A model-free no-arbitrage price bound for variance options

J. Frédéric BONNANS * Xiaolu TAN †

February 12, 2013

Abstract

We suggest a numerical approximation for an optimization problem, motivated by its applications in finance to find the model-free no-arbitrage bound of variance options given the marginal distributions of the underlying asset. A first approximation restricts the computation to a bounded domain. Then we propose a gradient projection algorithm together with the finite difference scheme to solve the optimization problem. We prove the general convergence, and derive some convergence rate estimates. Finally, we give some numerical examples to test the efficiency of the algorithm.

Key words. Variance option, model-free price bound, gradient projection algorithm.

AMS 2000 subject classifications. 60G40, 60H10.

1 Introduction

In financial mathematics, an underlying security is usually modeled as a one-dimensional continuous process, and a derivative option is defined by a payoff function on the underlying security's path. In a model where the underlying security is a continuous martingale, by the the fundamental theorem of pricing, the expectation value of the derivative option is a no-arbitrage price of the option. In the absence of other information on the underlying security, the no-arbitrage bound is then the supremum (or infimum) of the expectation value of the derivative option among all models where the security is a continuous martingale. An equivalent way is to model the underlying as canonical process in the canonical space, and then consider the collection of all martingale measures (i.e. the probability measures under which the canonical process is a

*INRIA-Saclay and CMAP, Ecole Polytechnique, Paris, frederic.bonnans@inria.fr.

†CMAP, Ecole Polytechnique, Paris, xiaolu.tan@polytechnique.edu. Research supported by the Chair *Financial Risks* of the *Risk Foundation* sponsored by Société Générale, the Chair *Derivatives of the Future* sponsored by the Fédération Bancaire Française, and the Chair *Finance and Sustainable Development* sponsored by EDF and CA-CIB.

martingale). Now, if we consider a market where the vanilla options (European call or put) are liquid, which can be used to hedge the path-dependent exotic options, then the no-arbitrage bound of the latter will change. A framework for this new bound was proposed recently by Galichon, Henry-Labordère and Touzi [12].

More precisely, let $\Omega^0 := C([0,T],\mathbb{R})$ be the canonical space with canonical process X ($X_t(\omega) := \omega_t$, $\forall \omega \in \Omega^0$) and canonical filtration $\mathbb{F}^0 = (\mathcal{F}_t^0)_{0\le t\le T}$, S_0 a constant. We denote by $\mathcal{P}(\delta_{S_0})$ the collection of all probability measures $\mathbb{P}$ on $(\Omega^0, \mathcal{F}_T^0)$ under which X is a $\mathbb{F}^0-$martingale and $X_0 = S_0$, $\mathbb{P}-$a.s. The canonical process X is a candidate of underlying security price process. Let the derivative options G be defined by $G(X_t, 0 \le t \le T) \in \mathcal{F}_T^0$. Then an upper bound of model-free no-arbitrage price of G is

$$\sup_{\mathbb{P}\in\mathcal{P}(\delta_{S_0})} \mathbb{E}^{\mathbb{P}}\big[G\big]. \tag{1.1}$$

Suppose in addition that in the financial market, the vanilla options of maturity T and of all strike are liquid, so that the investor can identify the marginal distribution μ of X_T. In other words, let $\phi \in \mathbb{L}^1(\mu)$, the $T-$maturity European option with payoff $\phi(X_T)$ has a unique no-arbitrage price $\mu(\phi) := \int_{\mathbb{R}^d} \phi(x)\mu(dx)$. We may use the vanilla option portfolio to hedge G. By buying a portfolio $\phi(X_T)$, we spend $\mu(\phi)$ and so the payoff at maturity T becomes $G - \phi(X_T)$. Therefore, we get a new upper bound for G given by $\sup_{\mathbb{P}\in\mathcal{P}(\delta_{S_0})} \mathbb{E}^{\mathbb{P}}\big[G - \phi(X_T)\big] + \mu(\phi)$. Following [12], by minimizing on the vanilla option portfolio ϕ, a new no-arbitrage upper bound of the option G is then given by

$$\inf_{\phi\in\mathbb{L}^1(\mu)} \sup_{\mathbb{P}\in\mathcal{P}(\delta_{S_0})} \Big\{\mathbb{E}^{\mathbb{P}}\big[G - \phi(X_T)\big] + \mu(\phi)\Big\}. \tag{1.2}$$

The above bound is stronger than the one in (1.1), since $\phi := 0 \in \mathbb{L}^1(\mu)$. In [12], the authors gave a duality result, showing that the above no-arbitrage upper bound (1.2) is equivalent to the minimum super-hedging cost for derivative option G. Moreover, they derived an explicit solution for this no-arbitrage bound when G is a lookback option.

For another specific class of payoff function G, Tan and Touzi [24] established another duality result, where the dual formulation can be viewed as a stochastic mass transportation problem. Namely, by exchanging the infimum and supremum, and observing that

$$\inf_{\phi\in\mathbb{L}^1(\mu)} \Big\{\mathbb{E}^{\mathbb{P}}\big[G - \phi(X_T)\big] + \mu(\phi)\Big\} \quad = \quad \begin{cases} \mathbb{E}^{\mathbb{P}}[G], & \text{if } X_T \sim^{\mathbb{P}} \mu, \\ -\infty, & \text{otherwise,} \end{cases}$$

it follows that a dual formulation of (1.2) is given by

$$\sup_{\mathbb{P}\in\mathcal{P}(\delta_{S_0})} \inf_{\phi\in\mathbb{L}^1(\mu)} \Big\{\mathbb{E}^{\mathbb{P}}\big[G - \phi(X_T)\big] + \mu(\phi)\Big\} \quad = \quad \sup_{\mathbb{P}\in\mathcal{P}(\delta_{S_0},\mu)} \mathbb{E}^{\mathbb{P}}\big[G\big], \tag{1.3}$$

where $\mathcal{P}(\delta_{S_0}, \mu)$ denotes the collection of all martingale probability measures $\mathbb{P} \in \mathcal{P}(\delta_{S_0})$ such that $X_T \sim^{\mathbb{P}} \mu$. Then under $\mathbb{P} \in \mathcal{P}(\delta_{S_0}, \mu)$, the martingale X can be viewed as a transportation plan from the distribution δ_{S_0} to the distribution μ.

The no-arbitrage bound problem of exotic options given marginals has been largely studied by the Skorohold Embedding Problem (SEP) approach. Given a Brownian motion W and a distribution μ, the SEP is to find a stopping time τ such that $W_\tau \sim \mu$ and $(W_{t\wedge\tau})_{t\geq 0}$ is uniformly integrable. By Dubins-Schwartz's time change theorem, a martingale can be represented as a time changed Brownian motion and T turns to be a stopping time w.r.t. the time-changed filtration. Then the problem (1.3) can be formally written as

$$\sup_{\tau\in\mathcal{T}(\mu)} \mathbb{E}\big[\tilde{G}(W_t, 0\leq t\leq \tau)\big], \tag{1.4}$$

for another corresponding payoff function $\tilde{G}$, where $\mathcal{T}(\mu)$ denotes the collection of all stopping times τ such that $W_\tau \sim \mu$. The connection between the SEP and no-arbitrage bound of exotic options was first observed in the seminal paper of Hobson [13]. Further, several solutions of SEP have been proved to have the optimality property, so that they induce optimal bounds for some exotic options. We also refer to Hobson [14] for a review of the SEP with applications in finance. However, the SEP approach to find the no-arbitrage bound is generally studied case by case.

In this paper, we are interested in particular in the no-arbitrage bound problem for a general class of variance options, whose payoff are given by $G = g(\langle X\rangle_T, X_T)$ for some function $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$, where $\langle X\rangle_T$ denotes the quadratic variation of X between 0 and T. As observed in Soner, Touzi and Zhang [21], we remark that there is a progressively measurable process $(\langle X\rangle_t)_{0\leq t\leq T}$ which is pathwisely defined on Ω^0 and coincides with the $\mathbb{P}$-quadratic variation of X, $\mathbb{P}$-a.s. for every $\mathbb{P}\in\mathcal{P}(\delta_{S_0})$. Then the no-arbitrage bound (1.2) turns to be

$$\inf_{\phi\in\mathbb{L}^1(\mu)} \sup_{\mathbb{P}\in\mathcal{P}(\delta_{S_0})} \Big\{\mathbb{E}^{\mathbb{P}}\big[g(\langle X\rangle_T, X_T) - \phi(X_T)\big] + \mu(\phi)\Big\}. \tag{1.5}$$

The main contribution of this paper is to provide a complete approximation for an optimization problem similar to (1.5), motivated by its application in finance to find the no-arbitrage bound. In a first part, we will convert the supremum part of (1.5) from an optimal control problem to an optimal stopping problem, using Dubins-Schwartz's time change theorem. Next, we explore the properties of the optimal stopping problem and restrict the computation to a bounded domain. Finally, we suggest a finite difference method to solve the optimal stopping problem as well as a gradient projection algorithm to solve the infimum part, where the gradient is also computed by a finite difference scheme.

The rest of the paper is organized as follows: In Section 2, we formulate a no-arbitrage upper bound for a general class of variance options with given marginals, which is the main problem of the paper. We then reformulate the problem as an optimal Brownian stopping problem and give an approximation optimization problem. In Section 3, we propose a numerical algorithm for the approximation problem, which combines the gradient projection algorithm and the finite difference method. We next provide some numerical examples in Section 4. Finally, we complete in Section 5 the proof of the convergence for the approximation problem.

Notations: We denote by Q the collection of all continuous functions of quadratic growth, i.e.

$$\mathrm{Q} \;:=\; \Big\{ \phi : \mathbb{R} \to \mathbb{R} \text{ be continuous and such that } \sup_{x\in\mathbb{R}} \frac{|\phi(x)|}{1+|x|^2} < \infty \Big\}. \tag{1.6}$$

Let μ be a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, denote

$$\mu(\phi) \;:=\; \int_{\mathbb{R}} \phi(x)\mu(dx), \quad \text{for every } \phi \in \mathbb{L}^1(\mu).$$

2 Main problem

We shall first formulate a no-arbitrage bound for a class of variance options, given two marginal distributions of the underlying at time T_0 and T_1. This provides the motivation of the main problem of the paper, which is then reformulated in Section 2.2 using the time change theorem. In Section 2.3, we give a first approximation which restrict the computation on a bounded domain.

2.1 A no-arbitrage bound for a class of variance options

We recall, as defined in the introduction section, that $\Omega^0 := C([0,T],\mathbb{R})$ is the canonical space with canonical filtration $\mathbb{F}^0$ and canonical process X. A progressive process $\langle X\rangle$ is defined on Ω^0 which coincides with the quadratic variation of X under every martingale measure $\mathbb{P}$. Then X is the candidate process of the underlying security. Let $0 \le T_0 \le T_1 \le T$, denote $\langle X\rangle_{T_0,T_1} := \langle X\rangle_{T_1} - \langle X\rangle_{T_0}$, we shall consider in this paper the forward variance option with payoff

$$G \;:=\; g\big(\langle X\rangle_{T_0,T_1}, X_{T_1}\big) \text{ at maturity } T_1 \text{ for a Lipschitz function } g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}.$$

Example 2.1. *The most popular variance option is the "variance swap", whose payoff function is $g(t,x) = t$. There are also "volatility swap" with payoff function $g(t,x) = \sqrt{t}$, and calls (puts) on variance or volatility, where the payoff functions are $(t-K)^+$ ($(K-t)^+$) or $(\sqrt{t}-K)^+$ ($(K-\sqrt{t})^+$). Another example is the "call sharpe" option with payoff function $g(t,x) = \frac{(x-K)^+}{t}$.*

We shall suppose that the vanilla options of maturities T_0, T_1 are liquid so that we can identify the marginal distributions μ_0 (resp. μ_1) for X_{T_0} (resp. X_{T_1}). Suppose in addition that the underlying process is a martingale, and the quadratic variation $\langle X\rangle_{T_0,T_1}$ conditioning on X_{T_0} is integrable. Equivalently, we shall only consider, for every $x \in \mathbb{R}$, the collection $\mathcal{P}^2(\delta_x)$ of probability measures $\mathbb{P}$ such that $\mathbb{P}(X_{T_0} = x) = 1$ and $\mathbb{E}^{\mathbb{P}}\big[\langle X\rangle_{T_0,T_1}\big] < \infty$. Given a function $\phi : \mathbb{R} \to \mathbb{R}$, define

$$\bar{\lambda}_0^{\phi}(x) \;:=\; \sup_{\mathbb{P}\in\mathcal{P}^2(\delta_x)} \mathbb{E}^{\mathbb{P}}\big[g(\langle X\rangle_{T_0,T_1}, X_{T_1}) - \phi(X_{T_1})\big].$$

Then following Galichon, Henry-Labordère and Touzi [12], we define a no-arbitrage upper bound of variance option $G = g(\langle X\rangle_{T_0,T_1}, X_{T_1})$ by

$$\overline{U} \;:=\; \inf_{\phi\in\mathrm{Q}} \big(\mu_0(\bar{\lambda}_0^{\phi}) + \mu_1(\phi)\big), \tag{2.1}$$

where Q is the collection of all continuous functions of quadratic growth defined by (1.6).

Remark 2.2. *In contrast to the static strategies set* $\mathbb{L}^1(\mu)$ *used in formulation* (1.5)*, we restrict our admissible strategies set to* Q*. Then* $\overline{U}$ *may not be the optimal no-arbitrage bound, but only a sub-optimal bound. The main reason to choose* Q *is, by the observation of Dupire [10], that variance swap (i.e. g(t,x) = t) is equivalent to a European option option with payoff* X_T^2*, see also Proposition 2.7 and Remark 2.9.*

Remark 2.3. *Here we only give the upper bound formulation. By the symmetry of the set* Q *defined in* (1.6)*, if we reverse the payoff function to* $-g(t,x)$*, then with the upper bound* $\overline{U}(-g)$ *associated to payoff* $-g$*, the value* $-\overline{U}(-g)$ *is the lower bound for the payoff* g*.*

The problem of no-arbitrage bound for variance options given marginal distributions has also been studied by the SEP approach in a dual formulation of the form (1.4). Suppose that $g(t,x) = f(t)$ for some function $f : \mathbb{R}^+ \to \mathbb{R}$, it is proved that the optimal upper bound can be induced by Root's embedding when f is concave and by Röst's embedding when f is convex (see Root [19] and Röst [20] and also Cox and Wang [9]). However, for general payoff functions $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$, there is no systematic approach to find the optimal no-arbitrage bound. When $g(t,x) = (t-K)^+$, i.e. the option is the variance call, Dupire [10], Carr and Lee [8] proposed a method to find a sub-optimal bound as well as the associated strategy ϕ in a similar context. In their implemented examples, they showed that their bounds are quite closed to the optimal bounds induced by Root's embedding solution. In our paper, we shall consider a general payoff function and provide a complete approximation of the bound $\overline{U}$ in (2.1) as well as the optimal static strategy ϕ.

2.2 A reformulation by optimal stopping problem

We would like reformulate the upper bound problem (2.1) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, equipped with a standard one-dimensional Brownian motion $B = (B_t)_{t\geq 0}$ such that $B_0 = 0$. Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the natural Brownian filtration, we define a set of $\mathbb{F}-$stopping times by

$$\mathcal{T}^\infty \quad := \quad \big\{\mathbb{F} - \text{stopping times } \tau \text{ such that } \mathbb{E}(\tau) < \infty\big\}. \tag{2.2}$$

Then for every function $\phi \in$ Q, we denote

$$g^\phi(t,x) \quad := \quad g(t,x) \; - \; \phi(x), \tag{2.3}$$

and define $\lambda^\phi : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ and $\lambda_0^\phi : \mathbb{R} \to \mathbb{R}$ by

$$\lambda^\phi(t,x) \quad := \quad \sup_{\tau\in\mathcal{T}^\infty} \; \mathbb{E}\left[\, g^\phi(t+\tau, x+B_\tau)\, \right] \;\; \text{and} \;\; \lambda_0^\phi(\cdot) \; := \; \lambda^\phi(0,\cdot), \tag{2.4}$$

and an optimization problem by

$$U \quad := \quad \inf_{\phi\in \mathrm{Q}} u(\phi), \quad \text{with} \quad u(\phi) \quad := \quad \mu_0(\lambda_0^\phi) \; + \; \mu_1(\phi). \tag{2.5}$$

Applying the time-change martingale theorem, it follows that U is an equivalent reformulation of the no-arbitrage bound $\overline{U}$ in (2.1).

Theorem 2.4. *For every $\phi \in$ Q, we have $\bar\lambda_0^\phi(x) = \lambda_0^\phi(x)$. And therefore, $\overline{U} = U$.*

Proof. Let us fix an arbitrary $\phi \in$ Q and $x \in \mathbb{R}$. First, given a stopping time $\tau \in \mathcal{T}^\infty$, we define a process Y by $Y_t := x + B_{\tau \wedge \frac{t-T_0}{T_1-t}}$ if $t \in [T_0, T_1]$, $Y_t := x$ if $t \in [0, T_0)$ and $Y_t := Y_{T_1}$ if $t \in [T_1, T]$. Then clearly Y is a continuous martingale between 0 and T such that $\langle Y \rangle_{T_0,T_1} = \tau$, and Y induces a probability measure in $\mathcal{P}^2(\delta_x)$. This implies that $\lambda_0^\phi(x) \le \bar\lambda_0^\phi(x)$.

Next, suppose that $\overline{\mathbb{P}} \in \mathcal{P}^2(\delta_x)$, then the canonical process X is a continuous martingale under $\overline{\mathbb{P}}$. It follows by the time-change martingale theorem (see e.g. Theorem 3.4.6 of Karatzas and Shreve [15]) that $X_t = x + W_{\langle X \rangle_t}$, where W is a standard Brownian motion and $\langle X \rangle_t$ is a stopping time w.r.t. the time-changed filtration. It is well-known that the supremum on stopping times w.r.t. the time-changed filtration is equivalent to the supremum on the stopping times w.r.t. the natural Brownian filtration (see Lemma 5.4). It follows that $\bar\lambda_0^\phi(x) \le \lambda_0^\phi(x)$, and we hence conclude the proof. □

To make the upper bound problem be wellposed, we now impose some assumptions on the marginal distributions μ_0 and μ_1.

Assumption 2.5. *The marginal distributions μ_0 and μ_1 have both finite second moment, i.e. $\mu_0(\phi_0) + \mu_1(\phi_0) < \infty$ with $\phi_0(x) := x^2$; and $\mu_0 \le \mu_1$ in the convex order, i.e.*

$$\mu_0(\phi) \le \mu_1(\phi), \quad \text{for every convex function } \phi \in \mathbb{L}^1(\mu_0) \cap \mathbb{L}^1(\mu_1). \tag{2.6}$$

Remark 2.6. (i) *It is shown in Strassen [22] that the convex order inequality* (2.6) *is a necessary and sufficient condition for the existence of a martingale X with marginal distributions μ_0 and μ_1 at time T_0 and T_1 such that $T_0 < T_1$.*
(ii) *Since the identity function I (where $I(x) := x$) and its opposite $-I$ are both convex, it follows from* (2.6) *that μ_0 and μ_1 have the same first moment, i.e. $\mu_0(I) = \mu_1(I)$.*

Proposition 2.7. *Let Assumption 2.5 hold true.*
(i) *Suppose that $\psi \in$ Q, $K \in \mathbb{R}$ and g is a Lipschitz payoff function. We define another payoff function $g_{K,\psi}$ by $g_{K,\psi}(t,x) := g(t,x) + Kt + \psi(x)$. Denote by $U(g)$ (resp. $U(g_{K,\psi})$) the no-arbitrage price upper bound defined in* (2.5) *associated with the payoff function g (resp. $g_{K,\psi}$). Then*

$$U(g_{K,\psi}) \;=\; U(g) \;+\; KC_0 \;+\; \mu_1(\psi), \tag{2.7}$$

where

$$C_0 \;:=\; \mu_1(\phi_0) - \mu_0(\phi_0), \quad \text{with} \quad \phi_0(x) \;:=\; x^2. \tag{2.8}$$

(ii) *For every Lipschitz payoff function g, the upper bound value is finite, i.e. $|U| < \infty$.*

The proof will be provided in Section 5.2 after some technical lemmas. We notice that, as a direct consequence, and the upper bound of a European option with payoff function $\psi(x)$ is given by $\mu_1(\psi)$, and the bound of "variance swap" option ($g(t,x) = t$) is C_0, which is consistent to Dupire's [10] observation that *variance swap* is equivalent to a European option with payoff function $g(x) = x^2$.

2.3 An approximation problem

The main purpose of the paper is to provide a complete approximation to the optimization problem (2.5), which needs further a numerical approximation. For this purpose, we shall introduce an approximation problem which restrict the computation to a bounded domain. Let us make some further assumptions on the payoff function.

Assumption 2.8. (i) *The payoff function $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is L_0-Lipschitz with constant $L_0 \in \mathbb{R}^+$.*
(ii) *For every $x \in \mathbb{R}$, $t \mapsto g(t,x)$ is increasing; and for every $t \in \mathbb{R}^+$, $x \mapsto g(t,x)$ is convex with minimum at $x = 0$.*
(iii) *There is some constant $M_0 \in \mathbb{R}^+$ such that for every $t \in \mathbb{R}$, $x \mapsto g(t,x)$ is affine on $[M_0, \infty)$ and $(-\infty, -M_0]$.*

Remark 2.9. *Let g be an arbitrary Lipschitz payoff function. By Proposition 2.7, it is equivalent to consider the transformed function $g_{K,\psi}(t,x) := g(t,x) + Kt + \psi(x)$ for some constant $K \in \mathbb{R}$ and function ψ, for which Assumption 2.8 may hold true.*

Example 2.10. *Let $g(t,x) = (K-t)^+$, we can consider $\hat{g}(t,x) := t + (K-t)^+$, which satisfies Assumption 2.8. When $g(t,x) = (K - \sqrt{t})^+$ which is not Lipschitz, we need to truncate the payoff function to $g_\varepsilon(t,x) = (K - \sqrt{t} \vee \varepsilon)^+$ to obtain the Lipschitz payoff function. For the "call sharpe" option $g(t,x) = \frac{(x-K)^+}{t}$, we also need truncate the denominator and consider $\hat{g}(t,x) = \frac{(x-K)^+}{t \vee \varepsilon}$.*

Let K, M, R and T are all positive constants, we denote

$$\mathrm{Q}_0 \;:=\; \Big\{ \phi \in \mathrm{Q} \text{ non negative, convex, such that } \phi(0) = 0, \\ \phi(x) \le K(|x| \vee x^2) \text{ and } \phi(x) = Kx^2 \text{ for } |x| \ge 2M \Big\} \tag{2.9}$$

and $\lambda_0^{\phi,0}(x) = \lambda^{\phi,0}(0,x)$ with

$$\lambda^{\phi,0}(t,x) \;:=\; \sup_{\tau \in \mathcal{T}^\infty,\ \tau \le \tau_x^R \wedge (T-t)} \mathbb{E}\big[g^\phi(t+\tau, x+B_\tau)\big], \tag{2.10}$$

where $\tau_x^R := \inf\{s\ : x + B_s \notin (-R,R)\}$. Clearly, $\tau_x^R = 0$ and $\lambda^{\phi,0}(t,x) = g(t,x) - \phi(x)$ whenever $|x| \ge R$. We then introduce the approximation optimization value U_0 by

$$U_0 \;:=\; \inf_{\phi \in \mathrm{Q}_0} u_0(\phi) \quad \text{with} \quad u_0(\phi) \;:=\; \mu_0\big(\lambda_0^{\phi,0}\big) \;+\; \mu_1(\phi). \tag{2.11}$$

For good choices of constants K, M, R and T, U_0 provides an approximation of the upper bound value U. Denote

$$\phi_{K,M}(x) \;:=\; 4KM\big(|x| - M\big)\mathbf{1}_{M \le |x| \le 2M} \;+\; Kx^2\mathbf{1}_{|x|>2M} \tag{2.12}$$

and

$$\delta \ := \ -\log(q(R)) > 0, \quad \text{where} \quad q(R) \ := \ \frac{1}{\sqrt{2\pi}} \int_{-2R}^{2R} e^{-x^2/2}\, dx.$$

Clearly, for every fixed $K > 0$, $\mu_1(\phi_{K,M}) \to 0$ as $M \to \infty$ since μ_1 has finite second order moment by Assumption 2.5.

Theorem 2.11. *Suppose that Assumptions 2.5 and 2.8 hold true. Let K, M, R and T turn to ∞ in such a way that $\mu_1(\phi_{K,M}) \to 0$, $R \geq \big(1 + \sqrt{\frac{K}{K-L_0}}\big)M$ and $2(K + 2L_0)(R^2 \vee 1)e^{-\delta(T-1)} \to 0$. Then $U_0 \to U$.*

The proof of the above theorem will be completed in Section 5.

The computation of U_0 is now restricted to a bounded domain, which permits further a numerical approximation in the next section. We finish this section by characterizing $\lambda^{\phi,0}$ as the unique viscosity solution of a variational inequality (see e.g. Theorem 6.7 of Touzi [25]).

Proposition 2.12. *The function $\lambda^{\phi,0}$ defined in* (2.10) *is the unique viscosity solution of variational inequality*

$$\min\Big(\ \lambda - g^{\phi},\ \ -\frac{1}{2}\frac{\partial^2 \lambda}{\partial x^2} - \frac{\partial \lambda}{\partial t}\ \Big)(t,x) \ = \ 0, \qquad \text{on} \ \ [0,T) \times (-R,R), \tag{2.13}$$

with boundary condition

$$\lambda(t,x) \ = \ g^{\phi}(t,x), \qquad \text{on} \ \ \big([0,T] \times \{\pm R\}\big) \ \cup \ \big(\{T\} \times [-R,R]\big).$$

3 Numerical approximation

We shall propose a numerical method to approximate U_0 in (2.11). It is easy to observe that $\phi \mapsto \lambda^{\phi,0}$ is convex since it is represented as the supremum of a family of linear mapping in (2.10). Thus $\phi \mapsto u_0(\phi)$ is a convex function and the problem of computing U_0 turns out to be a minimization problem of a convex function.

Our main idea is to compute $\lambda^{\phi,0}$ with a finite difference numerical scheme, and then to solve the minimization problem (2.11) with an iterative algorithm. Concretely, we shall first provide a discrete grid characterized by $h = (\Delta t, \Delta x)$, on which there is a discrete optimization problem with value U_h close to U_0. Then we use the gradient projection algorithm to solve the discrete optimization problem of U_h. Throughout the section, we fix the constants K, M, R and T, and suppose that $R \geq 2M$.

3.1 A finite difference approximation

Let $(l,m) \in \mathbb{N}^2$, $h = (\Delta x, \Delta t) \in (\mathbb{R}^+)^2$ such that $l\Delta t = T$ and $m\Delta x = M$. Without loss of generality, we suppose that there is $r \in \mathbb{N}$ such that $r\Delta x = R$. Denote $x_i := i\Delta x$ and $t_k := k\Delta t$ and we define the discrete grid:

$$\mathcal{N} \ := \ \{x_i \ : \ i \in \mathbb{Z}\}, \qquad \mathcal{N}_R \ := \ \mathcal{N} \cap [-R,R],$$

$$\mathcal{M} \;:=\; \big\{(t_k,x_i)\;:\;(k,i)\in\mathbb{Z}^+\times\mathbb{Z}\big\}\cap\big([0,T]\times[-R,R]\big),$$

The terminal set, boundary set as well as interior set of $\mathcal{M}$ are denoted by

$$\partial_T\mathcal{M} \;:=\; \big\{\;(T,x_i)\;:-r\le i\le r\;\big\},\qquad \partial_R\mathcal{M}\;:=\;\big\{\;(t_k,\pm R)\;:\;0\le k\le l\;\big\},$$
$$\mathring{\mathcal{M}} \;:=\; \mathcal{M}\setminus\big(\partial_R\mathcal{M}\cup\partial_T\mathcal{M}\big).$$

Let $w(t,x)$ be a function defined on $\mathcal{M}$, denote $w_i^k := w(t_k,x_i)$, we introduce the discrete derivative of w by

$$D^2 w(t_k,x_i) \;:=\; \frac{w_{i+1}^k - 2w_i^k + w_{i-1}^k}{\Delta x^2}.$$

Let $\theta\in[0,1]$, φ be a function defined on $\mathcal{N}_R$, denote

$$g^{\varphi}(t_k,x_i) \;:=\; g(t_k,x_i) \;-\; \varphi(x_i). \tag{3.1}$$

Following Barles, Daher and Romano [1], we define λ_h^{φ} as the solution of the finite difference scheme of variational inequality (2.13) on $\mathcal{M}$:

$$\begin{cases} \lambda_h(t_{k+1},x_i) \;-\; \tilde{\lambda}_h(t_k,x_i) \\ \qquad +\;\frac{1}{2}\Delta t\,\Big(\theta\; D^2\tilde{\lambda}_h(t_k,x_i) \;+\; (1-\theta)\;D^2\lambda_h(t_{k+1},x_i)\Big) \;=\; 0, \\ \lambda_h(t_k,x_i) \;=\; \max\Big(\;g^{\varphi}(t_k,x_i)\;,\;\tilde{\lambda}_h(t_k,x_i)\;\Big), \qquad (t_k,x_i)\in\mathring{\mathcal{M}}, \\ \lambda_h(t_k,x_i) \;=\; g^{\varphi}(t_k,x_i), \qquad\qquad (t_k,x_i)\in\partial_T\mathcal{M}\cup\partial_R\mathcal{M}. \end{cases} \tag{3.2}$$

The above $\theta-$scheme has a unique solution, and is a consistent approximation for (2.13) in sense of Barles and Souganidis [3]. Indeed, since the second equation of (3.2) is equivalent to $\min(\;\lambda_h - g^{\varphi},\;\;\frac{\lambda_h-\tilde{\lambda}_h}{\Delta t}\;\;)(t_k,x_i) \;=\; 0$, it follows with the first equation in (3.2) that

$$\begin{aligned} \min\Big(\;\lambda_h - g^{\varphi},\;\; & \frac{\lambda_h(t_k,x_i)-\lambda_h(t_{k+1},x_i)}{\Delta t} \\ +\;\tfrac{1}{2}\Big(\theta\;D^2\tilde{\lambda}_h(t_k,x_i) \;+\; & (1-\theta)\;D^2\lambda_h(t_{k+1},x_i)\Big)\;\Big) \;=\; 0. \end{aligned}$$

We shall assume in addition that the discretization parameters $h=(\Delta t,\Delta x)$ satisfy the CFL condition

$$(1-\theta)\,\frac{\Delta t}{\Delta x^2} \;\le\; 1. \tag{3.3}$$

Then the finite difference scheme (3.2) is consistent and monotone in sense of the monotone convergence scheme in [3], and the numerical solution λ_h^{φ} converges to $\lambda^{\phi,0}$ given $\varphi := \phi|_{\mathcal{N}}$ by the results of [3] (see also Barles, Daher and Romano [1]).

Remark 3.1. *The discrete system* (3.2) *is the θ-scheme for variational inequality* (2.13) *with Dirichlet boundary condition $g(x,t)-\varphi(x)$ on $\partial_T\mathcal{M}\cup\partial_R\mathcal{M}$. It is well-known that when the finite difference scheme is explicit (i.e. $\theta=0$) and the CFL condition $\frac{\Delta t}{\Delta x^2}\le 1$ holds, it can be interpreted as the dynamic programming principle for a system on a Markov chain Λ (see e.g. Kushner and Dupuis [17]). This interpretation holds also true for general θ-scheme under the monotone condition, as we shall show later in the proof of Proposition 3.5.*

We next introduce a natural approximation of $u_0(\phi)$ in (2.11):

$$u_h(\varphi) \quad := \quad \mu_0\big(\mathrm{lin}^R[\lambda^{\varphi}_{h,0}]\big) \; + \; \mu_1\big(\mathrm{lin}^R[\varphi]\big), \tag{3.4}$$

where $\lambda^{\varphi}_{h,0}(\cdot) := \lambda^{\varphi}_h(0,\cdot)$, and $\mathrm{lin}^R[\varphi]$ denotes the linear interpolation of φ extended by zero outside $[-R,R]$ for every function φ defined on $\mathcal{N}_R$.

Assumption 3.2. *There are constants* $(\rho_1, \rho_2, L_{K,M,T}) \in (\mathbb{R}^+)^3$ *which are independent of* $h = (\Delta t, \Delta x)$ *such that*

$$\mu_0\Big(\, \Big|\lambda^{\phi,0}_0 \mathbf{1}_{[-R,R]} - \mathrm{lin}^R[\lambda^{\varphi}_{h,0}]\Big| \, \Big) \quad \le \quad L_{K,M,T} \, \big(\Delta x^{\rho_1} + \Delta t^{\rho_2}\big), \tag{3.5}$$

$$\textit{for every } \phi \in \mathrm{Q}_0 \textit{ and } \varphi = \phi|_{\mathcal{N}_R}.$$

Remark 3.3. (i) *When* $\theta = 1$, (3.2) *is the implicit scheme for* (2.13)*, then Assumption 3.2 holds true with* $\rho_1 = \frac{1}{2}$ *and* $\rho_2 = \frac{1}{4}$ *in spirit of the analysis of Krylov [16].*
(ii) *When* $\theta = 0$ *and the CFL condition* (3.3) *holds,* (3.2) *is a monotone explicit scheme, then in spirit of Barles and Jakobsen [2], Assumption 3.2 holds with* $\rho_1 = \frac{1}{10}$ *and* $\rho_2 = \frac{1}{5}$.

Let Q_h be the collection of all functions on the grid $\mathcal{N}_R$ defined as restrictions of functions in Q_0 given by (2.9), i.e.

$$\mathrm{Q}_h \quad := \quad \big\{ \; \varphi := \phi|_{\mathcal{N}_R} \text{ for some } \phi \in \mathrm{Q}_0 \; \big\}, \tag{3.6}$$

we can then provide a discrete approximation for U_0 in (2.11):

$$U_h \quad := \quad \inf_{\varphi \in \mathrm{Q}_h} \; u_h(\varphi). \tag{3.7}$$

We notice further that

$$\mathrm{Q}_h = \Big\{\varphi \in B(\mathcal{N}_R) \text{ nonnegative, convex satisfying } \varphi(0) = 0, \; \varphi(x_i) = K x_i^2,$$
$$\text{for all } 2m \le |i| \le r, \text{ and } |\varphi(x_{i+1}) - \varphi(x_i)| \le 4KM\Delta x, \;\; -2m < i \le 2m\Big\}, \tag{3.8}$$

where $B(\mathcal{N}_R)$ denote the set of all functions defined on the grid $\mathcal{N}_R$.

Proposition 3.4. *Let Assumption 3.2 hold. Then we have, with the same constants* $L_{K,M,T}$*,* ρ_1*,* ρ_2 *introduced in Assumption 3.2,*

$$|U_0 - U_h| \;\; \le \;\; L_{K,M,T}\big(\Delta x^{\rho_1} + \Delta t^{\rho_2}\big) + 4KR\Delta x + (\mu_0 + \mu_1)(\phi^R_K), \tag{3.9}$$

where $\phi^R_K(x) := K x^2 \mathbf{1}_{|x|>R}$.

Proof. First, given $\phi \in \mathrm{Q}_0$ which is $4KR-$Lipschitz, we introduce $\varphi := \phi|_{\mathcal{N}_R} \in \mathrm{Q}_h$ so that $\big|\mathrm{lin}^R[\varphi] - \phi\big|_{L^\infty([-R,R])} \le 4KR\Delta x$. Then it follows by Assumption 3.2 that $|u_0(\phi) - u_h(\varphi)| \; \le L_{K,M,T}(\Delta x^{\rho_1} + \Delta t^{\rho_2}) + 4KR\Delta x + \; (\mu_0 + \mu_1)(\phi^R_K)$, and hence

$$U_0 - U_h \quad \le \quad L_{K,M,T} \, \big(\Delta x^{\rho_1} + \Delta t^{\rho_2}\big) \; + \; 4KR\Delta x + \; (\mu_0 + \mu_1)(\phi^R_K).$$

Next, given $\varphi \in \mathrm{Q}_h$, we define $\phi \; := \; \mathrm{lin}^R[\varphi] \, + \, \phi^R_K \, \in \mathrm{Q}_0$. It follows by Assumption 3.2 that $|u_0(\phi) - u_h(\varphi)| \; \le L_{K,M,T}(\Delta x^{\rho_1} + \Delta t^{\rho_2}) + \; (\mu_0 + \mu_1)(\phi^R_K)$, and therefore,

$$U_h - U_0 \quad \le \quad L_{K,M,T} \, \big(\Delta x^{\rho_1} + \Delta t^{\rho_2}\big) + \; (\mu_0 + \mu_1)(\phi^R_K).$$

We hence conclude the proof. □

3.2 Gradient projection algorithm

As we can easily observe from its definition in (2.5) that $\phi \mapsto u_0(\phi)$ is convex since it can be represented as the supremum of a family of linear maps, we shall show that $\varphi \mapsto u_h(\varphi)$ is also convex for u_h defined in (3.4). Then a natural candidate for the resolution of $U_h = \inf_{\varphi \in \mathrm{Q}_h} u_h(\varphi)$ in (3.7) is the gradient projection algorithm. The gradient projection algorithm is a classical iterative method for solving convex optimization problems. We would like to refer to Ben-Tal and Nemirovski [4], Bertsekas [5] for detailed presentations. We recall that $B(\mathcal{N}_R)$ denotes the collection of all functions on $\mathcal{N}_R$.

Proposition 3.5. *Under the CFL condition* (3.3)*, the function* $\varphi \mapsto u_h(\varphi)$ *is convex.*

Proof. Let us first rewrite the finite differences scheme (3.2) into a vector system. Denote $\alpha := \frac{\Delta t}{2\Delta x^2}$, $\lambda_k := \big(\lambda_h^\varphi(t_k, x_i)\big)_{-r\le i\le r}$, $\tilde\lambda_k := \big(\tilde\lambda_h^\varphi(t_k, x_i)\big)_{-r\le i\le r}$ and $q_k := \big(g^\varphi(t_k, x_i)\big)_{-r\le i\le r} \in \mathbb{R}^{2r+1}$. Let I_{2r+1} denote the $(2r+1)\times(2r+1)$ identity matrix, Π and $b_k \in \mathbb{R}^{2r+1}$ be defined by

$$\Pi := \begin{pmatrix} 0 & 0 & 0 & 0 & & & 0 \\ 1 & -2 & 1 & 0 & & & \\ 0 & 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 & 0 \\ & & & 0 & 1 & -2 & 1 \\ 0 & & & 0 & 0 & 0 & 0 \end{pmatrix}, \quad b_k := \begin{pmatrix} q_k(-r) - \lambda_{k+1}(-r) \\ 0 \\ \vdots \\ 0 \\ q_k(r) - \lambda_{k+1}(r) \end{pmatrix},$$

and $\Theta := \big[I_{2r+1} - \theta\alpha\Pi\big]^{-1}\big[I_{2r+1} + (1-\theta)\alpha\Pi\big]$, then scheme (3.2) can be rewritten as

$$\tilde\lambda_k \;=\; \Theta\lambda_{k+1} \;+\; b_k, \quad \text{and} \quad \lambda_k \;=\; \tilde\lambda_k \vee q_k. \tag{3.10}$$

Under CFL condition (3.3), one can verify that the above scheme is monotone, i.e. every element of Θ is positive, and moreover, $\Theta\mathbf{1} = \mathbf{1}$, where $\mathbf{1} := (1, \cdots, 1)^T \in \mathbb{R}^{2r+1}$. It follows that Θ can be the probability transition matrix of some Markov chain Λ, whose state space is the grid $\mathcal{N}_R$ with absorbing boundary. Let $\mathcal{T}_h^R$ denote the collection of all stopping times τ_h on Λ, then λ_h^φ can be represented as solutions of an optimal stopping problem on Λ:

$$\lambda_h^\varphi(t_k, x_i) \;=\; \sup_{\tau_h \in \mathcal{T}_h^R,\ \tau_h \ge t_k} \mathbb{E}\ \big[\ g^\varphi(\Lambda_{\tau_h}, \tau_h)\ \big|\ \Lambda_{t_k} = x_i\ \big].$$

Now given a stopping time $\tau_h \in \mathcal{T}_h^R$, we introduce the function $\lambda_{h,0}^{\varphi,T,R,\tau_h}$ defined on $\mathcal{N}_R$:

$$\lambda_{h,0}^{\varphi,\tau_h}(x_i) \;:=\; \mathbb{E}\ \big[\ g^\varphi(\Lambda_{\tau_h}, \tau_h)\ \big|\ \Lambda_0 = x_i\ \big].$$

Then u_h has an equivalent representation:

$$u_h(\varphi) = \sup_{\tau_h \in \mathcal{T}_h^R} \bar u_h^{\tau_h}(\varphi) \;:=\; \sup_{\tau_h \in \mathcal{T}_h^R} \mu_0\big(\mathrm{lin}^R[\lambda_{h,0}^{\varphi,\tau_h}]\big) \;+\; \mu_1\big(\mathrm{lin}^R[\varphi]\big). \tag{3.11}$$

Clearly, for every τ_h, the map $\varphi \mapsto \bar u_h^{\tau_h}(\varphi)$ is linear. It follows by (3.11) that $\varphi \mapsto u_h(\varphi)$ is convex. □

Remark 3.6. *In the above Markov chain system* (3.11)*, given* $\varphi \in B(\mathcal{N}_R)$*, one can define a stopping time* $\tau_h(\varphi)$ *by*

$$\tau_h(\varphi) \;:=\; \inf\left\{\, t_k \;:\; \lambda_h^{\varphi,\tau_h}(t_k, \Lambda_{t_k}) \;=\; g^{\varphi}(t_k, \Lambda_{t_k}) \,\right\}, \tag{3.12}$$

which is clearly an optimal stopping time, i.e.

$$u_h(\varphi) \;=\; \sup_{\tau_h \in \mathcal{T}_R^h} \bar{u}_h^{\tau_h}(\varphi) \;=\; \bar{u}_h^{\tau_h(\varphi)}(\varphi). \tag{3.13}$$

Now we are ready to give the gradient projection algorithm for U_h in (3.7). Given $\varphi \in B(\mathcal{N}_R)$, we denote by $P_{\mathrm{Q}_h}\big[\varphi\big]$ its projection on Q_h. Of course, such a projection depends on the norm equipped on $B(\mathcal{N}_R)$, which is an important issue to be discussed later.

Let $\gamma = (\gamma_n)_{n\geq 0}$ be a sequence of positive real numbers, we propose the following algorithm:

Algorithm 3.7. *For optimization problem* (3.7)*:*

- *1, Let* $\varphi_0 := \phi_{K,M}|_{\mathcal{N}_R}$*, where* $\phi_{K,M}$ *is defined in* (2.12)*.*
- *2, Given* φ_n*, compute* $u_h(\varphi_n)$ *and a sub-gradient* $\nabla u_h(\varphi_n)$*.*
- *3, Let* $\varphi_{n+1} := P_{\mathrm{Q}_h}\big[\varphi_n - \gamma_n \nabla u_h(\varphi_n)\big]$*.*
- *4, Go back to step 2.*

In the following, we shall discuss mainly three issues: the computation of sub-gradient $\nabla u_h(\varphi)$, the projection from $B(\mathcal{N}_R)$ to Q_h and the convergence of the above gradient projection algorithm.

3.2.1 Computation of sub-gradient

We notice that u_h can be represented as the supremum of a family functions in (3.13), and hence is convex. A natural method to obtain its sub-gradient is then first to identify the optimal stopping time which gives the supremum value in (3.13), and then to compute the gradient of the linear map associated with this optimal stopping time.

Let us fix $\varphi \in B(\mathcal{N}_R)$, we then denote by $(p^j, \tilde{p}^j)$ the unique solution of the following linear system on $\mathcal{M}$:

$$\begin{cases} p^j(t_k, x_i) \;=\; -\,\delta_{i,j}, & (t_k, x_i) \in \partial_T\mathcal{M} \cup \partial_R\mathcal{M}, \\ p^j(t_{k+1}, x_i) - \tilde{p}^j(t_k, x_i) + \frac{1}{2}\Delta t\,\big(\theta D^2\tilde{p}^j(t_k, x_i) + (1-\theta)D^2 p^j(t_{k+1}, x_i)\big) = 0, & \\ p^j(t_k, x_i) = \begin{cases} \tilde{p}^j(t_k, x_i), & \text{if } \lambda_h^{\varphi,T,R}(t_k, x_i) > g^{\varphi}(t_k, x_i), \\ -\,e_j(x_i), & \text{otherwise.} \end{cases} & (t_k, x_i) \in \mathring{\mathcal{M}}. \end{cases} \tag{3.14}$$

where $e_j \in B(\mathcal{N}_R)$ is defined by $e_j(x_i) := \delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$ Denote $p_0^j := p^j(0, \cdot)$.

Proposition 3.8. *Let CFL condition* (3.3) *hold true, then the vector*

$$\nabla u_h(\varphi) \;\;:=\;\; \Big(\; \mu_0(\mathrm{lin}^R[p_0^j]) \;+\; \mu_1(\mathrm{lin}^R[e_j]) \;\Big)_{-2m\le j\le 2m} \tag{3.15}$$

is a sub-gradient of map $\varphi \mapsto u_h(\varphi)$.

Proof. Consider the Markov chain Λ introduced in the proof of Proposition 3.5. By (3.13), we have for every perturbation $\Delta\varphi \in B(\mathcal{N}_R)$,

$$u_h(\varphi+\Delta\varphi) \;\;=\;\; \bar{u}_h^{\tau_h(\varphi+\Delta\varphi)}(\varphi+\Delta\varphi) \;\;\ge\;\; \bar{u}_h^{\tau_h(\varphi)}(\varphi+\Delta\varphi),$$

where $\tau_h(\varphi)$ and $\tau_h(\varphi+\Delta\varphi)$ are defined in (3.12). It follows still by (3.13) that

$$u_h(\varphi+\Delta\varphi) \;-\; u_h(\varphi) \;\;\ge\;\; \bar{u}_h^{\tau_h(\varphi)}(\varphi+\Delta\varphi) \;-\; \bar{u}_h^{\tau_h(\varphi)}(\varphi),$$

which implies that

$$\Big(\; \bar{u}_h^{\tau_h(\varphi)}(\varphi+e_j) \;-\; \bar{u}_h^{\tau_h(\varphi)}(\varphi) \;\Big)_{-r\le j\le r} \tag{3.16}$$

is a sub-gradient of u_h at φ since $\psi \mapsto \bar{u}_h^{\tau(\varphi)}(\psi)$ is linear by its definition in (3.11).

Finally, by the definition of $\tau_h(\varphi)$ in (3.12) as well as (3.2) and (3.14), it follows that

$$p^j(t_k, x_i) \;\;=\;\; -\; \mathbb{E}\left[\; e_j\big(\Lambda_{\tau_h(\varphi)}\big) \;\middle|\; \Lambda_{t_k} = x_i \;\right].$$

And hence the sub-gradient (3.16) coincides with $\nabla u_h(\varphi)$ defined in (3.15). □

3.2.2 Projection

To compute the projection P_{Q_h} from $B(\mathcal{N}_R)$ to Q_h, we still need to specify the norm equipped on $B(\mathcal{N}_R)$. The simplest norm can be the common one defined by $|\varphi|^2 := \sum_{i=-r}^{r} \varphi_i^2$. However, the computation of the projection may be too complicated. In order to make the projection algorithm simpler, we shall introduce an invertible linear map $\mathcal{L}_R$ from $B(\mathcal{N}_R)$ to $\mathbb{R}^{2r+1}$, then equip on $B(\mathcal{N}_R)$ the norm $|\cdot|_R$ induced by the classical $\mathbb{L}^2-$norm on $\mathbb{R}^{2r+1}$. Let $\mathcal{L}_R : B(\mathcal{N}_R) \to \mathbb{R}^{2r+1}$ be defined by

$$\xi_i \;\;=\;\; \begin{cases} \varphi(x_i) \;-\; \varphi(x_{i-1}), & \text{for } \;0 < i \le r, \\ \varphi(x_0), & \text{for } \; i = 0, \\ \varphi(x_i) \;-\; \varphi(x_{i+1}), & \text{for } \;-r \le i < 0. \end{cases} \tag{3.17}$$

We define the norm $|\cdot|_R$ on $B(\mathcal{N}_R)$ (easily be verified) by

$$|\varphi|_R \;\;:=\;\; \big|\xi\big|_{\mathbb{L}^2(\mathbb{R}^{2r+1})}, \quad \text{with } \xi := \mathcal{L}_R(\varphi), \quad \forall \varphi \in B(\mathcal{N}_R).$$

Denote

$$\begin{aligned} E_0 \;\;&:=\;\; \Big\{\; \mathcal{L}_R\varphi \;\;:\;\; \varphi \in \mathrm{Q}_0 \;\Big\} \\ &=\;\; \Big\{\; \xi \in \mathbb{R}^{2r+1} \;:\; 0 = \xi_0 \le \xi_{\pm 1} \le \cdots \le \xi_{\pm 2m} \le 4KM\Delta x, \\ &\qquad\qquad \xi_{\pm i} = K(x_{i+1}^2 - x_i^2),\; \forall 2m < i \le r \text{ and } \sum_{i=1}^{2m}\xi_i = \sum_{i=-1}^{-2m}\xi_i = 4KM^2 \;\Big\}. \end{aligned}$$

Then the projection P_{Q_h} from $B(\mathcal{N}_R)$ to Q_h under norm $|\cdot|_R$ is equivalent to the projection from $\mathbb{R}^{2r+1}$ to E_0 under the $\mathbb{L}^2-$norm, which consists in solving a quadratic minimization problem:

$$\xi^z \quad := \quad \arg\min_{\xi \in E_0} \sum_{i=-r}^{r} (z_i - \xi_i)^2, \quad \text{for a given } z \in \mathbb{R}^{2r+1}. \tag{3.18}$$

Clearly, for every $z \in \mathbb{R}^{2r+1}$, $\xi_0^z = 0$ and the above optimization problem (3.18) can be decomposed into two optimization problems:

$$\min_{\xi \in E_{0,+}} \sum_{i=1}^{2m} (z_i - \xi_i)^2 \quad \text{and} \quad \min_{\xi \in E_{0,-}} \sum_{i=-1}^{-2m} (z_i - \xi_i)^2, \tag{3.19}$$

where

$$E_{0,\pm} \ := \ \Big\{ \xi = (\xi_{\pm i})_{1 \le i \le 2m} \ : \ 0 \le \xi_{\pm 1} \le \cdots \le \xi_{\pm 2m} \le 4KM\Delta x, \ \sum_{i=1}^{2m} \xi_{\pm i} = 4KM^2 \Big\},$$

The optimization problem (3.19) can be solved, and we propose an algorithm in Appendix.

3.2.3 Convergence rate

We shall provide a convergence rate for the gradient projection algorithm. In preparation, let us first give an estimate on the norm of the sub-gradients ∇u_h.

Lemma 3.9. *Let $\varphi_1, \varphi_2 \in B(\mathcal{N}_R)$, then under the CFL condition* (3.3)*,*

$$\big|\ u_h(\varphi_1) \ - \ u_h(\varphi_2)\ \big| \quad \le \quad 2\ |\varphi_1 - \varphi_2|_\infty, \tag{3.20}$$

and it follows that

$$\big|\nabla u_h(\varphi)\big|_R \quad \le \quad 2\sqrt{2m+1} \quad = \quad 2\sqrt{\frac{2M}{\Delta x} + 1}, \qquad \forall \varphi \in B(\mathcal{N}_R). \tag{3.21}$$

Proof. Under the CFL condition (3.3), the $\theta-$scheme is monotone, which implies that $|\lambda_h^{\varphi_1} - \lambda_h^{\varphi_2}|_\infty \le |\varphi_1 - \varphi_2|_\infty$. Hence by the definition of u_h in (3.4), the inequality (3.20) holds true.

Next, denote $\xi^i := \mathcal{L}_R(\varphi_i)$, $i = 1, 2$, then by Cauchy-Schwarz inequality,

$$|\varphi_1 - \varphi_2|_\infty \le \max\Big(\sum_{i=0}^{2m} \big|\xi_i^1 - \xi_i^2\big|, \sum_{i=0}^{-2m} \big|\xi_i^1 - \xi_i^2\big| \Big) \le \sqrt{2m+1} \cdot \big\|\xi^1 - \xi^2\big\|_{\mathbb{L}^2},$$

which implies immediately (3.21). □

Finally, let us finish this section by providing a convergence rate of the proposed gradient projection algorithm (Algorithm 3.7). Denote

$$\Phi \quad := \quad \max_{\varphi_1, \varphi_2 \in \mathrm{Q}_h} |\varphi_1 - \varphi_2|_R^2 \quad \le \quad 4m\ (4KM\Delta x)^2 \quad \le \quad 64K^2M^3\Delta x,$$

it follows from Section 5.3.1 of Ben-Tal and Nemirovski [4] that one has the convergence rate:

$$\min_{n\le N}\ u_h(\varphi_n)\ -\ U_h\ \le\ \frac{\Phi\ +\ \sum_{i=n}^N \gamma_n^2\big|\nabla u_h(\varphi_n)\big|_R^2}{2\sum_{n=1}^N\gamma_n}$$
$$=\ \frac{32K^2M^3\Delta x\ +\ \big(4\frac{M}{\Delta x}+2\big)\ \sum_{i=n}^N\gamma_n^2}{\sum_{n=1}^N\gamma_n}. \qquad (3.22)$$

There are several choices for the sequence $\gamma=(\gamma_n)_{n\ge1}$:

- Divergent Series: $\gamma_n\ge0$, $\sum_{n=1}^\infty\gamma_n=+\infty$ and $\sum_{n=1}^\infty\gamma_n^2<+\infty$. Clearly, (3.22) converges to 0 as $N\to\infty$.
- Optimal stepsizes: $\gamma_n=\frac{\sqrt{\Phi}}{\big|\nabla u_h(\varphi_n)\big|_R\sqrt{n}}$, we have by [4] that

$$\min_{n\le N}u_h(\varphi_n)\ -\ U_h\ \le\ O(1)\frac{16KM\sqrt{2M^2+M\Delta x}}{\sqrt{N}}.$$

4 Numerical example

Finally, we implement the above algorithm and test it on several options, including the variance swap, call on variance. In all examples, we suppose that μ_0 and μ_1 are of log-normal distribution given as follows. Let $S_t:=S_0\exp(-\frac{1}{2}\sigma^2t+\sigma W_t)$ with some constants S_0, σ and a standard Brownian motion $(W_t)_{t\ge0}$, we suppose that $\mu_0\sim S_{\frac{1}{2}}$ and $\mu_1\sim S_1$. Clearly, for all constants σ, μ_0 and μ_1 satisfies Assumption 2.5.

4.1 Variance swap

We first test the algorithm on "variance swap", whose payoff function is given by $g(t,x)=t$. It follows by Proposition 2.7 that the model-free price upper bound of variance swap is given by C_0 in (2.8), i.e.

$$C_0\ =\ \int_{\mathbb{R}}x^2\mu_1(dx)-\int_{\mathbb{R}}x^2\mu_0(dx)\ =\ \mathbb{E}\ \big(\ S_1^2-S_{\frac{1}{2}}^2\ \big)\ =\ S_0^2\big(e^{\sigma^2}-e^{\sigma^2/2}\big).$$

In our implemented example, we set $\sigma=0.25$, $S_0=1$, hence $C_0\approx0.0327511$. For the approximation and discretization parameters, we set $T=0.15$, $K=1$, $M=2.5$, $R=2.6$, $\Delta t=0.003$, $\Delta x=0.1$ and $\gamma_n=\sqrt{n}$.

With a 2.40GHz CPU computer, it takes 84.89 seconds to finish 4×10^4 iterations, and we get the numerical upper bound 0.0328511, i.e. the relative error is less than 1%, see also Figure 1.

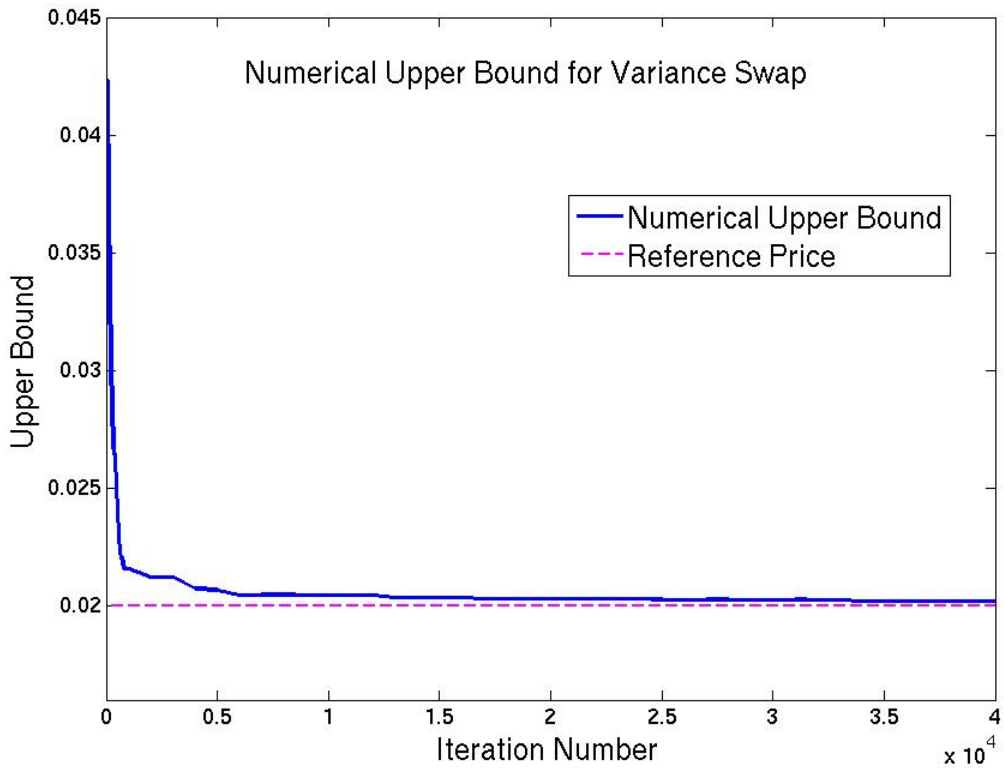

Figure 1: Numerical result for variance swap with approximation parameters: $T = 0.15$, $K = 1$, $M = 1$, $R = 2.6$, $\Delta t = 0.003$, $\Delta x = 0.1$ and $\gamma_n = \sqrt{n}$.

4.2 Call on variance

We next give some numerical tests on the option "call on variance", whose payoff function is $g(t, x) = (t - K_v)^+$ for some positive constant K_v. In our implemented examples, we set $\sigma = 0.2$, $S_0 = 1$. In a first example, we fix $K = 1$ and obtain the numerical solutions with different approximation parameters T. The result illustrated in Figure 2 is consistent to the convergence result in Theorem 2.11 as well as that in Proposition 5.8 below, i.e. when $T \to \infty$, the convergence is of exponential order.

In a second example, we fix $T = 0.5$ and test the numerical algorithm with different approximation parameters K. We notice that for both options, the minimum upper bounds are given by the case $K \approx 1$. When K is too small, the approximation strategy class Q_0 defined by (2.9) is not rich enough, and hence the approximated upper bound is greater than the real bound. When K is too large, another error term given by $\mu_1(\phi_{K,M})$ in Theorem 2.11 (see also Proposition 5.6 below) becomes too important. Therefore, for those calls on variance, the optimal approximation parameter K to choose is around 1. (see Figure 3).

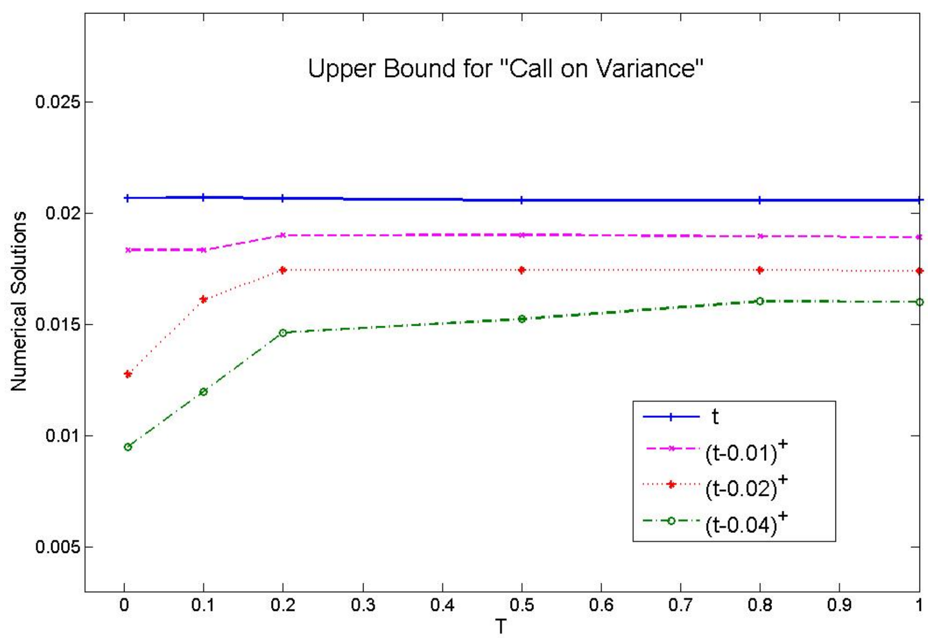

Figure 2: Numerical result for call on variance with approximation parameters: $K = 1$, $M = 1$, $R = 2.2$, $\Delta t = 0.001$, $\Delta x = 0.1$ and $\gamma_n = \sqrt{n}$.

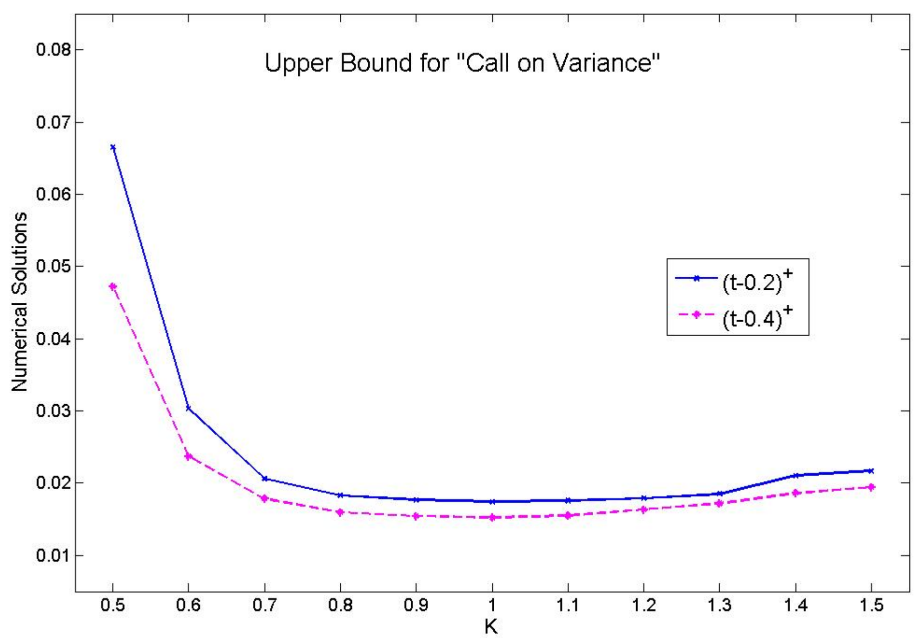

Figure 3: Numerical result for call on variance with approximation parameters: $T = 0.5$, $M = 1$, $R = 2.2$, $\Delta t = 0.001$, $\Delta x = 0.1$ and $\gamma_n = \sqrt{n}$.

5 Proof of the approximation result

In this section, we shall complete the proof of Proposition 2.7 as well as the convergence result in Theorem 2.11. We first give some technical lemmas about the stopping times on a Brownian motion in Section 5.1. Then we complete the proof of Proposition 2.7 in Section 5.2, and the proof of Theorem 2.11 in Section 5.3.

5.1 Technical lemmas

We recall that $B = (B_t)_{t\geq 0}$ is a standard Brownian motion in probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with natural Brownian filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, and $\mathcal{T}^\infty$ denotes the collection of all $\mathbb{F}-$stopping times τ such that $\mathbb{E}[\tau] < \infty$. Let $T > 0$, we also denote by $\mathcal{T}^T$ the collection of all $\mathbb{F}-$stopping times taking value in $[0, T]$, i.e.

$$\mathcal{T}^T \;\; := \;\; \big\{\tau \wedge T : \tau \in \mathcal{T}^\infty\big\}. \tag{5.1}$$

Lemma 5.1. *Let $\psi : (t, x) \in \mathbb{R}^+ \times \mathbb{R} \mapsto \psi(t, x) \in \mathbb{R}$ be a function Lipschitz in t, satisfying $\sup_{(t,x)\in\mathbb{R}^+\times\mathbb{R}} \frac{|\psi(t,x)|}{1+x^2} < \infty$. Then for every $\tau \in \mathcal{T}^\infty$,*

$$\mathbb{E}\,\Big[\;\psi(\tau, B_\tau)\;\Big] \;\; = \;\; \lim_{t\to\infty}\,\mathbb{E}\,\Big[\;\psi(\tau\wedge t, B_{\tau\wedge t})\;\Big]. \tag{5.2}$$

In particular,

$$\mathbb{E}[B_\tau^2] \;=\; \lim_{t\to\infty}\mathbb{E}[B_{\tau\wedge t}^2] \;=\; \lim_{t\to\infty}\mathbb{E}[\tau\wedge t] \;=\; \mathbb{E}[\tau] \;\; \textit{and} \;\; \mathbb{E}[B_\tau] \;=\; 0. \tag{5.3}$$

Proof. Given a stopping time $\tau \in \mathcal{T}^\infty$, let $Y_t := B_{\tau\wedge t}$. Then by assumptions on ψ, there is a constant $C > 0$ such that

$$\psi(B_{\tau\wedge t}, \tau\wedge t) \;\leq\; C\big(1 + Y_t^2 + \tau\big) \;\leq\; C\Big(1 + \sup_{s\geq 0} Y_s^2 + \tau\Big), \quad \forall t \geq 0.$$

We notice that $(Y_t)_{t\geq 0}$ is a continuous uniformly integrable martingale by its definition, and $\mathbb{E}\big[\sup_{s\geq 0} Y_s^2\big] \;\leq\; 4\mathbb{E}[\tau] \;<\; \infty$ by Doob's inequality. And hence it follows by the dominated convergence theorem that (5.2) holds true. □

Lemma 5.2. *Let $\psi \in \mathrm{Q}$ and denote by ψ^{conv} its convex envelope, then*

$$\inf_{\tau\in\mathcal{T}^T}\mathbb{E}\;\psi(B_\tau) \;\to\; \inf_{\tau\in\mathcal{T}^\infty}\mathbb{E}\;\psi(B_\tau) \;=\; \psi^{conv}(0), \quad \textit{as} \quad T\to\infty.$$

Proof. Let $a \leq 0 \leq b$ be two constants and $\tau_{a,b} := \inf\big\{t : B_t \notin (a, b)\big\}$. We first notice that $\tau_{a,b} \in \mathcal{T}^\infty$ since $\mathbb{E}[\tau_{a,b}] = \lim_{t\to\infty}\mathbb{E}[\tau_{a,b}\wedge t] = \lim_{t\to\infty}\mathbb{E}[B^2_{\tau_{a,b}\wedge t}] \leq (a^2 + b^2) < \infty$. Hence by (5.3), $\mathbb{E}[B_{\tau_{a,b}}] = 0$, which implies that $\mathbb{P}(B_{\tau_{a,b}} = a) = \frac{b}{b-a}$ and $\mathbb{P}(B_{\tau_{a,b}} = b) = \frac{-a}{b-a}$. Therefore,

$$\inf_{\tau\in\mathcal{T}^\infty}\mathbb{E}\psi(B_\tau) \;\leq\; \inf_{a<0<b}\mathbb{E}\psi(B_{\tau_{a,b}}) \;=\; \inf_{a<0<b}\Big(\frac{b}{b-a}\psi(a) + \frac{-a}{b-a}\psi(b)\Big) \;=\; \psi^{conv}(0).$$

On the other hand, for every $\tau \in \mathcal{T}^\infty$, it follows by Jensen's inequality and $\mathbb{E}[B_\tau] = 0$ from (5.3) that $\psi^{conv}(x) \leq \mathbb{E}[\psi^{conv}(x + B_\tau)] \leq \mathbb{E}[\psi(x + B_\tau)]$. Therefore,

$$\inf_{\tau\in\mathcal{T}^\infty}\;\mathbb{E}\psi(B_\tau) \;\;=\;\; \psi^{conv}(0).$$

Finally, the convergence of $\inf_{\tau\in\mathcal{T}^T}\mathbb{E}\psi(B_\tau)$ to $\inf_{\tau\in\mathcal{T}^\infty}\mathbb{E}\psi(B_\tau)$ as $T\to\infty$ is a direct consequence of (5.2). □

Corollary 5.3. *Let $\phi\in\mathrm{Q}$ and $(a,b)\in\mathbb{R}^2$. Then for function u defined in (2.5), we have $u(\phi)=u(\phi_{a,b})$, where $\phi_{a,b}$ is given by $\phi_{a,b}(x):=\phi(x)+ax+b$.*

Proof. By the definition of λ_0^ϕ in (2.4) together with Lemma 5.1, it follows that $\lambda_0^{\phi_{a,b}}(x)=\lambda_0^\phi(x)+ax+b$. Moreover, as discussed in Remark 2.6, $\mu_0(I)=\mu_1(I)$ for the identity function I. Then we get $u(\phi)=u(\phi_{a,b})$ by the definition of u in (2.5). □

It is also interesting to consider the stopping time w.r.t. a larger filtration. Let $\tilde{\mathbb{F}}$ be another filtration in $(\Omega,\mathcal{F},\mathbb{P})$ w.r.t which B is still a standard Brownian motion, denote

$$\tilde{\mathcal{T}}^\infty \;:=\; \big\{\tilde{\mathbb{F}}-\text{stopping times }\tau\text{ such that }\mathbb{E}(\tau)<\infty\big\}.$$

Lemma 5.4. *For all $\phi\in\mathrm{Q}$, we have*

$$\lambda^\phi(t,x) \;:=\; \sup_{\tau\in\mathcal{T}^\infty}\mathbb{E}\big[\,g^\phi(t+\tau,x+B_\tau)\,\big] \;=\; \sup_{\tau\in\tilde{\mathcal{T}}^\infty}\mathbb{E}\big[\,g^\phi(t+\tau,x+B_\tau)\,\big]. \quad (5.4)$$

Proof. By the same arguments as in Lemma 5.1, (5.2) holds still true for every $\tau\in\tilde{\mathcal{T}}^\infty$. Then, to prove (5.4), it is enough to prove that for every $T>0$,

$$\sup_{\tau\in\mathcal{T}^\infty,\tau\le T}\mathbb{E}\big[\,g^\phi(t+\tau,x+B_\tau)\,\big] \;=\; \sup_{\tau\in\tilde{\mathcal{T}}^\infty,\tau\le T}\mathbb{E}\big[\,g^\phi(t+\tau,x+B_\tau)\,\big].$$

Since the family of random variables $(g^\phi(t+\tau,x+B_\tau))_{\tau\in\tilde{\mathcal{T}}^\infty,\tau\le T}$ is clearly of class D, we then conclude the proof by Theorem 5 of Szpirglas and Mazziotto [23]. □

5.2 Proof of Proposition 2.7

(i) Given $\phi\in\mathrm{Q}$, we denote $\phi_{K,\psi}(x):=\phi(x)+Kx^2+\psi(x)$ which also belongs to Q. Then by (5.3), for all $\tau\in\mathcal{T}^\infty$,

$$\mathbb{E}\big[g_{K,\psi}(\tau,x+B_\tau)-\phi_{K,\psi}(x+B_\tau)\big] \;=\; \mathbb{E}\big[g^\phi(\tau,x+B_\tau)\big]-Kx^2.$$

It follows by the definition of U in (2.5) that $U(g_{K,\psi})\ge U(g)+KC_0+\mu_1(\psi)$. And moreover, by the arbitrariness of $K\in\mathbb{R}$, $\psi\in\mathrm{Q}$ and symmetric relationship between g and $g_{K,\psi}$, we conclude the proof of (2.7).

(ii) For the second assertion, we first claim that $u(g^0)=0$ for $g^0\equiv 0$. Indeed, with the payoff function $g^0\equiv 0$, we get immediately from (2.4) as well as Lemma 5.2 that

$$u(\phi) \;=\; -\;\mu_0(\phi^{conv})\;+\;\mu_1(\phi) \;\ge\; \mu_1(\phi^{conv})-\mu_0(\phi^{conv}) \;\ge\; 0.$$

where the last inequality follows from Assumption 2.5. It follows that that $U(g^0)=0$. Let us take the positive constant L_0 given in Assumption 2.8, then

$$g(0,x) \;\le\; g(t,x) \;\le\; g(0,x)\;+\;L_0t.$$

Further, it is clear that U is monotone w.r.t. the payoff function g by its definition in (2.5). Then it follows that

$$\mu_1(g(0,\cdot)) \;\le\; U \;\le\; \mu_1(g(0,\cdot))\;+\;L_0C_0, \quad \text{with } C_0 \text{ defined in } \quad (2.8).$$

Hence we conclude the proof. □

5.3 The proof of Theorem 2.11

We shall decompose the proof of Theorem 2.11 intro four steps.

In a first step, we show that in optimization problem (2.5), it is equivalent to minimize among all non negative convex functions. Denote

$$\mathrm{Q}^0 \;\; := \;\; \big\{ \; \phi \in \mathrm{Q} \text{ non negative, convex, such that } \phi(0) = 0 \; \big\}.$$

Proposition 5.5. *Let Assumptions 2.5 and 2.8 hold true, then*

$$U \;\; = \;\; \inf_{\phi \in \mathrm{Q}^0} \; u(\phi).$$

Proof. Let $T \in \mathbb{R}^+$, $\tau_0 \in \mathcal{T}^T$ and $\phi \in \mathrm{Q}$. By the dominated convergence theorem, it is clear that $x \mapsto \inf_{\tau \in \mathcal{T}^T} \mathbb{E}\phi(x + B_\tau)$ is continuous. This, together with the weak dynamic programming in Theorem 4.1 of Bouchard and Touzi [6], implies the dynamic programming principle:

$$\inf_{\tau_0 \le \tau \le T} \; \mathbb{E}\phi(x + B_\tau) \;\; = \;\; \mathbb{E}\Big[\mathrm{ess} \inf_{\tau_0 \le \tau \le T} \mathbb{E}\big[\phi(x + B_\tau) \big| \mathcal{F}_{\tau_0}\big]\Big].$$

Then for every constant $\hat{T} > T$,

$$\lambda_0^\phi(x) \;\; = \;\; \sup_{\tau \in \mathcal{T}^\infty} \; \mathbb{E}\big[\; g^\phi(\tau, x + B_\tau) \; \big] \; \ge \; \sup_{\tau_0 \le \tau \le \hat{T}} \; \mathbb{E}\big[\; g(\tau, x + B_\tau) - \phi(x + B_\tau) \; \big].$$

Since g increases in t and is convex in x from Assumption 2.8, we have

$$\mathbb{E}\big[g(\tau, x + B_\tau) \big| \mathcal{F}_{\tau_0}\big] \;\; \ge \;\; \mathbb{E}\big[g(\tau_0, x + B_\tau) \big| \mathcal{F}_{\tau_0}\big] \;\; \ge \;\; g(\tau_0, x + B_{\tau_0}),$$

and hence

$$\lambda_0^\phi(x) \;\; \ge \;\; \mathbb{E}\big[g(\tau_0, x + B_{\tau_0})\big] \; - \; \mathbb{E}\Big[\inf_{\tau_0 \le \tau \le \hat{T}} \mathbb{E}\big[\phi(x + B_\tau) \big| \mathcal{F}_{\tau_0}\big]\Big].$$

Sending $\hat{T}$ to $+\infty$, it follows by Lemma 5.2 that

$$\lambda_0^\phi(x) \;\; \ge \;\; \mathbb{E} \; \big[\; g(\tau_0, x + B_{\tau_0}) \; - \; \phi^{conv}(x + B_{\tau_0}) \; \big].$$

Thus, by arbitrariness of τ_0 in $\mathcal{T}^T$ as well as that of $T \in \mathbb{R}^+$, we get

$$\begin{aligned} \lambda_0^\phi(x) \;\; &\ge \;\; \lim_{T \to \infty} \sup_{\tau_0 \in \mathcal{T}^T} \; \mathbb{E} \; \big[\; g(\tau_0, x + B_{\tau_0}) \; - \; \phi^{conv}(x + B_{\tau_0}) \; \big], \\ &= \;\; \sup_{\tau_0 \in \mathcal{T}^\infty} \; \mathbb{E} \; \big[\; g(\tau_0, x + B_{\tau_0}) \; - \; \phi^{conv}(x + B_{\tau_0}) \; \big], \end{aligned}$$

where the last equality is a direct consequence of Lemma 5.1 since ϕ^{conv} is either of quadratic growth or equals to $-\infty$.

Finally, since $\phi \ge \phi^{conv}$, by the definitions of u and U in (2.5), it is clear that the infimum in (2.5) can be taken over the collection of all convex functions in Q. Moreover, by the property of $u(\phi)$ in Corollary 5.3, the infimum can be then taken over the collection of all positive convex functions ϕ in Q such that $\phi(0) = 0$, i.e. $U = \inf\{u(\phi) \; : \phi \in \mathrm{Q}^0\}$. We then conclude the proof. □

Our second step is on the growth coefficient of ϕ in Q^0. Let K be a positive constant, we denote

$$U^K \;:=\; \inf_{\phi \in \mathrm{Q}^{0,K}} u(\phi) \;\text{ with }\; \mathrm{Q}^{0,K} \;:=\; \big\{\phi \in \mathrm{Q}^0 \;:\; \phi(x) \le K(|x| \vee x^2)\big\}.$$

By the convexity of functions in Q^0, we see that every $\phi \in \mathrm{Q}^0$ is in fact locally Lipschitz continuous, and hence one can easily deduce that $\mathrm{Q}^0 = \cup_{K>0} \mathrm{Q}^{0,K}$. Then it follows immediately that

$$U^K \;\searrow\; U \;\text{ as }\; K \;\longrightarrow\; \infty. \tag{5.5}$$

The third step of the approximation is to fix the tail of functions in $\mathrm{Q}^{0,K}$. Given a constant $M \ge M_0$, where M_0 is given in Assumption 2.8, we denote

$$\mathrm{Q}^{0,K,M} \;:=\; \big\{\; \phi \in \mathrm{Q}^{0,K} \text{ such that } \phi(x) = Kx^2 \text{ for } |x| \ge 2M \;\big\} \tag{5.6}$$

and the approximation value

$$U^{K,M} \;:=\; \inf_{\phi \in \mathrm{Q}^{0,K,M}} u(\phi).$$

Proposition 5.6. *Suppose that Assumptions 2.5 and 2.8 hold true. Then*

$$0 \;\le\; U^{K,M} - U^K \;\le\; \mu_1(\phi_{K,M}), \tag{5.7}$$

where $\phi_{K,M}$ is defined by (2.12).

Proof. Let us first recall that every function $\phi \in \mathrm{Q}^{0,K}$ is nonnegative, convex such that $\phi(0) = 0$ and $\phi(x) \le K(|x| \vee x^2)$. Given $\phi \in \mathrm{Q}^{0,K}$, we denote $\phi_M := \phi \vee \phi_{K,M}$ with $\phi_{K,M}$ defined in (2.12). Clearly, ϕ_M lies in $\mathrm{Q}^{0,K,M}$ and $\lambda^{\phi_M} \le \lambda^{\phi}$ since $\phi_M \ge \phi$. It follows from the definition of $u(\phi)$ in (2.5) and the positivity of ϕ that

$$u(\phi_M) \;-\; u(\phi) \;\le\; \mu_1(\phi_M) \;-\; \mu_1(\phi) \;\le\; \mu_1(\phi_{K,M}).$$

This, together with the arbitrariness of $\phi \in \mathrm{Q}^{0,K}$ and the fact that $\phi_M \in \mathrm{Q}^{0,K,M}$, concludes the proof for (5.7). □

For the fourth step of the analytic approximation, we first introduce for every T, $R > 0$,

$$\lambda^{\phi,T}(t,x) \;:=\; \sup_{\tau \in \mathcal{T}^\infty,\; \tau \le T-t} \mathbb{E}\big[g^\phi(t+\tau, x+B_\tau)\big], \qquad \lambda_0^{\phi,T}(\cdot) \;:=\; \lambda^{\phi,T}(0,\cdot),$$

$$\lambda^{\phi,\tau_R}(t,x) \;:=\; \sup_{\tau \in \mathcal{T}^\infty,\; \tau \le \tau_x^R} \mathbb{E}\big[g^\phi(t+\tau, x+B_\tau)\big], \tag{5.8}$$

and

$$\lambda^{\phi,T,R}(t,x) \;:=\; \sup_{\tau \in \mathcal{T}^\infty,\; \tau \le \tau_x^R \wedge (T-t)} \mathbb{E}\big[g^\phi(t+\tau, x+B_\tau)\big], \tag{5.9}$$

where $\tau_x^R \;:=\; \inf\{s \;:\; x + B_s \notin (-R,R)\}$.

Proposition 5.7. *Let Assumption 2.8 hold true with constants L_0, M_0. Suppose that $K > L_0$, $M \geq M_0$ and $R \geq \big(1 + \sqrt{\frac{K}{K-L_0}}\big)M$. Then for every $\phi \in \mathrm{Q}^{0,K,M}$, we have*

$$\lambda^{\phi}(t,x) = \lambda^{\phi,\tau_R}(t,x) \quad \text{and} \quad \lambda^{\phi,T}(t,x) = \lambda^{\phi,T,R}(t,x), \quad \forall (t,x) \in [0,T] \times \mathbb{R}.$$

With the equivalence between λ^{ϕ} ($\lambda^{\phi,T}$) and λ^{ϕ,τ_R} ($\lambda^{\phi,T,R}$), we can now make an approximation on coefficient T. Let

$$U^{K,M,T} \quad := \quad \inf_{\phi \in \mathrm{Q}^{0,K,M}} \; u^T(\phi) \quad \text{with} \quad u^T(\phi) \; := \; \mu_0(\lambda_0^{\phi,T}) \; + \; \mu_1(\phi). \tag{5.10}$$

Proposition 5.8. *Let Assumptions 2.5 and 2.8 hold, M_0 and L_0 be constants given in Assumption 2.8. Let $K > L_0$, $M \geq M_0$, $R = \big(1 + \sqrt{\frac{K}{K-L_0}}\big)M$ and $L = 2(K + 2L_0)(R^2 \vee 1)$, we denote*

$$\delta \; := \; -\log(q(R)) > 0, \quad \text{where} \;\; q(R) \; := \; \frac{1}{\sqrt{2\pi}} \int_{-2R}^{2R} e^{-x^2/2}\, dx.$$

Then

$$0 \quad \leq \quad U^{K,M} - \; U^{K,M,T} \quad \leq \quad Le^{-\delta(T-1)}. \tag{5.11}$$

In preparation of the proof for Propositions 5.7 and 5.8, we first give a property for functions in $\mathrm{Q}^{0,K,M}$ defined by (5.6).

Lemma 5.9. *Let Assumption 2.8 hold true with constants L_0 and M_0, $K > L_0$, $M \geq M_0$ and $R = \big(1 + \sqrt{\frac{K}{K-L_0}}\,\big)M$. Given fixed $t \in \mathbb{R}^+$ and $\phi \in \mathrm{Q}^{0,K,M}$, we denote*

$$\psi(x) \quad := \quad - \; g^{\phi}(t,x) \; - \; L_0 x^2 \quad = \quad \phi(x) \; - \; g(t,x) \; - \; L_0 x^2.$$

Then $\psi^{conv}(x) = \psi(x)$ when $x \notin [-R,R]$.

Proof. By Assumption 2.8, we know that there are constants C_1, C_2 such that $x \mapsto g(t,x)$ is affine with derivative C_1 when $x \geq M$, and affine with derivative C_2 when $x \leq -M$. For fixed $t \in \mathbb{R}^+$, let χ be a continuous function defined on $\mathbb{R}$ by the following: χ is affine on intervals $[-2M,-M]$, $[-M,0]$, $[0,M]$, $[M,2M]$ and

$$\begin{cases} \chi(0) \;\; := \;\; -\, g(t,0), \\ \chi(\pm M) \;\; := \;\; -\, L_0 M^2 \; - \; g(t,\pm M), \\ \chi(\pm 2M) \; := \; 4(K-L_0)M^2 \; - \; g(t,\pm 2M), \\ \chi(x) \;\; := \;\; (K-L_0)x^2 \; - \; g(t,2M) \; - \; C_1(x-2M), \qquad x \geq 2M, \\ \chi(x) \;\; := \;\; (K-L_0)x^2 \; - \; g(t,-2M) - C_2(x+2M), \qquad x \leq -2M. \end{cases}$$

By Assumption 2.8, we can verify that for every $\phi \in \mathrm{Q}^{0,K,M}$ and the corresponding ψ defined in the statement of the lemma,

$$\psi(x) \; \begin{cases} \geq \; \chi(x), & \text{when } x \in [-2M,2M], \\ = \; \chi(x), & \text{when } x \notin [-2M,2M]. \end{cases}$$

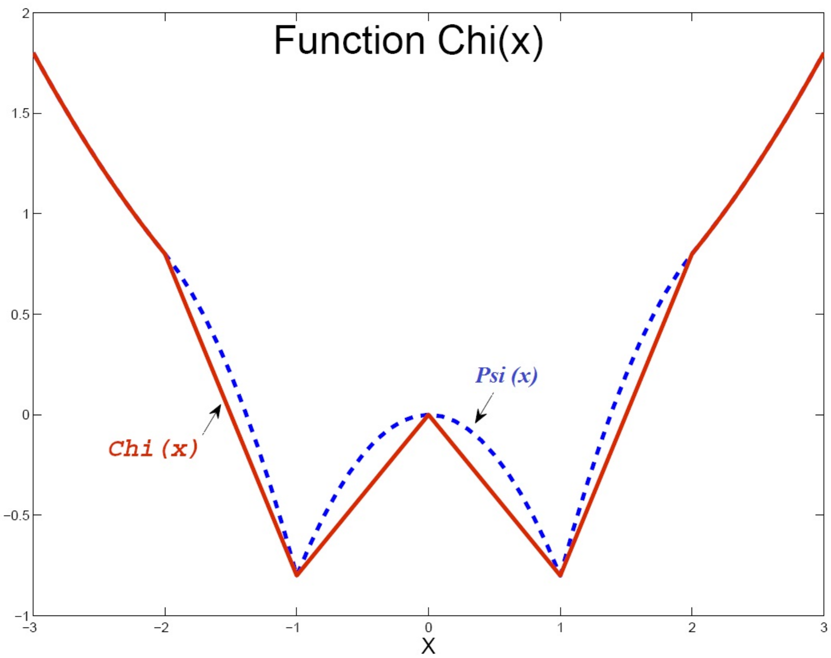

Figure 4: An example of functions χ and ψ, where $g \equiv 0$, $\phi = \phi_{K,M}$ with constants $K = 1$, $M = 1$ and $L_0 = 0.8$.

Then given $x \notin [-R, R]$, it follows by a simple computation that $\chi(y) \geq \chi(x) + \chi'(x)(y - x)$ for every $y \in \mathbb{R}$, which implies that $\chi^{conv}(x) = \chi(x)$. And hence $\psi(x) \geq \psi^{conv}(x) \geq \chi^{conv}(x) = \chi(x) = \psi(x)$ for $x \notin [-R, R]$. $\square$

Now, we can complete the proofs of Propositions 5.7 and 5.8.

Proof of Proposition 5.7. We shall only show that $\lambda^{\phi} = \lambda^{\phi,\tau_R}$ since the other equality $\lambda^{\phi,T} = \lambda^{\phi,T,R}$ can be proved by the same arguments. Moreover, to prove $\lambda^{\phi} = \lambda^{\phi,\tau_R}$, it is enough to show that $\lambda^{\phi} \leq \lambda^{\phi,\tau_R}$ since its inverse inequality is obvious from the definition of λ^{ϕ,τ_R} in (5.8).

First, let us fix $t \in \mathbb{R}^+$ and $x \notin (-R, R)$, we denote $\psi_x(y) := -g^{\phi}(t, y) - L_0 y^2 + L_0 x^2$. Then by Lemma 5.9, we have $\psi_x^{conv}(x) = \psi_x(x) = -g^{\phi}(t, x)$. Then for every $\tau \in \mathcal{T}^{\infty}$,

$$
\begin{aligned}
\mathbb{E}\left[\ g^{\phi}(t+\tau, x+B_{\tau})\ \right] &\leq \mathbb{E}\left[\ g^{\phi}(t, x+B_{\tau}) \ + \ L_0 \tau\ \right] \\
&= \mathbb{E}\left[\ g^{\phi}(t, x+B_{\tau}) \ + \ L_0(x+B_{\tau})^2 \ - \ L_0 x^2\ \right] \\
&= -\ \mathbb{E}\ \psi_x(x+B_{\tau}) \ \leq \ -\ \psi_x^{conv}(x) \ = \ g^{\phi}(t, x). \quad (5.12)
\end{aligned}
$$

It follows that $\lambda^{\phi}(t, x) \leq \lambda^{\phi,\tau_R}(t, x)$ for every $x \notin (-R, R)$ since in this case $\tau_x^R = 0$.

Next, for every $\tau \in \mathcal{T}^{\infty}$ and $x \in [-R, R]$, we have according to (5.12) that

$$
\begin{aligned}
& \mathbb{E}\left[\ g^{\phi}(t+\tau, x+B_{\tau})\ \right] \\
= \ & \mathbb{E}\left[g^{\phi}(t+\tau, x+B_{\tau}) \mathbf{1}_{\tau \leq \tau_x^R} \right] \ + \ \mathbb{E}\left[\ \mathbb{E}\left[g^{\phi}(t+\tau, x+B_{\tau}) \mathbf{1}_{\tau > \tau_x^R}\ \middle|\ \mathcal{F}_{\tau \wedge \tau_x^R}\ \right] \right] \\
\leq \ & \mathbb{E}\left[\ g^{\phi}(t+\tau \wedge \tau_x^R, x+B_{\tau \wedge \tau_x^R})\ \right],
\end{aligned}
$$

which implies that $\lambda^{\phi}(t,x) \le \lambda^{\phi,\tau_R}(t,x)$ for all $x \in [-R,R]$. □

Proof of Proposition 5.8. We first derive an estimate on stopping times inferior to τ_x^R, borrowed from Carlier and Galichon's [7] Lemma 5.2. Let $x \in [-R,R]$, then for every stopping time $\tau \le \tau_x^R$, we have

$$\mathbb{P}(\tau \ge T) \;\le\; \mathbb{P}\big(\tau_x^R \ge T\big) \;\le\; \Pi_{1\le n\le T}\mathbb{P}\big(|B_n - B_{n-1}| \le 2R\big) \;\le\; e^{-\delta(T-1)}. \quad (5.13)$$

Recall that $\mathbb{E}\big[(x+B_\tau)^2\big] = x^2 + \mathbb{E}[\tau]$, $\forall \tau \le \tau_x^R$ from (5.3). Then by the definitions of λ^{ϕ,τ_R} and $\lambda^{\phi,T,R}$ in (5.9), for every $\phi \in \mathrm{Q}^{0,K,M}$,

$$\begin{aligned}\lambda^{\phi,\tau_R}(0,x) - \lambda^{\phi,T,R}(0,x) &\le \sup_{\tau\le\tau_x^R} \mathbb{E}\left[\, g^{\phi}(\tau, x+B_\tau) \;-\; g^{\phi}(\tau\wedge T, x+B_{\tau\wedge T})\,\right] \\ &= \sup_{\tau\le\tau_x^R} \mathbb{E}\left[\, \psi(\tau\wedge T, x+B_{\tau\wedge T}) - \psi(x+B_\tau, \tau)\,\right],\end{aligned}$$

where $\psi(t,x) := -g^{\phi}(t,x) - L_0 x^2 + L_0 t$. Clearly, ψ increases in t and $|\psi(t,x_1) - \psi(t,x_2)| \le 2(K+2L_0)(R^2 \vee 1)$, $\forall x_1,\; x_2 \in [-R,R]$ by Assumption 2.8. Therefore,

$$\begin{aligned}\lambda^{\phi,\tau_R}(0,x) - \lambda^{\phi,T,R}(0,x) &\le \sup_{\tau\le\tau_x^R} \mathbb{E}\left[\, \big|\psi(\tau\wedge T, x+B_{\tau\wedge T}) \;-\; \psi(\tau\wedge T, x+B_\tau)\big|\,\right] \\ &= \sup_{\tau\le\tau_x^R} \mathbb{E}\left[\, \big|\psi(T, x+B_T) \;-\; \psi(T, x+B_\tau)\big|\mathbf{1}_{\tau\ge T}\,\right] \\ &\le \sup_{\tau\le\tau_x^R} 2\,(K+2L_0)\,(R^2\vee 1)\,\mathbb{P}(\tau\ge T) \\ &\le Le^{-\delta(T-1)},\end{aligned}$$

where the last inequality is from (5.13). Finally, by the arbitrariness of $\phi \in \mathrm{Q}^{0,K,M}$ together with Proposition 5.7, we prove (5.11). □

Finally, we complete the proof for Theorem 2.11.

Proof of Theorem 2.11. We notice that for fixed constants K, M, R and T, the functions set Q_0 and function $\lambda^{0,\phi}$ defined in Section 2.3 is exactly $\mathrm{Q}^{0,K,M}$ and $\lambda^{\phi,T,R}$ defined above. We can then conclude the proof by Propositions 5.5, 5.6, 5.7 and 5.8 as well as (5.5). □

A A projection algorithm

We would like to propose an algorithm for the optimization problem (3.19) associated with the projection P_{Q_h}. In place of the problem (3.19), let us consider a more general problem.

Let $a = (a_i)_{1\le i\le m} \in \mathbb{N}^m$ and $A \in \mathbb{R}^+$ such that $0 < A < |a|$, where $|a| := \sum_{i=1}^m a_i$. We define a cone of nondecreasing vectors in $\mathbb{R}^n$ by

$$\mathcal{K}_m^a \;:=\; \Big\{\, \xi = (\xi_i)_{1\le i\le m} \in \mathbb{R}^m \;:\; \xi_1 \le \cdots \le \xi_m \,\Big\},$$

$$\mathcal{K}_m^A \;:=\; \Big\{\, \xi = (\xi_i)_{1\le i\le m} \in [0,1]^m \;:\; \sum_{i=1}^m a_i\xi_i = A \,\Big\}, \text{ and } \mathcal{K}_m^{a,A} \;:=\; \mathcal{K}_m^a \,\cap\, \mathcal{K}_m^A.$$

The projection $P_{\mathcal{K}_m^{a,A}}(z)$ of $z \in \mathbb{R}^m$ to $\mathcal{K}_m^{a,A}$ is to solve the optimization problem

$$\xi_m^{a,A,z} \quad := \quad \arg \min_{\xi \in \mathcal{K}_m^{a,A}} \ \sum_{i=1}^m a_i (z_i - \xi_i)^2. \tag{A.1}$$

Similarly, one can also define the projection $P_{\mathcal{K}_m^a}$(resp. $P_{\mathcal{K}_m^A}$) by the optimization problem (A.1), where $\mathcal{K}_m^{a,A}$ in the formula is replaced by $\mathcal{K}_m^a$ (resp. $\mathcal{K}_m^A$), and the projected element $\xi_m^{a,A,z}$ is replaced by $\xi_m^{a,z}$ (resp. $\xi_m^{A,z}$).

In the following, we shall show that

$$P_{\mathcal{K}_m^{a,A}} \quad = \quad P_{\mathcal{K}_m^{a,A}} \circ P_{\mathcal{K}_m^a} \quad = \quad P_{\mathcal{K}_m^A} \circ P_{\mathcal{K}_m^a},$$

and give the algorithms for both $P_{\mathcal{K}_m^a}$ and $P_{\mathcal{K}_m^A}$. With these algorithms, one can deduce easily an algorithm for the projections $P_{E^+_{K,M}}$ and $P_{E^-_{K,M}}$ in (3.19). We just remark that similar algorithms are discussed in Edelsbrunner [11, P. 143-145] in order to compute the convex envelope of functions defined on a discrete grid.

Given $a \in \mathbb{N}^m$ and $z \in \mathbb{R}^m$, we define $S^{a,z} \in \mathbb{R}^{|a|}$ by $S_k^{a,z} := z_j$ for $\sum_{i=1}^{j-1} a_i < k \le \sum_{i=1}^{j} a_i$, and define the function $F^{a,z}$ on the grid $\mathbb{N} \cap [0, 1 + |a|]$ by

$$F^{a,z}(0) \quad := \quad 0 \quad \text{and} \quad F^{a,z}(k) \quad := \quad \sum_{i=1}^k S_i^{a,z} \quad \text{for} \quad k = 1, \cdots, |a|. \tag{A.2}$$

Lemma A.1. *Suppose that we are given $z \in \mathbb{R}^m$ such that $z_k \ge z_{k+1}$, denote $\xi_m^{a,z} := P_{\mathcal{K}_m^a}(z)$ and $\xi_m^{a,A,z} := P_{\mathcal{K}_m^{a,A}}(z)$. Then $(\xi_m^{a,z})_k = (\xi_m^{a,z})_{k+1}$ and $(\xi_m^{a,A,z})_k = (\xi_m^{a,A,z})_{k+1}$. In particular, the projections $P_{\mathcal{K}_m^a}(z)$ and $P_{\mathcal{K}_m^{a,A}}(z)$ are equivalent to $P_{\mathcal{K}_{m-1}^{\tilde{a}}}(\tilde{z})$ and $P_{\mathcal{K}_{m-1}^{\tilde{a},A}}(\tilde{z})$ for*

$$\tilde{a}_i \ = \ \begin{cases} a_i, \quad 1 \le i \le k-1, \\ a_k + a_{k+1}, \quad i = k, \\ a_{i+1}, \quad k+1 \le i \le m-1, \end{cases} \qquad \text{and} \quad \tilde{z}_i \ = \ \begin{cases} z_i, \quad 1 \le i \le k-1, \\ \frac{a_k z_k + a_{k+1} z_{k+1}}{a_k + a_{k+1}}, \quad i = k, \\ z_{i+1}, \quad k+1 \le i \le m-1, \end{cases} \tag{A.3}$$

in sense that $S^{a,\xi_m^{a,z}} = S^{\tilde{a},\xi_{m-1}^{\tilde{a},\tilde{z}}}$ and $S^{a,\xi_m^{a,A,z}} = S^{\tilde{a},\xi_{m-1}^{\tilde{a},A,\tilde{z}}}$, where $\xi_{m-1}^{\tilde{a},\tilde{z}} := P_{\mathcal{K}_{m-1}^{\tilde{a}}}(\tilde{z})$ and $\xi_{m-1}^{\tilde{a},A,\tilde{z}} := P_{\mathcal{K}_{m-1}^{\tilde{a},A}}(\tilde{z})$.

Proof. Given an arbitrary $\xi \in \mathbb{R}^m$ such that $\xi_{k+1} > \xi_k$, then there is $\varepsilon > 0$ satisfying that $\xi_{k+1} = \xi_k + (1 + \frac{a_k}{a_{k+1}})\varepsilon$. Define $\hat{\xi} \in \mathbb{R}^m$ by $\hat{\xi}_i = \begin{cases} \xi_k + \varepsilon, & i = k, k+1, \\ \xi_i, & \text{otherwise}, \end{cases}$ one can show that

$$\sum_{i=1}^m \ a_i (\hat{\xi}_i - z_i)^2 \quad < \quad \sum_{i=1}^m \ a_i (\xi_i - z_i)^2. \tag{A.4}$$

Thus ξ cannot be the projection of z since $\xi \in \mathcal{K}_m^a$ (resp. $\mathcal{K}_m^{a,A}$) implies that $\hat{\xi} \in \mathcal{K}_m^a$ (resp. $\mathcal{K}_m^{a,A}$), and hence (A.4) contradicts the definition of the projection in (A.1). It follows that $(\xi_m^{a,z})_k = (\xi_m^{a,z})_{k+1}$ and $(\xi_m^{a,A,z})_k = (\xi_m^{a,A,z})_{k+1}$.

To show the inequality (A.4), it is enough to verify that

$$\begin{aligned}
&\sum_{i=1}^{m} a_i\ (\xi_i - z_i)^2 \quad - \quad \sum_{i=1}^{m} a_i\ (\hat{\xi}_i - z_i)^2 \\
= \quad & a_k\ (\xi_k - z_k)^2 \quad + \quad a_{k+1}\ \big(\xi_k + (1 + \frac{a_k}{a_{k+1}})\varepsilon - z_{k+1}\big)^2 \\
& \quad - \quad a_k\ (\xi_k + \varepsilon - z_k)^2 \quad - \quad a_{k+1}\ (\xi_k + \varepsilon - z_{k+1})^2 \\
= \quad & \frac{a_k}{a_{k+1}}\ (a_k + a_{k+1})\ \varepsilon^2 \ + \ 2\ a_k\ \varepsilon\ (z_k - z_{k+1}) \quad > \quad 0.
\end{aligned}$$

Finally, the equivalence between $P_{\mathcal{K}_m^a}(z)$ (resp. $P_{\mathcal{K}_m^{a,A}}(z)$) and $P_{\mathcal{K}_{m-1}^{\tilde{a}}}(\tilde{z})$ (resp. $P_{\mathcal{K}_{m-1}^{\tilde{a},A}}(\tilde{z})$) is from the fact that for every ξ such that $\xi_k = \xi_{k+1}$, one has the decomposition

$$\sum_{i=1}^{m} a_i(z_i - \xi_i)^2 \ = \ \sum_{i=1}^{m-1} \tilde{a}_i(\tilde{z}_i - \tilde{\xi}_i)^2 \ + a_k z_k^2 + a_{k+1} z_{k+1}^2 \ - \ (a_k + a_{k+1})\frac{(z_k + z_{k+1})^2}{4},$$

where $\tilde{\xi}_i \ = \begin{cases} \xi_i, & i \le k-1, \\ \xi_k, & i = k, k+1, \\ \xi_{i-1}, & k+2 \le i \le m-1. \end{cases}$ □

Lemma A.1 gives an algorithm for projection $P_{\mathcal{K}_m^a}$ which finishes within less than m steps. The algorithm also simplifies the projection $P_{\mathcal{K}_m^{a,A}}$, as we can see later in Proposition A.3.

Algorithm A.2. *For projection $P_{\mathcal{K}_m^a}(z)$:*

- *1, Given system parameters (m, a, z), stop if $m = 1$.*
- *2, Find k such that $z_k \ge z_{k+1}$, stop if it does not exist.*
- *3, With the found k in step 2, reduce parameters (m, a, z) to $(m-1, \tilde{a}, \tilde{z})$ as in equation* (A.3)*.*
- *4, Go to 1.*

Proposition A.3. *$P_{\mathcal{K}_m^{a,A}} = P_{\mathcal{K}_m^{a,A}} \circ P_{\mathcal{K}_m^a}$, and for every $z \in \mathbb{R}^m$. Moreover, $F^{a,\xi}$ (with $\xi := P_{\mathcal{K}_m^a}(z)$) is the convex envelope of $F^{a,z}$, where the functions $F^{a,\xi}$ and $F^{a,z}$ are define in* (A.2)

Proof. Suppose that the entrance data of Algorithm A.2 is (m_1, a_1, z_1) and the exit data is (m_2, a_2, z_2), then clearly $P_{\mathcal{K}_{m_2}^{a_2}}(z_2) = z_2$. And by Lemma A.1, we have $S^{a_1,\xi_1} = S^{a_2,z_2}$ (with $\xi_1 := P_{\mathcal{K}_{m_1}^{a_1}}(z_1)$) and $S^{a_1,\xi_1^A} = S^{a_2,\xi_2^A}$ (with $\xi_1^A := P_{\mathcal{K}_{m_1}^{a_1,A}}(z_1)$ and $\xi_2^A := P_{\mathcal{K}_{m_2}^{a_2,A}}(z_2)$), from which we deduce that $P_{\mathcal{K}_m^{a,A}} = P_{\mathcal{K}_m^{a,A}} \circ P_{\mathcal{K}_m^a}$.

To see that $F^{a,\xi}$ (with $\xi := P_{\mathcal{K}_m^a}(z)$) is the convex envelope of $F^{a,z}$, it is enough to verify that at every step in Algorithm A.2, $F^{\tilde{a},\tilde{z}}$ is greater than the convex envelope of $F^{a,z}$. And at the exit, $F^{a,\xi}$ is a convex function. □

Now, we shall prove that $P_{\mathcal{K}_m^{a,A}} \circ P_{\mathcal{K}_m^a} = P_{\mathcal{K}_m^A} \circ P_{\mathcal{K}_m^a}$, for this purpose, it is enough to show that for every $z \in \mathcal{K}_m^a$, $P_{\mathcal{K}_m^{a,A}}(z) = P_{\mathcal{K}_m^A}(z)$. In fact, we shall give an algorithm of projection $P_{\mathcal{K}_m^A}(z)$ for $z \in \mathcal{K}_m^a$, and then verify that $P_{\mathcal{K}_m^A}(z) \in \mathcal{K}_m^{a,A}$.

Given $\nu \in \mathbb{R}$ and $x \in \mathbb{R}^m$, denote by $z - \nu$ the sequence $(z_i - \nu)_{1 \le i \le m}$, and by z^ν the sequence $(z_i^\nu)_{1 \le i \le m} := (0 \vee (z_i - \nu) \wedge 1)_{1 \le i \le m}$.

Lemma A.4. *Given $\nu \in \mathbb{R}$ and $z \in \mathbb{R}^m$, we have $P_{\mathcal{K}_m^{a,A}}(z) = P_{\mathcal{K}_m^{a,A}}(z - \nu)$ and $P_{\mathcal{K}_m^A}(z) = P_{\mathcal{K}_m^A}(z - \nu)$. If in addition $z \in \mathcal{K}_m^a$, then there is $\hat{\nu} \in \mathbb{R}$ such that $\sum_{i=1}^m a_i z_i^{\hat{\nu}} = A$ and $P_{\mathcal{K}_m^A}(z) = P_{\mathcal{K}_m^{a,A}}(z) = z^{\hat{\nu}}$. And it follows that $P_{\mathcal{K}_m^{a,A}} = P_{\mathcal{K}_m^{a,A}} \circ P_{\mathcal{K}_m^a} = P_{\mathcal{K}_m^A} \circ P_{\mathcal{K}_m^a}$.*

Proof. To prove that $P_{\mathcal{K}_m^{a,A}}(z) = P_{\mathcal{K}_m^{a,A}}(z - \nu)$ or $P_{\mathcal{K}_m^A}(z) = P_{\mathcal{K}_m^A}(z - \nu)$, we note that for every $\xi \in \mathbb{R}^m$ satisfying $\sum_{i=1}^m a_i \xi_i = A$,

$$\sum_{i=1}^m a_i (z_i - \nu - \xi_i)^2 \;=\; \sum_{i=1}^m a_i (z_i - \xi_i)^2 \;+\; \nu^2 \sum_{i=1}^m a_i \;-\; 2\nu \Big(\sum_{i=1}^m a_i z_i - A \Big).$$

Then for the existence of $\hat{\nu}$, it is enough to see that $\nu \mapsto \sum_{i=1}^m a_i z_i^\nu$ is continuous, and that $0 < A < \sum_{i=1}^m a_i$ is supposed at the beginning of the section. Clearly, by its definition, z^ν is the projected element of $z - \nu$ to $[0,1]^m$ in sense that $\xi_0 = z^\nu$ minimizes $\sum_{i=1}^m a_i (z_i - \nu - \xi_i)^2$ among all $\xi \in [0,1]^m$. Then for $z \in \mathcal{K}_m^a$, it is easy to verify that $z^{\hat{\nu}} \in \mathcal{K}_m^{a,A} \subset \mathcal{K}_m^A \subset [0,1]^m$ with the found $\hat{\nu}$. Therefore $P_{\mathcal{K}_m^A}(z) = P_{\mathcal{K}_m^{a,A}}(z) = P_{\mathcal{K}_m^A}(z - \hat{\nu}) = P_{\mathcal{K}_m^{a,A}}(z - \hat{\nu}) = z^{\hat{\nu}}$. □

Algorithm A.5. *To find $\hat{\nu}$ such that $\sum_{i=1}^m a_i z_i^{\hat{\nu}} = A$:*

- *1, Set $z_0 = -\infty$ and $z_{m+1} = \infty$.*
- *2, Find k such that $\sum_{i=1}^m a_i z_i^{z_{k-1}} \ge A$ and $\sum_{i=1}^m a_i z_i^{z_k} < A$, then $z_{k-1} \le \hat{\nu} < z_k$.*
- *3, Find j such that $\sum_{i=1}^m a_i z_i^{z_{j+1}-1} < A$ and $\sum_{i=1}^m a_i z_i^{z_j - 1} \ge A$, then $z_j - 1 \le \hat{\nu} < z_{j+1} - 1$.*
- *4, Set $\hat{\nu} = \frac{\sum_{i=j+1}^m a_i + \sum_{i=k}^j a_i z_i - A}{\sum_{i=k}^j a_i}$ when $k \le j$, or $\hat{\nu} = z_{k-1}$ when $k = j+1$.*

By the way how to find k and j, it follows that $z_{k-1} \le \hat{\nu} < z_{j+1} - 1 < z_{j+1}$, hence $k \le j+1$. Then step 4 of Algorithm A.5 gives the right $\hat{\nu}$ since $z_i^{\hat{\nu}} = \begin{cases} 0, & \text{if } i \le k-1, \\ 1, & \text{if } i \ge j+1, \\ z_i - \hat{\nu}, & \text{otherwise.} \end{cases}$ for k, j found in step 2 and 3, and hence for $k \le j$,

$$\sum_{i=k}^j a_i (z_i - \hat{\nu}) \;+\; \sum_{i=j+1}^m a_i = A \quad \Longrightarrow \quad \hat{\nu} \;=\; \frac{\sum_{i=j+1}^m a_i + \sum_{i=k}^j a_i z_i - A}{\sum_{i=k}^j a_i}.$$

Acknowledgments

The authors thank Nicole El Karoui, Nizar Touzi and Pierre Henry-Labordère for fruitful discussions, and an anonymous referee for his/her useful comments.

References

[1] G. Barles, C. Daher, and M. Romano, *Convergence of numerical schemes for parabolic equations arising in finance theory*, Math. Models Methods Appl. Sci., 5(1):125-143, 1995.

[2] G. Barles and E.R. Jakobsen, *Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations*, Math. Comp., 76(260):1861-1893 (electronic), 2007.

[3] G. Barles and P.E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic Anal., 4(3):271-283, 1991.

[4] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[5] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Cambridge, MA., 1999.

[6] B. Bouchard and N. Touzi, *Weak Dynamic Programming Principle for Viscosity Solutions*, SIAM J. Control Optim., 49(3), 948-962, 2011.

[7] G. Carlier and A. Galichon, *Exponential convergence for a convexifying equation and a non-autonomous gradient flow for global minimization*, ESAIM: Control, Optimisation and Calculus of Variations (COCV), 18(3), 611-620, 2011.

[8] P. Carr and R. Lee, *Hedging variance options on continuous semimartingales*, Finance Stoch., 14(2):179-207, 2010.

[9] A.M.G. Cox and J. Wang, *Root's Barrier: Construction, Optimality and Applications to Variance Options*, Ann. Appl. Probab., to appear.

[10] B. Dupire, *Volatility derivatives modeling*, Bloomberg LP, 2005.

[11] H. Edelsbrunner, *Algorithms in combinatorial geometry*, volume 10 of *EATCS Monographs on Theoretical Computer Science*, Springer-Verlag, Berlin, 1987.

[12] A. Galichon, P. Henry-Labordère, and N. Touzi, *A stochastic control approach to no-arbitrage bounds given marginals, with an application to Lookback options*, preprint.

[13] D. Hobson, *Robust hedging of the lookback option*, Finance and Stochastics, 2:329-347, 1998.

[14] D. Hobson, *The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices*, Paris-Princeton Lectures on Mathematical Finance, 2010.

[15] I. Karatzas and S.E. Shreve, *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*, Springer-Verlag, New York, second edition, 1991.

[16] N. Krylov, *The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients*, Appl. Math. Optim., 52(3):365-399, 2005.

[17] H.J. Kushner and P. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Second edition. Applications of Mathematics (New York), 24. Stochastic Modelling and Applied Probability. Springer-Verlag, New York, 2001

[18] J. Obłój, *The Skorokhod embedding problem and its offspring*, Probab. Surv., 1:321-390 (electronic), 2004.

[19] D.H. Root, *The existence of certain stopping times on Brownian motion*, Ann. Math. Statist., 40:715-718, 1969.

[20] H. Rost, *Skorokhod stopping times of minimal variance*, in Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire), pages 194-208. Lecture Notes in Math., Vol 511. Springer, Berlin, 1976.

[21] M. Soner, N. Touzi and J. Zhang, *Wellposedness of second order backward SDEs*, Probability Theory and Related Fields, 153, 149-190, 2012.

[22] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist., 36:423-439, 1965.

[23] J. Szpirglas and G. Mazziotto, *Théorème de séparation dans le problème d'arrêt optimal*, Séminaires de Probabilités, XIII: 32, 378-384, LNM 721, 1979.

[24] X. Tan and N. Touzi, *Optimal Transportation under Controlled Stochastic Dynamics*, Annals of Probability, to appear.

[25] N. Touzi, *Optimal Stochastic Control, Stochastic Target Problem, and Backward SDEs*, Lecture notes at Fields Institute, 2010.