Electron. Commun. Probab. **24** (2019), no. 59, 1–10.
https://doi.org/10.1214/19-ECP256
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Improved order 1/4 convergence for piecewise constant policy approximation of stochastic control problems

Espen R. Jakobsen* Athena Picarelli† Christoph Reisinger‡

Abstract

In *N. V. Krylov, Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies, Electron. J. Probab., 4(2), 1999*, it is proved under standard assumptions that the value functions of controlled diffusion processes can be approximated with order 1/6 error by those with controls which are constant on uniform time intervals. In this note we refine the proof and show that the provable rate can be improved to 1/4, which is optimal in our setting. Moreover, we demonstrate the improvements this implies for error estimates derived by similar techniques for approximation schemes, bringing these in line with the best available results from the PDE literature.

Keywords: optimal control; Bellman's equation; error estimates; piecewise constant policies.
AMS MSC 2010: 93E20; 35K55; 65M15; 49L20.
Submitted to ECP on January 4, 2019, final version accepted on July 12, 2019.

1 Introduction

In this paper we derive improved error estimates for approximations of value functions of stochastic optimal control problems. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space, $(W_t)_{t\geq 0}$ a p-dimensional $\{\mathcal{F}_t\}$-Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{A}$ the set of progressively measurable processes with values in a set $A \subseteq \mathbb{R}^m$. For any $\alpha \in \mathcal{A}$, $x \in \mathbb{R}^d$, $t \in [0, T]$ (with $T > 0$), let $X_\cdot = X_\cdot^{\alpha,t,x}$ be the (controlled) Itô diffusion which satisfies

$$X_s = x + \int_0^s b_{\alpha_r}(t+r, X_r)\,\mathrm{d}r + \int_0^s \sigma_{\alpha_r}(t+r, X_r)\,\mathrm{d}W_r \qquad \text{for} \qquad s \geq t. \tag{1.1}$$

Here we use the notation $\varphi_a(\cdot,\cdot) = \varphi(\cdot,\cdot,a)$ for any $a \in A$ and function φ. For a given terminal cost function g and running cost f, the optimal control problem consists of maximizing over $\alpha \in \mathcal{A}$ the expected total cost

$$J^\alpha(t,x) := \mathbb{E}^\alpha_{t,x}\Big[\int_0^{T-t} f_{\alpha_r}(t+r, X_r)\,\mathrm{d}r + g(X_{T-t})\Big]. \tag{1.2}$$

*Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway.
E-mail: erj@math.ntnu.no
†Department of Economics, University of Verona, via Cantarane 24, 37129 Verona, Italy.
E-mail: athena.picarelli@univr.it
‡Mathematical Institute, University of Oxford, Andrew Wiles Building, OX2 6GG, Oxford, UK.
E-mail: christoph.reisinger@maths.ox.ac.uk

The indices on the expectation $\mathbb{E}$ indicate that the law of the process depends on the starting point and control. Finally, the value function of the optimal control problem is defined by

$$v(t,x) := \sup_{\alpha\in\mathcal{A}} J^{\alpha}(t,x). \tag{1.3}$$

We consider the following set of assumptions:

(H1) A is a compact set;

(H2) $b : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}^{d\times p}$ are continuous functions. For $\varphi \in \{b,\sigma\}$, there exists $C_0 \geq 0$ such that for every $t,s \in [0,T], x,y \in \mathbb{R}^d, a \in A$:

$$|\varphi(t,x,a) - \varphi(s,y,a)| \leq C_0\left(|x-y| + |t-s|^{1/2}\right) \quad \text{and} \quad |\varphi(t,x,a)| \leq C_0;$$

(H3) $g : \mathbb{R}^d \to \mathbb{R}$ and $f : [0,T]\times\mathbb{R}^d\times A \to \mathbb{R}$ are continuous functions. There exists $C_1 \geq 0$ such that for every $t,s \in [0,T], x,y \in \mathbb{R}^d, a \in A$:

$$\begin{aligned}&|g(x) - g(y)| \leq C_1|x-y|,\\ &|f(t,x,a) - f(s,y,a)| \leq C_1\left(|x-y| + |t-s|^{1/2}\right) \quad \text{and} \quad |f(t,x,a)| \leq C_1.\end{aligned}$$

Observe that under assumptions (H1), (H2), and for any $\alpha \in \mathcal{A}$, there exists a unique strong solution of equation (1.1). For simplicity, we assume data and coefficients to be Lipschitz continuous in space and $1/2$-Hölder continuous in time, and have included no discount factor, but it is not difficult to extend our results to include discounting and a lower Hölder regularity for f and g. Moreover, as in [8], we assume the boundedness of data. It should be possible to extend the results to the case of linear growth in x of b, σ, f and g, obtaining local estimates growing polynomially in the space variable.

We aim to estimate the error introduced by approximating the set of measurable controls $\mathcal{A}$ by piecewise constant controls. Let $h > 0$ be the discretization parameter and $\mathcal{A}_h$ the subset of $\mathcal{A}$ of processes which are constant in the intervals $[nh,(n+1)h)$ for $n \in \mathbb{N}$.[1] The value function associated with this restricted set of controls is defined by

$$v_h(t,x) := \sup_{\alpha\in\mathcal{A}_h} J^{\alpha}(t,x). \tag{1.4}$$

Note that the definition of v_h in (1.4) under the "shifted" dynamics in (1.2) and (1.1) implies that the control discretisation is always centered at t. This will be important for establishing a dynamic programming principle. This is not, though, how one would compute v_h in practice, as discussed in the penultimate paragraph of this section.

From a probabilistic perspective, it is clear that 0 is a lower bound for $v - v_h$ since $\mathcal{A}_h \subseteq \mathcal{A}$. Under our assumptions, an upper bound on $v - v_h$ of order $h^{\frac{1}{6}}$ is given in [8].

An indication that the order 1/6 from [8] might be improved is the fact that under the same regularity assumptions as above it is shown in [5] that a fully discrete semi-Lagrangian scheme applied to the corresponding HJB equation has order 1/4 in the timestep for an Euler approximation. This scheme does not distinguish between constant or other controls over individual timesteps. It would therefore be somewhat surprising if the scheme which employs further approximations was closer to the original problem than the one which only holds the policies constant over timesteps.

A slightly different angle to the problem is provided in [3], where the authors construct from (1.4) a subsolution to the HJB equation corresponding to (1.3) by a second

[1] Note that in [8] the length of intervals is h^2, however, in absence of further discretisations, we use h for simplicity.

order local expansion in t. This results in an order 1 error bound in the case of smooth solutions, in contrast to 1/2 which would be obtained in the smooth case by the method in [8] (see also Section 2.3 below). However, in the general non-regular case, the order in [3] is limited by a switching system approximation of order $\varepsilon^{1/3}$ (for a switching cost chosen of order ε), which, combined with an error term of the regularised system of order h/ε^3 (for regularisation parameter ε), results in an order $1/10$ error by optimisation of ε.

In this paper, we combine the advantages of both methods to obtain order 1/4. The reason we can improve the error estimates of Krylov is that we use a higher order expansion when we derive the truncation error. Our discussion (see Subsection 2.3) also shows that no further improvement can be obtained in this way: our new proof uses the maximal possible order of the truncation error.

Piecewise constant policy time stepping has been used in a numerical method for solving Hamilton-Jacobi-Bellman equations in [13], where the computational advantage comes from the fact that over the time intervals in which the policy is constant, only linear PDEs have to be solved. This has been extended to mixed optimal stopping and control problems with nonlinear expectations and jumps in [6]. A further benefit lies in the inherent parallelism so that the linear problems with different controls can be solved on parallel processors. A proof of convergence is given in these works using pure viscosity solution arguments, but no rate of convergence is provided. Early results on this type of approximation can be found in [10] and an extension with "predicted" controls is proposed in [7].

In the remainder of this article, we give in Section 2 a proof of the order 1/4 convergence of the piecewise constant policy approximation, and deduce the linear convergence in the case of sufficiently regular solutions and data. We then outline in Section 3 the improved orders which can be derived for approximation schemes by similar techniques.

2 Main result

We begin by stating the main result. Throughout the entire section we work under assumptions (H1)–(H3).

Theorem 2.1. *For any $s \in [0,T]$, $x \in \mathbb{R}^d$, and $h > 0$, we have*

$$0 \;\le\; v(s,x) - v_h(s,x) \;\le\; Ch^{1/4}, \tag{2.1}$$

where the constant C only depends on the constants in Assumptions (H2) and (H3).

A major difficulty in the proof of Theorem 2.1 is the fact that typically v and v_h are not smooth. Even in the non-degenerate case where v is $C^{2+\delta}$, v_h is still not smooth in general. A simple example is the Black-Scholes-Barenblatt equation resulting from an uncertain volatility model (see [11]). Here, the control is of bang-bang type and the optimal control problem for piecewise constant policies reduces to taking the maximum of two smooth functions at the end of each time interval, so that for t on the time mesh, $v_h(t,\cdot)$ will only be Lipschitz (in the spatial argument).

Since the proof of Theorem 2.1 relies on repeated use of the Itô formula, we need to work with smooth functions, both for the coefficients and value functions v and v_h. This means that we need to introduce several regularization arguments and use Krylov's method of shaking the coefficients.

2.1 Background results and regularisation

In this section, we introduce Krylov's regularization and give related preliminary results. Some of the proofs are given in [8] and not repeated here; see also [1, 2] for

analogous results proved with PDE arguments. In order to apply Itô's formula twice, σ, b, f, g, v, and v_h must be regularized. Let $\varepsilon > 0$ and the mollifier ρ_ε be defined as

$$\rho_\varepsilon(t,x) := \frac{1}{\varepsilon^{d+2}}\rho\left(\frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right), \tag{2.2}$$

where

$$\rho \in C^\infty(\mathbb{R}^{d+1}), \quad \rho \geq 0, \quad \operatorname{supp}\rho = (0,1) \times \{|x| < 1\}, \quad \int_{\operatorname{supp}\rho} \rho(e)\,\mathrm{d}e = 1.$$

For any function $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, we define $\varphi^{(\varepsilon)} \in C^\infty([0,T] \times \mathbb{R}^d)$ to be the mollification of a suitable extension of φ to $[-\varepsilon^2, T]$

$$\varphi^{(\varepsilon)}(t,x) := (\varphi * \rho_\varepsilon)(t,x) = \int_{0 \leq s \leq \varepsilon^2} \int_{|y| \leq \varepsilon} \varphi(t-s, x-y)\rho_\varepsilon(s,y)\,\mathrm{d}s\,\mathrm{d}y.$$

We can always take an extension which preserves the Hölder continuity in time and Lipschitz continuity in space of φ. Then standard estimates for mollifiers imply that

$$\|\varphi - \varphi^{(\varepsilon)}\|_\infty \leq C\varepsilon \qquad \text{and} \qquad \left\|\partial_t^m D_x^k \varphi^{(\varepsilon)}\right\|_\infty \leq C\varepsilon^{1-2m-k} \qquad \text{for} \qquad k+m \geq 1. \tag{2.3}$$

Let $\tilde{X}_\cdot$ be the solution of (1.1) with coefficients replaced by $b^{(\varepsilon)}$ and $\sigma^{(\varepsilon)}$. Then we denote by $\tilde{v}$ and $\tilde{J}^\alpha$ the solution and cost function of the optimal control problem (1.1)–(1.3) where $X_\cdot$ is replaced by $\tilde{X}_\cdot$ and f, g by $f^{(\varepsilon)}, g^{(\varepsilon)}$.

Proposition 2.2. *There exists $C \geq 0$ such that for any $t \in [0,T], x \in \mathbb{R}^d$*

$$|v(t,x) - \tilde{v}(t,x)| \leq C\varepsilon.$$

Proof. The result follows from the definitions of v and $\tilde{v}$ since by standard continuous dependence results for SDEs and Lipschitz and Hölder continuity of f, g, b, σ,

$$\mathbb{E}^\alpha_{t,x}\Big[\sup_{s \in [0,T-s]} |X_s - \tilde{X}_s|^2\Big] \leq C(\|b - b^{(\varepsilon)}\|^2_\infty + \|\sigma - \sigma^{(\varepsilon)}\|^2_\infty) \leq C\varepsilon^2$$

for some constant C independent of the control α. □

To avoid heavy notation, we will use (f, g, b, σ) instead of $(f^{(\varepsilon)}, g^{(\varepsilon)}, b^{(\varepsilon)}, \sigma^{(\varepsilon)})$ in the rest of the paper, keeping in mind estimates (2.3) for their derivatives. We now proceed with the regularisation of the value function v_h. Let $\mathcal{E}_h$ be the set of progressively measurable processes $e \equiv (e_1, e_2)$ with values in $(-\varepsilon^2, 0) \times B_\varepsilon(0)$ (where $B_\varepsilon(0)$ denotes the ball of radius ε in $\mathbb{R}^d$) which are constant in each time interval $[nh, (n+1)h)$. Letting $S = T + \varepsilon^2$, we define for any $s \in [0,S], x \in \mathbb{R}^d$ the following "perturbed" value function

$$u_h(s,x) := \sup_{\alpha \in \mathcal{A}_h, e \in \mathcal{E}_h} \mathbb{E}^{(\alpha,e)}_{s,x}\Big[\int_0^{S-s} f_{\alpha_r}(s+r, \hat{X}_r)\,\mathrm{d}r + g(\hat{X}_{S-s})\Big], \tag{2.4}$$

where $\hat{X}_\cdot = \hat{X}^{(\alpha,e),s,x}_\cdot$ is the solution of the following SDE with (mollified and) "shaken coefficients":

$$\hat{X}_\cdot = x + \int_0^\cdot b_{\alpha_r}(s+r+e_{1,r}, \hat{X}_r + e_{2,r})\,\mathrm{d}r + \int_0^\cdot \sigma_{\alpha_r}(s+r+e_{1,r}, \hat{X}_r + e_{2,r})\,\mathrm{d}W_r. \tag{2.5}$$

Proposition 2.3. *There exists a constant $C \geq 0$ such that*

$$|v_h(t,x) - u_h(t,x)| \leq C\varepsilon$$

for any $t \in [0,T], x \in \mathbb{R}^d$*, and*

$$|u_h(t,x) - u_h(s,y)| \le C(|x-y| + |t-s|^{1/2})$$

for any $t, s \in [0,S]$ *and* $x, y \in \mathbb{R}^d$*. Moreover, for any* $s \in [0, S-h]$*,* u_h *satisfies the following dynamic programming principle (DPP):*

$$u_h(s,x) = \sup_{\substack{a\in A \\ 0\le\eta\le\varepsilon^2, |\xi|\le\varepsilon}} \mathbb{E}_{s,x}^{(a,(\eta,\xi))}\Big[\int_0^h f_a(s+r,\hat{X}_r)\,\mathrm{d}r + u_h(s+h,\hat{X}_h)\Big]. \tag{2.6}$$

Proof. These are standard results. The first two inequalities can be found e.g. in [8, Corollary 3.2], while (2.6) is a consequence of [8, Lemma 3.3]. □

Following the notation introduced above we consider the regularised (mollified) function $u_h^{(\varepsilon)}$.

Proposition 2.4. *The function* $u_h^{(\varepsilon)}$ *belongs to* $C^\infty([0,T]\times\mathbb{R}^d)$*. There exists a constant* $C \ge 0$ *such that*

$$\big|u_h(t,x) - u_h^{(\varepsilon)}(t,x)\big| \le C\varepsilon \tag{2.7}$$

for $t \in [0,T], x \in \mathbb{R}^d$*, and*

$$\Big\|\partial_t^m D_x^k u_h^{(\varepsilon)}\Big\|_\infty \le C\varepsilon^{1-2m-k} \qquad \text{for} \qquad k+m \ge 1. \tag{2.8}$$

Moreover, $u_h^{(\varepsilon)}$ *satisfies the following super-dynamic programming principle*

$$u_h^{(\varepsilon)}(t,x) \ge E_{t,x}^a\Big[\int_0^h f_a(t+r,\tilde{X}_r)\,\mathrm{d}r + u_h^{(\varepsilon)}(t+h,\tilde{X}_h)\Big] \tag{2.9}$$

for any $a \in A$*,* $0 \le \eta \le \varepsilon^2, |\xi| \le \varepsilon$*,* $t \in [0, T-h], x \in \mathbb{R}^d$*.*

Proof. The first part follows from Proposition 2.3 and (2.3), while (2.9) follows by the definitions of $u_h^{(\varepsilon)}$, $\hat{X}_t$, $\tilde{X}_t$, and the inequality $\int \sup(\cdots) \ge \sup \int(\cdots)$. See [8, bottom of page 9] for more details. Here $\alpha_t \equiv a$ constant over $t \in [0,h]$ by a slight abuse of notation. □

2.2 Proof of Theorem 2.1

1) *Upper bound on* $L_a u_h^{(\varepsilon)} + f_a$*.* By two applications of the Itô (or Dynkin) formula,

$$\begin{aligned}
&\mathbb{E}_{s,x}^a[u_h^{(\varepsilon)}(s+h,\tilde{X}_h)] \\
&\quad = u_h^{(\varepsilon)}(s,x) + \mathbb{E}_{s,x}^a\Big[\int_0^h (L_a u_h^{(\varepsilon)})(s+t,\tilde{X}_t)\,\mathrm{d}t\Big] \\
&\quad = u_h^{(\varepsilon)}(s,x) + h(L_a u_h^{(\varepsilon)})(s,x) + \mathbb{E}_{s,x}^a\Big[\int_0^h\int_0^t L_a(L_a u_h^{(\varepsilon)})(s+r,\tilde{X}_r)\,\mathrm{d}r\,\mathrm{d}t\Big]
\end{aligned}$$

for $s \le T-h$, $x \in \mathbb{R}^d$, $a \in A$, where the generator L_a of the diffusion process is defined as

$$L_a := \partial_t + b_a^T D_x + \frac{1}{2} tr[\sigma_a \sigma_a^T D_x^2].$$

Inserting this equality into the dynamic programming inequality (2.9) in Proposition 2.4, applying Itô once to the f_a-term, and dividing by h, we find that

$$(L_a u_h^{(\varepsilon)})(s,x) + f_a(s,x) \le \frac{1}{h}\sup_{a\in A}\Big(\|L_a f_a\|_\infty + \|L_a L_a u_h^{(\varepsilon)}\|_\infty\Big)\int_0^h\int_0^t \mathrm{d}r\,\mathrm{d}t. \tag{2.10}$$

Since the leading term $L_a L_a u_h^{(\varepsilon)}$ is a sum of terms of the form $\phi_1(\partial_t^m \phi_2)(D_x^k \phi_3)$ with $\phi_i \in \{\mu, \sigma\sigma^T, u_h^{(\varepsilon)}\}$ and $2m + k \le 4$, by (2.3) and (2.8),

$$(L_a u_h^{(\varepsilon)})(s,x) + f_a(s,x) \le C\varepsilon^{-3} h. \tag{2.11}$$

2) *Upper bound on $\tilde{v} - v_h$ for $s \in [0, T-h)$.* Let $\alpha \in \mathcal{A}$, $s \in [0, T-h]$, and $x \in \mathbb{R}^d$. By Itô's formula and part 1),

$$\begin{aligned}\mathbb{E}_{s,x}^{\alpha}[u_h^{(\varepsilon)}(T-h, \tilde{X}_{T-h-s})] &= u_h^{(\varepsilon)}(s,x) + \mathbb{E}_{s,x}^{\alpha}\Big[\int_0^{T-h-s} (L_{\alpha_t} u_h^{(\varepsilon)})(s+t, \tilde{X}_t)\,\mathrm{d}t\Big] \\ &\le u_h^{(\varepsilon)}(s,x) - \mathbb{E}_{s,x}^{\alpha}\Big[\int_0^{T-s} f_{\alpha_t}(s+t, \tilde{X}_t)\,\mathrm{d}t\Big] + TC\varepsilon^{-3}h.\end{aligned}$$

From (2.7) in Proposition 2.4 and the first part of Proposition 2.3, it then follows that

$$\begin{aligned}\mathbb{E}_{s,x}^{\alpha}[u_h(T-h, \tilde{X}_{T-h-s})] &\le u_h(s,x) - \mathbb{E}_{s,x}^{\alpha}\Big[\int_0^{T-s} f_{\alpha_t}(s+t, \tilde{X}_t)\,\mathrm{d}t\Big] + C(\varepsilon + \varepsilon^{-3}h) \\ &\le v_h(s,x) - \mathbb{E}_{s,x}^{\alpha}\Big[\int_0^{T-s} f_{\alpha_t}(s+t, \tilde{X}_t)\,\mathrm{d}t\Big] + C(\varepsilon + \varepsilon^{-3}h),\end{aligned}$$

for a generic constant C. Since by definition (2.4) and the regularity of u_h (Proposition 2.3),

$$\begin{aligned}\mathbb{E}_{s,x}^{\alpha}[(u_h(T-h, \tilde{X}_{T-h-s})] &= \mathbb{E}_{s,x}^{\alpha}[u_h(T-h, \tilde{X}_{T-h-s}) - u_h(S, \tilde{X}_{T-s}) + g(\tilde{X}_{T-s}))] \\ &\ge \mathbb{E}_{s,x}^{\alpha}[g(\tilde{X}_{T-s}))] - C(h^{1/2} + \varepsilon),\end{aligned}$$

we conclude that

$$\tilde{J}^{\alpha}(s,x) = \mathbb{E}_{s,x}^{\alpha}\Big[\int_0^{T-s} f_{\alpha_t}(s+t, \tilde{X}_t)\,\mathrm{d}t + g(\tilde{X}_{T-s})\Big] \le v_h(s,x) + C(\varepsilon + h^{1/2} + \varepsilon^{-3}h).$$

Since $\alpha \in \mathcal{A}$ was arbitrary, by the definition of $\tilde{v}$ (see just before Proposition 2.2),

$$\tilde{v}(s,x) - v_h(s,x) \le C(\varepsilon + h^{1/2} + \varepsilon^{-3}h).$$

3) *Upper bound on $\tilde{v} - v_h$ for $s \in [T-h, T]$.* By the definition of $\tilde{J}^{\alpha}$ (see just before Proposition 2.2), Itô's formula, the regularity of f and g, and using (2.3), there is a constant $C > 0$ such that for every $\alpha \in \mathcal{A}$ and $s \in [T-h, T]$,

$$|\tilde{J}^{\alpha}(s,x) - g(x)| = \Big|\mathbb{E}_{s,x}^{\alpha}\Big[\int_0^{T-s} \Big(f_{\alpha_r}(s+r, \tilde{X}_r) + L_{\alpha_r} g(\tilde{X}_r)\Big)\,\mathrm{d}r\Big]\Big| \le C(1+\varepsilon^{-1})h.$$

Then it follows from the definitions of $\tilde{v}$ and v_h that

$$|\tilde{v}(s,x) - g(x)| + |v_h(s,x) - g(x)| \le C\varepsilon^{-1}h,$$

and hence also $|\tilde{v}(s,x) - v_h(s,x)| \le 2C\varepsilon^{-1}h$ for $s \in [T-h, T]$.

4) *Conclusion:* Using Proposition 2.2 and parts 2) and 3), we have that

$$v(s,x) - v_h(s,x) \le \tilde{v}(s,x) - v_h(s,x) + C\varepsilon \le C(\varepsilon + h^{1/2} + \varepsilon^{-3}h)$$

for $s \in [s, T]$ and $x \in \mathbb{R}^d$. Taking $\varepsilon = h^{1/4}$ then concludes the proof of the right-hand inequality in (2.1). The left-hand inequality is immediate since $\mathcal{A}_h \subseteq \mathcal{A}$.

2.3 The maximal rate and comparison with [8]

If the data and value functions are smooth enough, we can adapt the proof of Theorem 2.1 to obtain the maximal rate of the approximation, which is 1. More specifically, if we assume v_h and f sufficiently smooth, we have in (2.10) $\sup_{a\in A}(\|L_a(L_a u_h^{(\varepsilon)})\|_\infty + \|L_a f\|_\infty) \leq C < \infty$ with C independent of ε. Therefore, instead of (2.11), the conclusion of step 1) in the previous proof gives

$$(L_a u_h^{(\varepsilon)})(s,x) + f_a(s,x) \leq Ch,$$

for some constant C independent of $a \in A$ and ε. Moreover, if we assume that b, σ and f are Lipschitz in t uniformly in x and a, and g belongs to $C_b^2(\mathbb{R}^d)$, then by standard results u_h will be Lipschitz in t. Hence, we find in step 2) that

$$\tilde{v}(s,x) - v_h(s,x) \leq C(\varepsilon + h).$$

Sending ε to zero then gives that $\tilde{v}$ converges to v, and we have the following result:

Proposition 2.5. *Additionally to assumptions (H1)-(H3), let b, σ and f be Lipschitz continuous in t uniformly with respect to x and a, and $g \in C_b^2(\mathbb{R}^n)$. If $\sup_{a\in A}(\|L_a(L_a v_h)\|_\infty + \|L_a f\|_\infty) < \infty$, then there exists $C > 0$ such that for any $s \in [0,T]$, $x \in \mathbb{R}^d$, and $h > 0$, we have*

$$0 \;\leq\; v(s,x) - v_h(s,x) \;\leq\; Ch. \tag{2.12}$$

This is the maximal rate that this approximation can reach. The reason is that the order obtained by applying Itô twice in step 1) of the proof cannot be improved. This can easily be checked by repeatedly applying Itô to obtain higher order error expansions and then noting that all such expansions contain terms of order h.

Step 1) of the proof also explains why Krylov in [8] got a less sharp result than ours. After one application of Itô, he used the moment bound $\mathbb{E}[|x - X_r|] \leq \sqrt{\mathbb{E}[|x - X_r|^2]} \leq C\sqrt{r}$ to get

$$\begin{aligned}\Big|\frac{1}{h}\mathbb{E}^a_{s,x}\Big[\int_0^h (L_a u_h^{(\varepsilon)})(s+t,\tilde{X}_t)\,\mathrm{d}t\Big] - (L_a u_h^{(\varepsilon)})(s,x)\Big| \\ \leq C\|D_x(L_a u_h^{(\varepsilon)})\|_\infty h^{1/2} + \|\partial_t(L_a u_h^{(\varepsilon)})\|_\infty h.\end{aligned}$$

This estimate requires only three derivatives in space of $u_h^{(\varepsilon)}$ but gives the lower rate 1/2. The conclusion of step 1) of the proof then becomes

$$L_a u_h^{(\varepsilon)}(s,x) + f_a(s,x) \leq C\left(\varepsilon^{-2}h^{1/2} + \varepsilon^{-3}h\right).$$

Completing the proof as in Section 2.2 then gives

$$\tilde{v}(s,x) - v_h(s,x) \leq C(\varepsilon + \varepsilon^{-2}h^{1/2} + \varepsilon^{-3}h),$$

and optimizing with respect to ε shows that $v(s,x) - v_h(s,x) \leq Ch^{1/6}$. Note that there is no need for regularization of the coefficients and data since Itô is applied only once. In the case of smooth enough solutions, this approach cannot give a higher rate than $1/2$.

3 Consequences on finite difference approximations

In this section, we outline the impact of the improved error bound for the control approximation on the achievable convergence order for numerical schemes, either by directly substituting the improved order (Section 3.1) or by applying adaptations of the steps here using higher order estimates (Section 3.2).

3.1 Improvement to Theorem 1.11 in [9]

Using the new bound for the control approximation from Section 2, one easily obtains a sharpening of the order from $1/39$ in [9, Theorem 1.11] and $1/21$ in [8, Theorem 5.4] to $1/15$, which holds for local, monotone schemes of consistency order $1/2$. Indeed, using Theorem 2.1 instead of [8, Theorem 2.3], the bound in the second inequality in the proof of [8, Theorem 5.4] (on top of page 14 in [8]) becomes

$$v \leq v_{\delta,1/n} + C(n\delta^{1/3} + n^{-1/4}),$$

where $\delta > 0$ is the time discretization step used in [8] for the approximation scheme for the value function, n the number of time intervals over which the policy is constant and $v_{\delta,1/n}$ is the obtained approximation of v.[2] Optimizing with respect to δ gives $n \sim \delta^{-4/15}$ and an estimate of order $1/15$ in δ.

Assuming order 1 consistency of the scheme used instead of order 1/2 as in [9, Theorem 1.11] and [8, Theorem 5.4], in conjunction with [9, Lemma 3.2], one gets

$$v \leq v_{\delta,1/n} + C(n\delta^{1/2} + n^{-1/4}),$$

and the rate improves further to $1/10$.

3.2 Improvement to Theorem 5.7 in [8]

For a wide class of numerical schemes, similar modifications as those used to prove Theorem 2.1 can be performed to improve the error estimates given in [8, Theorem 5.7]. Following as much as possible the notation in [8], let us define for any $s \geq 0$, $x \in \mathbb{R}^d$, $a \in A$ the random variable

$$Y^{a,s,x} := x + b(s,x,a)h + \sigma(s,x,a)\zeta,$$

where ζ is an $\mathbb{R}^p$-valued random variable such that

$$\mathbb{E}[\zeta] = 0, \quad |\mathbb{E}[\zeta_i\zeta_j] - h\delta_{ij}| \leq Ch^2 \quad \text{and} \quad \mathbb{E}[\zeta^k] \leq Ch^2 \text{ for any } k \geq 3. \tag{3.1}$$

It is easy to check, by Taylor expansion, that for any smooth function ϕ the estimate in [8, Lemma 5.10] for the truncation error of the generator becomes

$$\left|L_a\phi(s,x) - h^{-1}\mathbb{E}\left[\phi(s+h, Y^{a,s,x}) - \phi(s,x)\right]\right| \leq Ch$$

for a constant C depending only on C_1 and C_2 in assumptions (H2)–(H3) and the bounds on the derivatives $\partial_t^m D_x^k \phi$ for $2m + k \leq 4$.

Observe that conditions (3.1) are slightly stronger than (5.4) in [8], who only assume accuracy of the moments to order $h^{3/2}$ instead of h^2 in (3.1), so that only order $1/2$ consistency results instead of order 1 above. However, the higher order assumptions are satisfied by very common schemes such as the classical semi-Lagrangian scheme [4, 5] corresponding to the choice

$$\mathbb{P}(\zeta_i = \pm h^{1/2}) = 1/2 \qquad \text{for } i = 1, \ldots, p. \tag{3.2}$$

The scheme considered in [8] is then recursively defined, for any $x \in \mathbb{R}^d$, by

$$\begin{aligned} \hat{v}_h(s,x) &= g(x) && \text{if} \quad s \in (T-h, T], \\ \hat{v}_h(s,x) &= \sup_{a\in A} \left\{f_a(s,x)h + \mathbb{E}\left[\hat{v}_h(s+h, Y^{a,s,x})\right]\right\} && \text{if} \quad s \leq T-h. \end{aligned}$$

[2] Note that in Section 5 of [8], our δ above is denoted by h^2. We introduce δ to avoid ambiguity with the parameter h used in the previous sections of this paper (corresponding to $h = 1/n$ in the present section).

Proceeding to a perturbation and regularization of $\hat{v}_h$ as in [8] (the notation follows the one in Section 2.2, i.e. $\hat{u}_h^{(\varepsilon)}$ is the mollification of $\hat{u}_h$, the solution of the scheme with perturbed "shaken" coefficients) we get the inequality

$$L_a \hat{u}_h^{(\varepsilon)} + f_a \leq Ch\varepsilon^{-3}$$

in $[0, T-h] \times \mathbb{R}^d$ for some constant C depending only on C_0, C_1 in assumptions (H2) and (H3). Arguing as in the proof of Theorem 2.1, one obtains

$$\hat{v}_h \leq v + Ch^{1/4}.$$

Similarly, an upper bound of order $1/4$ for $v - \hat{v}_h$ can be obtained. This aligns the bounds for the scheme (3.2) with those obtained in [5] by PDE techniques.

4 Discussion and conclusions

In this short paper, we show a convergence rate of $1/4$ for piecewise constant control approximations to value functions of stochastic optimal control problems. This result is robust and holds for degenerate problems with non-smooth, merely Lipschitz continuous value functions. If the data and value function are smoother, we show that the approximation has rate 1 and explain why this is the maximal rate.

Our rate 1/4 in (2.1) improves both the order 1/6 in [8] and the rate 1/10 achieved in [3] by different (PDE) techniques. We also carefully explain why we can improve the result in [8]. It is an interesting open question if the same rate could be obtained purely by PDE techniques.

This work also opens up the possibility of improving the error estimates for other approximation schemes as outlined in Section 3. Moreover, it enables a purely probabilistic error analysis for semi-Lagrangian schemes for HJB equations with results that are in line with the best available results by PDE methods. We refer to [12] for the details.

References

[1] Barles, G. and Jakobsen, E.R.: On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.*, **36**, (2002), 33–54. MR-1916291

[2] Barles, G. and Jakobsen, E.R.: Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, **43(2)**, (2005), 540–558. MR-2177879

[3] Barles, G. and Jakobsen, E.R.: Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Math. Comput.*, **74(260)**, (2007), 1861–1893. MR-2336272

[4] Camilli, F. and Falcone, M.: An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.*, **29(1)**, (1995), 97–122. MR-1326802

[5] Debrabant, K. and Jakobsen, E.R.: Semi-Lagrangian schemes for linear and fully non-linear diffusion equations. *Math. Comp.*, **82(283)**, (2012), 1433–1462. MR-3042570

[6] Dumitrescu, R. and Reisinger, C. and Zhang, Y.: Approximation schemes for mixed optimal stopping and control problems with nonlinear expectations and jumps. *Appl. Math. Optim.*, (2019), to appear.

[7] Kossaczký, I. and Ehrhardt, M. and Günther, M.: Modifications of the PCPT method for HJB equations. In *AIP Conference Proceedings*, **1773(1)**, (2016).

[8] Krylov, N.V.: Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies. *Electron. J. Probab.*, **4(2)**, (1999), 1–19. MR-1668597

[9] Krylov, N.V.: On the rate of convergence of finite-difference approximations for Bellmans equations with variable coefficients. *Probab. Theory Relat. Fields*, **117(1)**, (2000), 1–16. MR-1759507

[10] Lions, P.L.: Formule de Trotter et equations de Hamilton–Jacobi–Bellman. *Calcolo*, **17(4)**, (1980), 321–331. MR-0631764

[11] Lyons, T.: Uncertain volatility and the risk-free synthesis of derivatives. *Appl. Math. Finance*, **2(2)**, (1995), 117–133.

[12] Picarelli, A. and Reisinger, C.: Probabilistic error analysis for some approximation schemes to optimal control problems, arXiv:1810.04691.

[13] Reisinger, C. and Forsyth, P.A.: Piecewise constant policy approximations to Hamilton–Jacobi–Bellman equations. *Appl. Numer. Math.*, **103**, (2016), 27–47. MR-3458022

Acknowledgments. The first author was supported by the Toppforsk (research excellence) project Waves and Nonlinear Phenomena (WaNP), grant no. 250070 from the Research Council of Norway.