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

Athena Picarelli, Christoph Reisinger

March 11, 2019

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

ATHENA PICARELLI AND CHRISTOPH REISINGER

Abstract. We consider a Markov chain approximation scheme for utility maximization problems in continuous time, which uses, in turn, a piecewise constant policy approximation, Euler-Maruyama time stepping, and a Gauß-Hermite approximation of the Gaußian increments. The error estimates previously derived in *A. Picarelli and C. Reisinger, Probabilistic error analysis for some approximation schemes to optimal control problems, arXiv:1810.04691* are asymmetric between lower and upper bounds due to the control approximation and improve on known results in the literature in the lower case only. In the present paper, we use duality results to obtain *a posteriori* upper error bounds which are empirically of the same order as the lower bounds. The theoretical results are confirmed by our numerical tests.

1. Introduction

We study the numerical approximation of a class of optimal control problems for diffusion processes arising in financial applications. It is well known that, under suitable assumptions, the value function associated to this kind of problems can be characterized as the solution of a second order Hamilton-Jacobi-Bellman (HJB) partial differential equation. To deal with the possible degeneracy of the diffusion component of the dynamics, it is in general necessary to consider solutions in the viscosity sense (see [7] for an overview). Furthermore, explicit solutions for this kind of nonlinear equations are rarely available, so that their numerical approximation becomes vital. In the framework of viscosity solutions, the basic theory of convergence for numerical schemes is established in [4]. The fundamental properties required are: monotonicity, consistency, and stability of the scheme. While standard finite difference schemes are in general non-monotone, semi-Lagrangian (SL) schemes (see [23, 6, 10]) are monotone by construction. The scheme considered in this paper belongs to this family and has been previously analyzed in [26].

We focus on theoretical bounds for the error associated to the scheme. By a technique pioneered by Krylov based on "shaking the coefficients" and mollification to construct smooth sub- and/or super-solutions, [20, 22, 1, 2, 3] prove certain fractional convergence orders significantly lower than one. These results are mainly derived by PDE techniques and strongly rely on the comparison principle between viscosity sub- and super-solutions of the HJB equations and the consistency properties of the scheme. For the scheme considered in the present paper, the probabilistic proof in [26] exploits the fact that the numerical scheme is based on a discrete approximation of the optimal control problem, specifically by a piecewise constant policy approximation, Euler-Maruyama time stepping, and a Gauß-Hermite approximation of the Gaußian increments. This yields the desired error bounds by a direct comparison between two value functions and leads to an improvement of the error contribution of the second and third of these approximations by avoiding the use of the truncation error. The piecewise constant policy approximation, however, introduces an asymmetry between the upper and the lower bound of the error and, as a result, the bounds in [26] give only a partial improvement of the classical PDE-based results.

For the class of convex optimal control problems studied here, namely typical utility maximization problems arising in financial applications, we propose to overcome this issue using information

coming from a dual problem. Indeed, an important part of the classical literature dealing with financial applications of optimal control theory (see the seminal work of Kramkov and Schachermayer [19]) applies duality techniques to solve utility maximization problems under suitable convexity assumptions. The basic idea of this method is to write the optimal control problem as a constrained optimization problem with respect to the state variable and then solve it by convex analysis techniques. A systematic approach to utility maximization problems admitting a dual formulation is discussed in [27]. Of these, the fairly general set-up of an optimal investment problem involving nonlinear dynamics given in [9] will be explicitly analyzed in this paper.

More specifically, a direct application of the results in [26] to this problem gives one-sided (lower) error bounds for the considered Markov chain approximation of order

$$h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)} \tag{1.1}$$

for timestep h, spatial mesh size Δx and number of Gaußian points M, for Lipschitz viscosity solutions. They coincide with the two-sided bounds in [10] for the standard linear-interpolation SL scheme, i.e. $M = 2$, and improve them for $M > 2$. In contrast, the piecewise constant policy approximation introduces an extra term in the upper bound of order $h^{1/4}$ (from a recent result in [17]), which strictly restricts the order for $M > 2$.

The main contribution of this paper is to analyse the error estimates in the case of optimal investment problems, which present a special structure that has not been exploited either by the classical literature on PDE-based error estimates for HJB equations nor by the analysis in [26], and prove that for the class of problems analyzed here, two-sided *a posteriori* bounds of the empirical order (1.1) can be obtained. As a side result, we complete the literature by deriving explicit values for the constants appearing in the error estimates in terms of the Lipschitz (resp. Hölder) regularity of the coefficients and the solution.

The paper is organised as follows. In Section 2, we define the problem set-up and state our assumptions. We define the scheme and give *a priori* lower error bounds for the primal problem in Section 3, and both *a priori* and *a posteriori* upper bounds, by way of the dual problem, in Section 4. We illustrate the theoretical results by numerical tests in Section 5, and offer conclusions and extensions in Section 6. In Appendix A, we derive explicit expressions for the constants in the error bounds.

2. Main assumptions and preliminary results

Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a probability space with filtration $\{\mathbb{F}_t, t \geq 0\}$ induced by a d-dimensional Brownian motion B and let $T > 0$. We consider a controlled (scalar) process governed by a dynamics of the following form, for $t \in [0, t)$,

$$\begin{cases} \mathrm{d}X_s = X_s\Big(r(s) + \alpha_s^\top (b(s) - r(s)\mathbb{1}) + g(s, \alpha_s)\Big)\,\mathrm{d}s + X_s \alpha_s^\top \sigma(s)\,\mathrm{d}B_s, & s \in (t, T) \\ X_t = x \geq 0, \end{cases} \tag{2.1}$$

where r, b, g and σ take values, respectively, in $\mathbb{R}, \mathbb{R}^d, \mathbb{R}$ and $\mathbb{R}^{d\times d}$ and $\mathbb{1} \equiv (1, \dots, 1)^\top \in \mathbb{R}^d$. Denote further by $\mathcal{A}$ the set of control policies, i.e. progressively measurable processes α taking values in a given set $A \subseteq \mathbb{R}^d$ such that $\int_0^T |\alpha_s|^2 \mathrm{d}s < +\infty$. This framework has been introduced and studied in [9], and encompasses a number of important optimal investment problems involving nonlinear dynamics, including the classical Merton problem [24], as special cases. In such models, the state $X_\cdot$ typically represents the wealth of an investor with initial endowment x at time t. The control vector $\alpha \equiv (\alpha_1, \dots, \alpha_d)^\top$ then determines the portion of wealth the investor puts in each stock. Here, the coefficient r is the return rate of a bond (riskless asset), while $b(\cdot) \equiv (b_1(\cdot), \dots, b_d(\cdot))^\top$ is the vector of the appreciation rates of the d considered stocks with volatility matrix $\sigma(\cdot)$. The nonlinearity in the investment strategy introduced by the function g models the effects of market frictions and trading constraints on the wealth (see [9, 8, 14]). We refer the reader to [27] for an overview of different utility maximization problems, including (2.1) and its special cases. We consider the following assumptions:

(H1) $A \subseteq \mathbb{R}^d$ is a convex set such that $0 \in A$.

(H2) (i) There exists $K_0 \geq 0$ such that

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

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

- there exists $K_1 \geq 0$ such that

$$|g(t, a) - g(t, a')| \leq K_1 |a - a'| \qquad \forall a, a' \in A, t \in [0, T];$$
$$|g(t, a) - g(s, a)| \leq K_1 (1 + |a|)|t - s|^{1/2} \qquad \forall\, t, s \in [0, T], a \in A;$$

- for each $t \in [0, T]$, $a \to g(t, a)$ is concave;
- $g(t, 0) = 0$ for all $t \in [0, T]$.

(H3) σ satisfies a uniform ellipticity condition, i.e. there exists $\eta > 0$ such that

$$\xi^\top \sigma \sigma^\top \xi \geq \eta |\xi|^2 \qquad \forall \xi \in \mathbb{R}^d.$$

One has the following existence and uniqueness result:

Lemma 2.1. *Let assumptions (H1)-(H3) be satisfied. For any choice of the control $\alpha \in \mathcal{A}$ and $x \geq 0$ there exists a unique strong solution to equation* (2.1).

Proof. For $x > 0$, a solution can be defined as $X_\cdot = \exp(Z_\cdot)$, where

$$Z_\cdot = z + \int_t^\cdot r(s) + \alpha_s^\top (b(s) - r(s)\mathbb{1}) + g(s, \alpha_s) - \frac{1}{2}(\alpha_s^\top \sigma)^2 \mathrm{d}s + \int_t^\cdot \alpha_s^\top \sigma(s)\, \mathrm{d}B_s,$$

for $z = \log x$, which is well defined under assumptions (H1)-(H3) for any $\alpha \in \mathcal{A}$. Moreover, for $x = 0$ the process $X \equiv 0$ is the unique solution to (2.1) for any $\alpha \in \mathcal{A}$. Uniqueness also follows by standard arguments. □

We denote by $X_\cdot^{t,x,\alpha}$ the unique solution of equation (2.1). To simplify the notation, where no ambiguities arise, we will indicate the starting point (t, x) of the processes involved as a subscript in the expectation, i.e. $\mathbb{E}_{t,x}[\cdot]$.

The value function $v : [0, T] \times [0, +\infty) \to \mathbb{R}$ of the optimal control problem is defined by

$$v(t, x) := \sup_{\alpha \in \mathcal{A}} \mathbb{E}_{t,x}\big[U(X_T^\alpha)\big], \tag{2.2}$$

where $U : [0, +\infty) \to \mathbb{R}$ is the so-called utility function of the investor and it is assumed to satisfy the following assumptions:

(H4) $U \in C^1((0, +\infty); \mathbb{R})$;
U is concave and strictly increasing;
$\lim_{x \to +\infty} U'(x) = 0$.

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

$$v(t, x) = \sup_{\alpha \in \mathcal{A}} \mathbb{E}_{t,x}\Big[v(t + \theta, X_{t+\theta}^\alpha)\Big], \tag{2.3}$$

from which, at least formally, one can show that the Hamilton-Jacobi-Bellman (HJB) equation associated with the optimal control problem (2.2) is

$$-v_t + \sup_{a \in A} \left(-x \left(r(t) + a^\top (b(t) - r(t)\mathbb{1}) + g(t, a) \right) v_x - \frac{1}{2} x^2 Tr[a(\sigma\sigma^\top)(t)a^\top] v_{xx} \right) = 0, \tag{2.4}$$

completed with the terminal condition $v(T, x) = U(x)$. We refer the reader to [30, Section 3, Chapter 4] and the references therein for a complete overview on the dynamic programming approach to optimal control problems. In the general case, v is not expected to have sufficient regularity to satisfy the previous equation in the classical sense and even if (2.4) admits a classical solution, it is rarely found explicitly. To handle the problem in its full generality, the notion of viscosity solution is needed (see [7] for an overview). Indeed, under suitable assumptions, it can be proved (see

for instance [30, Theorems 5.2 and 6.1]) that v defined in (2.2) is the unique continuous viscosity solution to (2.4).

3. The numerical scheme

We consider here the scheme analyzed in [26]. It belongs to the family of the so-called semi-Lagrangian (SL) schemes (see [6, 11, 21, 23] for their earlier introdution) which are based on discretization of the control set $\mathcal{A}$ and a Markov chain approximation of the associated optimal control problem. For completeness, we briefly discuss below the main features of the scheme. We refer the reader to [26] for further details.

3.1. Description of the scheme.

We start by introducing a discretization in time. Let $N \geq 1$,

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

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

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

In what follows, we identify any element $\alpha \in \mathcal{A}_h$ by the sequence of random variables a_i taking values in A (denoted by $a_i \in A$ 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 (2.2) to controls in $\mathcal{A}_h$, that is

$$v_h(t,x) := \sup_{\alpha \in \mathcal{A}_h} \mathbb{E}_{t,x} \left[U(X_T^\alpha) \right]. \tag{3.1}$$

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

$$v(t,x) \geq v_h(t,x), \tag{3.2}$$

for any $t \in [0,T]$, $x \geq 0$. Un upper bound of order $1/6$ for the error related to this approximation was first obtained by Krylov in [21]. Recently, this estimate has been improved to the order $1/4$ in [17], so that one has

$$v(t,x) \leq v_h(t,x) + Ch^{1/4} \tag{3.3}$$

for some constant $C \geq 0$. We point out that the results in [21] and [17] require some additional assumptions on the coefficients of the dynamics and do not directly apply to problem (2.1) to (2.2). It is possible that analogous estimates hold also in the setting of the present paper, but since we do not make use of (3.3) here, we did not check this point in detail. Indeed, a main objective of the present paper is to by-pass the estimate (3.3), which turns out to be a bottleneck in the provable approximation order, while still using the piecewise constant policy approximation itself by building an approximation to v_h. The more important observation from [17] is therefore that a better order than $1/4$ is not provable in the general case of Lipschitz viscosity solutions. Then no matter how precise the estimates obtained for the error of the final approximation to v_h are, without any further information the upper error bounds to v cannot be more accurate than $O(h^{1/4})$. Section 4 will show how the use of the information coming from the definition of a dual problem can eliminate this asymmetry.

For any given $\alpha \equiv (a_0, \dots, a_{N-1}) \in \mathcal{A}_h$, we consider the Euler-Maruyama approximation of the process $X^{t,x,\alpha}_\cdot$ given by the following recursive relation:

$$X_{t_{i+1}} = X_{t_i} + h\, X_{t_i} \left(r(t_i) + a_i^\top (b(t_i) - r(t_i)\mathbb{1}) + g(t_i, a_i) \right) + X_{t_i} a_i^\top \sigma(t_i) \Delta B_i \tag{3.4}$$

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

$$\Delta B_i \sim \sqrt{h}\, \mathcal{N}(0, I_d) \qquad \forall i = 0, \dots, N-1. \tag{3.5}$$

	ξ_i	λ_i
$M=2$	± 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 1. 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 $\{(z_i,\omega_i)\}_{i=1,\dots,M}$ for larger M.

We denote by $\overline{X}_{\cdot}^{t_n,x,\alpha}$ the solution to (3.4) with the control $\alpha\equiv(a_0,\dots,a_{N-1})\in\mathcal{A}_h$ and such that $\overline{X}_{t_n}^{t_n,x,\alpha}=x$. In the next step, we work towards a Markov chain approximation of $\overline{X}_{\cdot}^{t_n,x,\alpha}$.

Let us start for simplicity with the case $d=1$. Let $M\geq 2$ and 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

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

With the definitions

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

one can make use of the following approximation (see, e.g., [16, p. 395])

$$\int_{-\infty}^{+\infty}f(y)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy\approx\sum_{i=1}^{M}\lambda_i f(\xi_i),\tag{3.6}$$

which holds for any smooth real-valued function f (say f at least C^{2M}). Observing that $\lambda_i\geq 0,\forall i=1,\dots,M$, and $\sum_{i=1}^{M}\lambda_i=1$, given the sequence $\{\zeta_n\}_{n=0,\dots,N-1}$ of i.i.d. random variables such that for any $n=0,\dots,N-1$

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

one has

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

For any control $\alpha\equiv(a_0,\dots,a_{N-1})\in\mathcal{A}_h$, we will denote by $\widehat{X}_{\cdot}^{t_n,x,\alpha}$ the Markov chain approximation of the process $\overline{X}_{\cdot}^{t_n,x,\alpha}$, i.e.

$$\begin{cases}\widehat{X}_{t_n} &= x,\\ \widehat{X}_{t_{i+1}}=\widehat{X}_{t_i}+h\,\widehat{X}_{t_i}\left(r(t_i)+a_i(b(t_i)-r(t_i))+g(t_i,a_i)\right)+\sqrt{h}\,\widehat{X}_{t_i}a_i\sigma(t_i)\,\zeta_i,\end{cases}\tag{3.7}$$

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

Applying to (2.3) with $\theta=h$ the piecewise control approximation, the Euler-Maruyama discretization and the Gauß-Hermite quadrature formula (3.6), we obtain the following recursive semidiscrete approximation of the value function

$$\begin{cases}V(t_n,x) &= \sup\limits_{a\in A}\sum_{i=1}^{M}\lambda_i V\left(t_{n+1},x+h\,x\left(r(t_n)+a(b(t_n)-r(t_n))+g(t_n,a)\right)+\sqrt{h}\,x\,a\sigma(t_n)\,\xi_i\right)\\ &= \sup\limits_{a\in A}\mathbb{E}_{t_n,x}\left[V(t_{n+1},\widehat{X}^{a}_{t_{n+1}})\right], & n=N-1,\dots,0,\\ V(t_{_N},x) &= U(x).\end{cases}\tag{3.8}$$

Iterating gives the following representation formula for V:

$$V(t_n,x)=\sup_{\alpha\in\mathcal{A}_h}\mathbb{E}_{t_n,x}\left[U(\widehat{X}_T^{\alpha})\right].$$

In the case of $d > 1$, it is possible to extend formula (3.6) by a tensor product approximation

$$\int_{\mathbb{R}^d} f(y) \frac{e^{-\frac{|y|^2}{2}}}{(2\pi)^{p/2}} \,\mathrm{d}y \approx \sum_{i_1,\dots,i_d=1}^{M} \lambda_{i_1} \cdot \ldots \cdot \lambda_{i_d} f(\xi_{i_1}, \dots, \xi_{i_d}). \tag{3.9}$$

For any $i \equiv (i_1, \dots, i_d) \in \{1, \dots, M\}^d$, with $\xi_i \equiv (\xi_{i_1}, \dots, \xi_{i_d})^T$ and $\lambda_i \equiv \lambda_{i_1} \cdot \ldots \cdot \lambda_{i_d}$, the scheme reads

$$\begin{cases} V(t_n, x) = \sup\limits_{a \in A} \sum\limits_{i \in \{1 \dots M\}^d} \lambda_i V\Big(t_{n+1}, x + h\, x \left(r(t_n) + a^\top (b(t_n) - r(t_n)\mathbb{1}) + g(t_n, a)\right) + \sqrt{h}\, x\, a^\top \sigma(t_n) \cdot \xi_i\Big), \\ V(t_N, x) = U(x). \end{cases}$$

This construction leads to an exponential growth of the computational complexity in the dimension d, as it requires at each time step and for each node the evaluation of the solution at M^d points. However, as discussed in [26], the provable order of the error is the same for any choice of weights $\hat{\lambda}_i \geq 0$ and nodes $\hat{\xi}_i$, $i = 1, \dots, \hat{M}$ (with $\hat{M}$ possibly lower than M^d) which integrate exactly all polynomials of degree lower or equal than $2M - 1$.

The existence of such nodes and weights with $\hat{M} \leq \binom{2M-1+d}{d}$ is guaranteed by Tchakaloff's Theorem (see [28]), while a constructive method for independent Gaußian random variables as in the present case is proposed in [13]. Moreover, an efficient procedure for the general, dependent case applied to the uniform measure is given in [29]. This gives a substantial reduction of the computational cost for large d and moderate M. In what follows, we will use the notation $\{(\hat{\lambda}_i, \hat{\xi}_i)\}_{i=1,\dots,\hat{M}}$ to generalise (3.8) to any $d \geq 1$.

We introduce now a discretization of the space variable. Let $\Delta x > 0$ and consider the space grid $\mathcal{G}_{\Delta x} := \{x_m = m\Delta x : m \in \mathbb{Z}\}$. We also write $\mathcal{G}^+_{\Delta x} := \{x_m = m\Delta x : m \in \mathbb{N}\}$. Let $\mathcal{I}[\cdot]$ denote the linear interpolation operator with respect to the space variable, satisfying for every Lipschitz function ϕ (with Lipschitz constant L_ϕ):

$$\begin{cases} (i)\ \mathcal{I}[\phi](x_m) = \phi(x_m),\ \forall m \in \mathbb{Z}, \\ (ii)\ |\mathcal{I}[\phi](x) - \phi(x)| \leq L_\phi \Delta x, \\ (iii)\ |\mathcal{I}[\phi](x) - \phi(x)| \leq C\Delta x^2 \|D^2\phi\|_\infty \ \text{ if } \phi \in C^2(\mathbb{R}), \\ (iv) \text{ for any functions } \phi_1, \phi_2 : \mathbb{R} \to \mathbb{R},\ \phi_1 \leq \phi_2 \Rightarrow \mathcal{I}[\phi_1] \leq \mathcal{I}[\phi_2]. \end{cases} \tag{3.10}$$

We define an approximation W on this fixed grid as follows:

$$\begin{cases} W(t_n, x_m) = \sup\limits_{a \in A} \sum_{i=1}^{\hat{M}} \hat{\lambda}_i\, \mathcal{I}[W]\Big(t_{n+1}, x_m + h\, x_m \left(r(t_n) + a^\top (b(t_n) - r(t_n)\mathbb{1}) + g(t_n, a)\right) + \sqrt{h}\, x_m\, a^\top \sigma(t_n) \cdot \hat{\xi}_i\Big), \\ W(t_N, x_m) = U(x_m), \end{cases} \tag{3.11}$$

for $n = N - 1, \dots, 0$ and $m \in \mathbb{N}$. We will refer to this as the fully discrete scheme. For $M = 2$, the scheme coincides with the one introduced by Camilli and Falcone in [6]. However, as explained in the next section, the error estimates derived in the present paper improve the state of the art for this class of schemes only when $M > 2$ is considered.

3.2. **An *a priori* lower bound for v.** Under suitable assumptions, *a* priori estimates of the following form are proven in [26]:

$$-C\left(h^{(M-1)/2M} + \frac{\Delta x}{h}\right) \leq v(t_n, x_m) - W(t_n, x_m) \leq C\left(h^{1/4} + h^{(M-1)/2M} + \frac{\Delta x}{h}\right) \tag{3.12}$$

for any $n = 0, \dots, N$, $m \in \mathbb{N}$ and a constant C, possibly depending on x_m (the dependency of C on x_m can be explicitly derived and one has $C \leq C_0(1 + x_m^{2M})$ for some constant C_0).

The result is obtained by a direct comparison between the optimal problems (3.1) and (3.11). Contributing to the error estimates above are: the Euler-Maruyama error of order $h^{1/2}$; the Gauß-Hermite quadrature error of order h^{M-1}; and the interpolation error of order Δx, accumulated

over $1/h$ steps to $\Delta x/h$. The bounds (3.12) are then the result of the use of (3.2) to (3.3) for the piecewise constant controls approximation, which introduces the aforementioned asymmetry in the estimates (given by the term $h^{1/4}$ in the right-hand side of (3.12)), and of a regularization procedure, the so-called "shaking coefficients" technique in [20] and subsequent works, which is the classical tool to deal with nonsmooth solutions.

In the analysis of [26], Lipschitz continuity of the utility function is required. This property is not satisfied by very common utility functions found in literature (for instance the typical power utility $U(x) = x^p/p$ with $p \in (0,1)$). For this reason, we introduce a further approximation of the problem and consequently have to estimate an additional error contribution. Letting $\rho, c_0 > 0$ and $x_\rho = c_0/\rho$, $y_\rho = \rho$, we define

$$U_\rho(x) := \begin{cases} U(0) + \frac{U(x_\rho)-U(0)}{x_\rho} x & \text{if} \quad 0 \le x \le x_\rho, \\ U(x) & \text{if} \quad x_\rho < x \le y_\rho, \\ U(y_\rho) & \text{if} \quad x > y_\rho, \end{cases}$$

so that U_ρ is Lipschitz with Lipschitz constant $L_\rho := (U(x_\rho) - U(0))/x_\rho$ and $U_\rho \to U$ as $\rho \to +\infty$. We denote by V_ρ and W_ρ the numerical approximation recursively obtained by (3.8) and (3.11), respectively, replacing U with U_ρ.

Adapting the arguments of [26] to the present problem, we can obtain the following *a* priori estimate for the lower bound of the quantities $v - V_\rho$ and $v - W_\rho$:

Proposition 3.1. *Let assumptions (H1) to (H3) be satisfied. Then, there exists a constant $C \ge 0$ such that for any $n = 0, \dots, N$, $x \ge 0$*

$$v(t_n, x) \ge V_\rho(t_n, x) - L_\rho C(1 + x^{2M}) h^{(M-1)/2M}, \tag{3.13}$$

and for any $n = 0, \dots, N$, $m \in \mathbb{N}$

$$v(t_n, x_m) \ge W_\rho(t_n, x_m) - L_\rho C(1 + x_m^{2M}) \left(h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)} \right). \tag{3.14}$$

Proof. Without loss of generality we can assume that the set of control values A is bounded. Indeed, when only the lower bound is considered, a potential further restriction on the set of admissible controls $\mathcal{A}$ does not affect the estimates. Let v_ρ the value function given by (2.2) replacing U with U_ρ. By the assumptions on U one simply has $v_\rho \le v$.

When only the time discretization is taken into account, the estimate (3.13) directly follows by [26, Section 4.2, equation (4.9)] observing that the additional requirement in [26] of working with bounded coefficient is uniquely required in order to apply the results on the piecewice constant control approximation in [17] and then it is not required to get the lower bound. Moreover by [26, Section 4.3] for the fully discrete scheme one has

$$v(t_n, x_m) - W_\rho(t_n, x_m) \ge -L_\rho C(1 + |x_m|^{2M}) \left(h^{(M-1)/2M} + \frac{\Delta x}{h} \right),$$

where C only depends on M, T, the constants K_0 and K_1 in assumption (H2) and the uniform bound on elements in A. Balancing the terms $h^{(M-1)/2M}$ and $\Delta x/h$ on the right-hand side leads to (3.14). □

The scheme we are considering is monotone, stable and it has order one of consistency (for smooth test functions) for any $M \ge 2$. For a scheme of this type, (upper and lower) error bounds of order $1/4$ in h have been provided in [3, 10] by PDE techniques. Splitting each contribution to the error, namely the control discretization and the Euler-Maruyama and Gauß-Hermite approximations, the probabilistic proof proposed in [26] gives an improvement to the lower bound of these estimates by increasing the value of $M > 2$, i.e. by using a more accurate quadrature formula. For large M, the order is arbitrarily close to $1/2$ in h and $1/3$ in Δx, improving the corresponding orders $1/4$ and $1/5$ obtained for $M = 2$. It is hence for $M > 2$ that the term $h^{1/4}$ in the upper bound becomes dominant and restricts the order to $1/4$, independently of M. The analysis that follows aims to

eliminate the asymmetry in (3.12), due to the control discretization, and provide an upper bound of the same order of the lower bounds obtained in Proposition 3.1.

4. Duality-based error estimates

In this section we discuss how duality theory can be employed to obtain an upper bound of the error associated with our approximation scheme. Assuming to be able to extend either the PDE-based error estimates in [20, 22, 1, 2, 3] or the probabilistic ones in [26] to the particular problem (2.1) and (2.2), this would result in both cases in an upper bound of order $1/4$ in h for any choice of $M \geq 2$, as explained at the end of the previous section. We show here that for our class of problems it is possible to pass through the definition of a dual problem to improve such estimates.

4.1. **The dual problem.** The dual problem associated with (2.1) and (2.2) is defined in [9] by the dynamics

$$\begin{cases} \mathrm{d}Y_s = -\big(r(s)+\tilde{g}(s,\nu_s)\big)Y_s\,\mathrm{d}s + (\sigma\sigma^T(s))^{-1}Y_s(r(s)\mathbb{1}-b(s)-\nu_s)\cdot\sigma(s)\,\mathrm{d}B_s, \quad s\in[t,T], \\ Y_t = y, \end{cases} \tag{4.1}$$

where

$$\widetilde{g}(t,\nu) := \sup_{a\in A}\big\{g(t,a) - a\cdot\nu\big\}, \tag{4.2}$$

and the dual utility function $\widetilde{U}$ is the convex conjugate of U, i.e.

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

The dual value function is defined by

$$\tilde{v}(t,y) = \inf_{\nu\in\mathcal{V}} \mathbb{E}_{t,y}\big[\widetilde{U}(Y_T^\nu)\big], \tag{4.3}$$

where $\mathcal{V}$ is the set of $\mathbb{R}^d$-valued progressively measurable processes such that $\int_0^T |\nu_s|^2\mathrm{d}s + \int_0^T \widetilde{g}(s,\nu_s)\mathrm{d}s < +\infty$. One has the following duality result:

Proposition 4.1 ([9], Theorem 2)**.** *Let assumptions (H1)-(H4) be satisfied. Then for any $t\in[0,T]$, $x\geq 0$, the primal and dual value functions, v and $\widetilde{v}$, satisfy conjugate relation*

$$v(t,x) = \inf_{y>0}\big\{\widetilde{v}(t,y) + xy\big\}. \tag{4.4}$$

Remark 1. *The results in* [9] *hold also if r, b, σ and g are stochastic processes. However, our approximation scheme makes use of the Markovian framework. We would have to add extra variables to the state space to account for this, which is outside the scope of this work.*

Remark 2. *As mentioned in* [27, Section 6.5]*, the usual Inada condition*

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

(requested in [9]*) is not necessary for proving the main duality results. Moreover, the discussion in* [27, Section 6.5] *can also be used to show that* (4.4) *also holds for v and $\widetilde{v}$ replaced by v_ρ and $\widetilde{v}_\rho$.*

4.2. **Approximation of the dual problem.** The scheme presented in Section 3.1 can be used to approximate the value function $\widetilde{v}$ associated with the dual problem (4.1)-(4.3). To this end, we define $\Gamma\subseteq\mathbb{R}^d$ a compact set approximating the range of ν.

For the discrete time scheme, one can recursively define

$$\begin{cases} \widetilde{V}(t_n,y) = & \inf\limits_{\gamma\in\Gamma} \mathbb{E}_{t_n,y}\big[\widetilde{V}(t_{n+1},\widehat{Y}^\gamma_{t_{n+1}})\big], \qquad\qquad n = N-1,\dots,0, \\ \widetilde{V}(t_{_N},y) = & \widetilde{U}(y), \end{cases} \tag{4.5}$$

where for any $n = 0, \dots, N-1$, $y \geq 0$, $\nu \equiv (\gamma_n, \dots, \gamma_{N-1})$, with $\gamma_i \in \Gamma \subseteq \mathbb{R}^d$ $(i = n, \dots, N-1)$, $\widehat{Y}^{t_n,y,\nu}_{\cdot}$ is the Markov chain recursively defined by

$$\begin{cases} \widehat{Y}_{t_n} = y, \\ \widehat{Y}_{t_{i+1}} = \widehat{Y}_{t_i} - h\big(r(t_i)+\tilde{g}(t_i,\gamma_i)\big)\widehat{Y}_{t_i} + \sqrt{h}(\sigma\sigma^T(t_i))^{-1}\widehat{Y}_{t_i}(r(t_i)\mathbb{1} - b(t_i) - \gamma_i)\cdot\sigma(t_i)\zeta_i \end{cases} \tag{4.6}$$

for $i = n, \dots, N-1$. The fully discrete version of the scheme is then given by

$$\begin{cases} \widetilde{W}(t_n, y_j) = \inf\limits_{\gamma\in\Gamma} \sum_{i=1}^{\hat{M}} \hat{\lambda}_i \, \mathcal{I}[\widetilde{W}]\left(t_{n+1}, y_j + h\, y_j\,(r(t_n)+\tilde{g}(t_n,\gamma)) + \sqrt{h} y_j (\sigma\sigma^T(t_n))^{-1}(r(t_n)\mathbb{1} - b(t_n) - \gamma)\cdot\sigma(t_n)\hat{\xi}_i\right) \\ \widetilde{W}(t_N, y_j) = \widetilde{U}(y_j), \end{cases} \tag{4.7}$$

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

We denote by $\widetilde{U}_\rho$ the convex conjugate of the approximated utility function U_ρ, i.e.

$$\widetilde{U}_\rho(y) := \sup_{x\geq 0}\{U_\rho(x) - xy\}$$

and by $\widetilde{V}_\rho$ and $\widetilde{W}_\rho$ the numerical approximations obtained respectively by (4.5) and (4.7), replacing $\widetilde{U}$ with $\widetilde{U}_\rho$. Observe that $\widetilde{U}_\rho : [0, +\infty) \to \mathbb{R}$ is decreasing and Lipschitz continuous with constant $\widetilde{L}_\rho := y_\rho$. Moreover, it follows by the very definition of U_ρ that $\widetilde{U}_\rho(y) = 0$ for $y \geq L_\rho$.

4.3. **An *a priori* upper bound for $\tilde{v}$.** The approximation scheme we defined for the dual problem is 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 [26] to obtain an accurate upper bound for the differences $\widetilde{v} - \widetilde{V}_\rho$ and $\widetilde{v} - \widetilde{W}_\rho$.

Proposition 4.2. *Let assumptions (H1)-(H4) be satisfied and let us additionally assume that A is a bounded set. Then, there exists a constant $\widetilde{C} \geq 0$, and $\delta : \mathbb{R}^+ \to \mathbb{R}^+$ with $\delta(\rho) = o(\rho^{-p})$ as $\rho \to \infty$ for all $p > 0$, such that for any $n = 0, \dots, N$, $y > 0$,*

$$\widetilde{v}(t_n, y) \leq \widetilde{V}_\rho(t_n, y) + \widetilde{L}_\rho\widetilde{C}(1 + y^{2M})h^{(M-1)/2M} + \delta(\rho), \tag{4.8}$$

and for any $n = 0, \dots, N$, $j \in \mathbb{N}^+$,

$$\widetilde{v}(t_n, y_j) \leq \widetilde{W}_\rho(t_n, y_j) + \widetilde{L}_\rho\widetilde{C}(1 + y_j^{2M})\left(h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)}\right) + \delta(\rho). \tag{4.9}$$

Proof. We can assume that for any $\nu \in \mathcal{V}$ one has $\nu_s \in \Gamma$ a.e., a.s. If this is not the case, the upper bound still applies (but is less sharp as the infimum is taken over a smaller set).

Let $\tilde{v}_\rho$ the value function given by (4.3) replacing $\widetilde{U}$ with $\widetilde{U}_\rho$. Considering different type of Lipschitz continuous approximations of U the difference between v and v_ρ can be estimated by using large deviations arguments (see for instance [12]) to bound

$$\mathbb{P}\Big[\big(X_T^{t,x,\alpha} \leq x_\rho\big) \cup \big(X_T^{t,x,\alpha} \geq y_\rho\big)\Big]. \tag{4.10}$$

This decreases exponentially fast as ρ goes to $+\infty$. Therefore, recalling Remark 2, one has

$$\widetilde{v}(t, y) \leq \widetilde{v}_\rho(t, y) + \delta(\rho)$$

with $\delta(\cdot)$ as desired.

The result then follows by [26] observing that under the assumptions (H1)-(H3) one has

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

(where C_0 depends on T, the constants K_0, K_1 and η in assumptions (H2)-(H3) and the uniform bounds on the elements of A and Γ), so that the dynamics (4.1) satisfies the assumptions in [26] leading to the desired estimates. □

Remark 3. *Under the compactness assumption on the set A, the set Γ is typically unbounded, so that in general if we restrict the numerical approximation to an arbitrary bounded subset of Γ convergence can be lost. In the numerical tests below we make use of the information we have on the optimal controls $\alpha^* \in \mathcal{A}$ and $\nu^* \in \mathcal{V}$ to derive uniform bounds on the elements of A and Γ. A more detail discussion of the general case is given in Remark 5.*

4.4. **Using duality in error estimates.** In the sequel, we will use the following notation: for any $n = 0, \dots, N$, $x \geq 0$

$$G^h(t_n, x) := \inf_{y>0} \Big\{ \widetilde{V}_\rho(t_n, y) + xy \Big\} - V_\rho(t_n, x) \tag{4.11}$$
$$I^h(t_n, x) := \arg\min_{y>0} \Big\{ \widetilde{V}_\rho(t_n, y) + xy \Big\}$$

and for any $n = 0, \dots, N$, $m \in \mathbb{N}$

$$G^{h,\Delta x}(t_n, x_m) := \inf_{j \in \mathbb{N}^+} \Big\{ \widetilde{W}_\rho(t_n, y_j) + x_m y_j \Big\} - W_\rho(t_n, x_m) \tag{4.12}$$
$$I^{h,\Delta x}(t_n, x_m) := \arg\min_{j \in \mathbb{N}^+} \Big\{ \widetilde{W}_\rho(t_n, y_j) + x_m y_j \Big\}$$

We refer to G^h and $G^{h,\Delta x}$ as the *numerical duality gap* of the semidiscrete and fully discrete scheme respectively.

One has the following result:

Theorem 4.1. *Let assumptions (H1)-(H4) be satisfied. Then, there exist some constants $C, \widetilde{C} \geq 0$ and $\delta : \mathbb{R}^+ \to \mathbb{R}^+$ with $\delta(\rho) = o(\rho^{-p})$ as $\rho \to \infty$ for all $p > 0$, such that for any $n = 0, \dots, N$*

$$-L_\rho C(1 + x^{2M}) h^{(M-1)/2M} \leq v(t_n, x) - V_\rho(t_n, x) \leq G^h(t_n, x) + \widetilde{L}_\rho \widetilde{C}(1 + (I^h(t_n, x))^{2M}) h^{(M-1)/2M} + \delta(\rho) \tag{4.13}$$

and for any $n = 0, \dots, N$, $m \in \mathbb{N}$

$$\begin{aligned} -L_\rho C(1 + x_m^{2M}) \Big(h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)} \Big) \leq v(t_n, x_m) - W_\rho(t_n, x_m) \\ \leq G^{h,\Delta x}(t_n, x_m) + \widetilde{L}_\rho \widetilde{C}(1 + (I^{h,\Delta x}(t_n, x_m))^{2M}) \Big(h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)} \Big) + \delta(\rho). \end{aligned} \tag{4.14}$$

Proof. The first inequalities in (4.13) and (4.14) follow directly by Proposition 3.1. It remains to prove the upper bounds. We prove the result for the semi-discrete scheme, while the proof for the fully discrete scheme follows by similar arguments. Let us first assume A is bounded. Thanks to Proposition 4.1, Proposition 4.2 and the definition of $I^h(\cdot, \cdot)$ one has

$$\begin{aligned} v(t_n, x) = \inf_{y>0} \{ \widetilde{v}(t_n, y) + xy \} &\leq \inf_{y>0} \Big\{ \widetilde{V}_\rho(t_n, y) + xy + \widetilde{L}_\rho \widetilde{C}(1 + y^{2M}) \Big\} + \delta(\rho). \\ &\leq \widetilde{V}_\rho(t_n, I^h(t_n, x)) + x\, I^h(t_n, x) + \widetilde{L}_\rho \widetilde{C} \left(1 + (I^h(t_n, x))^{2M}\right) + \delta(\rho). \\ &= \inf_{y>0} \Big\{ \widetilde{V}_\rho(t_n, y) + xy \Big\} + \widetilde{L}_\rho \widetilde{C} \left(1 + (I^h(t_n, x))^{2M}\right) + \delta(\rho). \end{aligned}$$

Therefore,

$$v(t_n, x) - V_\rho(t_n, x) \leq \inf_{y>0} \Big\{ \widetilde{V}_\rho(t_n, y) + xy \Big\} - V_\rho(t_n, x) + \widetilde{L}_\rho \widetilde{C} \left(1 + (I^h(t_n, x))^{2M}\right) + \delta(\rho),$$

which gives the desired result. If A is not bounded we will consider a subset of $\mathcal{A}$ with uniformly bounded controls and the result still holds true. However, if the optimal control does not belong to this restricted set of controls this additional approximation will be captured by the duality gap (see also the discussion in Remark 5). □

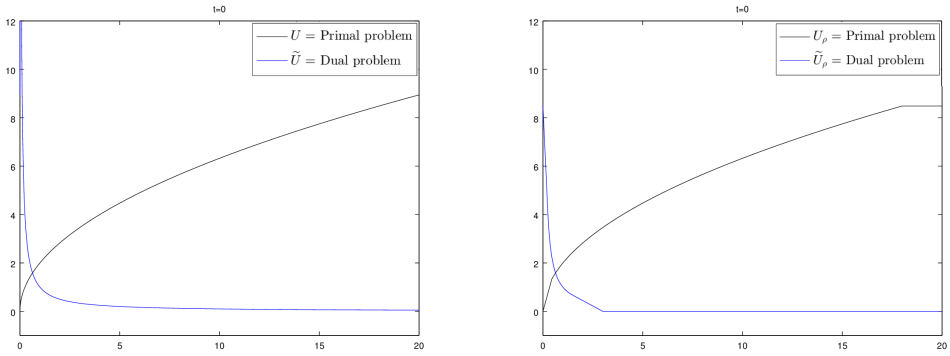

FIGURE 2. The power utility function U with its conjugate $\widetilde{U}$ (left) and the Lipschitz continuous approximation U_ρ with its conjugate $\widetilde{U}_\rho$ (right). Here, $x_{\max} = 20$, $\rho = 18$ and $c_0 = 8$.

Observe that due to the particular convexity feature of the dual problem, the quantity $I(x)$ typically increases as x approaches 0.

The duality gap for the fully discrete scheme is computable efficiently, see e.g. [15, Section 3.4], so that (4.14) provides a practically useful *a posteriori* bound.

A priori bounds could be obtained by proving that the numerical duality gap G^h (resp. $G^{h,\Delta x}$) decays with order at most $h^{(M-1)/2M}$ (resp. $h^{(M-1)/2M} + \Delta x^{(M-1)/(3M-1)}$). This requires a proof that V_ρ and $\widetilde{V}_\rho$ (resp. W_ρ and $\widetilde{W}_\rho$) satisfy an approximated duality relation. Indeed, the key feature of dynamics (2.1) and (4.1) leading to the conjugate relation (4.4) is the following so called "polar property"

$$\sup_{\nu\in\mathcal{V}} \mathbb{E}\left[X_T^{t,x,\alpha} Y_T^{t,y,\nu}\right] = xy \qquad \forall x, y \geq 0, t \in [0,T], \alpha \in \mathcal{A}.$$

For the discrete time dynamics $\widehat{X}_\cdot$ and $\widehat{Y}_\cdot$ defined in (3.7) and (4.6) respectively a straightforward computation shows that

$$\sup_{\nu\in\mathcal{V}} \mathbb{E}\left[\widehat{X}_T^{t_n,x,\alpha} \widehat{Y}_T^{t_n,y,\nu}\right] \leq xy\,(1+Ch) \qquad \forall x, y \geq 0, n = 0, \ldots, N,\ \alpha \in \mathcal{A}$$

for some constant $C \geq 0$. We conjecture that also the other inequality holds. This finds a confirmation in our numerical tests (see Tables 2 and 5 in Section 5) where at least first order of convergence in h of the numerical duality gap is observed. However the rigorous prove of the result involves delicate convex analysis arguments and we plan to investigate this point in future work.

5. NUMERICAL TESTS

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

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

We study the case of a power utility function:

$$U(x) = \frac{x^p}{p} \qquad \text{for some } p \in (0,1). \tag{5.1}$$

We consider the modification U_ρ of the utility function obtained for $\rho = 18$ and $c_0 = 8$. The utility function U for $p = 0.5$ and its conjugate $\widetilde{U}$, as well as its Lipschitz continuous approximation U_ρ and its conjugate $\widetilde{U}_\rho$ are shown in Figure 2.

In our tests, we take $M = 4$ with $\Delta x \sim h^{11/8}$ obtained from (4.14), more specifically $J \sim \lceil N^{11/8} \rceil$.

Remark 4. *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, even in the case of smooth solutions, the use of the Euler-Maruyama scheme reduces the order of consistency of the overall scheme to one (noting that a modified proof utilising the higher weak order 1 of the Euler-Maruyama scheme, compared to the strong order 1/2, can be used in the smooth case), regardless of the value of M. 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* [18].

Remark 5. *For the optimization over the controls in our computations, we truncate A and Γ first to a finite interval and then discretise the interval by N_a and N_γ 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 Δx.*

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_γ^{-1} since we have a smooth dependence of the Hamiltonian on the control. In our tests, we take $N_a \sim N_\gamma \sim N$, such that the control discretisation error becomes eventually negligible.

Remark 6. *It is clear that as the point x_m approaches 0 or $x_{\max}$ it may happen that $\widehat{X}_{t_{n+1}}^{t_n,x_m}$ oversteps the domain. In this case, we use linear extrapolation in order to define W_ρ and $\widetilde{W}_\rho$ outside the domain.*

Test 1: Merton problem. We first study the classical Merton problem. This corresponds to the dynamics (2.1) with $g \equiv 0$, constant coefficients b, r, σ and $A = \mathbb{R}$. It is well known that for this problem there exists a closed-form solution given by (see, e.g. [25])

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

where U and p are given in (5.1), and

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

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	σ	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 W_ρ 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 ρ, we cannot get convergence to v. In fact, the probability in (4.10), even if small at points x far from the boundaries of the domain, is different from zero everywhere (see also Figure 3, left). To reduce the contribution to the error coming from the term in ρ we compute the error locally, away from the boundary of the computational domain.

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 3 (right) we show the numerical solutions W_ρ and $\widetilde{W}_\rho$ of the primal and the dual problem, together with the convex conjugate of $\widetilde{W}_\rho$. Of course, this region can be shrunk by choosing ρ bigger.

J	N	Error L^1	Order L^1	Error L^2	Order L^2	Error L^∞	Order L^∞	CPU (s)
18	8	1.96E-01	-	1.86E-01	-	1.77E-01	-	0.30
46	16	1.44E-01	0.44	1.12E-01	0.74	1.05E-01	0.75	1.05
118	32	5.85E-02	1.30	4.54E-02	1.30	5.86E-02	0.84	3.91
305	64	1.52E-02	1.94	1.14E-02	2.00	1.52E-02	1.95	15.54
790	128	5.70E-03	1.42	4.11E-03	1.47	4.76E-03	1.67	61.95
2048	256	2.35E-03	1.28	1.68E-03	1.29	1.74E-03	1.45	467.54
5312	512	1.12E-03	1.07	8.14E-04	1.04	9.18E-04	0.92	2169.45

TABLE 2. Test 1: Local ($x \in [1,2]$) errors and convergence order comparing W_ρ with the exact solution v, for $M = 4$ (Gauß-Hermite quadrature points), $N = 4 \cdot 2^k$ (time steps), $J = \lceil N^{11/8} \rceil$ (space steps), $N_a = 2^k + 1$ (discrete controls), for $k = 1, 2, \ldots, 8$.

J	N	Gap L^1	Order L^1	Gap L^2	Order L^2	Gap L^∞	Order L^∞	CPU (s)
18	8	2.17E+01	-	7.17E+00	-	3.22E+00	-	0.56
46	16	1.24E+01	0.80	4.04E+00	0.83	1.65E+00	0.96	1.41
118	32	7.24E+00	0.78	2.31E+00	0.80	9.24E-01	0.88	4.70
305	64	3.92E+00	0.89	1.26E+00	0.88	5.06E-01	0.87	17.98
790	128	1.87E+00	1.07	6.03E-01	1.06	2.43E-01	1.06	110.56
2048	256	7.16E-01	1.38	2.37E-01	1.35	1.00E-01	1.28	656.69
5312	512	1.72E-01	2.05	5.53E-02	2.10	2.20E-02	2.19	2813.47
13778	1024	5.97E-02	1.53	1.94E-02	1.51	8.05E-03	1.45	17059.00

TABLE 3. Test 1: Global ($x \in [0, x_{\max}]$) duality gap $G^{h,\Delta x}$ from (4.12) and related convergence order, for $M = 4$ (Gauß-Hermite quadrature points), $N = 4 \cdot 2^k$ (time steps), $J = \lceil N^{11/8} \rceil$ (space steps), $N_a = 2^k + 1$ (discrete controls), for $k = 1, 2, \ldots, 8$.

From the results in Table 3 we deduce that (given the choice of Δx in relation to h)

$$|G^{h,\Delta x}(t,x)| \leq C\left(h + \Delta x^{8/11}\right),$$

which, combined with (4.14) and taking $M = 4$, gives the *a posteriori* bounds

$$-C\left(h^{3/8} + \Delta x^{3/11}\right) \leq v(t_n, x_i) - W_\rho(t_n, x_i) \leq C\left(h + \Delta x^{8/11} + h^{3/8} + \Delta x^{3/11}\right), \tag{5.2}$$

which in conclusion is a symmetric bound of order 3/8 in time and 3/11 in space.

We illustrate the different contributions to the error, together with the actual error, in Figure 4. The figure shows the order (at least) one for the empirical error and for the numerical duality gap, as one would have expected from the first order error of the scheme for sufficiently smooth solutions. We also plot the theoretical error bounds, which hold in the general non-smooth case, for the Euler-Maruyama scheme given in (A.1), of order 1/2, and for the Gauß-Hermite approximation from (A.2), of order 3/8. The big constants appearing in the theoretical *a priori* bounds, which are not sharp, put the magnitude of these theoretical errors far from that of the empirical one.

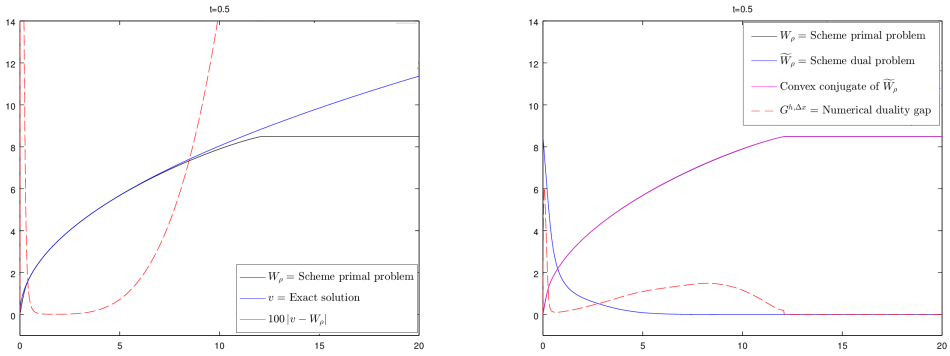

FIGURE 3. Test 1: Numerical solution W_ρ (in black) compared with the exact solution (blue, left) and the convex conjugate of $\widetilde{W}_\rho$ (magenta, right). The dashed red line represents the error (left) and the numerical duality gap (right).

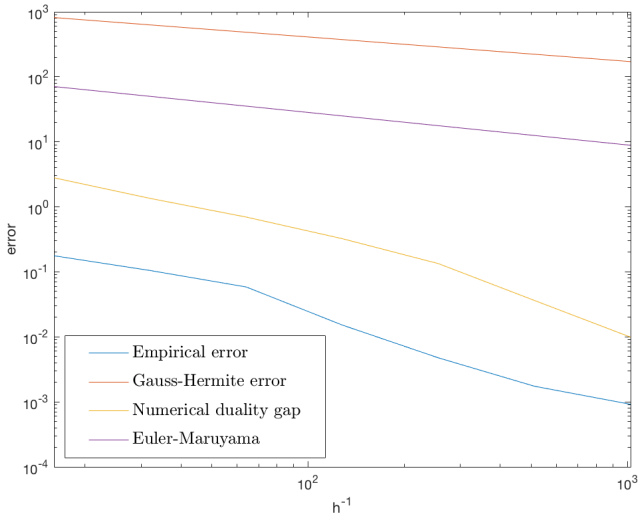

FIGURE 4. Test 1. Local ($x \in [1,2]$) empirical error $|v(0,x) - W_\rho(0,x)|$ as reported in Table 3, global numerical duality gap $G^{h,\Delta x}$ reported in Table 2, theoretical error estimate for the Euler-Maruyama and Gauß-Hermite approximation given by (A.1) and (A.2), respectively.

For this problem, the optimal control is constant over time, so there is no error coming from the piecewise control approximation and theoretical bounds as those provided by (A.1) and (A.2) can be used for both the upper and lower bound. The numerical duality gap in this case contains the sum of the numerical approximation errors for the primal and the dual problem as well as the error coming from the approximation in ρ and the computation of the numerical convex conjugate.

Test 2: Cuoco and Liu example. This example is taken from [9]. In this paper, the authors consider the nonlinear dynamics in (2.1) (i.e. $g \not\equiv 0$) and portfolio constraints (i.e. $A \subsetneq \mathbb{R}$). We still consider a power utility and $d = 1$. Let A be defined by

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

p	r	R	b	σ	T	$x_{\max}$	ι	λ_+	λ_-
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^∞	Order L^∞	CPU (s)
18	8	2.26E+01	-	7.44E+00	-	3.59E+00	-	0.79
46	16	1.09E+01	1.05	3.48E+00	1.10	1.47E+00	1.29	2.51
118	32	5.59E+00	0.96	1.74E+00	1.00	6.87E-01	1.10	9.83
305	64	2.82E+00	0.99	8.79E-01	0.99	3.47E-01	0.98	45.94
790	128	1.38E+00	1.03	4.35E-01	1.01	1.77E-01	0.97	552.49
2048	256	5.75E-01	1.26	1.83E-01	1.25	7.49E-02	1.24	6305.33
5312	512	1.56E-01	1.88	5.00E-02	1.87	2.08E-02	1.85	54006.70

TABLE 5. Test 2: Global ($x \in [0, x_{\max}]$) duality gap $G^{h,\Delta x}$ from (4.12 and related convergence order, for $M = 4$ (Gauß-Hermite quadrature points), $N = 4 \cdot 2^k$ (time steps), $J = \lceil N^{11/8} \rceil$ (space steps), $N_a = 2^k + 1$ (discrete controls), for $k = 1, 2, \ldots, 8$.

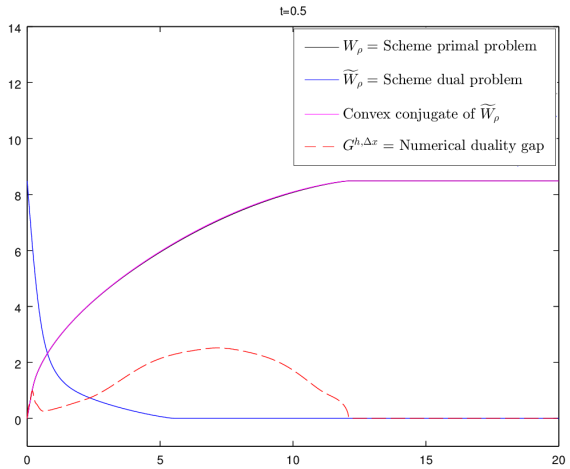

FIGURE 5. Test 2: Numerical solution W_ρ (in black) compared with the convex conjugate of $\widetilde{W}_\rho$ (in magenta). The dashed red line represents the numerical duality gap. In blue the numerical approximation of the dual problem.

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

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

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 Γ, we use the explicit expression given in [9, Section 5.2] for the optimal control. 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 W_ρ and $\widetilde{W}_\rho$ of the primal and the dual problem, together with the convex conjugate of $\widetilde{W}_\rho$ are shown in Figure 5.

The results in Table 5 give once again an estimate of the form

$$|G^{h,\Delta x}(t,x)| \leq C\left(h + \Delta x^{8/11}\right)$$

for the duality gap, leading to the *a posteriori* bounds (5.2).

6. Conclusion and perspectives

For a suitable class of convex optimal control problems, we obtained in this paper *a posteriori* error bounds 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, 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.

Appendix A. Explicit computation of the constants

In this section, we explicitly compute the constants C which appears in the lower bound of (4.13). Analogous estimates can be used to derive the constant $\widetilde{C}$ appearing in the upper bound. In what follows we denote for $t \in [0,T]$, $a \in A$, $x \in \mathbb{R}$:

$$\mu(t,x,a) := \left(r(t) + a^\top \cdot (b(t) - r(t)\mathbb{1}) + g(t,a)\right)x, \qquad \psi(t,x,a) := a^\top \sigma(t) x.$$

Let $C_\mu, C_\psi \geq 0$ such that for $t,s \in [0,T]$, $a \in A$, $x,y \in \mathbb{R}$:

$$|\mu(t,x,a) - \mu(s,y,a)| \leq C_\mu\left(|x-y| + (1+|x|)|t-s|^{1/2}\right)$$
$$|\psi(t,x,a) - \psi(s,y,a)| \leq C_\psi\left(|x-y| + (1+|x|)|t-s|^{1/2}\right)$$

and

$$|\mu(t,x,a)| \leq C_\mu(1+|x|) \qquad |\psi(t,x,a)| \leq C_\psi(1+|x|).$$

A.1. **Explicit bounds for the Euler-Maruyama approximation.** We consider the Euler-Maruyama approximation given by (3.4) for $\alpha \equiv (a_0, \ldots, a_{N-1}) \in \mathcal{A}_h$. This leads to the following expression for $\overline{X}_\cdot^{t_n,x,\alpha}$:

$$\overline{X}_{t_k}^{t_n,x,\alpha} = x + \sum_{i=n}^{k-1} \int_{t_i}^{t_{i+1}} \mu(t_i, \overline{X}_{t_i}^{t_n,x,\alpha}, a_i)\,\mathrm{d}s + \int_{t_i}^{t_{i+1}} \psi(t_i, \overline{X}_{t_i}^{t_n,x,\alpha}, a_i)\,\mathrm{d}W_s.$$

Moreover, by the very definition of $X_\cdot^{t_n,x,\alpha}$:

$$X_{t_k}^{t_n,x,\alpha} = x + \sum_{i=n}^{k-1} \int_{t_i}^{t_{i+1}} \mu(s, X_s^{t_n,x,\alpha}, a_i)\,\mathrm{d}s + \int_{t_i}^{t_{i+1}} \psi(s, X_s^{t_n,x,\alpha}, a_i)\,\mathrm{d}W_s.$$

Therefore, using the Cauchy-Schwartz inequality and Itô isometry together with classical estimates, one has

$$\begin{aligned}\mathbb{E}\left[|\overline{X}_{t_k}^{t_n,x,\alpha}-X_{t_k}^{t_n,x,\alpha}|^2\right] &\leq 2T\sum_{i=n}^{k-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}\left|\mu(t_i,\overline{X}_{t_i}^{t_n,x,\alpha},a_i)-\mu(s,X_s^{t_n,x,\alpha},a_i)\right|^2\mathrm{d}s\right]\\ &\quad+2\sum_{i=n}^{k-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}\left|\psi(t_i,\overline{X}_{t_i}^{t_n,x,\alpha},a_i)-\psi(s,X_s^{t_n,x,\alpha},a_i)\right|^2\mathrm{d}s\right]\\ &\leq 8K_1h\sum_{i=n}^{k-1}\left(\mathbb{E}\left[\left|\overline{X}_{t_i}^{t_n,x,\alpha}-X_{t_i}^{t_n,x,\alpha}\right|^2\right]+h+h\mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha}\right|^2\right]\right.\\ &\quad\left.+\mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha}-X_{t_i}^{t_n,x,\alpha}\right|^2\right]\right),\end{aligned}$$

where we denoted $K_1:=(C_\mu^2T+C_\psi^2)$. By classical estimates on the process $X^{t_n,x}_\cdot$ and denoting $K_2(\xi):=(C_\mu^2\xi+4C_\psi^2)$, one has

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha}\right|^2\right]&\leq 3\left(|x|^2+2K_2(T)\right)e^{6K_2(T)T},\\ \mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha}-X_{t_i}^{t_n,x,\alpha}\right|^2\right]&\leq 4K_2(h)h\left(1+3\left(|x|^2+4K_2(T)\right)e^{6K_2(T)T}\right).\end{aligned}$$

Putting these estimates together:

$$\mathbb{E}\left[|\overline{X}_{t_k}^{t_n,x,\alpha}-X_{t_k}^{t_n,x,\alpha}|^2\right]\leq 8K_1h\sum_{i=n}^{k-1}\mathbb{E}\left[\left|\overline{X}_{t_i}^{t_n,x,\alpha}-X_{t_i}^{t_n,x,\alpha}\right|^2\right]+8K_1Th(1+2K_2(h))(1+K_3(x))$$

with $K_3(x):=3\left(|x|^2+2K_2(T)T\right)e^{6K_2(T)T}$, so that, using Gronwall's lemma, one obtains

$$\begin{aligned}\mathbb{E}\left[|\overline{X}_{t_k}^{t_n,x,\alpha}-X_{t_k}^{t_n,x,\alpha}|^2\right]&\leq 8K_1h(1+2K_2(h))(1+K_3(x))\left(1+e^{\sum_{i=n}^{k-1}8K_1h}\left(\sum_{i=n}^{k-1}8K_1h\right)\right)\\ &\leq 8K_1h(1+2K_2(h))(1+K_3(x))\left(1+8K_1Te^{8K_1T}\right).\end{aligned}$$

Using the Lipschitz continuity of U_ρ, one has

$$\left|\sup_{\alpha\in\mathcal{A}_h}\mathbb{E}\left[|U_\rho(\overline{X}_T^{t_n,x,\alpha})\right]-\sup_{\alpha\in\mathcal{A}_h}\mathbb{E}\left[U_\rho(X_T^{t_n,x,\alpha})\right]\right|\leq L_\rho\sup_{\alpha\in\mathcal{A}_h}\mathbb{E}\left[|\overline{X}_T^{t_n,x,\alpha}-X_T^{t_n,x,\alpha}|\right].$$

In conclusion, the contribution to the error coming from the Euler-Maruyama approximation can be bounded by

$$L_\rho\Big(8K_1(1+2K_2(h))(1+K_3(x))(1+8K_1Te^{8K_1T})\Big)^{1/2}h^{1/2}.$$

For a linear (in the state), time independent dynamics as the one considered in Section 5, one simply has

$$|\mu(x,a)-\mu(y,a)|\leq C_\mu|x-y|,\qquad|\psi(x,a)-\psi(y,a)|\leq C_\psi|x-y|$$

and

$$|\mu(x,a)|\leq C_\mu|x|,\qquad|\psi(x,a)|\leq C_\psi|x|.$$

It is possible to verify that this leads to

$$\mathbb{E}\left[|\overline{X}_{t_k}^{t_n,x,\alpha} - X_{t_k}^{t_n,x,\alpha}|^2\right] \leq 4K_1 h \sum_{i=n}^{k-1}\left(\mathbb{E}\left[\left|\overline{X}_{t_i}^{t_n,x,\alpha} - X_{t_i}^{t_n,x,\alpha}\right|^2\right] + \mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha} - X_{t_i}^{t_n,x,\alpha}\right|^2\right]\right)$$
$$\leq 4K_1 h \sum_{i=n}^{k-1}\left(\mathbb{E}\left[\left|\overline{X}_{t_i}^{t_n,x,\alpha} - X_{t_i}^{t_n,x,\alpha}\right|^2\right] + 2K_2(h)h\mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha}\right|^2\right]\right)$$

with

$$\mathbb{E}\left[\sup_{s\in[t_i,t_{i+1}]}\left|X_s^{t_n,x,\alpha}\right|^2\right] \leq 3|x|^2 e^{3K_2(T)T}.$$

Neglecting the infinitesimal terms, one has

$$\mathbb{E}\left[|\overline{X}_{t_k}^{t_n,x,\alpha} - X_{t_k}^{t_n,x,\alpha}|^2\right] \leq 4K_1 h\left(\sum_{i=n}^{k-1}\mathbb{E}\left[\left|\overline{X}_{t_i}^{t_n,x,\alpha} - X_{t_i}^{t_n,x,\alpha}\right|^2\right] + 24TC_\psi^2|x|^2 e^{3K_2(T)T}\right)$$

which leads to the sharper error estimate

$$L_\rho\left(96K_1TC_\psi^2|x|^2e^{3K_2(T)T}\left(1+4K_1Te^{4K_1T}\right)\right)^{1/2}\, h^{1/2}.$$

We point out that in the estimates plotted in Section 5 we consider

$$L_\rho\left(24K_1TC_\psi^2|x|^2\left(1+4K_1Te^{4K_1T}\right)\right)^{1/2}\, h^{1/2} \tag{A.1}$$

since being interested in a local error we can approximate the second order moment of $X_\cdot$ by x^2.

A.2. **Explicit bounds for the Gauß-Hermite approximation.** We consider the case of a one-dimensional Brownian motion. Given a function $f \in C^{2M}(\mathbb{R})$, the analysis in [26, Proposition 3.2] shows that

$$\begin{aligned}
&\left|\ \mathbb{E}_{t_n,x}\left[f(\overline{X}^a_{t_{n+1}})\right] - \mathbb{E}_{t_n,x}\left[f(\widehat{X}^a_{t_{n+1}})\right]\ \right| \\
&\leq \left|\int_{-\infty}^{+\infty}\frac{f^{(2M)}(\hat{z})}{(2M)!}(\sqrt{2h}\psi(t_n,x,a)y)^{2M}\frac{e^{-y^2}}{\sqrt{\pi}}dy - \sum_{i=1}^M\lambda_i\frac{f^{(2M)}(\tilde{z})}{(2M)!}(\sqrt{h}\psi(t_n,x,a)\xi_i)^{2M}\right| \\
&\leq 2\|f^{2M}\|_\infty\frac{(2h)^M}{2M!}(\psi(t_n,x,a))^{2M}\int_{-\infty}^{\infty}y^{2M}\frac{e^{-y^2}}{\sqrt{\pi}}\mathrm{d}y \\
&\quad + \|f^{2M}\|_\infty\frac{h^M}{2M!}(\psi(t_n,x,a))^{2M}\left|2^M\int_{-\infty}^{\infty}y^{2M}\frac{e^{-y^2}}{\sqrt{\pi}}\mathrm{d}y - \sum_{i=1}^M\lambda_i{\xi_i}^{2M}\right| \\
&\leq \|f^{2M}\|_\infty\frac{h^M}{2M!}C_\psi^{2M}(1+|x|)^{2M}\left(2(2M-1)!! + \left|(2M-1)!! - \sum_{i=1}^M\lambda_i{\xi_i}^{2M}\right|\right),
\end{aligned}$$

where in the last inequality we have used that

$$2^M\int_{-\infty}^{\infty}y^{2M}\frac{e^{-y^2}}{\sqrt{\pi}}\mathrm{d}y = (2M-1)!!$$

The estimate above corresponds to the error associated with the Gauß-Hermite approximation at each time step, i.e. considering the error at time t_{n+1} starting from t_n. Our scheme being iterative in time, the overall contribution to the error will be

$$\left\|\frac{\partial f}{\partial x^{2M}}\right\|_\infty\frac{h^{M-1}}{2M!}2^{2M-1}C_\psi^{2M}\left((2M-1)!! + \left|(2M-1)!! - \sum_{i=1}^M\frac{\omega_i}{\sqrt{\pi}}{z_i}^{2M}\right|\right)\left(1+\sup_{\substack{\alpha\in\mathcal{A}_h\\ k=n\dots N}}\mathbb{E}_{t_n,x}\left[(\widehat{X}^\alpha_{t_k})^{2M}\right]\right),$$

where we also used the classical inequality $|a+b|^{2M} \leq 2^{2M-1}(|a|^{2M}+|b|^{2M})$. It remains to estimate $\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_k})^{2M}\right]$. By the recursive definition of $\widehat{X}$, one has for any $k=n,\dots,N$

$$\begin{aligned}
\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_{k+1}})^{2M}\right] &= \mathbb{E}_{t_n,x}\left[\left(\widehat{X}^{\alpha}_{t_k}+h\mu(t_k,\widehat{X}_k,a_k)+\sqrt{h}\psi(t_k,\widehat{X}_k,a_k)\zeta_k\right)^{2M}\right] \\
&= \mathbb{E}_{t_n,x}\left[\sum_{i=0}^{2M}\sum_{j=0}^{i}\binom{2M}{i}\binom{i}{j}h^{i-j}(\widehat{X}^{\alpha}_{t_k})^{j}(\mu(t_k,\widehat{X}_k,a_k))^{i-j}\left(\sqrt{h}\psi(t_k,\widehat{X}_k,a_k)\zeta_k\right)^{2M-i}\right] \\
&= \mathbb{E}_{t_n,x}\left[\sum_{i=0}^{M}\sum_{j=0}^{2i}\binom{2M}{2i}\binom{2i}{j}h^{2i-j}(\widehat{X}^{\alpha}_{t_k})^{j}(\mu(t_k,\widehat{X}_k,a_k))^{2i-j}\left(\sqrt{h}\psi(t_k,\widehat{X}_k,a_k)\zeta_k\right)^{2M-2i}\right],
\end{aligned}$$

where the last equality follows observing that $\mathbb{E}[(\dots)\zeta_k^{2j+1}]=0$ for $j=0,\dots,M-1$ for any quantity, represented by "$(\dots)$", independent of ζ_k. Therefore, thanks to the linear growth of μ and ψ (taking for simplicity $C_1 := \max(C_\mu, C_\psi)$):

$$\begin{aligned}
\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_{k+1}})^{2M}\right] &= \mathbb{E}_{t_n,x}\left[\sum_{i=0}^{M}\sum_{j=0}^{2i}\binom{2M}{2i}\binom{2i}{j}h^{M+i-j}C_1^{2M-j}(\widehat{X}^{\alpha}_{t_k})^{j}(1+|\widehat{X}^{\alpha}_{t_k}|)^{2i-j}\zeta_k^{2M-2i}\right] \\
&\leq \sum_{i=0}^{M}\sum_{j=0}^{2i}\binom{2M}{2i}\binom{2i}{j}h^{M+i-j}C_1^{2M-j}\mathbb{E}_{t_n,x}\left[|\widehat{X}^{\alpha}_{t_k}|^{j}(1+|\widehat{X}^{\alpha}_{t_k}|)^{2i-j}\right]\mathbb{E}\left[\zeta_k^{2M-2i}\right] \\
&= \mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_k})^{2M}\right]+\sum_{j=0}^{2M-1}\binom{2M}{j}h^{2M-j}C_1^{2M-j}\mathbb{E}_{t_n,x}\left[|\widehat{X}^{\alpha}_{t_k}|^{j}(1+|\widehat{X}^{\alpha}_{t_k}|)^{2M-j}\right] \\
&\quad+\sum_{i=0}^{M-1}\sum_{j=0}^{2i}\binom{2M}{2i}\binom{2i}{j}h^{M+i-j}C_1^{2M-j}\mathbb{E}_{t_n,x}\left[|\widehat{X}^{\alpha}_{t_k}|^{j}(1+|\widehat{X}^{\alpha}_{t_k}|)^{2i-j}\right]\mathbb{E}\left[\zeta_k^{2M-2i}\right].
\end{aligned}$$

For $0\leq i\leq M$ and $0\leq j\leq 2i$, one has

$$\mathbb{E}_{t_n,x}\left[|\widehat{X}^{\alpha}_{t_k}|^{j}(1+|\widehat{X}^{\alpha}_{t_k}|)^{2i-j}\right] \leq 2^{2i}\left(1+\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_k})^{2M}\right]\right).$$

This gives:

$$\begin{aligned}
\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_{k+1}})^{2M}\right] &\leq \mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_k})^{2M}\right]+\left(1+\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_k})^{2M}\right]\right)h\left\{\sum_{j=0}^{2M-1}\binom{2M}{j}h^{2M-j-1}C_1^{2M-j}2^{2M}\right. \\
&\quad\left.+\sum_{i=0}^{M-1}\sum_{j=0}^{2i}\binom{2M}{2i}\binom{2i}{j}h^{M+i-j-1}C_1^{2M-j}2^{2i}\mathbb{E}\left[\zeta_k^{2M-2i}\right]\right\}.
\end{aligned}$$

Neglecting the infinitesimal terms and denoting

$$K_4 := 2MC_1 2^{2M}+\frac{2M(2M-1)}{2}C_1^2 2^{2M-2},$$

we have

$$\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_{k+1}})^{2M}\right] \leq (1+K_4h)\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_k})^{2M}\right]+K_4h.$$

Iterating, this leads to

$$\mathbb{E}_{t_n,x}\left[(\widehat{X}^{\alpha}_{t_{k+1}})^{2M}\right] \leq (1+K_4h)^k x^{2M}+khK_4$$

for any $n \leq k \leq N-1$, with K_4 not depending on k and $\alpha \in \mathcal{A}_h$. Therefore, we can conclude that

$$\sup_{\substack{\alpha\in\mathcal{A}_h \\ k=n\dots N}} \mathbb{E}_{t_n,x}\left[(\widehat{X}^\alpha_{t_k})^{2M}\right] \leq x^{2M}e^{K_4 T} + K_4 T.$$

To avoid an exponential growth in M of the constants and motivated by the fact that in Section 5 we empirically computed a local error, we can strongly simplify our estimates by approximating

$$\sup_{\substack{\alpha\in\mathcal{A}_h \\ k=n\dots N}} \mathbb{E}_{t_n,x}\left[(\widehat{X}^\alpha_{t_k})^{2M}\right] \approx x^{2M}.$$

The presence of the $2M$-th derivative in the error bound requires to pass by a mollification of the original value function. For a given regularization parameter ε and mollified value function v_ε it is possible to show that an estimate of the form

$$\left\| \frac{\partial^{2M} v_\varepsilon}{\partial x^{2M}} \right\|_\infty \leq L_\rho K_5 \varepsilon^{1-2M}$$

holds with $K_5 := (3+9K_1 T e^{3K_1 T})^{1/2}$. The balancing between the Gauß-Hermite and regularization error (the last one giving an extra error term of order ε) leads to the choice of optimal order $\varepsilon = h^{(M-1)/2M}$. Therefore, we get

$$L_\rho K_5 h^{(M-1)/2M} \frac{2^{2M-1}}{2M!} C_\psi^{2M} \left((2M-1)!! + \left| (2M-1)!! - \sum_{i=1}^{M} \frac{\omega_i}{\sqrt{\pi}} {z_i}^{2M} \right| \right) (1 + x^{2M}). \tag{A.2}$$

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 J. Cvitanič. Optimal consumption choices for a 'large' investor. Journal of Economic Dynamics and Control, 22(3):401–436, 1998.

[9] D. Cuoco and H. Liu. A martingale characterization of consumption choices and hedging costs with margin requirements. Math. Financ., 10:355–385, 2000.

[10] K. Debrabant and E.R. Jakobsen. Semi-Lagrangian schemes for linear and fully non-linear diffusion equations. Math. Comp., 82(283):1433–1462, 2012.

[11] K. Debrabant and E.R. Jakobsen. Semi-Lagrangian schemes for parabolic equations. In T. Gerstner and P. Kloeden, editors, Recent Developments in Computational Finance: Foundations, Algorithms and Applications, pages 279–297. World Scientific, 2013.

[12] A. Dembo and O. Zeitouni. Large Deviations Techniques and Applications, volume 38 of Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, 2010.

[13] E.A. DeVuyst and P.V. Preckel. Gaussian cubature: A practitioner's guide. Math. Comput. Model., 45(7-8):787–794, 2007.

[14] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. Mathematical Finance, 7(1):1–71, 1997.

[15] M. Falcone and R. Ferretti. Semi-Lagrangian Approximation Schemes for Linear and Hamilton-Jacobi Equations, volume 133. SIAM, Philadelphia, 2014.

[16] F.B. Hildebrand. Introduction to Numerical Analysis. New York: McGraw-Hill, 1956.

[17] E.R. Jakobsen, A. Picarelli, and C. Reisinger. Improved order 1/4 convergence for piecewise constant policy approximation of stochastic control problems. Preprint, arXiv:1901.01193, 2018.

[18] P.E. Kloeden and E. Platen. Numerical Solution of Stochastic Differential Equations. Berlin, New York, Springer-Verlag, 1992.
[19] 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.
[20] N.V. Krylov. On the rate of convergence of finite-difference approximations for Bellman's equations. St. Petersburg Math. J., 9:639–650, 1997.
[21] 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.
[22] 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.
[23] J.L. Menaldi. Some estimates for finite difference approximations. SIAM J. Control Optim., 27:579–607, 1989.
[24] R.C. Merton. Optimal consumption and portfolio rules in continuous time. J. Economic Theory, 3:373–413, 1971.
[25] H. Pham. Continuous-time Stochastic Control and Optimization with Financial Applications, volume 61. Series Stochastic Modeling and Applied Probability, Springer, 2009.
[26] A. Picarelli and C. Reisinger. Probabilistic error analysis for some approximation schemes to optimal control problems. Preprint, arXiv:1810.04691, 2018.
[27] L.C.G. Rogers. Duality in Constrained Optimal Investment Problems: A Synthesis. Number 1814 in Paris–Priceton Lectures on Mathematical Finance. Springer, 2002.
[28] V. Tchakaloff. Formules de cubatures mécaniques à coefficients non négatifs. Bull. Sci. Math, 81(2):123–134, 1957.
[29] M. Tchernychova. Carathéodory cubature measures. PhD thesis, University of Oxford, 2015.
[30] J. Yong and X.Y. Zhou. Stochastic Controls: Hamiltonian Systems and HJB Equations, volume 43 of Applications of Mathematics. Springer-Verlag, New York, 1999.

Department of Economical Sciences, University of Verona, Via Cantarane 24, 37129, Verona, Italy
E-mail address: athena.picarelli@univr.it

Mathematical Institute, University of Oxford, Andrew Wiles Building, OX2 6GG, Oxford, UK
E-mail address: christoph.reisinger@maths.ox.ac.uk