$

Recall that $\mathcal{Q}_n$ is the set of all probability measures on $(\Omega, \mathcal{F}_n)$. Finally, we combine (4.13) and (4.14) to conclude that

$$V_n = \sup_{\mathbb{P}\in\mathcal{Q}_n} \mathbb{E}^{\mathbb{P}} \Bigg(F_n - \sum_{k=0}^{n-1} G\bigg(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}(k)\big), \mathbb{E}^{\mathbb{P}}\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k) \bigg)\Bigg). \quad \square$$

5 Proof of Theorem 3.5

In this section, we prove Theorem 3.5. However, the proofs of several technical results needed in this proof are relegated to Sect. 7. Also, Kusuoka's construction of discrete martingales is outlined in the next section.

We start with some definitions. Let B be the canonical map on the space $\mathcal{C}[0,1]$, i.e., for each $t \in [0,1]$, $B(t) : \mathcal{C}[0,1] \to \mathbb{R}$ is given by $B(t)(x) = x(t)$. Next, let M be a strictly positive, continuous martingale defined on some probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ and satisfying

$$M(0) = s_0 \quad \text{and} \quad \frac{d\langle \ln M\rangle(t)}{dt} \le c\, \mathcal{L} \otimes \tilde{P}\text{-a.s.} \tag{5.1}$$

for some constant c. For a martingale M satisfying (5.1), we define several related quantities. Let $\widehat{G}$ be as in Assumption 3.2, and σ the constant volatility in the definition of the discrete market, cf. (2.2). Set

$$A(t:M) := \frac{\langle \ln M\rangle(t) - \sigma^2 t}{2\sigma}, \qquad a(t:M) := \frac{d}{dt} A(t:M), \tag{5.2}$$

$$J(M) = \tilde{E}\left(F(M) - \int_0^1 \widehat{G}\big(t, M, a(t:M)M(t)\big)\, dt \right), \tag{5.3}$$

where $\tilde{E}$ is the expectation with respect to $\tilde{P}$. Notice that the notation a is consistent with the already introduced function $a(t:S_\alpha)$ in (3.4) and that $J(M)$ agrees with the function defined in (3.5). Also, from (2.1), (3.1), and (5.1) it follows that the right-hand side of (5.3) is well defined.

Upper bound We start with the proof of the upper bound of Theorem 3.5,

$$\limsup_{n\to\infty} V_n^c \le \sup_{\alpha \in \mathcal{A}^c} J(S_\alpha). \tag{5.4}$$

In what follows, to simplify the notation, we assume that indices have been renamed so that the whole sequence converges. Let $n \in \mathbb{N}$.

By Theorem 3.1, we construct probability measures P_n on $(\Omega, \mathcal{F}_n)$ such that

$$\begin{aligned} V_n^c \le \frac{1}{n} + E_n\Bigg(& F\big(\mathcal{W}_n\big(S^{(n)}\big)\big) \\ & - \sum_{k=0}^{n-1} G_n^c\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), E_n\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k)\right)\Bigg), \end{aligned}$$

where E_n denotes the expectation with respect to P_n. Since V_n^c is nonnegative, the right-hand side of the above inequality is not minus infinity. This, together with (2.7), yields that for any $0 \le k < n$,

$$\big|E_n\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k)\big| \le \frac{c}{\sqrt{n}}\, S^{(n)}(k) \quad P_n\text{-a.s.} \tag{5.5}$$

and

$$V_n^c \le E_n\left(F\big(\mathcal{W}_n\big(S^{(n)}\big)\big) - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), E_n\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k)\right)\right) + \frac{1}{n}.$$

For $0 \le k \le n$, set

$$M^{(n)}(k) := E_n\big(S^{(n)}(n) \,\big|\, \mathcal{F}_k\big),$$
$$\alpha_n(k) := \frac{\sqrt{n}\xi_k(M^{(n)}(k) - S^{(n)}(k))}{S^{(n)}(k)},$$
$$A_n(t) := \int_0^t \alpha_n\big([nu]\big)\,du = \frac{1}{n}\sum_{k=0}^{[nt]-1} \alpha_n(k) + \frac{nt-[nt]}{n}\,\alpha_n\big([nt]\big).$$

Let Q_n be the joint distribution of the stochastic processes $(\mathcal{W}_n(S^{(n)}), A_n)$ under the measure P_n. In view of (5.5), the hypothesis of Lemma 7.1 is satisfied. Hence, there exist a subsequence (denoted by n again) and a probability measure P on the probability space $\mathcal{C}[0,1]$ such that

$$Q_n \Longrightarrow Q \quad \text{on the space } \mathcal{C}[0,1] \times \mathcal{C}[0,1],$$

where Q is the joint distribution under P of the canonical process B and the process $A(\cdot : B)$ defined in (5.2). From the Skorohod representation theorem (Theorem 3 of [11]), it follows that there exists a probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ on which

$$\big(\mathcal{W}_n\big(S^{(n)}\big), A_n(\cdot)\big) \longrightarrow \big(M, A(\cdot : M)\big) \quad \tilde{P}\text{-a.s.} \tag{5.6}$$

on the space $\mathcal{C}[0,1] \times \mathcal{C}[0,1]$, where M is a positive martingale. More precisely, the $\mathcal{C}[0,1] \times \mathcal{C}[0,1]$-valued random variables $(\mathcal{W}_n(S^{(n)}), A_n(\cdot))$ (which are defined on the probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$) converge a.s. to (the $\mathcal{C}[0,1] \times \mathcal{C}[0,1]$-valued random variable) $(M, A(\cdot : M))$. As usual, the topology on $\mathcal{C}[0,1]$ is the sup topology. Furthermore, (5.5) implies that Lemma 6.5 applies to this sequence. Hence, we have the pointwise estimate

$$\big|a(t : M)\big| = \big|A'(t : M)\big| \le c.$$

Observe that we can redefine the processes α_n and $M^{(n)}$, $n \in \mathbb{N}$, on the probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ by setting

$$M^{(n)}(k) = \tilde{E}\big(S^{(n)}(n) \,\big|\, \sigma\big\{S^{(n)}(1), \ldots, S^{(n)}(k)\big\}\big),$$
$$\alpha_n(k) = n\big(A_n\big(((k+1)/n) \wedge 1\big) - A_n(k/n)\big), \quad k = 0, 1, \ldots, n.$$

Since for any n, the joint distribution of the processes $S^{(n)}$, $M^{(n)}$, α_n remains the same as before, we keep the same notations. In particular, we get for any $k \le n$ that

$$\frac{\sqrt{n}(M^{(n)}(k) - S^{(n)}(k))}{\alpha_n(k) S^{(n)}(k)} \in \{-1, 1\} \quad \tilde{P}\text{-a.s.} \tag{5.7}$$

Next, we replace the sequence $\{\alpha_n\}$ (which converges only weakly) by a pointwise convergent sequence. Indeed, by Lemma A1.1 in [10], we construct a sequence

$$\eta_n \in \operatorname{conv}(\tilde{\alpha}_n, \tilde{\alpha}_{n+1}, \ldots), \quad \text{where } \tilde{\alpha}_n(t) := \alpha_n\big([nt]\big),$$

such that $\{\eta_n\}$ converges almost surely with respect to $\mathcal{L} \otimes \tilde{P}$ to a stochastic process η. We now use (5.5) together with the dominated convergence theorem. The result is

$$\int_0^t \eta(u)\,du = \lim_{n\to\infty} \int_0^t \eta_n(u)\,du = \lim_{n\to\infty} \int_0^t \alpha_n\big([nu]\big)\,du$$
$$= A(t : M) = \int_0^t a(u : M)\,du \quad \mathcal{L} \otimes \tilde{P}\text{-a.s.}$$

Hence, we conclude that

$$\eta(t) = a(t : M) \quad \mathcal{L} \otimes \tilde{P}\text{-a.s.}$$

From (3.2) and (5.6)–(5.7) we conclude that

$$\lim_{n\to\infty} \left| nG\left(\frac{[nt]}{n}, \mathcal{W}_n\big(S^{(n)}\big), M^{(n)}\big([nt]\big) - S^{(n)}\big([nt]\big)\right) - \widehat{G}\big(t, M, \alpha_n\big([nt]\big)M(t)\big) \right|$$
$$= 0 \quad \mathcal{L} \otimes \tilde{P}\text{-a.s.}$$

Estimate (6.5) and the growth assumption (3.1) imply that the above sequences are uniformly integrable. Therefore,

$$\begin{aligned}
I &:= \lim_{n\to\infty} E_n\left(\sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), E_n\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k)\right)\right) \\
&= \lim_{n\to\infty} \tilde{E}\left(\int_0^1 nG\left(\frac{[nt]}{n}, \mathcal{W}_n\big(S^{(n)}\big), M^{(n)}\big([nt]\big) - S^{(n)}\big([nt]\big)\right) dt\right) \\
&= \lim_{n\to\infty} \tilde{E}\left(\int_0^1 \widehat{G}\big(t, M, \alpha_n\big([nt]\big)M(t)\big)\, dt\right),
\end{aligned}$$

where again, without loss of generality (by passing to a subsequence), we assume that the above limits exist. We now use the convexity of $\widehat{G}$ with respect to the third variable (in fact, $\widehat{G}$ is quadratic in y) together with the uniform integrability (which again follows from (3.1) and Lemma 6.5) and the Fubini theorem. The result is

$$\begin{aligned}
I &= \lim_{n\to\infty} \tilde{E}\left(\int_0^1 \widehat{G}\big(t, M, \alpha_n\big([nt]\big)M(t)\big)\, dt\right) \\
&\geq \lim_{n\to\infty} \tilde{E}\left(\int_0^1 \widehat{G}\big(t, M, \eta_n(t)M(t)\big)\, dt\right) \\
&= \tilde{E}\left(\int_0^1 \widehat{G}\big(t, M, \eta(t)M(t)\big)\right) dt = \tilde{E}\left(\int_0^1 \widehat{G}\big(t, M, a(t : M)M(t)\big)\, dt\right).
\end{aligned}$$

The growth assumption on F, namely (2.1) and Lemma 6.5, also imply that the sequence $F(\mathcal{W}_n(S^{(n)}))$ is uniformly integrable. Then, by (5.6),

$$\lim_{n\to\infty} E_n\big(F\big(\mathcal{W}_n\big(S^{(n)}\big)\big)\big) = \tilde{E}\big(F(M)\big).$$

Hence, we have shown that

$$\begin{aligned}
\limsup_{n\to\infty} V_n^c &\leq \limsup_{n\to\infty} E_n\Bigg(F\big(\mathcal{W}_n\big(S^{(n)}\big)\big) \\
&\qquad - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), E_n\big(S^{(n)}(n) \mid \mathcal{F}_k\big) - S^{(n)}(k)\right)\Bigg) \\
&\leq \tilde{E}\left(F(M) - \int_0^1 \widehat{G}\big(t, M, a(t : M)M(t)\big)\, dt\right) = J(M).
\end{aligned}$$

The above, together with Lemma 7.2, yields (5.4).

Lower bound Let $\mathcal{L}(c)$ be the class of all adapted volatility processes given in Definition 6.1 below. In Lemma 7.3 below, it is shown that this class is dense in $\mathcal{A}^c$ (with respect to convergence in probability). Hence, for the lower bound, it is sufficient to prove that for any $\alpha \in \mathcal{L}(c)$,

$$\lim_{n\to\infty} V_n^c \geq J(S_\alpha). \tag{5.8}$$

Our main tool is the Kusuoka construction which is summarized in Theorem 6.2.

We fix $\alpha \in \mathcal{L}(c)$. Let $P_n^{(\alpha)}$, $\kappa_n^{(\alpha)}$, and $M_n^{(\alpha)}$ be as in Theorem 6.2 below. In view of the definition of $M_n^{(\alpha)}$, (6.2), and the bounds on $\kappa_n^{(\alpha)}$, for all sufficiently large n, we have the estimate

$$\left|M_n^{(\alpha)}(k) - S^{(n)}(k)\right| \leq \frac{c}{\sqrt{n}} S^{(n)}(k) \quad \forall k \ P_n^{(\alpha)}\text{-a.s.}$$

By the dual representation and the above estimate,

$$\begin{aligned}\lim_{n\to\infty} V_n^c \geq \limsup_{n\to\infty} E_n^{(\alpha)}\Bigg(& F\big(\mathcal{W}_n\big(S^{(n)}\big)\big) \\ & - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), M_n^{(\alpha)}(k) - S^{(n)}(k)\right)\Bigg),\end{aligned} \tag{5.9}$$

where $E_n^{(\alpha)}$ denotes the expectation with respect to $P_n^{(\alpha)}$. From Theorem 6.2 and the Skorohod representation theorem, it follows that there exists a probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ on which

$$\big(\mathcal{W}_n\big(S^{(n)}\big), \mathcal{W}_n\big(\kappa_n^{(\alpha)}\big)\big) \longrightarrow \big(S_\alpha, a(\cdot : S_\alpha)\big) \quad \tilde{P}\text{-a.s.} \tag{5.10}$$

on the space $\mathcal{C}[0,1] \times \mathcal{C}[0,1]$. Recall that the quadratic variation density a is defined in (3.4) and also in (5.2). We argue exactly as in the upper bound to show that

$$\lim_{n\to\infty} E_n^{(\alpha)} F\big(\mathcal{W}_n\big(S^{(n)}\big)\big) = \tilde{E}\big(F(S_\alpha)\big). \tag{5.11}$$

Similarly to the upper bound case, we redefine the stochastic processes $M_n^{(\alpha)}$, $n \in \mathbb{N}$, on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. Namely for any $k \leq n$, define

$$M_n^{(\alpha)}(k) = S^{(n)}(k) \exp\big(\xi_k^{(n)}(\omega)\kappa_n^{(\alpha)}(k) n^{-1/2}\big), \tag{5.12}$$

where

$$\xi_k^{(n)} := \frac{\sqrt{n}}{\sigma}\big(\ln S^{(n)}(k) - \ln S^{(n)}(k-1)\big).$$

Again, the joint distribution of $M^{(n)}$, $S^{(n)}$, α_n remains as before. Finally, we need to connect the difference $M_n^{(\alpha)} - S^{(n)}$ to $\kappa_n^{(\alpha)}$ and therefore to $a(\cdot : S_\alpha)$ through (5.10). Indeed, in view of definition (5.12),

$$\begin{aligned}\sqrt{n}\xi_k^{(n)}\big(M_n^{(\alpha)}(k) - S^{(n)}(k)\big) &= \sqrt{n}\xi_k^{(n)} S^{(n)}(k)\big(\exp\big(\xi_k^{(n)}\kappa_n^{(\alpha)}(k) n^{-1/2}\big) - 1\big) \\ &= S^{(n)}(k)\kappa_n^{(\alpha)}(k) + O\big(n^{-1/2}\big).\end{aligned}$$

In the approximation above, we used the fact that the $\kappa^{(\alpha)}$ are uniformly bounded by construction and that $\xi_k^{(n)} = \frac{1}{\xi_k^{(n)}}$. We now use (5.10) to arrive at

$$\lim_{n\to\infty} \sqrt{n}\xi_{[nt]}^{(n)}\big(M_n^{(\alpha)}([nt]) - S^{(n)}([nt])\big) = a(t : S_\alpha)S_\alpha(t) \quad \mathcal{L}\otimes \tilde{P}\text{-a.s.} \tag{5.13}$$

As in the upper bound case, the growth condition (3.1) and Lemma 6.5 imply that the sequence $nG([nt]/n, \mathcal{W}_n(S^{(n)}), M_n^{(\alpha)}([nt]) - S^{(n)}([nt]))$ is uniformly integrable in $\mathcal{L}\otimes \tilde{P}$. Since $\widehat{G}$ is continuous, by Fubini's theorem and (3.2), (5.10), (5.13), we obtain

$$\begin{aligned}
\tilde{I} &:= \lim_{n\to\infty} E_n\left(\sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathcal{W}_n\big(S^{(n)}\big), M_n^{(\alpha)}(k) - S^{(n)}(k)\right)\right)\\
&= \lim_{n\to\infty} \tilde{E}\left(\int_{[0,1]} nG\left(\frac{[nt]}{n}, \mathcal{W}_n\big(S^{(n)}\big), M_n^{(\alpha)}\big([nt]\big) - S^{(n)}\big([nt]\big)\right) dt\right)\\
&= \tilde{E}\left(\int_{[0,1]} \widehat{G}\big(t, S_\alpha, a(t : S_\alpha)S_\alpha(t)\big)\, dt\right).
\end{aligned}$$

We combine the above equality together with (5.9) and (5.11). The resulting inequality is exactly (5.8). Hence, the proof of the lower bound is also complete. □

6 Kusuoka's construction

In this section, we fix a martingale S_α given by (3.3). Then the main goal of this section is to construct a sequence of martingales on the discrete space that approximate S_α. We also require the quadratic variation of S_α to be approximated as well.

In [16] Kusuoka provides an elegant approximation for a sufficiently smooth volatility process α. Here we only state the results of Kusuoka and refer to [16] for the construction. We start by defining the class of "smooth" volatility processes. As before, let $(\Omega_W, \mathcal{F}^W, \mathbb{P}^W)$ be a Brownian probability space, and W a standard Brownian motion.

Definition 6.1 For a fixed constant $c > 0$, let $\mathcal{L}(c) \subseteq \mathcal{A}^c$ be the set of all adapted processes α on the Brownian space $(\Omega_W, \mathcal{F}^W, \mathbb{P}^W)$ given by

$$\alpha(t) := \alpha(t, \omega) = f\big(t, W(\omega)\big), \quad (t, \omega) \in [0, 1] \times \Omega_W,$$

where $f : [0, 1] \times \mathcal{C}[0, 1] \to \mathbb{R}_+$ is a bounded function which satisfies the following conditions:

(i) For any $t \in [0, 1]$, if $w, \tilde{w} \in \mathcal{C}[0, 1]$ satisfy $w(u) = \tilde{w}(u)$ for all $u \in [0, t]$, then $f(t, w) = f(t, \tilde{w})$. (This simply means that α is adapted.)

(ii) There is $\delta(f) > 0$ such that for all $(t, w) \in [0, 1] \times \mathcal{C}[0, 1]$,

$$f^2(t, w) \in \big[0 \vee \big(\sigma(\sigma - 2c)\big) + \delta(f), \sigma(\sigma + 2c) - \delta(f)\big]$$

and

$$f(t, w) = \sigma, \quad \text{if } t > 1 - \delta(f). \tag{6.1}$$

(iii) There is $L(f)>0$ such that for all $(t_1,t_2)\in[0,1]$, $w,\tilde{w}\in\mathcal{C}[0,1]$,

$$\big|f(t_1,w)-f(t_2,\tilde{w})\big|\le L(f)\big(|t_1-t_2|+\|w-\tilde{w}\|_\infty\big).$$

In Kusuoka's construction, condition (6.1) is not needed. However, this regularity will allow us control the behavior of the discrete-time martingales at maturity.

Recall from Sect. 2 that $\Omega=\{-1,1\}^{\mathbb{N}}$, ξ is the canonical map (i.e., $\xi_k(\omega)=\omega_k$), and $\mathbb{Q}$ is the symmetric product measure. The martingales constructed in [16] are of the form

$$M_n^{(\alpha)}(k,\omega):=S^{(n)}(k,\omega)\exp\big(\xi_k(\omega)\kappa_n^{(\alpha)}(k,\omega)n^{-1/2}\big),\quad 0\le k\le n, \tag{6.2}$$

where the sequence of discrete *predictable* processes $\kappa_n^{(\alpha)}$ needs to be constructed. Now let $P_n^{(\alpha)}$ be a measure on Ω such that the process $M_n^{(\alpha)}$ is a $P_n^{(\alpha)}$-martingale. Since κ_n^{α} will be constructed as predictable processes, a direct calculation shows that on the σ-algebra $\mathcal{F}_n$, this martingale measure is given by

$$\frac{dP_n^{(\alpha)}}{d\mathbb{Q}}\bigg|_{\mathcal{F}_n}(\omega)=2^n\prod_{k=1}^{n}\tilde{q}_n^{(\alpha)}(k,\omega),$$

where for $0\le k\le n$, $\omega\in\Omega$,

$$\begin{aligned}
\tilde{q}_n^{(\alpha)}(k,\omega)&=q_n^{(\alpha)}(k,\omega)\mathbb{I}_{\{\xi_k(\omega)=1\}}+\big(1-q_n^{(\alpha)}(k,\omega)\big)\mathbb{I}_{\{\xi_k(\omega)=-1\}},\\
q_n^{(\alpha)}(k,\omega)&=\frac{\exp(\xi_{k-1}\kappa_n^{(\alpha)}(k-1,\omega)n^{-1/2})-(\exp(\sigma n^{-1/2})e_n^{(\alpha)}(k,\omega))^{-1}}{\exp(\sigma n^{-1/2})e_n^{(\alpha)}(k,\omega)-(\exp(\sigma n^{-1/2})e_n^{(\alpha)}(k,\omega))^{-1}},\\
e_n^{(\alpha)}(k,\omega)&=\exp\big(\kappa_n^{(\alpha)}(k,\omega)n^{-1/2}\big).
\end{aligned}$$

We require that $\kappa_n^{(\alpha)}$ is constructed to satisfy

$$\begin{aligned}
&\big|\kappa_n^{(\alpha)}(k,\omega)\big|<c-\delta,\qquad \kappa_n^{(\alpha)}(k,\omega)>\delta-\frac{\sigma}{2},\\
&\big|\kappa_n^{(\alpha)}(k-1,\omega)-\kappa_n^{(\alpha)}(k,\omega)\big|\le\frac{L}{\sqrt{n}},\qquad 1\le k\le n,
\end{aligned} \tag{6.3}$$

with constants $L,\delta>0$ independent of n and ω. These regularity conditions on $\kappa_n^{(\alpha)}$ imply that for all sufficiently large n, $q_n^{(\alpha)}(k,\omega)\in(0,1)$ for all $k\le n$ and $\omega\in\Omega=\{-1,1\}^{\mathbb{N}}$. Hence, $P_n^{(\alpha)}$ is indeed a probability measure.

We also require

$$\kappa_n^{(\alpha)}(n,\omega)=0\quad\text{for sufficiently large } n, \tag{6.4}$$

to ensure $M_n^{(\alpha)}(n)=S_n^{(n)}(n)$. This is exactly the reason why we added condition (6.1) in Definition 6.1.

Let $Q_n^{(\alpha)}$ be the joint distribution of the pair $(\mathcal{W}_n(S^{(n)}),\mathcal{W}_n(\kappa_n^{(\alpha)}))$ under $P_n^{(\alpha)}$ on the space $\mathcal{C}[0,1]\times\mathcal{C}[0,1]$ with uniform topology. Recall once again that the probability space is $\Omega=\{-1,1\}^{\mathbb{N}}$, the filtration $\{\mathcal{F}_k\}_{k=0}^n$ is the usual one generated by the canonical map, and the quadratic variation density process $a(\cdot:S_\alpha)$ is given in (3.4) as

$$a(t:S_\alpha)=\frac{\alpha^2(t)-\sigma^2}{2\sigma}.$$

Theorem 6.2 (Kusuoka [16]) *Let $c > 0$ and $\alpha \in \mathcal{L}(c)$. Then there exists a sequence of predictable processes $\kappa_n^{(\alpha)}$ on $(\Omega, \{\mathcal{F}_k\}_{k=0}^n)$ satisfying (6.3) and (6.4). Hence, there also exist sequences of martingales $M_n^{(\alpha)}$ and martingale measures $P_n^{(\alpha)}$ such that*

$$Q_n^{(\alpha)} \Longrightarrow Q^{(\alpha)} \quad \text{on the space } \mathcal{C}[0,1] \times \mathcal{C}[0,1],$$

where $Q^{(\alpha)}$ is the joint distribution of $(S_\alpha, a(\cdot : S_\alpha))$ under the Wiener measure $\mathbb{P}^W$.

For the construction of $\kappa_n^{(\alpha)}$, we refer the reader to Proposition 5.3 in [16].

Remark 6.3 It is clear that one constructs the process $\kappa_n^{(\alpha)}$ by an appropriate discrete approximation of $a(\cdot : S_\alpha)$. However, this discretization is not only in time, but also in the probability space. Namely, the process α is a process on the canonical probability space $\mathcal{C}[0,1]$, while $\kappa_n^{(\alpha)}$ lives on the discrete space Ω. This difficulty is resolved by Kusuoka in [16].

Remark 6.4 Although the tightness of the processes $\mathcal{W}_n(\kappa_n^{(\alpha)})$ is not stated in Proposition 5.3 of [16], it follows directly from the proof of that result. Indeed, Kusuoka shows in the proof that the joint distributions of the constructed processes $\mathcal{W}_n(M_n)$ and $\mathcal{W}_n(B_n)$ converge weakly to the distribution of a pair of the form (M, B), where B is a Brownian motion, and M is an exponential martingale,

$$M(t) = s_0 \exp\left(\int_0^t g(u, B)\, dB(u) - \frac{1}{2}\int_0^t g^2(u, B)\, du\right), \quad t \in [0,1].$$

Since $g : [0,1] \times \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ is a continuous map, we conclude that the joint distributions of $\mathcal{W}_n(M_n)$ and $\{(g^2(k/n, \mathcal{W}_n(B_n)) - \sigma^2)/(2\sigma)\}_{k=0}^n$ converge weakly to $(M, a(\cdot : M))$. This is exactly the statement of Theorem 6.2.

We complete this section by stating (without proof) a lemma which summarizes the main results from Sect. 4 in [16]; see in particular Propositions 4.8 and 4.27 in [16]. In our analysis, the lemma below provides the crucial tightness result which is used in the proof of the upper bound of Theorem 3.5. Furthermore, inequality (6.5) is essential in establishing the uniform integrability of several sequences.

Let $(\Omega, \mathcal{F}_k)$ be the probability space introduced in Sect. 2.

Lemma 6.5 (Kusuoka [16]) *Let $\{M^{(n)}\}$ be a sequence of positive martingales with respect to probability measures P_n on $(\Omega, \mathcal{F}_n)$. Suppose that there exists a constant $c > 0$ such that for any $k \le n$,*

$$\left|S^{(n)}(k) - M^{(n)}(k)\right| \le \frac{cS^{(n)}(k)}{\sqrt{n}} \quad P_n\text{-a.s.}$$

Then from Proposition 4.8 in [16] and the inequality $S_k^{(n)} \le (1 + c/\sqrt{n}) M_k^{(n)}$ it follows that for any $p > 0$,

$$\sup_n E_n\Big(\max_{0\le k\le n}\big(S^{(n)}(k)\big)^p\Big) \le (1+c)^p \sup_n E_n\Big(\max_{0\le k\le n}\big(M^{(n)}(k)\big)^p\Big) < \infty, \quad (6.5)$$

where E_n is the expectation with respect to P_n. Moreover, the distributions Q_n on $\mathcal{C}[0,1]$ of $\mathcal{W}_n(S^{(n)})$ under P_n form a tight sequence, and under any limit point Q of this sequence, the canonical process B is a strictly positive martingale in its own filtration. Furthermore, the quadratic variation density of B under Q satisfies

$$\big|a(t:B)\big| \le c.$$

Observe that the last inequality follows directly from Proposition 4.27 in [16].

7 Auxiliary lemmas

In this section, we prove several results that are used in the proof of our convergence result. Lemmas 7.2 and 7.3 are related to the optimal control (3.5). The first result, Lemma 7.1, is related to the properties of a sequence of discrete-time martingales $M^{(n)}$. Motivated by (5.5) and Lemma 6.5, we assume that these martingales are sufficiently close to the price process $S^{(n)}$. Then, in Lemma 7.1 below, we prove that the processes α_n defined below converge weakly. The structure that we outline below is very similar to the one constructed in Theorem 6.2. However, below the martingales $M^{(n)}$ are given, while in the previous section, they are constructed. The limit theorem in Lemma 7.1 is the main tool in the proof of the upper bound of Theorem 3.5.

Let $(\Omega, \mathcal{F}_n)$ be the discrete probability structure given in Sect. 2. For a probability measure P_n on $(\Omega, \mathcal{F}_n)$ and $k \le n$, set

$$M^{(n)}(k) := E_n\big(S^{(n)}(n) \,\big|\, \mathcal{F}_k\big),$$
$$\alpha_n(k) := \frac{\sqrt{n}\xi_k(M^{(n)}(k) - S^{(n)}(k))}{S^{(n)}(k)}.$$

Suppose that there exists a constant $c > 0$ such that for any $k \le n$,

$$\big|\alpha_n(k)\big| \le c \quad P_n\text{-a.s.} \quad (7.1)$$

Let Q_n be the distribution of $\mathcal{W}_n(S^{(n)})$ under the measure P_n. Then by Lemma 6.5, this sequence $\{Q_n\}$ is tight. Without loss of generality, we assume that the whole sequence $\{Q_n\}_{n=1}^{\infty}$ converges to a probability measure Q on $\mathcal{C}[0,1]$. Moreover, the canonical map B under Q is a strictly positive martingale. Then Lemma 6.5 also implies that the process $A(\cdot : B)$ given in (5.2) is well defined. The next lemma proves the convergence of the sequence $\{\alpha_n\}$ as well.

Lemma 7.1 *Assume (7.1). Let $\hat{Q}_n$ be the joint distribution under P_n of $\mathcal{W}_n(S^{(n)})$ and $\int_0^t \alpha_n([nu])\,du$. Then*

$$\hat{Q}_n \Longrightarrow \hat{Q} \quad \text{on the space } \mathcal{C}[0,1]\times\mathcal{C}[0,1],$$

where $\hat{Q}$ is the joint distribution of the canonical process B and $A(\cdot : B)$ under Q.

Proof Hypothesis (7.1) implies that Lemma 6.5 applies to the sequence $\{P_n\}$. Hence, under this sequence of measures, estimate (6.5) holds. Let Y_n be the piecewise constant process defined by

$$Y_n(t) = \sum_{j=1}^{[nt]} \frac{M^{(n)}(j) - M_n(j-1)}{S^{(n)}(j-1)}, \quad t \in [0,1], \tag{7.2}$$

with $Y_n(t) = 0$ if $t < \frac{1}{n}$. In view of (7.1), there exists a constant c_1 such that for any $k < n$,

$$\left| M^{(n)}(k+1) - M^{(n)}(k) \right| \le \frac{c_1}{\sqrt{n}} S^{(n)}(k) \quad P_n\text{-a.s.}$$

We use this together with (6.5) to arrive at

$$\lim_{n\to\infty} E_n \Big(\max_{1\le k\le n} \left| M^{(n)}(k) - M^{(n)}(k-1) \right| \Big) = 0. \tag{7.3}$$

Let $\mathcal{D}[0,1]$ be the space of all càdlàg functions equipped with the Skorohod topology. Let $\hat{P}_n$ be the distribution on the space $\mathcal{D}[0,1] \times \mathcal{D}[0,1]$ of the piecewise constant process $\{(1/S^{(n)}([nt]), M^{(n)}([nt]))\}_{0\le t\le 1}$. We use (7.1) and Lemma 6.5 to conclude that

$$\hat{P}_n \Longrightarrow \hat{P} \quad \text{on the space } \mathcal{D}[0,1] \times \mathcal{D}[0,1], \tag{7.4}$$

where the measure $\hat{P}$ is the distribution of the process $(1/B, B)$ under Q. In fact, for this convergence, we extend the definition of B so that it is still the canonical process on the space $\mathcal{D}[0,1]$, and the measure Q is extended to a probability measure on $\mathcal{D}[0,1]$.

Since the canonical process B is a strictly positive continuous martingale under Q, we apply Theorem 4.3 of [12] and use (7.3), (7.4). This gives the convergence

$$\hat{\mathbb{Q}}_n \Longrightarrow \hat{\mathbb{Q}} \quad \text{on the space } \mathcal{D}[0,1] \times \mathcal{D}[0,1] \times \mathcal{D}[0,1],$$

where $\hat{\mathbb{Q}}_n$ is the distribution of the triple $\{(1/S^{(n)}([nt]), M^{(n)}([nt]), Y_n([nt]))\}_{0\le t\le 1}$ under P_n, and $\hat{\mathbb{Q}}$ is the distribution of $\{(1/B(t), B(t), \int_0^t dB(u)/B(u))\}_{0\le t\le 1}$ under the measure Q.

In view of the Skorohod representation theorem, we may assume without loss of generality that there exist a probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ and a strictly positive continuous martingale M such that

$$\left\{ \left(\frac{1}{S^{(n)}([nt])}, M^{(n)}\big([nt]\big), Y_n\big([nt]\big) \right) \right\}_{0\le t\le 1} \longrightarrow \left\{ \left(\frac{1}{M(t)}, M(t), \int_0^t \frac{dM(u)}{M(u)} \right) \right\}_{0\le t\le 1}$$

$\tilde{P}$-a.s. on the space $\mathcal{D}[0,1] \times \mathcal{D}[0,1] \times \mathcal{D}[0,1]$. Now set $Y(t) = \int_0^t dM(u)/M(u)$, so that $dM = M dY$. Therefore,

$$M(t) = M(0) \exp\left(Y(t) - \frac{\langle Y\rangle(t)}{2} \right) \quad \text{and} \quad \langle \ln M\rangle(t) = \langle Y\rangle(t).$$

Hence, to complete the proof, it is sufficient to show that

$$\left\{\int_0^t \alpha_n\big([nu]\big)\,du\right\}_{0\le t\le 1} \longrightarrow \left\{\frac{\langle Y\rangle(t)-\sigma^2 t}{2\sigma}\right\}_{0\le t\le 1} \qquad \tilde{P}\text{-a.s. on the space } \mathcal{D}[0,1].$$

Observe that the process α_n can be redefined on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ by

$$\alpha_n(k) = \frac{\sqrt{n}\xi_k^{(n)}(M^{(n)}(k) - S^{(n)}(k))}{S^{(n)}(k)},$$

where $\xi_k^{(n)} := \frac{\sqrt{n}}{\sigma}(\ln S^{(n)}(k) - \ln S^{(n)}(k-1))$. From the definitions of α_n and $\xi_k^{(n)}$ we have (again using that $\xi^{(n)} = 1/\xi^{(n)}$)

$$\begin{aligned} M^{(n)}(k) &= S^{(n)}(k)\big(1+\xi_k^{(n)}\alpha_n(k)n^{-1/2}\big) \\ &= S^{(n)}(k-1)\exp\big(\sigma\xi_k^{(n)}n^{-1/2}\big)\big(1+\xi_k^{(n)}\alpha_n(k)n^{-1/2}\big). \end{aligned}$$

Then, by a Taylor expansion, there exists a constant c_2 such that for any $1\le j\le n$,

$$\begin{aligned} &\Bigg|\frac{M^{(n)}(j)-M^{(n)}(j-1)}{S^{(n)}(j-1)} - \frac{1}{\sqrt{n}}\big(\big(\sigma+\alpha_n(j)\big)\xi_j^{(n)} - \alpha_n(j-1)\xi_{j-1}^{(n)}\big) \\ &\quad - \frac{\sigma}{2n}\big(\sigma+2\alpha_n(j)\big)\Bigg| \le \frac{c_2}{n^{3/2}} \qquad \text{a.s.} \end{aligned}$$

This, together with (7.2), yields that for any $n\in\mathbb{N}$ and $t\in[0,1]$, we have

$$\left|Y_n(t) - \frac{\sigma}{\sqrt{n}}\sum_{j=1}^{[nt]}\xi_j^{(n)} - \frac{\sigma}{2n}\left(\sigma[nt] + 2\sum_{j=1}^{[nt]}\alpha_n(j)\right)\right| \le \frac{c_3}{\sqrt{n}} \qquad \text{a.s.}$$

for some constant c_3. Since $\frac{\sigma}{\sqrt{n}}\sum_{j=1}^k \xi_j^{(n)} = \ln(S^{(n)}(k)/s_0)$, the above calculations imply that

$$\int_0^t \alpha_n\big([nu]\big)\,du \longrightarrow \frac{1}{\sigma}\left(Y(t) - \ln\big(M(t)/s_0\big) - \frac{t}{2}\right) = \frac{\langle Y\rangle(t)-\sigma^2 t}{2\sigma} \qquad \tilde{P}\text{-a.s.} \quad \square$$

Next, let $c>0$ be a constant, and M a strictly positive, continuous martingale defined on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ satisfying the conditions

$$M(0) = s_0 \quad \text{and} \quad \big|a(t:M)\big| \le c \quad \tilde{P}\text{-a.s.} \tag{7.5}$$

In fact, a nonnegative volatility process α is in $\mathcal{A}^c$ if and only if the corresponding process S_α satisfies the above condition. However, S_α is defined on the canonical space $(\Omega_W, \mathcal{F}^W, \mathbb{P}^W)$, and M is defined on a general space. In the next lemma, we show that maximization of the function $J(M)$ defined in (5.3) over all martingales M satisfying constraint (7.5) is the same as maximizing $J(S_\alpha)$ over $\alpha\in\mathcal{A}^c$. The main difficulty in the proof of this statement is that a priori the optimal martingale (for the functional J) can generate a filtration larger than the Brownian filtration. The proof follows the ideas of Lemma 5.2 in [16] and uses a randomization technique.

Lemma 7.2 *Let M be a strictly positive, continuous martingale on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ satisfying (7.5). Then*

$$J(M) \le \sup_{\alpha\in\mathcal{A}^c} J(S_\alpha).$$

Proof Set

$$Y(t)=\int_0^t \frac{dM(u)}{M(u)},\qquad t\in[0,1],$$

so that

$$M(t)=s_0\exp\left(Y(t)-\frac{\langle Y\rangle(t)}{2}\right),\qquad t\in[0,1].$$

By enlarging the space if necessary, we may assume that $(\tilde\Omega,\tilde F,\tilde P)$ is rich enough to support a Brownian motion $\hat W$ which is independent of M. For $\lambda\in[0,1]$, define

$$Y_\lambda=\sqrt{1-\lambda}Y+\sigma\sqrt{\lambda}\hat W \quad\text{and}\quad M_\lambda=s_0\exp\left(Y_\lambda-\frac{\langle Y_\lambda\rangle}{2}\right).$$

Notice that for all λ, M_λ satisfies the conditions of (7.5). Hence, the family

$$F(M_\lambda)-\int_0^1 \widehat G\big(t,M_\lambda,a(t:M_\lambda)M_\lambda(t)\big)\,dt,\qquad \lambda\in[0,1],$$

is uniformly integrable, and the continuity of $\widehat G$ implies that

$$J(M)=\lim_{\lambda\to 0}J(M_\lambda).$$

Hence, it suffices to show that

$$J(M_\lambda)\le \sup_{\alpha\in\mathcal A^c}J(S_\alpha)$$

for all $\lambda>0$. Since $d\langle Y\rangle(t)\ge\lambda\sigma^2dt$ for any $\lambda>0$, without loss of generality we may assume that for the initial processes M,Y, we have

$$Z(t):=\frac{d\langle Y\rangle(t)}{dt}\ge\epsilon\quad \mathcal L\otimes\tilde P\text{-a.s.}$$

for some $\epsilon>0$. Set

$$\tilde W(t)=\int_0^t\frac{dY(u)}{\sqrt{Z}(u)},\qquad t\in[0,1],$$

$$\kappa_n(0)=\sigma\quad\text{and}\quad \kappa_n(k)=n\int_{(k-1)/n}^{k/n}\sqrt{Z(u)}\,du\quad\text{for }0<k<n, \tag{7.6}$$

$$M^{(n)}(t)=s_0\exp\left(\int_0^t\kappa_n\big([nu]\big)\,d\tilde W(u)-\frac12\int_0^t\kappa_n^2\big([nu]\big)\,du\right),\qquad t\in[0,1],\ n\in\mathbb N.$$

By Lévy's theorem, $\tilde W$ is a Brownian motion with respect to the filtration of M. Therefore, the martingale $M^{(n)}$ satisfies (7.5). Also, from (7.6) it is clear that

$$\lim_{n\to\infty}\kappa_n\big([nt]\big)=\sqrt{Z(t)}$$

in probability for the measure $\mathcal L\otimes\tilde P$. On the other hand, Itô's isometry and the Doob–Kolmogorov inequality imply that

$$\lim_{n\to\infty}\max_{0\le t\le 1}\left|\int_0^t\kappa_n\big([nu]\big)\,d\tilde W(u)-Y(t)\right|=0$$

in probability with respect to $\tilde{P}$. Thus,

$$M^{(n)} \longrightarrow M \quad \text{and} \quad a\big(\cdot : M^{(n)}\big) \longrightarrow a(\cdot : M) \quad \text{a.s.}$$

on the space $\mathcal{D}[0,1]$. We use these convergence results and the uniform integrability of

$$F\big(M^{(n)}\big) - \int_0^1 \widehat{G}\big(t, M^{(n)}, a\big(t : M^{(n)}\big)M^{(n)}(t)\big)\,dt, \quad \lambda \in [0,1],\ n \in \mathbb{N}$$

(which follows from the growth assumption (3.1)) to conclude that

$$J(M) = \lim_{n\to\infty} J\big(M^{(n)}\big).$$

Hence, it suffices to prove for any $n \in \mathbb{N}$ that

$$J\big(M^{(n)}\big) \le \sup_{\alpha\in\mathcal{A}^c} J(S_\alpha). \tag{7.7}$$

We prove the above inequality by a randomization technique. Fix $n \in \mathbb{N}$. From the existence of regular conditional distributions, there exists for any $1 \le k < n$ a function $\rho_k : \mathbb{R} \times \mathcal{C}[0,1] \times \mathbb{R}^k \to [0,1]$ such that for any y, $\rho_k(y,\cdot) : \mathcal{C}[0,1] \times \mathbb{R}^k \to [0,1]$ is measurable and satisfies $\tilde{P}$-a.s.

$$\tilde{P}\big(\kappa_n(k) \le y \,\big|\, \sigma\big\{\tilde{W}, \kappa_n(0), \dots, \kappa_n(k-1)\big\}\big) = \rho_k\big(y, \tilde{W}, \kappa_n(0), \dots, \kappa_n(k-1)\big).$$

Furthermore, $\tilde{P}$-almost surely, $\rho_k(\cdot, \tilde{W}, \kappa_n(0), \dots, \kappa_n(k-1))$ is a distribution function on $\mathbb{R}$. Let W be the Brownian motion on our canonical space $(\Omega_W, \mathcal{F}^W, P^W)$. We extend this space so that it contains a family $\Xi_1, \dots, \Xi_{n-1}$ of i.i.d. random variables which are uniformly distributed on the interval $(0,1)$ and independent of W. Let $(\tilde{\Omega}_W, \tilde{F}^W, \tilde{P}^W)$ be the extended probability space. We assume that it is complete.

Next, we recursively define the random variables $U_0 = \sigma$ and, for $1 \le k < n$,

$$U_k = \sup\big\{y \,\big|\, \rho_k(y, W, U_1, \dots, U_{k-1}) < \Xi_k\big\}. \tag{7.8}$$

In view of the properties of the functions ρ_i, we can show that $U_1, \dots, U_{n-1}$ are measurable. Furthermore, U_i is independent of Ξ_k for any $i < k$. This property, together with (7.8), yields that for any $y \in \mathbb{R}$ and $1 \le k < n$,

$$\begin{aligned}
&\tilde{P}^W\big(U_k \le y \,\big|\, \sigma\{W, U_0, \dots, U_{k-1}\}\big)\\
&\quad = \tilde{P}^W\big(\rho_k(y, W, U_0, \dots, U_{k-1}) \ge \Xi_k \,\big|\, \sigma\{W, U_0, \dots, U_{k-1}\}\big)\\
&\quad = \rho_k(y, W, U_0, \dots, U_{k-1}).
\end{aligned}$$

Thus, we conclude that the vector $(W, U_0, \dots, U_{n-1})$ has the same distribution as $(\tilde{W}, \kappa_n(0), \dots, \kappa_n(n-1))$. Also note that for any k and $t \ge k/n$, $\kappa_n(k)$ is independent of $\tilde{W}(t) - \tilde{W}(k/n)$. Furthermore, since for any k, $\kappa_n(k)$ takes values in the interval $[\sqrt{0 \vee \sigma(\sigma - 2c)}, \sqrt{\sigma(\sigma + 2c)}]$, for $1 \le k < n$, there exist functions

$$\Theta_k : \mathcal{C}[0, k/n] \times (0,1)^k \to \big[\sqrt{0 \vee \sigma(\sigma - 2c)}, \sqrt{\sigma(\sigma + 2c)}\big]$$

satisfying

$$U_k = \Theta_k(W, \Xi_1, \dots, \Xi_k), \quad 1 \le k < n,$$

where in the expression above we consider the restriction of W to the interval $[0, k/n]$. Next, we introduce the martingale

$$S_U(t) := s_0 \exp\left(\sum_{i=0}^{[nt]}\left(U_i\left(W\left(\frac{i+1}{n}\right) - W\left(\frac{i}{n}\right)\right) - \frac{U_i^2}{2n}\right)\right), \quad t \in [0, 1].$$

Finally, for any $z := (z_1, \dots, z_{n-1}) \in (0, 1)^{n-1}$, define a stochastic process by

$$U^{(z)}(t) = \sigma \text{ if } t = 0 \quad \text{and} \quad U^{(z)}(t) = \Theta_{[nt]}(W, z_1, \dots, z_{[nt]}) \quad \text{for } t \in (0, 1].$$

Observe that $U^{(z)}$ is in $\mathcal{A}^c$. Recall the definition of S_α in (3.3). Since $U^{(z)}$ is piecewise constant, we have

$$S_{U^{(z)}}(t) = s_0 \exp\left(\sum_{i=0}^{[nt]}\left(\Theta_i(W, z_1, \dots, z_i)\left(W\left(\frac{i+1}{n}\right) - W\left(\frac{i}{n}\right)\right)\right.\right.$$
$$\left.\left. - \frac{\Theta_i^2(W, z_1, \dots, z_i)}{2n}\right)\right).$$

We now use Fubini's theorem to conclude that

$$J\big(M^{(n)}\big) = J(S_U) = \int_{z\in(0,1)^n} J(S_{U^{(z)}})dz_1 \cdots dz_n \le \sup_{\alpha\in\mathcal{A}^c} J(S_\alpha), \tag{7.9}$$

and (7.7) follows. □

Our final result is the denseness from Definition 6.1 in $\mathcal{A}^c$ of the subset $\mathcal{L}(c)$. The following result is proved by standard arguments. Since we could not find a direct reference, we provide a self-contained proof.

Lemma 7.3 *For any $c > 0$,*

$$\sup_{\alpha\in\mathcal{A}^c} J(S_\alpha) = \sup_{\tilde{\alpha}\in\mathcal{L}(c)} J(S_{\tilde{\alpha}}).$$

Proof Let $\{\phi_n\}_{n=1}^\infty \subseteq \mathcal{L}(c)$ be a sequence which converges in probability (with respect to $\mathcal{L} \otimes P^W$) to some $\alpha \in \mathcal{A}^c$. By the Itô isometry and the Doob–Kolmogorov inequality, we directly conclude that $\{S_{\phi_n}\}$ converges to S_α in probability on the space $\mathcal{C}[0, 1]$. Then, invoking the uniform integrability of

$$F(S_{\phi_n}) - \int_0^1 \widehat{G}\big(t, S_{\phi_n}, a(t :, S_{\phi_n}), S_{\phi_n}(t)\big)\,dt, \quad \lambda \in [0, 1], \quad n \in \mathbb{N},$$

once again gives $\lim_{n\to\infty} J(S_{\phi_n}) = J(S_\alpha)$.

Therefore, to prove the lemma, we need to construct, for any $\alpha \in \mathcal{A}^c$, a sequence $\{\phi_n\}_{n=1}^\infty \subseteq \mathcal{L}(c)$ which converges in probability to α. Thus, take $\alpha \in \mathcal{A}^c$ and $\delta > 0$. It is well known (see [15], Chap. 4, part b of Lemma 2.4) that there exists a continuous process ϕ adapted to the Brownian filtration which satisfies

$$\mathcal{L} \otimes P^W\big(|\alpha - \phi| > \delta\big) < \delta. \tag{7.10}$$

Since the process ϕ is continuous, for all sufficiently large m, we have

$$P^W\Big(\max_{0\le k\le m-2}\sup_{k/m\le t\le (k+2)/m}\big|\phi(t)-\phi(k/m)\big|>\delta\Big)<\delta. \tag{7.11}$$

Clearly, for any $1\le k\le m$, there exists a measurable function $\theta_k:\mathcal{C}[0,k/m]\to\mathbb{R}$ for which

$$\theta_k(W)=\phi(k/m),\quad 1\le k\le m,$$

where in the expression above we consider the restriction of W to the interval $[0,k/m]$. Fix k. It is well known (see, for instance, [3], Chap. 1, Theorem 1.2) that any simple function (i.e., a linear combination of indicator functions) on a separable metric space can be represented as a limit (in probability) of Lipschitz-continuous functions. From the denseness of simple functions in the space of bounded functions (with respect to convergence a.s.) it follows that we can find a sequence of bounded Lipschitz-continuous functions $\vartheta_n:\mathcal{C}[0,k/m]\to\mathbb{R}$, $n\in\mathbb{N}$, such that $\lim_{n\to\infty}\vartheta_n=\theta_k$ a.s. with respect to the Wiener measure on the space $\mathcal{C}[0,k/m]$. We conclude that there exist a constant $\mathcal{H}>0$ and a sequence of functions $\Theta_k:\mathcal{C}[0,1]\to\mathbb{R}$, $1\le k\le m-3$, such that for any $z_1,z_2\in\mathcal{C}[0,1]$ and $1\le k\le m-3$,

$$\Theta_k(z_1)=\Theta_k(z_2)\quad\text{if } z_1(s)=z_2(s)\text{ for all } s\le k/m,$$

$$\big|\Theta_k(z_1)\big|\le\mathcal{H}, \tag{7.12}$$

$$\big|\Theta_k(z_1)-\Theta_k(z_2)\big|\le\mathcal{H}\big(\|z_1-z_2\|\big), \tag{7.13}$$

$$P^W\big(\big|\Theta_k(W)-\phi(k/m)\big|>\delta\big)<\delta/m. \tag{7.14}$$

Let $\Theta_{-1},\Theta_0,\Theta_{m-2}:\mathcal{C}[0,1]\to\mathbb{R}$ be given by $\Theta_{-1}=\Theta_0\equiv\phi(0)$ and $\Theta_{m-2}\equiv\sigma$. Define $f_1:[0,1]\times\mathcal{C}[0,1]\to\mathbb{R}$ by

$$f_1(t,z)=\begin{cases}([mt]+1-mt)\Theta_{[mt]-1}(z)+(mt-[mt])\Theta_{[mt]}(z), & t<1-1/m,\\ f_1(t,z)=\sigma, & t\ge 1-1/m.\end{cases}$$

Denote $a=\sqrt{0\vee\sigma(\sigma-2c)}$ and $b=\sqrt{\sigma(\sigma+2c)}$. Without loss of generality we assume that $\delta<\min(\sigma-a,b-\sigma)$. Set

$$f(t,z)=\big((a+\delta)\vee f_1(t,z)\big)\wedge(b-\delta),\quad t\in[0,1],\ z\in\mathcal{C}[0,1].$$

Using (7.12) and (7.13), we conclude that for any $0\le k\le m-2$, $z_1,z_2\in\mathcal{C}[0,1]$, and $t_1,t_2\in[k/m,(k+1)/m]$,

$$\begin{aligned}\big|f(t_2,z_2)-f(t_1,z_1)\big|&\le\big|f_1(t_2,z_2)-f_1(t_1,z_2)\big|+\big|f_1(t_1,z_2)-f_1(t_1,z_1)\big|\\&\le m|t_1-t_2|\big(\big|\Theta_{k-1}(z_2)\big|+\big|\Theta_k(z_2)\big|\big)\\&\quad+\big|\Theta_{k-1}(z_2)-\Theta_{k-1}(z_1)\big|+\big|\Theta_k(z_2)-\Theta_k(z_1)\big|\\&\le 2(\mathcal{H}+\sigma)(m+1)\big(|t_1-t_2|+\|z_1-z_2\|\big).\end{aligned}$$

Define the process $\{\Theta(t)\}_{0\le t\le 1}$ by $\Theta(t)=f(t,W)$, $t\in[0,1]$. By the choice of δ, it follows that $\Theta\in\mathcal{L}(c)$. Next, observe that for any $t\in[1/m,1-1/m]$, we have

$$\big|\Theta(t)-\phi(t)\big|\le\max\big(\big|\phi(t)-\Theta_{[mt]}(W)\big|,\big|\phi(t)-\Theta_{[mt]-1}(W)\big|\big).$$

Thus, for any $t \in [1/m, 1-1/m]$,

$$\big|\Theta(t)-\alpha(t)\big| \le \Big(\max_{0\le k\le m-3}\sup_{k/m\le t\le (k+2)/m}\big|\phi(t)-\phi(k/m)\big|\Big) + \Big(\max_{0\le k\le m-3}\big|\phi(k/m)-\Theta_k(W)\big|\Big) + \big|\alpha(t)-\phi(t)\big|. \tag{7.15}$$

Finally, by combining (7.10), (7.11), (7.14), and (7.15) we get

$$\mathcal{L}\otimes P^W\big(|\Theta-\alpha|>3\delta\big) < \frac{2}{m}+3\delta<5\delta.$$

Since $\delta>0$ was arbitrary, we complete the proof. □

Acknowledgements Research supported by the European Research Council Grant 228053-FiRM, the Swiss Finance Institute and the ETH Foundation. The authors would like to thank Prof. Kusuoka and Marcel Nutz for insightful discussions.

References

1. Bank, P., Baum, D.: Hedging and portfolio optimization in financial markets with a large trader. Math. Finance **14**, 1–18 (2004)
2. Barles, G., Soner, H.M.: Option pricing with transaction costs and a nonlinear Black–Scholes equation. Finance Stoch. **2**, 369–397 (1998)
3. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
4. Çetin, U., Rogers, L.C.G.: Modelling liquidity effects in discrete time. Math. Finance **17**, 15–29 (2007)
5. Çetin, U., Jarrow, R., Protter, P.: Liquidity risk and arbitrage pricing theory. Finance Stoch. **8**, 311–341 (2004)
6. Çetin, U., Soner, H.M., Touzi, N.: Option hedging under liquidity costs. Finance Stoch. **14**, 317–341 (2010)
7. Chalasani, P., Jha, S.: Randomized stopping times and American option pricing with transaction costs. Math. Finance **11**, 33–77 (2001)
8. Dolinsky, Y.: Hedging of Game Options with the Presence of Transaction Costs (2012), submitted. arXiv:1103.1165,v3
9. Dolinsky, Y., Nutz, M., Soner, H.M.: Weak approximation of G-expectations. Stoch. Process. Appl. **122**, 664–675 (2012)
10. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**, 463–520 (1994)
11. Dudley, R.M.: Distances of probability measures and random variables. Ann. Math. Stat. **39**, 1563–1572 (1968)
12. Duffie, D., Protter, P.: From discrete to continuous time finance: weak convergence of the financial gain process. Math. Finance **2**, 1–15 (1992)
13. Gökay, S., Soner, H.M.: Liquidity in a binomial market. Math. Finance **22**, 250–276 (2012)
14. Guasoni, P., Rasonyi, M., Schachermayer, W.: Consistent price systems and face-lifting pricing under transaction costs. Ann. Appl. Probab. **18**, 491–520 (2008)
15. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, New York (1991)
16. Kusuoka, S.: Limit theorem on option replication cost with transaction costs. Ann. Appl. Probab. **5**, 198–221 (1995)
17. Levental, S., Skorohod, A.V.: On the possibility of hedging options in the presence of transaction costs. Ann. Appl. Probab. **7**, 410–443 (1997)
18. Peng, S.: Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. Stoch. Process. Appl. **118**, 2223–2253 (2008)
19. Pennanen, T., Penner, I.: Hedging of claims with physical delivery under convex transaction costs. SIAM J. Financ. Math. **1**, 158–178 (2010)
20. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)

21. Schachermayer, W.: The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. Math. Finance **14**, 19–48 (2004)
22. Soner, H.M., Shreve, S.E., Cvitanić, J.: There is no nontrivial hedging portfolio for option pricing with transaction costs. Ann. Appl. Probab. **5**, 327–355 (1995)
23. Soner, H.M., Touzi, N.: The dynamic programming equation for second order stochastic target problems. SIAM J. Control Optim. **48**, 2344–2365 (2009)
24. Soner, H.M., Touzi, N., Zhang, J.: Dual formulation of second order target problems. Ann. Appl. Prob. (2012), in press. arXiv:1003.6050