



**HAL**  
open science

# DUALITY-BASED A POSTERIORI ERROR ESTIMATES FOR SOME APPROXIMATION SCHEMES FOR CONVEX OPTIMAL CONTROL PROBLEMS

Athena Picarelli, Christoph Reisinger

► **To cite this version:**

Athena Picarelli, Christoph Reisinger. DUALITY-BASED A POSTERIORI ERROR ESTIMATES FOR SOME APPROXIMATION SCHEMES FOR CONVEX OPTIMAL CONTROL PROBLEMS. 2017. hal-01538617v1

**HAL Id: hal-01538617**

**<https://hal.science/hal-01538617v1>**

Preprint submitted on 13 Jun 2017 (v1), last revised 11 Mar 2019 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# DUALITY-BASED A POSTERIORI ERROR ESTIMATES FOR SOME APPROXIMATION SCHEMES FOR CONVEX OPTIMAL CONTROL PROBLEMS

ATHENA PICARELLI AND CHRISTOPH REISINGER

ABSTRACT. We introduce a class of numerical schemes for Hamilton-Jacobi-Bellman equations based on a novel Markov chain approximation of the associated optimal control problem, which uses, in turn, a piecewise constant policy approximation, Euler-Maruyama time stepping, and a Gauß-Hermite approximation of the Gaussian increments. We provide one-sided (lower) error bounds of order arbitrarily close to 1/2 in time and 1/3 in space for Lipschitz viscosity solutions, coupling probabilistic arguments with regularization techniques as introduced by Krylov. The order for sufficiently regular solutions is 1 in both time and space. For a class of convex problems arising in optimal investment, we use duality results to derive also *a posteriori* upper error bounds which are empirically of the same order as the lower bounds, as confirmed in our numerical tests.

## 1. INTRODUCTION

Let  $(\Omega, \mathbb{F}, \mathbb{P})$  be a probability space with filtration  $\{\mathbb{F}_t, t \geq 0\}$  induced by an  $\mathbb{R}^p$ -Brownian motion  $\mathcal{B}(\cdot)$  ( $p \geq 1$ ). We consider a controlled process governed by

$$\begin{cases} dX(s) = \mu(s, X(s), \alpha_s)ds + \sigma(s, X(s), \alpha_s) \cdot d\mathcal{B}(s), & s \in (t, T), \\ X(t) = x, \end{cases} \quad (1.1)$$

where  $\mu$  and  $\sigma$  take values, respectively, in  $\mathbb{R}$  and  $\mathbb{R}^p$ . We assume that the control vector process  $\alpha \in \mathcal{A}$  is progressively measurable and takes values in the set  $A \subseteq \mathbb{R}^d$ . For any  $x \in \mathbb{R}$ , we will denote by  $X_{t,x}^\alpha(\cdot)$  the strong solution of (1.1).

Given a utility function  $U$ , the aim is to maximize for any  $x \geq 0$  the quantity

$$J(t, x) := \mathbb{E}[U(X_{t,x}^\alpha(T))].$$

It is well known that this problem is related to the solution of a Hamilton-Jacobi-Bellman (HJB) equation, of the form

$$\begin{cases} -\partial_t v + \inf_{a \in A} \left\{ \mu(t, x, a) \partial_x v - \frac{1}{2} \sigma \sigma^T(t, x, a) \partial_{xx} v \right\} = 0, & (t, x) \in [0, T) \times \mathbb{R}, \\ v(T, x) = U(x), & x \in \mathbb{R}. \end{cases} \quad (1.2)$$

The possible degeneracy of the diffusion term makes it necessary to consider solutions in the viscosity sense (see [7]). Furthermore, explicit solutions for this kind of nonlinear equations are rarely available, so that their numerical approximation becomes vital. The seminal work by Barles and Souganidis [4] establishes the basic framework for convergence of numerical schemes to the viscosity solution. The fundamental properties required are: monotonicity, consistency, and stability of the scheme. We recall that, in multiple dimensions, standard finite difference schemes are in general non-monotone. As an alternative to finite difference schemes, semi-Lagrangian (SL) schemes [20, 6, 9] are monotone by construction. The scheme we present in this paper belongs to this family.

In general, the provable order of convergence for second order HJB equations is significantly less than one. By a technique pioneered by Krylov based on “shaking the coefficients” and mollification to construct smooth sub- and/or super-solutions, [15, 17, 1, 2, 3] prove certain fractional convergence orders, mainly using PDE-based techniques i.e., a comparison principle and the consistency error. Here we study a family of SL schemes based on a discrete time approximation of the optimal

control problem. We aim to provide error estimates making use of purely probabilistic techniques and directly comparing the two optimal control problems. An important step in order to define our scheme is to approximate the set of controls  $\mathcal{A}$  by piecewise constant controls. This introduces an asymmetry between the upper and the lower bound of the error. For a particular class of convex optimal control problems, typically utility maximisation problems arising in financial applications, we propose to overcome this issue using information coming from a dual problem.

The main contributions of this paper are as follows. We propose new discrete approximations of controlled diffusion processes based on piecewise constant controls over intervals of length  $h$  and  $M$  Gauß-Hermite points. We present a novel analysis technique for the resulting semi-discrete approximations by probabilistic techniques and direct use of the dynamic programming principle. This allows us to derive one-sided error bounds of order  $h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)}$  for timestep  $h$  and spatial mesh size  $\Delta x$ , for Lipschitz viscosity solutions (assumptions (H1) to (H3) below). They coincide with the two-sided bounds in [9] for the standard linear-interpolation semi-Lagrangian scheme, i.e.  $M = 2$ , and improve them for  $M > 2$ . For sufficiently smooth solutions, the corresponding error bounds are of order 1 in both  $h$  and  $\Delta x$ . Moreover, for convex control problems for which duality holds (such as under assumptions (H4) to (H7) below), two-sided *a posteriori* bounds of the same order are obtained.

The paper is organised as follows. We present the setting and main assumptions for the optimal control problem in Section 2. For a very general class of discrete time schemes, in Section 3 we describe the main idea of the paper and explain how duality techniques can be used for obtaining error estimates. In Section 4 the Markov-chain approximation scheme is introduced and one-sided (lower) *a priori* error bounds are obtained by using purely probabilistic arguments. The dual problem, its numerical approximation and the *a posteriori* error bounds are discussed in Sections 6 and 7. Numerical tests are presented in Section 8.

## 2. MAIN ASSUMPTIONS AND PRELIMINARY RESULTS

We consider standard assumptions on the optimal control problem:

**(H1)**  $A$  is a compact set;

**(H2)**  $\mu : [0, T] \times \mathbb{R} \times A \rightarrow \mathbb{R}$  and  $\sigma : [0, T] \times \mathbb{R} \times A \rightarrow \mathbb{R}^p$  are continuous functions and there exists  $K_0 \geq 0$  such that

$$\begin{aligned} |\mu(t, x, a) - \mu(t, y, a)| + \|\sigma(t, x, a) - \sigma(t, y, a)\| &\leq K_0|x - y| \quad \forall t \in [0, T], x, y \in \mathbb{R}, a \in A, \\ |\mu(t, x, a) - \mu(s, x, a)| + \|\sigma(t, x, a) - \sigma(s, x, a)\| &\leq K_0|t - s|^{1/2} \quad \forall t, s \in [0, T], x \in \mathbb{R}, a \in A; \end{aligned}$$

**(H3)**  $U$  is a Lipschitz continuous function with Lipschitz constant  $L$ .

For any choice of the control  $\alpha \in \mathcal{A}$  we will denote by  $X_{t,x}^\alpha(\cdot)$  the unique strong solution of (1.1). The value function of the optimal control problem we consider is defined by

$$v(t, x) := \sup_{\alpha \in \mathcal{A}} \mathbb{E}[U(X_{t,x}^\alpha(T))]. \quad (2.1)$$

Under these assumptions one can prove the following regularity result on  $v$ :

**Proposition 2.1** ([24, Proposition 3.1, Chapter IV]). *Let (H1)-(H3) be satisfied. There exists  $K \geq 0$  (depending only on  $T$  and the constant  $K_0$  in (H2)) such that*

$$|v(t, x) - v(s, y)| \leq LK(|x - y| + (1 + |x||y|)|t - s|^{1/2})$$

for any  $x, y \in \mathbb{R}$  and  $t, s \in [0, T]$  (where  $L$  denotes the Lipschitz constant of  $U$  and  $K$  only depends on  $T$  and  $K_0$  in assumption (H2)).

For any  $[0, T - t]$ -valued stopping time  $\theta$ ,  $v$  satisfies the Dynamic Programming Principle (DPP)

$$v(t, x) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}\left[v(t + \theta, X_{t,x}^\alpha(t + \theta))\right], \quad (2.2)$$

which can be used to characterize  $v$  as the viscosity solution of (1.2).

### 3. THE MAIN IDEA

In this section, we are going to heuristically describe the main idea of our approach for a very general class of discrete-time schemes. To this end, we introduce a time mesh for  $N \geq 1$  by

$$h = T/N \quad \text{and} \quad t_n = nh,$$

for  $n = 0, \dots, N$ .

**3.1. Piecewise constant control approximation.** The first step in our approximation is to introduce a time discretization of the control set. We consider the set  $\mathcal{A}_h$  of controls  $\alpha \in \mathcal{A}$  which are constant in each interval  $[t_n, t_{n+1}]$ , for  $n = 0, \dots, N-1$ , i.e.

$$\mathcal{A}_h := \left\{ \alpha \in \mathcal{A} : \alpha_s \equiv \sum_{i=0}^{N-1} a_i \mathbb{1}_{s \in [t_i, t_{i+1})} \text{ s.t. } a_i \in A, \quad i = 0, \dots, N-1 \right\}.$$

In what follows, we will identify any element of  $\alpha \in \mathcal{A}_h$  by the sequence of random variables  $a_i$  taking values in  $A$  (denoted  $a_i \in A$  for simplicity) and will write  $\alpha \equiv (a_0, \dots, a_{N-1})$ . We denote by  $v_h$  the value function obtained by restricting the supremum in (2.1) to controls in  $\mathcal{A}_h$ , that is

$$v_h(t, x) := \sup_{\alpha \in \mathcal{A}_h} \mathbb{E}[U(X_{t,x}^\alpha(T))]. \quad (3.1)$$

Clearly, since  $\mathcal{A}_h \subseteq \mathcal{A}$ , one has

$$v(t, x) \geq v_h(t, x), \quad (3.2)$$

for any  $t \in [0, T]$ ,  $x \in \mathbb{R}$ . An upper bound for the error related to this first approximation is obtained (under suitable assumptions, and for a bounded set  $A$ ) by Krylov in [18],

$$v(t, x) \leq v_h(t, x) + Ch^{1/6} \quad (3.3)$$

for some constant  $C$  possibly depending on  $x$ . The arguments we are going to develop in Sections 6 and 7 aim to derive error bounds using purely probabilistic arguments and without making use of such a (pessimistic, in many cases) result.

Taking  $\theta \equiv h$  in (2.2), the DPP for the value function  $v_h$  reads

$$v_h(t, x) = \sup_{a \in A} \mathbb{E}[v_h(t+h, X_{t,x}^a(t+h))]. \quad (3.4)$$

In particular, the restriction of the control set to  $\mathcal{A}_h$  implies that the supremum in (3.4) is taken over the set of control values  $A$  (compare this with (2.2)). The family of schemes we have in mind are recursively defined by an approximation of (3.4) and lead to the definition of a numerical solution  $V$  approximating  $v_h$ .

**3.2. Markov chain approximation.** Let us consider a numerical approximation  $\mathcal{S}$  of the argument of the supremum in (3.4) such that

$$\mathcal{S}[f(t_{n+1}, \cdot)](t_n, x, a) \approx \mathbb{E}\left[f(t_{n+1}, X_{t_n, x}^a(t_{n+1}))\right],$$

for any smooth function  $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$  and  $a \in A$ . Then we consider  $V$  defined by

$$\begin{cases} V(t_N, x) &= U(x), \\ V(t_n, x) &= \sup_{a \in A} \mathcal{S}[V(t_{n+1}, \cdot)](t_n, x, a), \quad n = N-1, \dots, 0. \end{cases}$$

If the set  $A$  is compact and the scheme  $\mathcal{S}$  is sufficiently accurate, we hope to obtain (after a regularization procedure, if necessary) the following bound for the error between  $v_h$  and  $V$ :

$$\begin{aligned} & \sup_{n=0,\dots,N} \left| v_h(t_n, x) - V(t_n, x) \right| \\ & \leq \sup_{n=0,\dots,N-1} \sup_{a \in A} \left| \mathbb{E} \left[ v_h(t_{n+1}, X_{t_n, x}^a(t_{n+1})) \right] - \mathcal{S}[V(t_{n+1}, \cdot)](t_n, x, a) \right| \\ & \leq \dots \leq Ch^\beta, \end{aligned}$$

for some  $\beta \gg 1/6$  and  $C$  eventually depending on  $x$ . Considering this together with (3.2)-(3.3), we can provide an accurate lower bound for the quantity  $(v - V)$ , whereas the upper bound on the error will remain restricted to order  $1/6$ , irrespective of the accuracy of the approximation of  $v_h$  we consider. The situation is illustrated in Figure 1.

In the absence of further information, this is the best order we can establish theoretically in this way, even if the practically observed order is typically much larger than  $1/6$ . We therefore propose a way to bypass this bottleneck.

**3.3. New error estimates using duality results.** For a class of optimal control problems which have a special convex structure, it is possible to define a dual problem of the form

$$\tilde{v}(t, y) := \inf_{\nu \in \mathcal{V}} \mathbb{E}[\tilde{U}(Y_{t,y}^\nu(T))], \quad (3.5)$$

for some given dynamics  $Y_{t,y}^\nu(\cdot)$  and cost function  $\tilde{U}$ , such that the primal and dual value functions are related by a conjugate relation:

$$v(t, x) = \inf_{y > 0} \{ \tilde{v}(t, y) + xy \}. \quad (3.6)$$

Applying the same steps as above to the dual problem and defining its numerical approximation  $\tilde{V}$ , the situation in terms of error estimates is reversed, due to the fact that we face a minimization problem. This means that if the duality result (3.6) holds true, we are in the situation shown in Figure 2. Here we also use the fact that if

$$|\tilde{v}(t_n, y) - \tilde{V}(t_n, y)| \leq Ch^\beta$$

for some  $C \equiv C(y)$ , then

$$\left| \inf_{y \geq 0} \{ \tilde{v}(t_n, y) + xy \} - \inf_{y \geq 0} \{ \tilde{V}(t_n, y) + xy \} \right| \leq \tilde{C}h^\beta$$

for some  $\tilde{C} \equiv \tilde{C}(x)$ .

These considerations can be used in order to provide:

- *A posteriori* error estimates based on the computation of the quantity

$$\left| V(t_n, x) - \inf_{y \geq 0} \{ \tilde{V}(t_n, y) + xy \} \right|.$$

In fact, if  $V$  and  $\inf_{y \geq 0} \{ \tilde{V}(t, y) + xy \}$  are seen empirically to be at a distance of order  $h^\beta$  or less, this will guarantee that the scheme considered has order of convergence at least  $\beta$ . Observe that the quantity above can be easily computed from  $V$  and  $\tilde{V}$  as discussed in [11, Section 3.4].

- *A priori* error estimates if a duality relation holds not only for the continuous problem, but also for the approximated value functions  $V$  and  $\tilde{V}$ . If we assume to be able to prove that the relation

$$V(t_n, x) = \inf_{y \geq 0} \{ \tilde{V}(t_n, y) + xy \}$$

holds up to some error of order greater or equal than  $\beta$ , we obtain that the *a priori* bound on the error  $|V - v|$  reduces to the narrow area around  $v$  that appears in Figure 2.

We pursue the first approach here and plan to investigate the second one in future work.

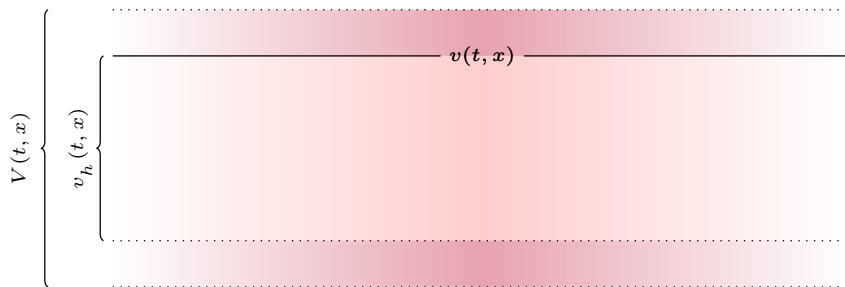


FIGURE 1. Error bounds without duality arguments.

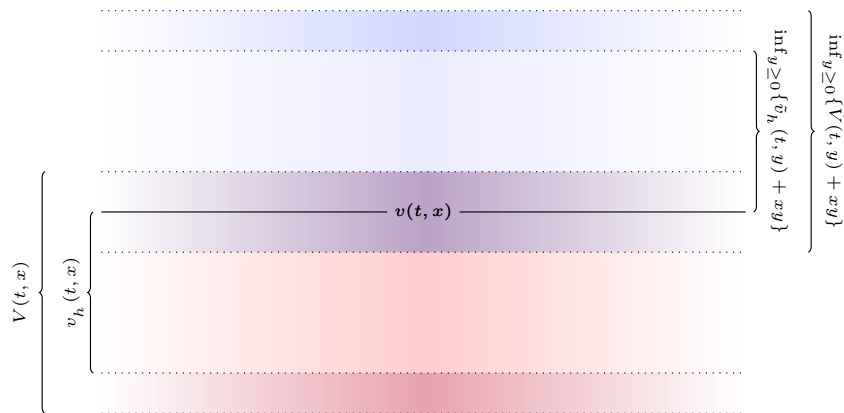


FIGURE 2. Error bounds with duality arguments.

#### 4. MARKOV-CHAIN APPROXIMATION SCHEMES

We now present a particular class of schemes which are based on a Markov chain approximation of the optimal control problem (3.1). This follows the classical philosophy presented in [19], although they take the opposite direction and use finite difference approximations to construct Markov chains, while here we use time stepping schemes and quadrature formulae to define semi-Lagrangian schemes. Similar probabilistic interpretations of such schemes have been given in [6] for the time-dependent case and in [20] for the infinite horizon case. What is new here is the construction of schemes with provable higher order error bounds, and the direct use of the dynamic programming principle for the discrete approximation to derive these bounds.

**4.1. Euler-Maruyama scheme.** We start with an approximation of the process  $X_{t,x}^\alpha(\cdot)$  by the Euler-Maruyama scheme. For any given  $\alpha \equiv (a_0, \dots, a_{N-1}) \in \mathcal{A}_h$ , we consider the following recursive relation:

$$X(t_{i+1}) = X(t_i) + \mu(t_i, X(t_i), a_i)h + \sigma(t_i, X(t_i), a_i) \cdot \Delta \mathcal{B}_i \quad (4.1)$$

for  $i = 0, \dots, N - 1$ . The increments  $\Delta \mathcal{B}_i := (\mathcal{B}(t_{i+1}) - \mathcal{B}(t_i))$  are independent, identically distributed random variables such that

$$\Delta \mathcal{B}_i \sim \sqrt{h} \mathcal{N}(0, I_p) \quad \forall i = 0, \dots, N - 1. \quad (4.2)$$

We will denote by  $X_{t_n,x}^{\alpha,EM}(\cdot)$  the solution to (4.1) associated with the control  $\alpha \equiv (a_n, \dots, a_{N-1}) \in \mathcal{A}_h$  and such that  $X_{t_n,x}^{\alpha,EM}(t_n) = x$ . Under assumptions (H1)-(H2), the rate of strong convergence of the scheme (4.1) is  $1/2$  [13]. Although the result from there is not directly applicable here as the coefficients are non-Lipschitz in time due to the jumps in the control process, we can follow

the same steps as in the proof of [21, Theorem 1.1, Chapter I], utilising the fact that the controls  $\alpha \in \mathcal{A}_h$  are constant over individual timesteps. Therefore we have:

**Proposition 4.1** ([13, Theorem 10.3.5]). *Let assumptions (H1)-(H2) be satisfied. Then there exists a constant  $K_{EM}$  (independent of  $h$ ) such that*

$$\mathbb{E} \left[ |X_{t_n, x}^\alpha(T) - X_{t_n, x}^{\alpha, EM}(T)| \right] \leq K_{EM}(1 + |x|)h^{1/2}$$

for any  $\alpha \in \mathcal{A}_h, n = 0, \dots, N, x \in \mathbb{R}$ .

As a consequence, thanks to the Lipschitz continuity of  $U$ , we can conclude that, denoting

$$v_{EM}(t_n, x) := \sup_{\alpha \in \mathcal{A}_h} \mathbb{E} \left[ U(X_{t_n, x}^{\alpha, EM}(T)) \right],$$

for any  $n = 0, \dots, N-1, x \geq 0$ ,

$$\begin{aligned} |v_h(t_n, x) - v_{EM}(t_n, x)| &\leq \sup_{\alpha \in \mathcal{A}_h} \left| \mathbb{E} [U(X_{t_n, x}^\alpha(T))] - \mathbb{E} [U(X_{t_n, x}^{\alpha, EM}(T))] \right| \\ &\leq LK_{EM}(1 + |x|)h^{1/2}. \end{aligned} \quad (4.3)$$

Moreover,  $v_{EM}$  still satisfies a DPP,

$$v_{EM}(t_n, x) = \sup_{a \in A} \mathbb{E} \left[ v_{EM}(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1})) \right], \quad (4.4)$$

for any  $n = 0, \dots, N-1$ .

**4.2. Gauß-Hermite quadrature.** Recalling that  $\Delta \mathcal{B}_i \sim \sqrt{h} \mathcal{N}(0, I_p)$ , we can also write (4.4) as

$$v_{EM}(t_n, x) = \sup_{a \in A} \int_{\mathbb{R}^p} v_{EM} \left( t_{n+1}, x + \mu(t_n, x, a)h + \sqrt{h}\sigma(t_n, x, a) \cdot y \right) \frac{e^{-\frac{|y|^2}{2}}}{(2\pi)^{p/2}} dy. \quad (4.5)$$

The discrete-time scheme we are going to define is based on the Gauß-Hermite approximation of the right-hand term in (4.5).

Let us consider for simplicity  $p = d = 1$ . Let  $M \geq 2$  and let us denote by  $\{z_i\}_{i=1, \dots, M}$  the zeros of the Hermite polynomial  $H_M$  of order  $M$  and by  $\{\omega_i\}_{i=1, \dots, M}$  the corresponding weights given by

$$w_i = \frac{2^{M-1} M! \sqrt{\pi}}{M^2 [H_{M-1}(z_i)]^2}, \quad i = 1, \dots, M.$$

Therefore, defining

$$\lambda_i := \frac{\omega_i}{\sqrt{\pi}} \quad \text{and} \quad \xi_i := \sqrt{2}z_i, \quad i = 1, \dots, M,$$

for any smooth real-valued function  $f$  (say  $f$  at least  $C^{2M}$ ) we can make use of the following approximation (see [12, p. 395] for instance):

$$\int_{-\infty}^{+\infty} f(y) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = \int_{-\infty}^{+\infty} f(\sqrt{2}y) \frac{e^{-y^2}}{\sqrt{\pi}} dy \approx \sum_{i=1}^M \frac{1}{\sqrt{\pi}} w_i f(\sqrt{2}z_i) = \sum_{i=1}^M \lambda_i f(\xi_i). \quad (4.6)$$

Observe that  $\lambda_i \geq 0, \forall i = 1, \dots, M$  and  $\sum_{i=1}^M \lambda_i = 1$ . Defining the sequence of i.i.d. random variables  $\{\zeta_n\}_{n=0, \dots, N-1}$  such that for any  $n = 0, \dots, N-1$

$$\mathbb{P}(\zeta_n = \xi_i) = \lambda_i, \quad i = 1, \dots, M,$$

we have

$$\mathbb{E}[\zeta_n] = 0 \quad \text{and} \quad \text{Var}[\zeta_n] = 1 \quad \forall n = 0, \dots, N-1.$$

	$\xi_i$	$\lambda_i$
$M = 2$	$\pm 1$	$1/2$
$M = 3$	$0$	$2/3$
	$\pm\sqrt{3}$	$1/6$
$M = 4$	$\pm\sqrt{3 - \sqrt{6}}$	$(3 + \sqrt{6})/12$
	$\pm\sqrt{3 + \sqrt{6}}$	$(3 - \sqrt{6})/12$

FIGURE 3. Analytical expressions of  $\{(\xi_i, \lambda_i)\}_{i=1, \dots, M}$  for  $M = 2, 3, 4$ . We refer to [5, p. 464] for numerical approximations of  $\{(z_i, w_i)\}_{i=1, \dots, M}$  for larger  $M$ .

For any control  $\alpha \equiv (a_n, \dots, a_{N-1}) \in \mathcal{A}_h$ , in the sequel we will denote by  $X_{t_n, x}^{\alpha, h}(\cdot)$  the Markov chain approximation of the process  $X_{t_n, x}^{\alpha, EM}(\cdot)$  recursively defined by

$$\begin{cases} X_{t_n, x}^{\alpha, h}(t_n) = x, \\ X_{t_n, x}^{\alpha, h}(t_{i+1}) = X_{t_n, x}^{\alpha, h}(t_i) + \mu(t_i, X_{t_n, x}^{\alpha, h}(t_i), a_i)h + \sqrt{h}\sigma(t_i, X_{t_n, x}^{\alpha, h}(t_i), a_i)\zeta_i, \end{cases} \quad (4.7)$$

for  $i = n, \dots, N-1$ . Therefore, starting from (4.5) and applying the Gauß-Hermite quadrature formula (4.6), our scheme will be defined by

$$\begin{cases} V(t_n, x) = \sup_{a \in A} \sum_{i=1}^M \lambda_i V(t_{n+1}, x + \mu(t_n, x, a)h + \sqrt{h}\sigma(t_n, x, a)\xi_i) \\ \quad = \sup_{a \in A} \mathbb{E} \left[ V(t_{n+1}, X_{t_n, x}^{\alpha, h}(t_{n+1})) \right], & n = N-1, \dots, 0, \\ V(t_N, x) = U(x). \end{cases} \quad (4.8)$$

Referring to the notation in Section 3, one has

$$\mathcal{S}[V(t_{n+1}, \cdot)](t_n, x, a) := \mathbb{E} \left[ V(t_{n+1}, x + h\mu(t_n, x, a) + \sqrt{h}\sigma(t_n, x, a)\zeta_n) \right].$$

**Remark 1.** We point out that for  $M = 2$ , (4.10) is the semi-Lagrangian (SL) scheme introduced by Camilli and Falcone in [6] (for now, without considering the interpolation on the space grid).

Iterating, we obtain the following representation formula for  $V$ :

$$V(t_n, x) = \sup_{\alpha \in \mathcal{A}_h} \mathbb{E} \left[ U(X_{t_n, x}^{\alpha, h}(T)) \right].$$

**Remark 2.** In the case of a multi-dimensional Brownian motion, one possibility to define an approximation is by a tensor product of the formula (4.6), that is

$$\int_{\mathbb{R}^p} f(\sqrt{2}y) \frac{e^{-|y|^2}}{\pi^{d/2}} dy \approx \sum_{i_1, \dots, i_p=1}^M \lambda_{i_1} \cdots \lambda_{i_p} f(\xi_{i_1}, \dots, \xi_{i_p}), \quad (4.9)$$

and  $\{\zeta_i\}_{i=n, \dots, N-1}$  are i.i.d. random variables in  $\mathbb{R}^p$  such that if  $\xi_j \equiv (\xi_{j_1}, \dots, \xi_{j_p})$ ,

$$\mathbb{P}(\zeta_i = \xi_j) = \lambda_{j_1} \cdots \lambda_{j_p}.$$

Then one can define an approximation to  $v$  by

$$\begin{cases} V(t_n, x) = \sup_{a \in A} \sum_{i_1, \dots, i_p=1}^M \lambda_{i_1} \cdots \lambda_{i_p} V(t_{n+1}, x + \mu(t_n, x, a)h + \sqrt{h}\sigma(t_n, x, a) \cdot \xi_i) \\ \quad = \sup_{a \in A} \mathbb{E} \left[ V(t_{n+1}, X_{t_n, x}^{\alpha, h}(t_{n+1})) \right], & n = N-1, \dots, 0, \\ V(t_N, x) = U(x). \end{cases} \quad (4.10)$$

In the following, we restrict ourselves to the case  $p = d = 1$  for brevity.



**4.3. The rate of weak convergence.** In this section we prove the rate of weak convergence of the random walk  $X_{t_n,x}^{\alpha,h}(\cdot)$  defined by (4.7) to the process  $X_{t_n,x}^{\alpha,EM}(\cdot)$  given by the Euler-Maruyama scheme (4.1). For a smooth function  $f : \mathbb{R} \rightarrow \mathbb{R}$  we aim to estimate

$$\sup_{a \in A} \left| \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,EM}(t_{n+1})) \right] - \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,h}(t_{n+1})) \right] \right|,$$

which leads to the weak convergence error for the approximation in (4.7).

**Proposition 4.2.** *Let assumptions (H1)-(H2) be satisfied and let  $M \geq 2$ . Then there exists a constant  $C \equiv C(M)$  such that for any function  $f \in C^{2M}(\mathbb{R}; \mathbb{R})$  one has*

$$\left| \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,EM}(t_{n+1})) \right] - \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,h}(t_{n+1})) \right] \right| \leq C \|f^{(2M)}\|_{\infty} (1 + x^{2M}) h^M,$$

for any  $x \in \mathbb{R}$ ,  $a \in A$ ,  $h \geq 0$  and  $n = 0, \dots, N-1$ .

*Proof.* We adapt a standard argument from numerical quadrature. Let us denote  $z = x + h\mu(t_n, x, a)$ . By Taylor expansion, we can write

$$\begin{aligned} \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,EM}(t_{n+1})) \right] &= \int_{-\infty}^{+\infty} f(z + \sqrt{2h}\sigma(t_n, x, a)y) \frac{e^{-y^2}}{\sqrt{\pi}} dy \\ &= \int_{-\infty}^{+\infty} \left\{ \sum_{k=0}^{2M-1} \frac{f^{(k)}(z)}{k!} (\sqrt{2h}\sigma(t_n, x, a)y)^k + \frac{f^{(2M)}(\hat{z})}{(2M)!} (\sqrt{2h}\sigma(t_n, x, a)y)^{2M} \right\} \frac{e^{-y^2}}{\sqrt{\pi}} dy, \end{aligned}$$

for some  $\hat{z}$ . In the same way we get

$$\mathbb{E} \left[ f(X_{t_n,x}^{\alpha,h}(t_{n+1})) \right] = \sum_{i=1}^M \frac{\omega_i}{\sqrt{\pi}} \left\{ \sum_{k=0}^{2M-1} \frac{f^{(k)}(z)}{k!} (\sqrt{2h}\sigma(t_n, x, a)\xi_i)^k + \frac{f^{(2M)}(\tilde{z})}{(2M)!} (\sqrt{2h}\sigma(t_n, x, a)\xi_i)^{2M} \right\},$$

for some  $\tilde{z}$ . At this point we recall that, by construction, the Gauß-Hermite quadrature formula is exact for any polynomial of degree  $\leq 2M-1$ , so for any  $k \in \{0, \dots, 2M-1\}$  we have

$$\frac{1}{\sqrt{\pi}} \frac{f^{(k)}(z)}{k!} (\sqrt{2h}\sigma(t_n, x, a))^k \left\{ \int_{-\infty}^{+\infty} y^k e^{-y^2} dy - \sum_{i=1}^M w_i z_i^k \right\} = 0.$$

This implies that

$$\begin{aligned} &\left| \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,EM}(t_{n+1})) \right] - \mathbb{E} \left[ f(X_{t_n,x}^{\alpha,h}(t_{n+1})) \right] \right| \\ &\leq \left| \int_{-\infty}^{+\infty} \frac{f^{(2M)}(\hat{z})}{(2M)!} (\sqrt{2h}\sigma(t_n, x, a)y)^{2M} \frac{e^{-y^2}}{\sqrt{\pi}} dy - \sum_{i=1}^M \frac{w_i}{\sqrt{\pi}} \frac{f^{(2M)}(\tilde{z})}{(2M)!} (\sqrt{2h}\sigma(t_n, x, a)z_i)^{2M} \right| \\ &\leq C \|f^{(2M)}\|_{\infty} h^M (1 + x^{2M}), \end{aligned}$$

where the constant  $C$  depends on  $M$  and the constants in assumption (H2) and we used the fact that  $|\sigma(t, x, a)| \leq C_0(1 + |x|)$  for some  $C_0 \geq 0$  depending on  $K_0$  in (H2).  $\square$

## 5. A LOWER BOUND FOR $v$

In order to obtain error estimates for the scheme described in Section 4, we will adapt the technique of “shaking coefficients” and regularization introduced by Krylov in [15, 17] and studied later by many authors (see for instance [1, 2, 3]) for obtaining the rate of convergence of monotone numerical scheme for second order HJB equations. We do so without passing by the PDE consistency error and work instead with the direct estimates we presented in the previous section. We refer to Remark 3 for a discussion on the regular case.

In this section we work for simplicity with  $p = d = 1$ .

**5.1. Regularization.** Let  $\varepsilon > 0$  and let  $\mathcal{E}_h$  be the set of progressively measurable processes  $e$  bounded by  $\varepsilon$  that are constant in each time interval  $[t_i, t_{i+1}]$ , that is,

$$\mathcal{E}_h := \{\text{prog. meas. process } e: e(s) \equiv e_i, \forall s \in [t_i, t_{i+1}] \text{ with } |e_i| \leq \varepsilon\}.$$

For any pair  $(\alpha, e) \in \mathcal{A}_h \times \mathcal{E}_h$ , let us consider the process  $X_{t_n, x}^{\alpha, e}(\cdot)$  defined as an  $\varepsilon$ -perturbation of the dynamics (4.1):

$$\begin{cases} X_{t_n, x}^{\alpha, e}(t_{i+1}) = X_{t_n, x}^{\alpha, e}(t_i) + \mu(t_i, X_{t_n, x}^{\alpha, e}(t_i) + e_i, a_i)h + \sigma(t_i, X_{t_n, x}^{\alpha, e}(t_i) + e_i, a_i)\Delta\mathcal{B}_i \\ X_{t_n, x}^{\alpha, e}(t_n) = x, \end{cases} \quad (5.1)$$

for  $i = n, \dots, N-1$ . We define the following ‘‘perturbed’’ value function:

$$v^\varepsilon(t_n, x) := \sup_{\alpha \in \mathcal{A}_h, e \in \mathcal{E}_h} \mathbb{E} \left[ U(X_{t_n, x}^{\alpha, e}(T)) \right] \quad (5.2)$$

for  $n = 0, \dots, N, x \in \mathbb{R}$ .

**Proposition 5.1.** *Let assumptions (H1)-(H3) be satisfied. Then there exists a constant  $C \geq 0$  such that*

$$|v_{EM}(t_n, x) - v^\varepsilon(t_n, x)| \leq LC\varepsilon$$

and

$$|v^\varepsilon(t_n, x) - v^\varepsilon(t_n, y)| \leq LC|x - y|,$$

for any  $n = 0, \dots, N$  and  $x, y \in \mathbb{R}$ .

*Proof.* Let us fix a control  $\alpha \in \mathcal{A}_h$  and  $e \in \mathcal{E}_h$ . For any  $i = n, \dots, N-1$ , we denote for simplicity

$$X_{t_n, x}^{\alpha, EM}(t_i) \equiv X_i^{EM} \quad \text{and} \quad X_{t_n, x}^{\alpha, e}(t_i) \equiv X_i^e.$$

By the definition of processes (4.1) and (5.1) one has

$$\begin{aligned} X_{i+1}^{EM} - X_{i+1}^e &= \left( X_i^{EM} - X_i^e \right) + h \left( \mu(t_i, X_i^{EM}, a_i) - \mu(t_i, X_i^e + e_i, a_i) \right) \\ &\quad + \sqrt{h} \left( \sigma(t_i, X_i^{EM}, a_i) - \sigma(t_i, X_i^e + e_i, a_i) \right) \Delta\mathcal{B}_i \end{aligned}$$

for any  $i = n, \dots, N-1$ . Then

$$\begin{aligned} \left( X_{i+1}^{EM} - X_{i+1}^e \right)^2 &= \left( X_i^{EM} - X_i^e \right)^2 + h^2 \left( \mu(t_i, X_i^{EM}, a_i) - \mu(t_i, X_i^e + e_i, a_i) \right)^2 \\ &\quad + h \left( \sigma(t_i, X_i^{EM}, a_i) - \sigma(t_i, X_i^e + e_i, a_i) \right)^2 \Delta\mathcal{B}_i^2 \\ &\quad + 2h \left( X_i^{EM} - X_i^e \right) \left( \mu(t_i, X_i^{EM}, a_i) - \mu(t_i, X_i^e + e_i, a_i) \right) \\ &\quad + 2\sqrt{h} \left( X_i^{EM} - X_i^e \right) \left( \sigma(t_i, X_i^{EM}, a_i) - \sigma(t_i, X_i^e + e_i, a_i) \right) \Delta\mathcal{B}_i \\ &\quad + 2h^{3/2} \left( \mu(t_i, X_i^{EM}, a_i) - \mu(t_i, X_i^e + e_i, a_i) \right) \left( \sigma(t_i, X_i^{EM}, a_i) - \sigma(t_i, X_i^e + e_i, a_i) \right) \Delta\mathcal{B}_i. \end{aligned}$$

Taking the expectation, a straightforward calculation shows that

$$\begin{aligned} &\mathbb{E} \left[ \left( X_{i+1}^{EM} - X_{i+1}^e \right)^2 \right] \\ &\leq \mathbb{E} \left[ \left( X_i^{EM} - X_i^e \right)^2 \right] + K_0^2 h^2 \mathbb{E} \left[ \left( X_i^{EM} - X_i^e - e_i \right)^2 \right] + K_0^2 h \mathbb{E} \left[ \left( X_i^{EM} - X_i^e - e_i \right)^2 \Delta\mathcal{B}_i^2 \right] \\ &\quad + h \mathbb{E} \left[ \left( X_i^{EM} - X_i^e \right)^2 + \left( \mu(t_i, X_i^{EM}, a_i) - \mu(t_i, X_i^e + e_i, a_i) \right)^2 \right] \\ &\leq \mathbb{E} \left[ \left( X_i^{EM} - X_i^e \right)^2 \right] \left( 1 + 2K_0^2 h^2 + 4K_0^2 h + h \right) + \varepsilon^2 \left( 2K_0^2 h^2 + 4K_0^2 h \right) \\ &\leq \mathbb{E} \left[ \left( X_i^{EM} - X_i^e \right)^2 \right] \left( 1 + Ch \right) + \varepsilon^2 Ch, \end{aligned}$$

where we also used the inequality  $2ab \leq (a^2 + b^2)$  and denoted by  $C$  any positive constant independent of  $h, e$  and  $\alpha$ .

By iteration (using the fact that  $1 + Ch \leq e^{Ch}$ ) we finally get

$$\mathbb{E} \left[ \left| X_{t_n, x}^{\alpha, EM}(t_i) - X_{t_n, x}^{\alpha, e}(t_i) \right|^2 \right] \leq \varepsilon^2 e^{CT} T,$$

for any  $i = n, \dots, N$  and we can conclude that

$$|v_{EM}(t_n, x) - v^\varepsilon(t_n, x)| \leq \sup_{\alpha \in \mathcal{A}_h, e \in \mathcal{E}_h} \mathbb{E} \left[ \left| U(X_{t_n, x}^{\alpha, EM}(T)) - U(X_{t_n, x}^{\alpha, e}(T)) \right| \right] \leq LC\varepsilon$$

so that the first property is proved. The Lipschitz continuity of  $v^\varepsilon$  follows by the estimate

$$\mathbb{E} \left[ \sup_{i=n, \dots, N} |X_{t_n, x}^{\alpha, e}(t_i) - X_{t_n, y}^{\alpha, e}(t_i)| \right] \leq K|x - y|$$

that leads to

$$|v^\varepsilon(t_n, x) - v^\varepsilon(t_n, y)| \leq LK|x - y|,$$

for any  $n = 0, \dots, N$ ,  $x, y \in \mathbb{R}$ , where the constant  $K$  is the same constant that appears in Proposition 2.1.  $\square$

Moreover, the value function  $v^\varepsilon$  satisfies the following DPP:

$$v^\varepsilon(t_n, x) = \sup_{a \in A, |e| \leq \varepsilon} \mathbb{E} \left[ v^\varepsilon(t_{n+1}, X_{t_n, x}^{a, e}(t_{n+1})) \right], \quad (5.3)$$

for any  $n = 0, \dots, N - 1$ .

The step that follows consists in a regularization of the function  $v^\varepsilon$ . For the regularization procedure, we consider a smooth function  $\delta : \mathbb{R} \rightarrow [0, +\infty)$  supported in the unit ball  $B_1(0)$  with  $\int_{\mathbb{R}} \delta(x) dx = 1$ , and we define  $\{\delta_\varepsilon\}_{\varepsilon > 0}$  as the following sequence of mollifiers:

$$\delta_\varepsilon(x) := \frac{1}{\varepsilon} \delta\left(\frac{x}{\varepsilon}\right).$$

Then define, for any  $n = 0, \dots, N$ ,

$$v_\varepsilon(t_n, x) := \int_{\mathbb{R}} v^\varepsilon(t_n, x - \xi) \delta_\varepsilon(\xi) d\xi. \quad (5.4)$$

**Proposition 5.2.** *Let assumptions (H1)-(H3) be satisfied. The function  $v_\varepsilon(t_n, \cdot)$  defined by (5.4) is  $C^\infty$  for any  $n = 0, \dots, N$ . Moreover,*

(i) *there exists a constant  $C \geq 0$  such that*

$$|v_\varepsilon(t_n, x) - v^\varepsilon(t_n, x)| \leq LC\varepsilon \quad n = 0, \dots, N, x \in \mathbb{R};$$

(ii) *for any  $k \geq 1$  one has*

$$\left\| \frac{\partial^k v_\varepsilon}{\partial x^k} \right\|_\infty \leq LC_k \varepsilon^{1-k}, \quad (5.5)$$

*for a certain constant  $C_k \geq 0$ ;*

(iii)  *$v_\varepsilon$  satisfies the following super-dynamic programming principle*

$$v_\varepsilon(t_n, x) \geq \sup_{a \in A} \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1})) \right] \quad (5.6)$$

*for any  $n = 0, \dots, N - 1, x \in \mathbb{R}$ .*

*Proof.* The regularity of  $v_\varepsilon(t_n, \cdot)$  and properties (i)-(ii) follow by the properties of mollifiers and the Lipschitz continuity of  $v^\varepsilon$  (Proposition 5.1). It remains to prove (iii). By the definition of  $v_\varepsilon$ , equality (5.3) and using the fact that

$$X_{t_n, x - \xi}^{a, \xi}(t_{n+1}) = X_{t_n, x}^{a, EM}(t_{n+1}) - \xi \quad \forall n = 0, \dots, N - 1, a \in A, |\xi| \leq \varepsilon,$$

we have

$$\begin{aligned}
v_\varepsilon(t_n, x) &= \int_{\mathbb{R}} v^\varepsilon(t_n, x - \xi) \delta_\varepsilon(\xi) d\xi \\
&= \int_{\mathbb{R}} \sup_{a \in A, |e| \leq \varepsilon} \mathbb{E} \left[ v^\varepsilon(t_{n+1}, X_{t_n, x - \xi}^{a, e}(t_{n+1})) \right] \delta_\varepsilon(\xi) d\xi \\
&\geq \int_{\mathbb{R}} \sup_{a \in A} \mathbb{E} \left[ v^\varepsilon(t_{n+1}, X_{t_n, x - \xi}^{a, \xi}(t_{n+1})) \right] \delta_\varepsilon(\xi) d\xi \\
&\geq \sup_{a \in A} \int_{\mathbb{R}} \mathbb{E} \left[ v^\varepsilon(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1}) - \xi) \right] \delta_\varepsilon(\xi) d\xi \\
&= \sup_{a \in A} \mathbb{E} \left[ \int_{\mathbb{R}} v^\varepsilon(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1}) - \xi) \delta_\varepsilon(\xi) d\xi \right] \\
&= \sup_{a \in A} \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1})) \right],
\end{aligned}$$

which concludes the proof.  $\square$

**5.2. A lower bound for the discrete time scheme.** Applying (4.3), Proposition 5.1 and Proposition 5.2(i), we obtain

$$v(t_n, x) \geq v_\varepsilon(t_n, x) - LK_{EM}(1 + |x|)h^{1/2} - LC\varepsilon, \quad (5.7)$$

for some new  $C$ . Recalling that  $v_\varepsilon$  satisfies the super-dynamic programming principle given by Proposition 5.2(iii) and using the definition of  $V$  we also have for any  $n = 0, \dots, N-1, x \in \mathbb{R}$

$$\begin{aligned}
&V(t_n, x) - v_\varepsilon(t_n, x) \\
&\leq \sup_{a \in A} \mathbb{E} \left[ V(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] - \sup_{a \in A} \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1})) \right] \\
&\leq \sup_{\alpha \in A} \left\{ \mathbb{E} \left[ V(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] - \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] \right. \\
&\quad \left. + \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] - \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, EM}(t_{n+1})) \right] \right\} \\
&\leq \sup_{a \in A} \left\{ \mathbb{E} \left[ V(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] - \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] \right\} + Cx^{2M} \|\partial_x^{2M} v_\varepsilon\|_\infty h^M \\
&\leq \sup_{a \in A} \left\{ \mathbb{E} \left[ V(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] - \mathbb{E} \left[ v_\varepsilon(t_{n+1}, X_{t_n, x}^{a, h}(t_{n+1})) \right] \right\} + LCx^{2M} \varepsilon^{1-2M} h^M, \quad (5.8)
\end{aligned}$$

where for the last two inequalities we applied, respectively, Proposition 4.2 and Proposition 5.2(ii). Observing that

$$\mathbb{E} \left[ \sup_{i=n, \dots, N} |X_{t_n, x}^{\alpha, h}(t_i)|^{2M} \right] \leq C(1 + x^{2M})$$

for some  $C \geq 0$  independent of  $\alpha$ , we can iterate inequality (5.8) and get

$$\begin{aligned}
V(t_n, x) - v_\varepsilon(t_n, x) &\leq \|V(t_N, \cdot) - v_\varepsilon(t_N, \cdot)\|_\infty + LC(1 + x^{2M})\varepsilon^{1-2M}h^{M-1} \\
&\leq LC\varepsilon + LC(1 + x^{2M})\varepsilon^{1-2M}h^{M-1}.
\end{aligned}$$

Therefore, by this last inequality and (5.7), we can conclude that for any  $n = 0, \dots, N, x \in \mathbb{R}$

$$v(t_n, x) \geq V(t_n, x) - LC(1 + x^{2M})\left(\varepsilon^{1-2M}h^{M-1} + h^{1/2} + \varepsilon\right).$$

Balancing the quantities  $\varepsilon$  and  $h$ , i.e. taking  $\varepsilon = h^{(M-1)/2M}$  and observing that  $1/2 > (M-1)/2M$  one has

$$v(t_n, x) \geq V(t_n, x) - LC(1 + x^{2M})h^{(M-1)/2M}. \quad (5.9)$$

**Remark 3.** *If the value function  $v$  can be shown to be sufficiently smooth, the regularization step is not necessary and it is also possible to consider the rate of weak convergence of the Euler-Maruyama scheme, which is one, and under differentiability assumptions on  $U$  this gives*

$$\sup_{i=n,\dots,N} \left| \mathbb{E}[U(X_{t_n,x}^\alpha(t_i)) - U(X_{t_n,x}^{\alpha,EM}(t_i))] \right| \leq \tilde{K}_{EM}h.$$

Thus, we obtain the following estimate

$$v(t_n, x) \geq V(t_n, x) + Cx^{2M}h^{M-1} + \tilde{K}_{EM}h,$$

which is of order one as we would expect in the regular case.

**5.3. The fully discrete scheme.** In order to be able to compute the numerical solution practically, we need to introduce also a discretization in space (otherwise the total number of nodes of all trajectories grows exponentially in  $N$ ). Let  $\Delta x > 0$  be the space step. We consider the space grid  $\mathcal{G}_{\Delta x} := \{x_m = m\Delta x : m \in \mathbb{Z}\}$ . Let  $\mathcal{I}[\cdot]$  denote the linear (P1) interpolation operator with respect to the space variable, satisfying for every Lipschitz function  $\phi$  (with Lipschitz constant  $L_\phi$ ):

$$\begin{cases} (i) \mathcal{I}[\phi](x_m) = \phi(x_m), \forall m \in \mathbb{Z}, \\ (ii) |\mathcal{I}[\phi](x) - \phi(x)| \leq L_\phi \Delta x, \\ (iii) |\mathcal{I}[\phi](x) - \phi(x)| \leq C\Delta x^2 \|D_x^2 \phi\|_\infty \text{ if } \phi \in C^2(\mathbb{R}), \\ (iv) \text{ for any functions } \phi_1, \phi_2 : \mathbb{R} \rightarrow \mathbb{R}, \phi_1 \leq \phi_2 \Rightarrow \mathcal{I}[\phi_1] \leq \mathcal{I}[\phi_2]. \end{cases} \quad (5.10)$$

The solution of the fully discrete scheme, denoted by  $V_F$ , is defined by:

$$\begin{cases} V_F(t_n, x_m) = \sup_{a \in A} \sum_{i=1}^M \lambda_i \mathcal{I}[V_F](t_{n+1}, X_{t_n, x_m}^{a,h}(t_{n+1})), \\ V_F(t_N, x_m) = U(x_m), \end{cases} \quad (5.11)$$

for  $n = N - 1, \dots, 0$  and  $m \in \mathbb{Z}$ .

**Lemma 5.1.** *Let assumptions (H1)-(H3) be satisfied. Then, there exists  $C_{\mathcal{I}} \geq 0$  such that*

$$\sup_{\substack{n=0,\dots,N, \\ m \in \mathbb{Z}}} |V(t_n, x_m) - V_F(t_n, x_m)| \leq LC_{\mathcal{I}} \frac{\Delta x}{h}.$$

*Proof.* The result follows by properties (5.10)(ii),(iv) and by the Lipschitz continuity of  $V$ , which can be easily proved under (H1)-(H3) (see also [9, Lemma 7.1]).  $\square$

Observe that, in absence of further regularity assumptions, this introduces the following ‘‘inverse’’ CFL condition for the convergence of the fully discrete scheme:

$$\Delta x/h \rightarrow 0 \quad \text{as } \Delta x, h \rightarrow 0.$$

The contribution from the interpolation error has to be added to (5.9), giving an overall error of

$$O\left(h^{(M-1)/2M} + \frac{\Delta x}{h}\right).$$

Optimising the choice of  $\Delta x$  with respect to  $h$  we get

$$\Delta x \sim h^{(3M-1)/2M}. \quad (5.12)$$

This effectively leads to order  $(M-1)/(2M)$  in time and  $(M-1)/(3M-1)$  in space, which can be made arbitrarily close to  $1/2$  and  $1/3$ , respectively, by choosing  $M$  large enough.

On the other hand, if  $v$  is sufficiently smooth, (5.10)(iii) can be applied. This, together with Remark 3, gives error estimates of order  $O(h + \frac{\Delta x^2}{h})$ . In many cases, this corresponds to the practically observed situation so that choosing  $\Delta x \sim h$  is sufficient to observe convergence, with order 1, of the fully discrete scheme.

**Remark 4** (Comparison with existing results). *By a Taylor expansion it is possible to compute the consistency error of the scheme with respect to the HJB equation (1.2). For simplicity let us consider the uncontrolled case with  $\mu \equiv 0$ . Using the fact that  $\sum_{i=1}^M \lambda_i = 1$ ,  $\sum_{i=1}^M \lambda_i \xi_i^2 = 1$  and  $\sum_{i=1}^M \lambda_i \xi_i^{2k+1} = 0$  ( $\forall k \in \mathbb{N}$ ), one gets*

$$\begin{aligned} & \frac{1}{h} \left( v(t_{n+1}, x) - \sum_{i=1}^M \lambda_i v(t_n, x + \sqrt{h} \sigma(t_n, x) \xi_i) \right) \\ &= v_t(t_n, x) - (\sigma(t_n, x))^2 v_{xx}(t_n, x) + \frac{h^2}{2} v_{tt}(t_n, x) - \frac{h^2}{4!} (\sigma(t_n, x))^4 v_{4x}(t_n, x) \sum_{i=1}^M \lambda_i \xi_i^4 + O(h^3), \end{aligned}$$

which shows that the scheme has order 1 consistency, for all  $M$ . Applying the results in [9], this would lead to error estimates of order  $h^{1/4} + \Delta x/h$ , i.e., with the optimal choice of  $\Delta x$  order 1/4 in  $h$  and 1/5 in  $\Delta x$ . The improvement we get for the lower bound is due to the fact that, splitting the two contributions of the error coming from the Euler-Maruyama time stepping and the Gauß-Hermite quadrature formula, we can reduce the second one increasing  $M$ , whereas for the first one the lower regularity requirement allows us to get order 1/2.

**Remark 5.** *The upper bound with this approach would be restricted from (3.3) by order 1/6 in  $h$  and consequently 1/7 in  $\Delta x$ . Hence the technique of [9] gives the sharper result. We will discuss in the subsequent sections how we can get improvements for certain convex problems.*

## 6. AN UPPER BOUND FOR THE DUAL PROBLEM

The previous section gives an *a priori* lower bound for the difference between  $v$  and  $V_F$  making use of purely probabilistic arguments. With the objective pointed out in Section 3 to obtain an accurate upper bound, we now restrict our attention on those problem for which it is possible to define a dual problem.

**6.1. Duality in continuous time.** An important part of the literature dealing with financial applications of optimal control theory applies, under suitable convexity assumptions, duality techniques for solving problems of the form (1.1)-(2.1). The basic idea of this method is to write the optimal control problem as a constrained optimization one with respect to the state variable and then solve it generalising convex analysis techniques (see for instance [22] for a clear presentation of the main concepts).

This approach leads to the definition of a dual problem, associated with the dynamics

$$\begin{cases} dY(s) = \tilde{\mu}(s, Y(s), \alpha_s) ds + \tilde{\sigma}(s, Y(s), \alpha_s) \cdot d\mathcal{B}(s), & s \in (t, T), \\ Y(t) = y, \end{cases} \quad (6.1)$$

and the dual value function as the solution of a minimization problem by

$$\tilde{v}(t, y) := \inf_{\nu \in \mathcal{V}} \mathbb{E}[\tilde{U}(Y_{t,y}^\nu(T))], \quad (6.2)$$

with suitably defined  $\tilde{\mu}$ ,  $\tilde{\sigma}$ ,  $\tilde{U}$ , and  $\mathcal{V}$ . Under certain assumptions, such as those detailed below, the primal and dual problem are linked by a conjugate relation:

$$v(t, x) = \inf_{y > 0} \{ \tilde{v}(t, y) + xy \}. \quad (6.3)$$

For the duality result we mainly refer to [8], where a quite general set-up of an optimal investment problem involving portfolio constraints and nonlinear dynamics is discussed (see (6.4) below). Moreover, an abstract formulation and framework for establishing duality results is given in [14] and [23], where also some concrete examples are discussed in detail. In particular, we consider a

special case of problem (1.1), (2.1) with dynamics

$$\begin{cases} dX(s) = X(s) \left( r(s) + \alpha_s \cdot (b(s) - r(s)\mathbb{1}) + g(s, \alpha_s) \right) ds + X(s) \alpha_s \cdot \psi(s) d\mathcal{B}(s), & s \in (t, T), \\ X(t) = x, \end{cases} \quad (6.4)$$

and with the following assumptions:

**(H4)**  $A \subseteq \mathbb{R}^d$  is a closed and convex set such that  $0 \in A$ .

**(H5)** (i) There exists  $K_0 \geq 0$  such that

$$|r(t) - r(s)| + \|b(t) - b(s)\| + \|\psi(t) - \psi(s)\| \leq K_0 |t - s|^{1/2} \quad \forall t, s \in [0, T].$$

(ii)  $g : [0, T] \times A \rightarrow \mathbb{R}$  satisfies:

- there exists  $K_1 \geq 0$  such that

$$|g(t, a) - g(t, a')| \leq K_1 |a - a'| \quad \forall a, a' \in A, t \in [0, T];$$

$$\|g\|_\infty \leq K_1;$$

$$|g(t, a) - g(s, a)| \leq K_1 |t - s|^{1/2} \quad \forall t, s \in [0, T], a \in A;$$

- for each  $t \in [0, T]$ ,  $a \rightarrow g(t, a)$  is concave;

-  $g(t, 0) = 0$  for all  $t \in [0, T]$ .

**(H6)**  $\psi$  satisfies the uniform ellipticity condition, i.e. there exists  $\eta > 0$  such that

$$\xi^T \psi \psi_t^T \xi \geq \eta |\xi|^2 \quad \forall t \in [0, T], \xi \in \mathbb{R}^d.$$

**(H7)**  $U \in C^1((0, +\infty); \mathbb{R})$ ;

$U$  is concave and strictly increasing;

$$\lim_{x \rightarrow +\infty} U'(x) = 0.$$

We can observe that, under (6.4), for any  $x > 0$  also  $X_{t,x}^\alpha(s) > 0, \forall s \in [t, T]$ .

**Remark 6.** We point out that, as mentioned in [23, Section 6.5], the usual Inada condition

$$\lim_{x \rightarrow 0^+} U'(x) = +\infty$$

(requested in [8]) is not necessary for proving the main duality results.

In this setting, it is possible to define a dual problem. This is given by the following dynamics

$$\begin{cases} dY(s) = -(r(s) + \tilde{g}(s, \nu_s)) Y(s) ds + (\psi(s) \psi^T(s))^{-1} Y(s) (r(s)\mathbb{1} - b(s) - \nu_s) \cdot \psi(s) d\mathcal{B}(s), & s \in [t, T], \\ Y(t) = y, \end{cases} \quad (6.5)$$

where

$$\tilde{g}(t, \nu) := \sup_{a \in A} \{g(t, a) - a\nu\} \quad (6.6)$$

and  $\nu$  is an  $\mathbb{R}^d$ -valued adapted process. We denote by  $\tilde{U}$  the convex conjugate of  $U$ , i.e.

$$\tilde{U}(x) := \sup_{y > 0} \{U(y) - xy\}.$$

We define the value of the dual problem by (6.2), where  $Y_{t,y}^\nu(\cdot)$  denotes the solution to (6.5) associated with the control  $\nu$  and  $\mathcal{V}$  is the set of adapted processes such that

$$\int_0^T \|\nu_t\|^2 dt + \int_0^T \tilde{g}(t, \nu_t) dt < +\infty.$$

We denote by  $\Gamma \in \mathbb{R}^d$  the set of values taken by the controls in  $\mathcal{V}$ . One has the following duality:

**Proposition 6.1** (Cuoco-Liu, [8]). *Let assumptions (H4)-(H7) be satisfied. Then for any  $t \in [0, T]$ ,  $x > 0$ , the primal and dual value functions,  $v$  and  $\tilde{v}$ , satisfy (6.3).*

**Remark 7.** We point out that the results in [8] hold also if  $r, b, \psi$  and  $g$  are stochastic processes. However, our approximation schemes make use of the Markovian framework.

**6.2. Approximation of the dual problem.** The scheme presented in Section 4 can be used for approximating the value function  $\tilde{v}$  associated with the dual problem (6.1)-(6.2), defining recursively

$$\begin{cases} \tilde{V}(t_n, y) = \inf_{\gamma \in \Gamma} \mathbb{E}[\tilde{V}(t_{n+1}, Y_{t_n, y}^{\gamma, h}(t_{n+1}))], & n = N-1, \dots, 0, \\ \tilde{V}(t_N, y) = \tilde{U}(y), \end{cases} \quad (6.7)$$

where for any  $\nu \equiv (\gamma_n, \dots, \gamma_{N-1})$ , with  $\gamma_i \in \Gamma \subseteq \mathbb{R}^d \forall i = n, \dots, N-1$ , we define

$$\begin{cases} Y_{t_n, y}^{\nu, h}(t_n) = y, \\ Y_{t_n, y}^{\nu, h}(t_{i+1}) = Y_{t_n, y}^{\nu, h}(t_i) + \tilde{\mu}(t_i, Y_{t_n, y}^{\nu, h}(t_i), \gamma_i)h + \sqrt{h}\tilde{\sigma}(t_i, Y_{t_n, y}^{\nu, h}(t_i), \gamma_i) \cdot \zeta_i, \end{cases} \quad (6.8)$$

for  $i = n, \dots, N-1$ . This leads to

$$\tilde{V}(t_n, y) = \inf_{\nu \in \mathcal{V}_h} \mathbb{E}[\tilde{U}(Y_{t_n, y}^{\nu, h}(T))],$$

where  $\mathcal{V}_h$  denotes the set of processes  $\nu \in \mathcal{V}$  such that  $\nu_s \equiv \gamma_i$  for any  $s \in [t_i, t_{i+1}]$ . In order to derive similar error estimates to the primal problem, we make analogous assumptions for the dual problem (but see Remark 8(3) for an important relaxation of (H3) and  $(\tilde{H}3)$ ):

**(H1)**  $\Gamma$  is a compact set;

**(H2)**  $\tilde{\mu} : [0, T] \times \mathbb{R} \times \Gamma \rightarrow \mathbb{R}$  and  $\tilde{\sigma} : [0, T] \times \mathbb{R} \times \Gamma \rightarrow \mathbb{R}^p$  are continuous functions and there exists  $\tilde{K}_0 \geq 0$  such that

$$\begin{aligned} |\tilde{\mu}(t, x, \gamma) - \tilde{\mu}(t, y, \gamma)| + \|\tilde{\sigma}(t, x, \gamma) - \tilde{\sigma}(t, y, \gamma)\| &\leq \tilde{K}_0|x - y| \quad \forall t \in [0, T], x, y \in \mathbb{R}, \gamma \in \Gamma, \\ |\tilde{\mu}(t, x, \gamma) - \tilde{\mu}(s, x, \gamma)| + \|\tilde{\sigma}(t, x, \gamma) - \tilde{\sigma}(s, x, \gamma)\| &\leq \tilde{K}_0|t - s|^{1/2} \quad \forall t, s \in [0, T], x \in \mathbb{R}, \gamma \in \Gamma; \end{aligned}$$

**(H3)**  $\tilde{U}$  is a Lipschitz continuous function with Lipschitz constant  $\tilde{L}$ .

Under these assumptions, the results of Propositions 2.1, 4.1, 4.2 also hold for the dual problem.

**Remark 8.** Assumptions (H1)-(H3) and  $(\tilde{H}1)$ - $(\tilde{H}3)$  are necessary in order to derive the error estimates of this and the following section. We discuss them here.

- (1) The compactness of  $A$  and  $\Gamma$  (assumptions (H1) and  $(\tilde{H}1)$ ) has to be checked example by example. In various cases, such as those in our numerical tests below, the boundedness of the optimal controls  $\alpha^* \in \mathcal{A}$  and  $\nu^* \in \mathcal{V}$  is known in advance. In other cases see Remark 9.
- (2) The assumptions (H2) and  $(\tilde{H}2)$  on the dynamics are satisfied in the examples we consider. For (6.4), assumptions (H5), (H6) and  $(\tilde{H}1)$  are sufficient to ensure that for some  $C_0 \geq 0$

$$\begin{aligned} \|(\psi\psi^T(t))^{-1}(r(t)\mathbb{1} - b(t) - \gamma) - (\psi\psi^T(s))^{-1}(r(s)\mathbb{1} - b(s) - \gamma)\| \\ + |\tilde{g}(t, \gamma) - \tilde{g}(s, \gamma)| + |r(t) - r(s)| \leq C_0|t - s|^{1/2}, \end{aligned}$$

for any  $\gamma \in \Gamma$ . The Lipschitzianity in  $x$  and  $y$  is automatically given by the linearity.

- (3) In general, (H3) and  $(\tilde{H}3)$  are not simultaneously satisfied due to the Inada conditions. Indeed, in our example of a power utility, only (H7) holds but  $U$  and  $\tilde{U}$  are not Lipschitz in 0. This forces us to introduce a further approximation of the problem and consequently an additional error. Let  $\rho, c_0 > 0$  and  $x_1 = c_0/\rho$ ,  $x_2 = \rho$ , we define

$$U_\rho(x) := \begin{cases} U(0) + \frac{U(x_1) - U(0)}{x_1}x & \text{if } 0 \leq x \leq x_1 \\ U(x) & \text{if } x_1 \leq x \leq x_2 \\ U(x_2) & \text{if } x \geq x_2, \end{cases}$$

so that  $U_\rho \rightarrow U$  as  $\rho \rightarrow +\infty$ . We denote by  $\tilde{U}_\rho$  its convex conjugate

$$\tilde{U}_\rho(x) := \sup_{y > 0} \{U_\rho(y) - xy\}.$$



These approximations are shown for a power utility in Figure 4. The difference between  $v$  and the value function associated with  $U_\rho$  can be estimated by using large deviations arguments (see [10]) to bound

$$\mathbb{P}\left[\left(X_{t,x}^\alpha(T) \leq x_1\right) \cup \left(X_{t,x}^\alpha(T) \geq x_2\right)\right]. \quad (6.9)$$

This decreases exponentially fast as  $\rho$  goes to 0, so it will not strongly modify our results.

Under assumptions  $(\widetilde{\text{H1}})$ - $(\widetilde{\text{H3}})$ , the regularization procedure presented in Section 5.1 can be applied to the value function (6.2) associated with the dual problem.

**Proposition 6.2.** *Let  $(\widetilde{\text{H1}})$ - $(\widetilde{\text{H3}})$  be satisfied. For any  $\varepsilon > 0$  there exists a function  $\tilde{v}_\varepsilon$  such that  $\tilde{v}_\varepsilon(t_n, \cdot)$  is  $C^\infty$  for any  $n = 0, \dots, N$  and*

(i) *there exists  $C \geq 0$  such that*

$$\left| \tilde{v}_\varepsilon(t_n, y) - \tilde{v}^{EM}(t_n, y) \right| \leq \tilde{L}C\varepsilon,$$

*for any  $n = 0, \dots, N$ ,  $y \in \mathbb{R}$ ;*

(ii) *for any  $k \geq 1$*

$$\left\| \frac{\partial^k \tilde{v}_\varepsilon}{\partial y^k} \right\|_\infty \leq \varepsilon^{1-k} \tilde{L}C_k,$$

*for some  $C_k \geq 0$ ;*

(iii)  *$\tilde{v}_\varepsilon$  satisfies the following sub-dynamic programming principle*

$$\tilde{v}_\varepsilon(t_n, y) \leq \inf_{\gamma \in \Gamma} \mathbb{E} \left[ \tilde{v}_\varepsilon(t_{n+1}, Y_{t_n, y}^{\gamma, EM}(t_{n+1})) \right],$$

*for any  $n = 0, \dots, N$ ,  $y \in \mathbb{R}$ , where  $Y_{t_n, y}^{\gamma, EM}(\cdot)$  denotes the Euler-Maruyama approximation of the process defined by (6.1).*

*Proof.* Under assumptions  $(\widetilde{\text{H1}})$ - $(\widetilde{\text{H3}})$  the proof follows the same arguments of Proposition 5.2.  $\square$

**6.3. An upper bound for  $\tilde{v}$ .** The approximation scheme we defined for the dual problem is almost the same we used for the primal one, with the only difference that we have to handle a minimization problem. Therefore, we can use the arguments in Section 5.2 to obtain an upper bound for the difference  $(\tilde{v} - \tilde{V})$ . For some constant  $C \geq 0$ , one has

$$\tilde{v}(t_n, y) \leq \tilde{V}(t_n, y) + \tilde{L}\tilde{C}(1 + y^{2M})h^{(M-1)/2M}. \quad (6.10)$$

If then the fully discrete solution of the dual problem is denoted by  $\tilde{V}_F$ , under assumptions  $(\widetilde{\text{H1}})$ - $(\widetilde{\text{H3}})$  one also has

$$\sup_{\substack{n=0, \dots, N, \\ m \in \mathbb{Z}}} |\tilde{V}(t_n, y_m) - \tilde{V}_F(t_n, y_m)| \leq \tilde{L}\tilde{C}_I \frac{\Delta x}{h}.$$

## 7. A POSTERIORI ERROR ESTIMATES FOR THE DISCRETE TIME SCHEME

We now assume all of  $(\text{H1})$ - $(\text{H7})$  and  $(\widetilde{\text{H1}})$ - $(\widetilde{\text{H3}})$ . Using the conjugate dual property between  $v$  and  $\tilde{v}$  (see Proposition 6.1, under  $(\text{H4})$ - $(\text{H7})$ ), together with (6.10) we get

$$v(t_n, x) = \inf_{y \geq 0} \left\{ \tilde{v}(t_n, y) + xy \right\} \leq \inf_{y \geq 0} \left\{ \tilde{V}(t_n, y) + xy + \tilde{L}\tilde{C}(1 + y^{2M})h^{(M-1)/2M} \right\}.$$

As the dynamics is linear, the dual value function is convex. Therefore, for any  $x > 0$  the quantity  $\{\tilde{V}(t_n, y) + xy\}$  attains its infimum, say at some  $I(x)$  (which can be big as  $x$  approaches zero). It

follows that for  $x > 0$

$$\begin{aligned} & \inf_{y \geq 0} \left\{ \tilde{V}(t_n, y) + xy + \tilde{L}\tilde{C}(1 + y^{2M})h^{(M-1)/2M} \right\} \\ & \leq \tilde{V}(t_n, I(x)) + xI(x) + \tilde{L}\tilde{C}(1 + I(x)^{2M})h^{(M-1)/2M} \\ & = \inf_{y \geq 0} \left\{ \tilde{V}(t_n, y) + xy \right\} + \tilde{L}\tilde{C}(1 + I(x)^{2M})h^{(M-1)/2M}. \end{aligned}$$

Considering now the fully discrete scheme and putting together the upper and the lower bound one finally obtains

$$-C(x_i) \left( h^{(M-1)/2M} + \frac{\Delta x}{h} \right) \leq v(t_n, x_i) - V_F(t_n, x_i) \leq G^{h, \Delta x}(t_n, x_i) + C(x_i) \left( h^{(M-1)/2M} + \frac{\Delta x}{h} \right), \quad (7.1)$$

where we defined the numerical duality gap

$$G^{h, \Delta x}(t_n, x_i) := \left| \inf_{y_j \in \mathcal{G}_{\Delta x}} \left\{ \tilde{V}_F(t_n, y_j) + x_i y_j \right\} - V_F(t_n, x_i) \right| \quad (7.2)$$

and  $C(x_i)$  denotes any quantity depending on constants and the point  $x_i$  (recall that such a quantity increases as  $x_i$  approaches 0 or  $+\infty$ ). It is important to note that the duality gap for the fully discrete scheme is computable efficiently, see e.g. [11], such that (7.1) is a practically useful *a posteriori* bound.

## 8. NUMERICAL TESTS

We test our results on some examples. We consider  $d = p = 1$  and the computational domain  $[0, x_{\max}]$ . We denote by  $N$  and  $J$  respectively the number of time and space steps, i.e.

$$h = \frac{T}{N} \quad \text{and} \quad \Delta x = x_{\max}/J.$$

We study the case of a power utility function:

$$U(x) = \frac{x^p}{p} \quad \text{for some } p \in (0, 1). \quad (8.1)$$

In order to satisfy assumption (H3), we consider a modification of the utility function in a neighborhood of 0, and truncate the function at some  $x = \rho$  (see Remark 8). The utility function  $U$  and its conjugate  $\tilde{U}$ , as well as its Lipschitz continuous approximation  $U_\rho$  and its conjugate  $\tilde{U}_\rho$  are shown in Figure 4. In order to not further complicate the notation, we simply denote by  $V$  and  $\tilde{V}$  (instead of  $V_{\rho, F}$ ,  $\tilde{V}_{\rho, F}$  or similar) the fully discrete numerical solutions of the primal and dual problem associated respectively with  $U_\rho$  and  $\tilde{U}_\rho$ .

**Remark 9.** *For the optimisation over the controls in our computations, we truncate  $A$  and  $\Gamma$  first to a finite interval and then discretise the interval by  $N_a$  and  $N_\nu$  equally spaced mesh points, respectively. This further approximation decreases the value of the discrete primal (maximisation) problem and increases the value of the discrete dual (minimisation) problem, in the same way as the piecewise constant (in time) control approximation does. This implies that this component of the error is captured in the duality gap which we compute a posteriori. The approximation can generally only be improved by increasing the size of the control intervals and decreasing the control mesh spacing, concurrently with decreasing  $h$  and  $\Delta x$ .*

*As the optimal control in our examples is bounded, the error of the control truncation is zero if the interval is chosen large enough. It is seen from the computations that the contribution of the control discretisation error is small, decreasing quadratically in  $N_a^{-1}$  and  $N_\nu^{-1}$  since we have a smooth dependence of the Hamiltonian on the control. In our tests, we take  $N_a \sim N_\nu \sim N$ , such that the control discretisation error becomes eventually negligible.*

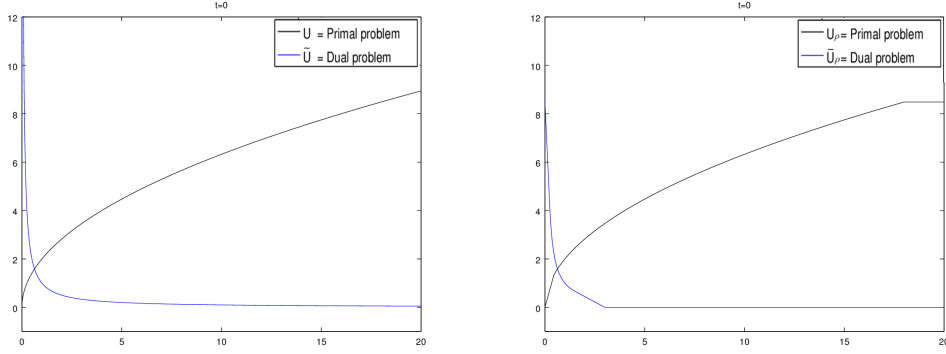


FIGURE 4. The power utility function  $U$  with its conjugate  $\tilde{U}$  (left) and the Lipschitz continuous approximation  $U_\rho$  with its conjugate  $\tilde{U}_\rho$  (right). Here,  $x_{\max} = 20$ ,  $\rho = 18$  and  $c_0 = 8$ .

**Remark 10.** *It is clear that as  $x_m$  approaches 0 or  $x_{\max}$  it may happen that  $X_{t_n, x_m}^{a, h}(t_{n+1})$  oversteps the domain. In this case, we use linear extrapolation in order to define  $V$  and  $\tilde{V}$  outside the domain.*

**Remark 11.** *In our tests, we take  $M = 2$  with  $\Delta x \sim h^{5/4}$  from (5.12), more specifically  $J \sim \lceil N^{5/4} \rceil$ . Taking  $M > 2$  has only (theoretical) advantages for non-smooth solutions, while we would observe order of convergence at most one for any choice of  $M \geq 2$  even in the smooth case. This is due to the fact that the use of the Euler-Maruyama scheme reduces the order of consistency of the overall scheme to one, regardless of the value of  $M$  (see also Remark 4). This has been confirmed numerically for  $M = 4$  with  $\Delta x = h^{11/8}$ , but we do not report the results here.*

*An improvement of the order of consistency might be achieved by combining higher values of  $M$  with the use of higher order time-stepping schemes, for instance the higher order Taylor schemes of [13]. This is beyond the scope of this paper and we leave it for future research.*

**Test 1: Merton problem.** We first study the classical Merton problem. This corresponds to the dynamics (6.4) with  $g \equiv 0$ , constant coefficients  $b, r, \psi$  and  $A = \mathbb{R}$ . It is well known that for this problem there exists a closed-form solution given by (see, e.g. [22])

$$v(t, x) = \exp \left\{ t \left( a^*(b - r) + r - \frac{1}{2}(a^*)^2(1 - p)\psi^2 \right) \right\} U(x),$$

where  $U$  and  $p$  are given in (8.1), and

$$a^* := \frac{(b - r)}{\psi^2(1 - p)}$$

is the optimal control. We recall that in this case the dual problem is linear and no optimisation is necessary since  $\Gamma = \{0\}$ . The values of the coefficients used in the test is given in Table 1. For these values, setting  $A = [-1, 1]$  is sufficient to have  $a^* \in A$ .

$p$	$r$	$b$	$\psi$	$T$	$x_{\max}$
0.5	0.8	1.2	1	0.5	20

TABLE 1. Test 1: Parameters used in numerical experiments.

Table 2 reports the error and the convergence rate of  $V$  to the exact solution  $v$  of the primal problem. As expected, the order of convergence is around 1. It is important to notice that continuing to refine the mesh without increasing  $\rho$ , we cannot get convergence to  $v$ . In fact, the probability in (6.9), even

if small at points  $x$  far from the boundaries of the domain, is different from zero everywhere (see also Figure 5, left). In Table 3, we report the numerical duality gap, i.e. the quantity  $G^{h,\Delta x}(T, x)$ . This quantity also decreases with order 1 or even slightly higher. In this case, the duality gap is bigger than the error, but of the same order. In Figure 5 (right) we show the numerical solutions  $V$  and  $\tilde{V}$  of the primal and the dual problem, together with the convex conjugate of  $\tilde{V}$ . Close to  $x = 0$  we do not observe convergence due to the local modification of  $U$ . Of course, this region can be shrunk by choosing  $\rho$  bigger.

$J$	$N$	Error $L^1$	Order $L^1$	Error $L^2$	Order $L^2$	Error $L^\infty$	Order $L^\infty$
14	8	2.90E-01	-	2.43E-01	-	2.03E-01	-
32	16	1.56E-01	0.90	1.40E-01	0.80	1.34E-01	0.60
77	32	6.64E-02	1.23	6.70E-02	1.06	8.51E-02	0.66
182	64	1.83E-02	1.86	1.94E-02	1.79	3.09E-02	1.46
431	128	6.70E-03	1.45	6.78E-03	1.51	9.61E-03	1.68
1024	256	2.52E-03	1.41	2.59E-03	1.39	3.90E-03	1.30
2436	512	1.02E-03	1.31	1.04E-03	1.31	1.63E-03	1.26
5793	1024	4.27E-04	1.25	4.42E-04	1.24	7.57E-04	1.11

TABLE 2. Test 1: Local ( $x \in [1, 2]$ ) errors and convergence order comparing  $V$  with the exact solution  $v$ , for  $M = 2$  (Gauß-Hermite quadrature),  $N = 4 \cdot 2^k$  (time steps),  $J = \lceil N^{5/4} \rceil$  (space steps),  $N_a = 2^k + 1$  (discrete controls), for  $k = 1, 2, \dots, 8$ .

$J$	$N$	Gap $L^1$	Order $L^1$	Gap $L^2$	Order $L^2$	Gap $L^\infty$	Order $L^\infty$
14	8	4.17E+00	-	3.49E+00	-	2.92E+00	-
32	16	1.99E+00	1.07	1.80E+00	0.95	1.84E+00	0.67
77	32	7.66E-01	1.38	7.61E-01	1.24	8.78E-01	1.07
182	64	4.00E-01	0.94	4.08E-01	0.90	5.10E-01	0.78
431	128	1.83E-01	1.12	1.87E-01	1.13	2.50E-01	1.03
1024	256	5.72E-02	1.68	6.00E-02	1.64	8.97E-02	1.48
2436	512	1.69E-02	1.76	1.73E-02	1.79	2.42E-02	1.89
5793	1024	6.40E-03	1.40	6.58E-03	1.40	9.08E-03	1.41

TABLE 3. Test 1: Local ( $x \in [1, 2]$ ) duality gap  $G^{h,\Delta x}$  from (7.2) and related convergence order, for  $M = 2$  (Gauß-Hermite quadrature),  $N = 4 \cdot 2^k$  (time steps),  $J = \lceil N^{5/4} \rceil$  (space steps),  $N_a = 2^k + 1$  (discrete controls), for  $k = 1, 2, \dots, 8$ .

**Test 2: Cuoco and Liu example.** This example is taken from [8]. In this paper, the authors consider the nonlinear dynamics in (6.4) (i.e.  $g \neq 0$ ) and portfolio constraints (i.e.  $A \subsetneq \mathbb{R}$ ). We still consider a power utility and  $d = 1$ . Let  $A$  be defined by

$$A = \left\{ a \in \mathbb{R} : \max(0, -a)\lambda_- + \max(0, a)\lambda_+ \leq 1 \right\}$$

for some  $\lambda_- \geq 0$  and  $\lambda_+ \in [0, 1]$ . The function  $g$  is defined by

$$g(a) = -r(1 + \iota\lambda_-) \max(0, -a) - (R - r)(1 - \max(0, a) - \iota\lambda_- \max(0, -a)),$$

where  $R \geq r$  and  $\iota \in [0, 1]$ . The values used in our numerical simulation are reported in Table 4. Observe that the choice  $\lambda_+ = \lambda_- = 1$  corresponds to  $A = [-1, 1]$ . In order to define  $\Gamma$ , we use the explicit expression given in [8, Section 5.2] for the optimal control. From this, for the data in Table 4, we can take  $\Gamma = [-1, 1]$  to guarantee  $\nu_t^* \in \Gamma$  for any  $t \in [0, T]$ . Table 5 reports the numerical duality gap and the corresponding convergence order. The numerical solutions  $V$  and  $\tilde{V}$  of the primal and the dual problem, together with the convex conjugate of  $\tilde{V}$  are shown in Figure 6. Still

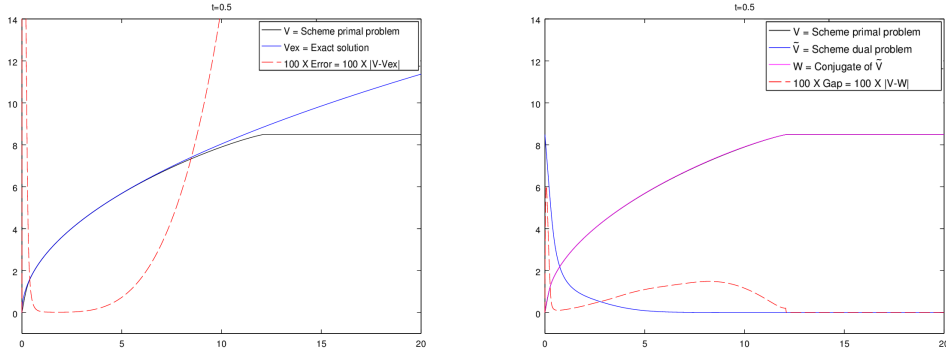


FIGURE 5. Test 1: Numerical solution  $V$  (in black) compared with the exact solution (blue, left) and the convex conjugate of  $\tilde{V}$  (magenta, right). The dashed red line represents the error (left) and the numerical duality gap (right).

$p$	$r$	$R$	$b$	$\psi$	$T$	$x_{\max}$	$\iota$	$\lambda_+$	$\lambda_-$
0.5	0.8	1	1.2	0.5	0.5	20	0.5	1	1

TABLE 4. Test 2: Parameters used in numerical experiments.

$J$	$N$	Gap $L^1$	Order $L^1$	Gap $L^2$	Order $L^2$	Gap $L^\infty$	Order $L^\infty$
14	8	2.30E+01	-	8.08E+00	-	4.10E+00	-
32	16	1.55E+01	0.57	5.09E+00	0.66	2.14E+00	0.94
77	32	7.01E+00	1.14	2.19E+00	1.22	9.01E-01	1.25
182	64	4.23E+00	0.73	1.31E+00	0.74	5.09E-01	0.82
431	128	2.49E+00	0.76	7.80E-01	0.75	3.21E-01	0.67
1024	256	1.43E+00	0.80	4.51E-01	0.79	1.83E-01	0.81
2436	512	6.33E-01	1.18	2.00E-01	1.18	8.05E-02	1.18
5793	1024	1.86E-01	1.76	6.00E-02	1.74	2.52E-02	1.68

TABLE 5. Test 2: Global duality gap  $G^{h,\Delta x}$  from (7.2) and related convergence order, for  $M = 2$  (Gauß-Hermite quadrature),  $N = 4 \cdot 2^k$  (time steps),  $J = \lceil N^{5/4} \rceil$  (space steps),  $N_a = N_\nu = 2^k + 1$  (discrete controls), for  $k = 1, 2, \dots, 8$ .

we can observe a peak of the gap close to  $x = 0$ , but in this case convergence to zero of the duality gap is globally achieved (see Table 5). However for small values of  $x$  we loose speed of convergence, which is in line with the local analysis in Section 7.

## 9. CONCLUSION AND PERSPECTIVES

This paper analyses numerical schemes for HJB equations based on a discrete time approximation of the optimal control problem. Using purely probabilistic arguments and under very general assumptions, in Section 5 we give an *a priori* lower bound for the solution generated by such an approximation. The error bound obtained in this way allows us to improve the results available in the literature. For a suitable class of convex optimal control problems, an upper bound is obtained *a posteriori* using the numerical approximation of the dual problem.

Our numerical tests confirm the results given by the theoretical analysis and suggest a convergence to zero with order one of the numerical duality gap. Establishing rigorously a duality relation between the numerical approximations of the primal and the dual problem seems to us an interesting direction of research that we would like to pursue. Beyond the independent theoretical interest,

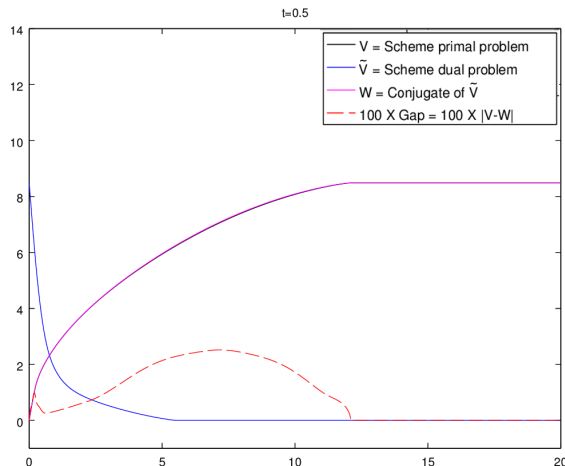


FIGURE 6. Test 2: Numerical solution  $V$  (black) compared with the convex conjugate of  $\tilde{V}$  (magenta). The dashed red line represents the numerical duality gap.

this would also allow us to obtain an *a priori* upper bound for the numerical error. The possibility of improving the order by higher order time stepping is also left for future research.

#### REFERENCES

- [1] G. Barles and E.R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.*, 36:33–54, 2002.
- [2] G. Barles and E.R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 43(2):540–558, 2005.
- [3] G. Barles and E.R. Jakobsen. Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Math. Comput.*, 74(260):1861–1893, 2007.
- [4] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4:271–283, 1991.
- [5] W.H. Beyer. *CRC Standard Mathematical Tables*. CRC Press, 28th edition, 1987.
- [6] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.*, 29(1):97–122, 1995.
- [7] M.G. Crandall, H. Ishii, and P.L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27(1):1–67, 1992.
- [8] D. Cuoco and H. Liu. A martingale characterization of consumption choices and hedging costs with margin requirements. *Math. Financ.*, 10:355–385, 2000.
- [9] K. Debrabant and E.R. Jakobsen. Semi-Lagrangian schemes for linear and fully non-linear diffusion equations. *Math. Comp.*, 82(283):1433–1462, 2012.
- [10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010.
- [11] M. Falcone and R. Ferretti. *Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations*, volume 133. SIAM, Philadelphia, 2014.
- [12] F.B. Hildebrand. *Introduction to Numerical Analysis*. New York: McGraw-Hill, 1956.
- [13] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Berlin, New York, Springer-Verlag, 1992.
- [14] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.*, 9(3):904–950, 1999.
- [15] N.V. Krylov. On the rate of convergence of finite-difference approximations for Bellman’s equations. *St. Petersburg Math. J.*, 9:639–650, 1997.
- [16] N.V. Krylov. Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies. *Electron. J. Probab.*, 4(2):1–19, 1999.
- [17] N.V. Krylov. On the rate of convergence of finite-difference approximations for Bellman’s equations with variable coefficients. *Probab. Theory Relat. Fields*, 117:1–16, 2000.

- [18] N.V. Krylov. Mean value theorem for stochastic integrals. *Ann. Probab.*, 29(1):385–410, 2001.
- [19] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer Verlag, New York, second edition, 2001.
- [20] J.L. Menaldi. Some estimates for finite difference approximations. *SIAM J. Control Optim.*, 27:579–607, 1989.
- [21] G.N. Milstein. *Numerical Integration of Stochastic Differential Equations*. Kluwer Academic Publishers, MA, 1995.
- [22] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*, volume 61. Series Stochastic Modeling and Applied Probability, Springer, 2009.
- [23] L.C.G. Rogers. *Duality in Constrained Optimal Investment Problems: A Synthesis*. Number 1814 in Paris–Princeton Lectures on Mathematical Finance. Springer, 2002.
- [24] J. Yong and X.Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*, volume 43 of *Applications of Mathematics*. Springer-Verlag, New York, 1999.

MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, ANDREW WILES BUILDING, OX2 6GG, OXFORD, UK  
*E-mail address:* athena.picarelli@maths.ox.ac.uk

MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, ANDREW WILES BUILDING, OX2 6GG, OXFORD, UK  
*E-mail address:* christoph.reisinger@maths.ox.ac.uk