Duality and Convergence for Binomial Markets with Friction *

Yan Dolinsky † H. Mete Soner ‡

June 9, 2011

Abstract

We prove limit theorems for the super-replication cost of European options in a Binomial model with friction. The examples covered are markets with proportional transaction costs and the illiquid markets. The dual representation for the super-replication cost in these models are obtained and used to prove the limit theorems. In particular, the existence of the liquidity premium for the continuous time limit of the model proposed in [6] is proved. Hence, this paper extends the previous convergence result of [13] 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 in [14].

Keywords: Super-replication, Liquidity, Binomial model, Limit theorems, G-expectation
AMS 2000 Subject Classifications: 91B28, 60F05, 60H30
JEL Classifications: G11, G13, D52

1 Introduction

We consider a one-dimensional Binomial model in which the size of the trade has an immediate but temporary effect on the price of the asset. Indeed, let $g(t,\nu)$ be the cost of trading ν shares at time t. We simply assume that g is adapted to the natural filtration and it is 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 a given constant $\lambda>0$. However, it also covers the illiquidity model considered in [5] and [13] which is the Binomial version of the model introduced by Cetin, Jarrow and Protter in [6] for continuous time. In this example, g is twice differentiable at $\nu=0$.

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 non-differentiable at the origin), the super-replication cost is very costly as proved in [22, 16, 8]. In several papers [3, 14] asymptotic problems with vanishing transaction costs are considered to obtain non-trivial pricing equations. On the other hand, when g is differentiable then any

*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.

†ETH Zurich, Dept. of Mathematics, yan.dolinsky@math.ethz.ch
‡ETH Zurich, Dept. of Mathematics, and Swiss Finance Institute, mete.soner@math.ethz.ch

continuous trading strategy which has finite variation, has no liquidity cost. Thus one may avoid the liquidity cost entirely as shown in [6] and also in [2]. However, in [7] 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

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

where t is time, s is the current stock price and H is a convex nonlinear function of the second derivative derived explicitly in [7, 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 the homogenization techniques for viscosity solutions. Thus it is limited to the 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 the classical theory. This derivation is an advantage of the discrete model as the derivation of the dual in continuous time is essentially an open problem. Although 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 step 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 [14] constructs a sequence of martingales on the discrete probability space $\{-1,1\}^\infty$ of a specific form which converge in law to M. Moreover, the quadratic variation of M is approximated through this powerful procedure of Kusuoka. This construction is our main tool in proving the lower bound (i.e., existence of 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 costly 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 smaller. Indeed, if in an n-step model, we take $\lambda_n=\Lambda/\sqrt{n}$ then the limit problem is the uncertain volatility model or equivalently G-expectation of Peng [17]. This is exactly the main result of Kusuoka in [14]. In fact, relatedly, the authors in joint work with M. Nutz [12] provides a different discretization of the G-expectation.

The paper is organized as following. In the next section we introduce the setup. In Section 3 we formulate the main results of this paper. In Section 4 we prove Theorem 3.1, that 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 Section 5. In Section 6 we state the main results from [14], which are used in this paper. In particular we give a short formulation of the main properties of Kusuoka construction, which is the main tool in proving the lower bound (liquidity premium) of Theorem 3.5. In Section 7 we derive auxiliary lemmas, Lemmas 7.1–7.2 and Lemma 7.3 that are used in the proof of Theorem 3.5.

2 Preliminaries and the model

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

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

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

For any $T > 0$ denote by $\mathscr{C}[0,T]$ 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 $F : \mathscr{C}[0,T] \to \mathbb{R}_+$ be a continuous map such that there are constant $C, p > 0$ for which

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

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). Namely, for any n consider the n–step binomial model of a financial market which is active at times $0, 1/n, 2/n, \ldots, 1$. It consists of a savings account, and of a stock. Without loss of generality (by discounting), we assume that the savings account price is a constant which equals to 1. The stock price at time k/n is given by

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

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

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

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

$$F_n := F\left(\mathscr{W}_n(S^{(n)})\right) \tag{2.3}$$

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

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

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. We assume that this function represents the cost of trading in this market. We assume the following.

Assumption 2.1 The *trading cost function*

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

is assumed to be non-negative, adapted with $g(t, S, 0) = 0$. Moreover we assume that $g(t, S, \cdot)$ is convex for every $(t, S) \in [0,1] \times \mathscr{C}[0,1]$.

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

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

A self financing portfolio π with an initial capital x is a pair $\pi = (x, \{\gamma(k)\}_{k=0}^{n})$ where $\gamma(0) = 0$ and for any $k \geq 1$, $\gamma(k)$ is a $\mathscr{F}_{k-1}$ 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

$$Y^{\pi}(k+1) = Y^{\pi}(k) + \gamma(k+1)\left(S^{(n)}(k+1) - S^{(n)}(k)\right) - g\left(\frac{k}{n}, \mathscr{W}_n(S^{(n)}), \gamma(k+1) - \gamma(k)\right), \tag{2.4}$$

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

Observe that $Y^{\pi}(k)$ is the portfolio value at the time (k/n) *before* a transfer is made at this time, and the last term in equation (2.4) represents the cost of trading and it is the only source of friction in the model. We would mostly use the notation $Y^{\pi}(k)$ when the dependence on the penalty function is clear.

Let $\mathscr{A}_n(x)$ be the set of all portfolios with an initial capital x. The problem we consider is the super-replication cost of a European claim whose pay-off is given in (2.3). Then, the problem is

$$V_n := V_n(g,F_n) = \inf\{x \mid \exists\ \pi \in \mathscr{A}_n(x) \ \text{ such that } Y^{\pi}(n:g) \geq F_n,\ \mathbb{Q}\text{-a.s.}\}. \tag{2.5}$$

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 the models with proportional transaction costs and models with price impact.

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

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

Observe that G may become infinite. It is well known that the following dual relation holds,

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

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 Cetin, Jarrow, Protter [6] that was studied in [13] (see Remark 2.4 below).

In [23], it is proved that the optimal trading strategies in continuous time 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 $\Lambda = g_{\nu\nu}(t,S,0)$.

b. This example which corresponds to the example of proportional transaction costs. For fix n recently there has been interesting results in relation to arbitrage. We refer to the paper of Schachermayer [20], Pennanen and Penner [18] and the references therein. But as remarked earlier, 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 exactly the model studied in [14] by Kusuoka. In this case, the dual function is given by

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

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

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

where G is the convex conjugate of the original g.

An important but a 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

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

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 2.1.b is the truncation of the following function

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

Example 2.1.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 [6], we introduce a path dependent supply curve,

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

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

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

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

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

∎

3 Main results

Our first result is the characterization of the dual problem. We believe that this simple result is quite interesting by itself. Also it will be the essential tool to study the asymptotic behavior of the super-replication costs.

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 $\mathscr{Q}_n$ be the set of all probability measures on $(\Omega, \mathscr{F}_n)$. Then*

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

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

The duality is proved in the next section.

In the limit theorem that we state below, 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 $\beta \geq 2$ such that

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

b). There exists a continuous function

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

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),

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

It is straightforward to show that $\widehat{G}$ is quadratic in the y-variable. Moreover, the above assumption (3.2) is essentially equivalent to assume that G is twice differentiable at the origin. Indeed, when G twice differentiable, Taylor approximation implies that

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

We give the following example to clarify the above assumption.

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

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

Then, for $\gamma > 1$

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

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 $(\Omega_W, \mathscr{F}^W, \mathbb{P}^W)$ be a complete probability space together with a standard one-dimensional Brownian motion W and the right continuous filtration $\mathscr{F}_t^W = \sigma\big\{\sigma\{W(s)|s \leq t\} \bigcup \mathscr{N}\big\}$, where $\mathscr{N}$ is the collection of all $\mathbb{P}^W$ null sets. For any α progressively measurable, bounded, real-valued process, let $S_\alpha(t)$ be the continuous martingale given by

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

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).

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

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 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, real-valued process $\alpha(\cdot)$ satisfying

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

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

As before g is the penalty function and g_n^c is the truncation of g at the level c as defined in Definition 2.3. Let F_n be a given claim and $V_n = V_n(g, F_n)$ be 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 Section 5, is the main result of the paper. It provides the asymptotic behavior of the truncated super-replication costs V_n^c. Since $V_n^c \le V_n$ for every c, the below result 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 the Assumption 3.2 and let $\widehat{G}$ be as in* (3.2)*. Then, for every $c > 0$,*

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

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

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

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

Corollary 3.6

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

where $\mathscr{A}$ is the set of all bounded, 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 pay-offs, [13] proves the equality.

Remark 3.7 (Liquidity Premium) It is an interesting question whether the limiting super-replication cost contain liquidity premium. Namely, whether the right hand side of (3.6) is strictly bigger than $V_{BS}(F)$. For Markovian non-affine pay-offs it was proved in [7]. 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 setting 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 liquidity premium. Fix $\varepsilon > 0$. From (3.1), one can prove the following estimate

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

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

$$\sup_{\alpha \in \mathscr{A}^\varepsilon} \mathbb{E}^W [F(S_\alpha)] \ge \mathbb{E}^W F(S_\sigma) + C\varepsilon.$$

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

We close this section by revisiting the 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 [7, 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, larger γ means less trading cost. Hence when γ is sufficiently large (i.e., $\gamma > 2$), then 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 Theorem 3.1. Fix $n \in \mathbb{N}$ and consider the n-step binomial model with the 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 $k = 0, \ldots, n-1$,

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

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

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

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

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}[Y^\pi(n)] &= x + \mathbb{E}^{\mathbb{P}} \left(\sum_{k=0}^{n-1} [\gamma(k+1) - \gamma(k)]\ [\mathbb{E}^{\mathbb{P}}(S^{(n)}(n)|\mathscr{F}_k) - S^{(n)}(k)] - g\left(\frac{k}{n}, \gamma(k+1) - \gamma(k)\right)\right) \\ &\leq x + \mathbb{E}^{\mathbb{P}} \left(\sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathbb{E}^{\mathbb{P}}(S^{(n)}(n)|\mathscr{F}_k) - S^{(n)}(k)\right)\right). \end{aligned}$$

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

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

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

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

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

Proof of Theorem 3.1.

We model the n-step binomial model as in [4]. Consider a tree whose nodes are sequences of the form $(a_1,...,a_k) \in \{-1,1\}^k$, $0 \le k \le n$. The set of all nodes will be denoted by $\mathbb{V}$. The empty sequence (corresponds to the case $k=0$) is the root of the tree and will be denoted by $\emptyset$. In our model each node of the form $u=(u_1,...,u_k) \in \{-1,1\}^k$, $k<n$ has two immediate successors $(u_1,...,u_k,1)$ and $(u_1,...,u_k,-1)$. Let $\mathbb{T} := \{-1,1\}^n$ be the set of all terminal nodes. For $u \in \mathbb{V}\setminus\mathbb{T}$, denote by u^+ the set which consists of the immediate successors of u. The unique immediate predecessor of a node $u=(u_1,...,u_k) \in \mathbb{V}\setminus\{\emptyset\}$ is denoted by $u^- := (u_1,...,u_{k-1})$. For $u=(u_1,...,u_k) \in \mathbb{V}\setminus\mathbb{T}$, let

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

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

$$S(u) = s_0 \exp\left(\tfrac{\sigma}{\sqrt{n}}\textstyle\sum_{i=1}^{l(u)} u_i\right), \ \ \hat{S}(u) = \mathscr{W}_n(\{S(u_1,...,u_{k\wedge l(u)})\}_{k=0}^n)$$
$$\hat{F}(v) = F(\hat{S}(v)), \ \ \forall u \in \mathbb{V}, \ \ v \in \mathbb{T}.$$

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

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

over all β,γ,Y subject to the constrains

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

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

$$Y(u) + g\left(\frac{l(u^-)}{n}, \hat{S}(u^-), \beta(u^-)\right) - \gamma(u)[S(u) - S(u^-)] - Y(u^-) \le 0, \ \ \forall u \in \mathbb{V}\setminus\{\emptyset\}, \tag{4.4}$$

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

Notice that (2.4) implies that the constraint (4.4) should be in fact an equality. However, this modification of the constraint does not alter the value of the optimization problem. The optimization problem which is 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 [19]) 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

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

We rearrange the above expressions to arrive at

$$L(\Upsilon,\Phi,\Theta,\beta,\gamma,Y) = Y(\emptyset)(1-\sum_{u\in\emptyset^+}\Phi(u)) + \sum_{u\in\mathbb{V}\setminus(\{\emptyset\}\cup\mathbb{T})} Y(u)\big(\Phi(u)-\sum_{\tilde{u}\in u^+}\Phi(\tilde{u})\big) \tag{4.6}$$

$$+\sum_{u\in\mathbb{T}} Y(u)(\Phi(u)-\Theta(u)) + \gamma(\emptyset)\big(\Upsilon(\emptyset)-\sum_{u\in\emptyset^+}\Upsilon(u)\big)$$

$$+\sum_{u\in\mathbb{V}\setminus\{\emptyset\}}\gamma(u)\left(\Upsilon(u)-\sum_{\tilde{u}\in u^+}\Upsilon(\tilde{u})\Phi(u)\big(S(u)-S(u^-)\big)\right)+\sum_{u\in\mathbb{T}}\Theta(u)\hat{F}(u)$$

$$+\sum_{u\in\mathbb{V}\setminus\mathbb{T}}\left(\sum_{\tilde{u}\in u^+}\Phi(\tilde{u})g\left(\frac{l(u)}{n},\hat{S}(u),\beta(u)\right)-\beta(u)\sum_{\tilde{u}\in u^+}\Upsilon(\tilde{u})\right).$$

By Theorem 28.2 in [19], 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

$$V_n = \sup_{(\Upsilon,\Phi,\Theta)\in D}\ \inf_{(\beta,\gamma,Y)\in\mathbb{R}^{\mathbb{V}\setminus\mathbb{T}}\times\mathbb{R}^{\mathbb{V}}\times\mathbb{R}^{\mathbb{V}}}\Bigg[\sum_{u\in\mathbb{T}}\Theta(u)\hat{F}(u) \tag{4.8}$$

$$+\sum_{u\in\mathbb{V}\setminus\mathbb{T}}\left(\sum_{\tilde{u}\in u^+}\Phi(\tilde{u})g\left(\frac{l(u)}{n},\hat{S}(u),\beta(u)\right)-\beta(u)\sum_{\tilde{u}\in u^+}\Upsilon(\tilde{u})\right)\Bigg]$$

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

$$\sum_{u\in\emptyset^+}\Phi(u)=1,\quad \sum_{\tilde{u}\in u^+}\Phi(\tilde{u})=\Phi(u),\quad \forall u\in\mathbb{V}\setminus(\mathbb{T}\cup\{\emptyset\}), \tag{4.9}$$

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

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

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

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

where we use the convention that $0/0=0$ (observe that if $\Phi(u)=0$ then $\sum_{\tilde{u}\in\mathbb{T}(u)}\Phi(\tilde{u})S(\tilde{u})=0$). 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),

$$V_n=\sup_{\Phi\in\mathbb{D}}\sum_{u\in\mathbb{T}}\Phi(u)\left(\hat{F}(u)-G\left(\frac{l(u)}{n},\hat{S}(u),\frac{\sum_{\tilde{u}\in\mathbb{T}(u)}\Phi(\tilde{u})S(\tilde{u})}{\Phi(u)}-S(u)\right)\right). \tag{4.13}$$

Clearly there is a natural bijection $\pi:\mathbb{D}\to\mathscr{Q}_n$ (where, recall $\mathscr{Q}_n$ is the set of all probability measures on $(\Omega,\mathscr{F}_n)$) such that for any $\Phi\in\mathbb{D}$ the probability measure $P:=\pi(\Phi)$ is given by

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

Finally we combine (4.13) and (4.14) to conclude that

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

∎

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 Section 7. Also the Kusuoka's construction of discrete martingales are outlined in the next section.

We start with some definitions. Let B be the canonical map on the space $\mathscr{C}[0,1]$, i.e., for each $t \in [0,1]$ $B(t) : \mathscr{C}[0,1] \to \mathbb{R}$ 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 satisfies

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

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 be the constant volatility in the definition of the discrete market, c.f., (2.2). Set

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

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

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 $J(M)$ 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.

For fix $c > 0$, we start by proving the upper bound of Theorem 3.5:

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

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 $(\Omega, \mathscr{F}_n)$ such that

$$\begin{aligned} V_n^c &\leq \frac{1}{n} + E_n\left[F\left(\mathscr{W}_n(S^{(n)})\right) - \sum_{k=0}^{n-1} G_n^c\left(\frac{k}{n}, \mathscr{W}_n(S^{(n)}), E_n\left[S^{(n)}(n)|\mathscr{F}_k\right] - S^{(n)}(k)\right)\right], \\ &= \frac{1}{n} + E_n\left[F\left(\mathscr{W}_n(S^{(n)})\right) - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathscr{W}_n(S^{(n)}), E_n\left[S^{(n)}(n)|\mathscr{F}_k\right] - S^{(n)}(k)\right)\right] \end{aligned} \tag{5.5}$$

where E_n denotes the expectation with respect to P_n. In the last identity we used the form of the dual function G_n^c. Indeed, (2.7) states that either $G_n^c = G$ or $G_n^c = +\infty$. This argument also shows that for any $0 \leq k < n$,

$$\left| E_n\left[S^{(n)}(n)|\mathscr{F}_k\right] - S^{(n)}(k) \right| \leq \frac{c}{\sqrt{n}}\, S^{(n)}(k), \quad P_n \text{ a.s.} \tag{5.6}$$

Indeed, if above does not hold, then in view of (2.7), we would conclude that the right hand side of (5.5) would be equal to negative infinity. But it is easy to show that V_n^c is non-negative.

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

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

Let Q_n be the joint distribution of the stochastic processes $(\mathscr{W}_n(S^{(n)}), A_n)$ under the measure P_n. In view of (5.6), the hypothesis of Lemma 7.1 is satisfied. Hence, there exists a subsequence (denoted by n again) and a probability measure P on the probability space $\mathscr{C}[0,1]$ such that

$$Q_n \Rightarrow Q \text{ on the space } \mathscr{C}[0,1] \times \mathscr{C}[0,1]$$

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 [9]) it follows that there exists a probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ on which

$$\Big(\mathscr{W}_n(S^{(n)}), A_n(\cdot)\Big) \to (M, A(\cdot : M)) \quad \tilde{P}\text{-a.s.} \tag{5.7}$$

on the space $\mathscr{C}[0,1] \times \mathscr{C}[0,1]$, where M is a strictly positive martingale. Furthermore, (5.6) implies that Lemma 6.4 applies to this sequence. Hence we have the following pointwise estimate,

$$|a(t:M)| = |A'(t:M)| \leq c \quad \mathscr{L} \otimes \tilde{P}\text{-a.s.}$$

Next, we will replace the sequence α_n (which converges only weakly) by a pointwise convergent sequence. Indeed, by Lemma A1.1 in [11], we construct a sequence

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

such that η_n converges almost surely in $\mathscr{L} \otimes \tilde{P}$ to a stochastic process η. We now use (5.7) together with the Lebesgue dominated convergence theorem. The result is

$$\begin{aligned}
\int_0^t \eta(u)du &= \lim_{n\to\infty}\int_0^t \eta_n(u)du = \lim_{n\to\infty}\int_0^t \alpha_n([nu])du \\
&= A(t:M) = \int_0^t a(u:M)du, \qquad \mathscr{L} \otimes \tilde{P} \text{ a.s.}
\end{aligned}$$

Hence, we conclude that

$$\eta(t) = a(t:M), \qquad \mathscr{L} \otimes \tilde{P} \text{ a.s.}$$

We are now ready to use the assumption (3.2). Indeed, by definition

$$M^{(n)}(k) - S^{(n)}(k) = \alpha_n(k)\frac{\xi_k}{\sqrt{n}}\, S^{(n)}(k) = \alpha_n(k)\frac{\xi_k}{\sqrt{n}}\, \mathscr{W}_n(S^{(n)})(k/n).$$

Also, by (5.7), $\mathscr{W}_n(S^{(n)})$ converges to M. Hence in view of (5.6), we can use (3.2) to conclude that

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

The estimate (6.3) and the growth assumption (3.1) imply that the above sequences are uniformly integrable. Therefore,

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

where again, without loss of generality (by passing to a subsequence) we assumed that the above limits exist. We now use the convexity of $\widehat{G}$ with respect to third variable (in fact, $\widehat{G}$ is quadratic in y) together with the uniform integrability (which again follows from (3.1) and Lemma 6.4) and the Fubini theorem. The result is

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

The growth assumption on F, namely (2.1) and Lemma 6.4, also imply that the sequence $F\left(\mathscr{W}_n(S^{(n)})\right)$ is uniformly integrable. Then, by (5.7),

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

Hence, we have shown that

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

The above together with Lemma 7.2 yields (5.4).

Lower Bound.

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

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

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

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

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

By the dual representation and the above estimate,

$$\lim_{n\to\infty} V_n^c \geq \limsup_{n\to\infty} E_n^{(\alpha)}\left[F\left(\mathscr{W}_n(S^{(n)})\right) - \sum_{k=0}^{n-1} G\left(\frac{k}{n}, \mathscr{W}_n(S^{(n)}), M_n^{(\alpha)}(k) - S^{(n)}(k)\right)\right] \tag{5.9}$$

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

$$\left(\mathscr{W}_n(S^{(n)}), \mathscr{W}_n(\kappa_n^{(\alpha)})\right) \to (S_\alpha, a(\cdot : S_\alpha)) \quad \tilde{P} \text{ a.s.} \tag{5.10}$$

on the space $\mathscr{C}[0,1] \times \mathscr{C}[0,1]$. 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

$$\lim_{n\to\infty} E_n^{(\alpha)} F\left(\mathscr{W}_n(S^{(n)})\right) = \tilde{E} F(S_\alpha).$$

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 the definition (6.2),

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

In the approximation above, we used the fact that $\kappa^{(\alpha)}$'s are uniformly bounded by construction. We now use (5.10) to arrive at

$$\lim_{n\to\infty} \sqrt{n}\xi_{[nt]}(M_n^{(\alpha)}([nt]) - S^{(n)}([nt])) = a(t : S_\alpha) S_\alpha(t), \ \ \mathscr{L} \otimes \tilde{P} \text{ a.s.} \tag{5.11}$$

As in the upper bound case, the growth condition (3.1) and Lemma 6.4 imply that the sequences

$$nG([nt]/n, \mathscr{W}_n(S^{(n)}), M_n^{(\alpha)}([nt]) - S^{(n)}([nt])) \quad \text{and} \quad \widehat{G}(t, S_\alpha, a(t : S_\alpha) S_\alpha(t)),$$

are uniformly integrable in $\mathscr{L} \otimes \tilde{P}$. Since $\widehat{G}$ is continuous by the Fubini's theorem and (3.2), (5.10), (5.11), we obtain,

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

We use the above limit results for $\tilde{I}$ and for F_n in (5.9). 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 [14] Kusuoka provides an elegant approximation for sufficiently smooth volatility process α. Here we will only state the results of Kusuoka and refer to [14] for the construction. We start by defining the class of "smooth" volatility processes. As before, let $(\Omega_W, \mathscr{F}^W, \mathbb{P}^W)$ be a Brownian probability space and W be the standard Brownian motion.

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

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

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

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

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

$$\left| \frac{f^2(t,S) - \sigma^2}{2\sigma} \right| \le c - \delta(f),$$

and

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

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

$$\left| f(t_1,S) - f(t_2,\tilde{S}) \right| \le L(f) \left(|t_1 - t_2| + \|S - \tilde{S}\|_\infty \right).$$

■

In Kusuoka's construction the condition (6.1) is not needed. However, this regularity allows us to control the behavior of the martingales near maturity .

Recall from Section 2 that $\Omega = \{1,1\}^\infty$, ξ is the canonical map (i.e., $\xi_k(\omega) = \omega_k$) and $\mathbb{Q}$ is the symmetric product measure. The martingales constructed in [14] are of the form

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

where the sequence of discrete *predictable* processes $\kappa_n^{(\alpha)}$ need 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 $\mathscr{F}_n$, this martingale measure is given by,

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

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

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

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

$$\begin{aligned} |\kappa_n^{(\alpha)}(k,\omega)| &< c-\delta, \;\; \kappa_n^{(\alpha)}(k,\omega) > \delta - \frac{1}{2}, \\ |\kappa_n^{(\alpha)}(k-1,\omega) - \kappa_n^{(\alpha)}(k,\omega)| &\leq \frac{L}{\sqrt{n}}, \;\; 1 \leq k \leq n, \end{aligned}$$

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

We also require

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

to ensure $M_n^{(\alpha)}(n) = S^{(n)}(n)$.

Let $Q_n^{(\alpha)}$ be the joint distribution of the pair $(\mathscr{W}_n(S^{(n)}), \mathscr{W}_n(\kappa_n^{(\alpha)}))$ under $P_n^{(\alpha)}$ on the space $\mathscr{C}[0,1] \times \mathscr{C}[0,1]$ with the uniform topology.

Recall once again that the probability space is $\Omega = \{-1,1\}^\infty$ and the filtration $\{\mathscr{F}_k\}_{k=0}^n$ is the usual one generated by the canonical map and that 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 [14]) *Let $c > 0$ and $\alpha \in \mathscr{L}(c)$. Then, on $\left(\Omega, \{\mathscr{F}_k\}_{k=0}^n\right)$ there exists a sequence of predictable processes $\kappa_n^{(\alpha)}$ satisfying the above conditions, hence there also exist sequences of martingales $M_n^{(\alpha)}$ and martingale measures $P_n^{(\alpha)}$ so that*

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

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 [14].

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 is also in the probability space. Namely, the process α is a process on the canonical probability space $\mathscr{C}[0,1]$ while $\kappa_n^{(\alpha)}$ lives in the discrete space Ω. This difficulty is resolved by Kusuoka in [14]. ■

We complete this section by stating (without proof) a lemma which summarizes the main results from Section 4 in [14]; see in particular, Lemma 4.2 and Proposition 4.27 in [14]. In our analysis the below lemma provides the crucial tightness result which is used in the proof of the upper bound of Theorem 3.5. Furthermore, the inequality (6.3) is essential in establishing the uniform integrability of several sequences.

Let $(\Omega, \mathbb{Q})$ be the probability space introduced in Section 2.

Lemma 6.4 (Kusuoka [14]) *Let $M^{(n)}$ be a sequence positive martingales with respect to probability measures P_n on $(\Omega, \mathscr{F}_n)$. Suppose that there exists a constant $c > 0$ such that for any $k \leq n$,*

$$\left| S^{(n)}(k) - M^{(n)}(k) \right| \leq \frac{cS^{(n)}(k)}{\sqrt{n}}, \;\; P_n \;\; a.s.$$

Then, for any $p > 0$

$$\sup_n E_n\big(\max_{0\le k\le n} S^{(n)}(k)\big)^p < \infty, \tag{6.3}$$

where E_n *is an expectation with respect to* P_n.

Moreover, the distribution Q_n *on* $\mathscr{C}[0,1]$ *of* $\mathscr{W}_n(S^{(n)})$ *under* P_n *is a tight sequence and under any limit point* Q *of this sequence, the canonical process* B *is a strictly positive martingale in its usual filtration. Furthermore, the quadratic variation density of* B *under* Q *satisfies,*

$$|a(t:B)| \le c, \quad \mathscr{L}\otimes Q\text{-a.s.}$$

7 Auxiliary lemmas

In this section, we prove several results that are used in the proof of our convergence result. Lemmas 7.2-7.3 are related to the optimal control (3.5). The first result, Lemma 7.1 is related to the properties of a sequence discrete time martingales $M^{(n)}$. Motivated by (5.6) and Lemma 6.4, we assume that these martingales are sufficiently close to the price process $S^{(n)}$. Then, in Lemma 7.1 below, we prove that the process α_n, defined below, converges 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.

This limit theorem is the main tool in the proof of the upper bound of Theorem 3.5.

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

$$\begin{aligned} M^{(n)}(k) &:= E_n(S^{(n)}(n)|\mathscr{F}_k), \\ \alpha_n(k) &:= \frac{\sqrt{n}\xi_k(M^{(n)}(k) - S^{(n)}(k))}{S^{(n)}(k)}. \end{aligned}$$

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

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

Let Q_n be the distribution of $\mathscr{W}_n(S^{(n)})$ under the measure P_n. Then, by Lemma 6.4 this sequence is tight. Without loss of generality we assume that the whole sequence $\{Q_n\}_{n=1}^\infty$ converges to a probability measure Q on $\mathscr{C}[0,1]$. Moreover, under Q the canonical map B is a strictly positive martingale. Then, Lemma 6.4 also implies that the process $A(\cdot : B)$ given in (5.2) is well defined. The next lemma proves the convergence of the process α_n as well.

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

$$\hat{Q}_n \Rightarrow \hat{Q} \text{ on the space } \mathscr{C}[0,1]\times\mathscr{C}[0,1]$$

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

Proof. Hypothesis (7.1) imply that Lemma 6.4 apply to the sequence P_n. Hence under this sequence of measures the estimate (6.3) holds.

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

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

We use this together with (6.3) to arrive at

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

Let $\mathscr{D}[0,1]$ be the space of all *càdlàg* functions equipped with the Skorohod topology (see [1]). Let $\hat{P}_n$ be the distribution on the space $\mathscr{D}[0,1]\times\mathscr{D}[0,1]$, of the piecewise constant process $\{(1/S^{(n)}([nt]), M^{(n)}([nt]))\}_{t=0}^1$ under the measure P_n. We use (7.1) and Lemma 6.4, to conclude that

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

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 $\mathscr{D}[0,1]$ and the measure Q is extended as a probability measure on $\mathscr{D}[0,1]$.

Since the canonical process B is a strictly positive continuous martingale under Q, we apply Theorem 4.3 of [10] and use (7.3), (7.4). The result is the following convergence,

$$\hat{\mathbb{Q}}_n \Rightarrow \hat{\mathbb{Q}} \text{ on thespace } \mathscr{D}[0,1]\times\mathscr{D}[0,1]\times\mathscr{D}[0,1],$$

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

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

$$\left\{ \left(\frac{1}{S^{(n)}([nt])}, M^{(n)}([nt]), Y_n([nt]) \right) \right\}_{t=0}^1 \to \left\{ \left(\frac{1}{M(t)}, M(t), \int_0^t \frac{dM(u)}{M(u)} \right) \right\}_{t=0}^1 \quad \tilde{P}\text{-a.s.}$$

on the space $\mathscr{D}[0,1]\times\mathscr{D}[0,1]\times\mathscr{D}[0,1]$.

Now set $Y(t) = \int_0^t dM(u)/M(u)$ so that $dM = MdY$. Therefore,

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

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

$$\left\{ \int_0^t \alpha_n([nu])du \right\}_{t=0}^1 \to \left\{ \frac{\langle Y\rangle(t) - \sigma^2 t}{2\sigma} \right\}_{t=0}^1 \quad \tilde{P}\text{-a.s. on the space } \mathscr{D}[0,T].$$

From (2.2) and and the definition of α_n, we have

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

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

$$\left| \frac{M^{(n)}(j) - M^{(n)}(j-1)}{S^{(n)}(j-1)} - \frac{1}{\sqrt{n}}\big((\sigma+\alpha_n(j))\xi_j - \alpha_n(j-1)\xi_{j-1} \big) - \frac{\sigma}{2n}(\sigma + 2\alpha_n(j)) \right| \le \frac{c_2}{n^{3/2}}, \text{ a.s.}$$

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

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

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

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

■

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

$$M(0) = s_0 \;\text{ and }\; |a(t:M)| \le c \;\; \mathscr{L}\otimes\tilde{P} \;\text{ a.s.} \tag{7.5}$$

In fact, a volatility process $\alpha \in \mathscr{A}^c$ if and only if the corresponding process S_α satisfies the above condition. However, S_α is defined on the canonical space $(\Omega_W, \mathscr{F}^W, \mathbb{P}^W)$ 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 martingale M's satisfying the constraint (7.5) is the same as maximizing $J(S_\alpha)$ over $\alpha \in \mathscr{A}^c$. The proof follows the ideas of Lemma 5.2 in [14] and uses the randomization technique.

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

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

Proof. Set

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

so that

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

If necessary, by enlarging the space, we may assume that the probability space $(\tilde{\Omega}, \tilde{F}, \tilde{P})$ is rich enough to contain a Brownian motion $\hat{W}(t)$ which is independent of M. For $\lambda \in [0,1]$ define

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

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

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

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

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

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

$$Z(t) := \frac{d\langle Y\rangle}{dt} \ge \varepsilon, \qquad \mathscr{L}\otimes\tilde{P}\text{-a.s.}$$

for some $\varepsilon > 0$. Set,

$$\tilde{W}(t) = \int_0^t \frac{dY(u)}{\sqrt{Z}(u)}, \ t \in [0,1], \tag{7.6}$$

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

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

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

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

in probability with the measure $\mathscr{L} \otimes \tilde{P}$. On the other hand, Ito's isometry and the Doob-Kolmogorov inequality, imply that

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

in probability with respect to $\tilde{P}$. We use these convergence results and the uniform integrability, to conclude that

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

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

$$J(M^{(n)}) \le \sup_{\alpha \in \mathscr{A}^C} J(S_\alpha). \tag{7.7}$$

We prove the above inequality by the randomization technique. Fix $n \in \mathbb{N}$. From the existence of the regular distribution function (for details see [21] page 227), for any $1 \le k < n$ there exists a function $\rho_k : \mathbb{R} \times \mathscr{C}[0,1] \times \mathbb{R}^k \to [0,1]$ such that for any y, $\rho_k(y,\cdot) : \mathscr{C}[0,1] \times \mathbb{R}^k \to [0,1]$ is measurable and satisfies

$$\tilde{E}\left(\kappa_n(k) \le y \middle| \sigma\{\tilde{W}, \kappa_n(0), ..., \kappa_n(k-1)\}\right) = \rho_k(y, \tilde{W}, \kappa_n(0), ..., \kappa_n(k-1)), \ \ \tilde{P} \ \text{ a.s.}$$

Furthermore, $\tilde{P}$ almost surely, $\rho_k(\cdot, \tilde{W}, \kappa_n(0), ..., \kappa_n(k-1))$ is a distribution function on $\mathbb{R}$. Let W be the Brownian motion in our canonical space $(\Omega_W, \mathscr{F}^W, P^W)$. We extend this space so that it contains a sequence $\Xi_1, ..., \Xi_{n-1}$ 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 its complete.

Next, we recursively define the random variables

$$U_0 = \sigma \ \text{ and for } 1 \le k < n \ \ U_k = \sup\{y | \rho_k(y, W, U_1, ..., U_{k-1}) < \Xi_k\}. \tag{7.8}$$

In view of the properties of the functions ρ_i, we can show that $U_1, ..., U_{n-1}$ 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 $1 \le k < n$,

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

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

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

satisfying

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

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

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

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

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

Observe that for any $z\in(0,1)^{n-1}$, the stochastic process $U^{(z)}\in\mathscr{A}^c$. We now use the Fubini's theorem to conclude that

$$J(M^{(n)})=J(S)=\int_{z\in(0,1)^n}J(S_{U^{(z)}})dz_1...dz_n\le\sup_{\alpha\in\mathscr{A}^C}J(S_\alpha) \tag{7.9}$$

and (7.7) follows. ■

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

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

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

Proof. Let $\{\phi_n\}_{n=1}^\infty\subset\mathscr{L}(c)$ be a sequence which converge in probability (with respect to $\mathscr{L}\otimes P^W$) to some $\alpha\in\mathscr{A}^C$. By the Ito's isometry and the Doob-Kolmogorov inequality, we directly conclude that S_{ϕ_n} converges to S_α in probability on the space $\mathscr{C}[0,1]$. Then, invoking the uniform integrability once again, we obtain $\lim_{n\to\infty}J(S_{\phi_n})=J(S_\alpha)$.

Therefore to prove the lemma, for any $\alpha\in\mathscr{A}^c$ we need to construct a sequence $\{\phi_n\}_{n=1}^\infty\subset\mathscr{L}(c)$ which converges in probability to α. Moreover, by the decomposition $\alpha=\alpha^+-\alpha^-$, without loss of generality, we may assume that α is a non negative stochastic process. Thus, let $\alpha\in\mathscr{A}^c$ be a non negative stochastic process and let $\delta>0$. It is well known (see [15]) that there exists a continuous processes ϕ adapted to the Brownian filtration, satisfying

$$\mathscr{L}\otimes P^W\{|\alpha-\phi|>\delta\}<\delta. \tag{7.10}$$

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

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

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

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

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 [1], Chapter 1) that we can find a sequence of bounded Lipschitz continuous functions $\vartheta_n : \mathscr{C}[0,k/m] \to \mathbb{R}$, $n \in \mathbb{N}$ such that $\lim_{n\to\infty} \vartheta_n = \theta_k$ a.s. with respect to the Wiener measure on the space $\mathscr{C}[0,k/m]$. We conclude that there exists a constant $\mathscr{H} > 0$ and a sequence of functions $\Theta_k : \mathscr{C}[0,1] \to \mathbb{R}$, $1 \le k \le m-3$ such that for any $z_1, z_2 \in \mathscr{C}[0,1]$ and $1 \le k \le m-3$

$$
\begin{aligned}
&\text{i.} && \Theta_k(z_1) = \Theta_k(z_2) \text{ if } z_1(s) = z_2(s) \ \text{ for any } s \le k/m, \\
&\text{ii.} && |\Theta_k(z_1)| \le \mathscr{H}, && (7.12)\\
&\text{iii.} && |\Theta_k(z_1) - \Theta_k(z_2)| \le \mathscr{H}\left(||z_1 - z_2||\right), && (7.13)\\
&\text{iv.} && P^W\left\{|\Theta_k(W) - \phi(k/m)| > \delta\right\} < \delta/m. && (7.14)
\end{aligned}
$$

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

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

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,

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

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

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

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

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

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

$$
\begin{aligned}
|\Theta(t) - \alpha(t)| \le &\Big(\max_{0\le k \le m-3} \sup_{k/m \le t \le (k+2)/m} |\phi(t) - \phi(k/m)|\Big) \qquad (7.15)\\
&+ (\max_{0 \le k \le m-3} |\phi(k/m) - \Theta_k(W)|) + |\alpha(t) - \phi(t)|.
\end{aligned}
$$

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

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

Since $\delta > 0$ was arbitrary small we complete the proof. ■

References

[1] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[2] P. Bank and D. Baum, *Hedging and portfolio optimization in financial markets with a large trader*, Math. Finance **14**, 1–18, (2004).

[3] G. Barles and H.M. Soner, *Option pricing with transaction costs and a nonlinear Black-Scholes equation*, Finance and Stochastics, **2**, 369–397, (1998).

[4] P. Chalasani and S. Jha, *Randomized stopping times and American option pricing with transaction costs*, Math. Finance. **11**, 33–77, (2001).

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

[6] U. Çetin, R. Jarrow and P. Protter, *Liquidity risk and arbitrage pricing theory*, Finance and Stoch. **8**, 311–341, (2004).

[7] U. Çetin, H.M. Soner and N. Touzi, *Option hedging under liquidity costs*, Finance and Stochastics. **14**, 317–341, (2010).

[8] Y. Dolinsky, *Hedging of Game Options With the Presence of Transaction Costs*, arXiv:1103.1165, (2011).

[9] R.M. Dudley, *Distances of Probability Measures and Random Variables*, Ann. Math. Statist. **39**, 1563-1572, (1968).

[10] D. Duffie and P. Protter, *From Discrete to Continuous Time Finance: Weak Convergence of the Financial Gain Process*, Mathematical Finance. **2**, 1–15, (1992).

[11] F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Math. Annalen. **300**, 463–520, (1994).

[12] Y. Dolinsky, M. Nutz and H.M. Soner, *Weak Approximation of G-Expectations*, arXiv:1103.0575v1, (2011).

[13] S. Gokay and H.M. Soner, *Liquidity in a Binomial market*, to appear in Mathematical Finance.

[14] S. Kusuoka, *Limit Theorem on Option Replication Cost with Transaction Costs*, Ann. Appl. Probab. **5**, 198–221, (1995).

[15] I Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, New York, (1991).

[16] S. Levental and A.V. Skorohod, *On the possibility of hedging options in the presence of transaction costs*. Ann. Appl. Probab. **7(2)**, 410–443, (1997).

[17] S. Peng. *Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation*, Stochastic Process. Appl., **118(12)**, 2223–2253, (2008).

[18] T. Pennanen and I. Penner, *Hedging of claims with physical delivery under convex transaction costs.* SIAM Journal on Financial Mathematics, **1**, 158–178, (2010).

[19] R.T. Rockafellar, *Convex Analysis*, Princton University Press, Princeton, NJ, (1970).

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

[21] A.N. Shiryaev, *Probability*, Springer-Verlag, New York, (1984).

[22] H.M. Soner, S.E. Shreve and J. Cvitanic, *There is no nontrivial hedging portfolio for option pricing with transaction costs*, Annals of Applied Prob., **5/2**, 327355, (1995).

[23] H.M. Soner and N. Touzi, *The dynamic programming equation for second order stochastic target problems*, SIAM Journal on Control and Optimization, **48/4**, 2344–2365, (2009).

[24] H.M. Soner, N. Touzi and J. Zhang *Dual Formulation of Second Order Target Problems*, preprint, 2010.