Error Estimates for Second Order Hamilton-Jacobi-Bellman Equations. Approximation of Probabilistic Reachable Sets

Mohamed Assellaou, Olivier Bokanowski, Hasnaa Zidani

Error estimates for second order Hamilton-Jacobi-Bellman equations. Approximation of probabilistic reachable sets*

Mohamed Assellaou[†] Olivier Bokanowski [‡] Hasnaa Zidani[§]

Abstract This work deals with numerical approximations of unbounded and discontinuous value functions associated to some stochastic control problems. We derive error estimates for monotone schemes based on a Semi-Lagrangian method (or more generally in the form of a Markov chain approximation). A motivation of this study consists in approximating chance-constrained reachability sets. The latters will be characterized as level sets of a discontinuous value function associated to an adequate stochastic control problem. A precise analysis of the level-set approach is carried out and some numerical simulations are given to illustrate the approach.

2010 Mathematics Subject Classification. Primary: 65M15, 93E20; Secondary: 49L25

Keywords. Error estimates for HJB equations, numerical approximation, unbounded and discontinuous value function, chance constraints, level-set approach, stochastic reachability analysis.

1 Introduction

Throughout this paper, we denote by $T > 0$ a fixed final horizon. Consider a controlled process $X^u_{t,x}$ satisfying :

$$\begin{cases} dX(s) = b(s, X(s), u(s))ds + \sigma(s, X(s), u(s))dW(s), & \forall s \in [t, T] \\ X(t) = x. \end{cases} \tag{1}$$

where the diffusion σ and drift b are two Lipschitz continuous functions, $W(\cdot)$ is the classical Brownian motion, and u is a control function that takes its values in a compact

*This work was partially supported by the EU under the 7th Framework Programme Marie Curie Initial Training Network "FP7-PEOPLE-2010-ITN", SADCO project, GA number 264735-SADCO. The second author would like to thank F. Delarue, L. Huang and S. Menozzi for useful discussions concerning Lemma3.3.

[†]Unité des mathématiques appliquées (UMA), ENSTA ParisTech, 828 Bd Maréchaux, 91120 Palaiseau. Mohamed.Assellaou@ensta-paristech.fr

[‡]Laboratoire Jacques-Louis Lions, UMR 7598, Université Paris-Diderot (Paris 7), UFR de Mathématiques - 5 rue Thomas Mann, 75205 Paris CEDEX 13. boka@math.jussieu.fr

[§]Unité des mathématiques appliquées (UMA), ENSTA ParisTech, 828 Bd Maréchaux, 91120 Palaiseau. Hasnaa.Zidani@ensta-paristech.fr

subset U of $\mathbb{R}^q$ ($q \geq 1$). Under suitable assumptions on b, σ and on $\mathcal{U}$, equation (1) admits a unique solution (see Section 2, for precise assumptions). Now, consider the following control problem

$$\vartheta(t,x) := \sup_{u \in \mathcal{U}} \mathbb{E}\big[\Phi(X^u_{t,x}(T))\big] \tag{2}$$

where $\Phi : \mathbb{R}^d \to \mathbb{R}$ is measurable, with linear growth. In this paper, we are interested in error estimates of numerical approximations of ϑ.

The first approximation that will be considered here is a very classical one that consists of introducing a family of Lipschitz continuous functions $(\Phi_\epsilon)_\epsilon$ converging pointwisely to Φ. Then the value function ϑ can be itself approximated by the value functions ϑ_ϵ defined as:

$$\vartheta_\epsilon(t,x) := \sup_{u \in \mathcal{U}} \mathbb{E}\big[\Phi_\epsilon(X^u_{t,x}(T))\big].$$

It is known that under quite general assumptions on the data and on Φ_ϵ, one can show that ϑ_ϵ converges pointwisely towards ϑ, when $\epsilon \to 0$. In this paper, we are also interested in the error estimate of $\vartheta - \vartheta_\epsilon$ depending on the measure of the set where the two functions Φ and Φ_ϵ differ. The result that will be studied here is obtained under an ellipticity condition of the diffusion matrix. An extension to the more general case whith degenerate matrices is still a challenging problem that is not covered in this paper.

The second step in the approximation of ϑ is to discretize the Hamilton-Jacobi-Bellman equation satisfied by ϑ_ϵ. Indeed ϑ_ϵ will be shown to be the unique continuous viscosity solution of:

$$\begin{aligned} -\partial_t \vartheta_\epsilon + \mathcal{H}(t,x,D\vartheta_\epsilon, D^2\vartheta_\epsilon) &= 0 \qquad \text{in } (0,T)\times\mathbb{R}^d \\ \vartheta_\epsilon(T,x) &= \Phi_\epsilon(x) \quad \text{in } \mathbb{R}^d \end{aligned}$$

where $\mathcal{H}(t,x,p,Q) := \sup_{a \in U}\big(-b(t,x,a)\cdot p - \mathrm{Tr}([\sigma\sigma^{\mathsf{T}}](t,x,a)Q)\big)$. In the case when the the drift b and the diffusion σ are bounded and when the value function ϑ_ϵ is itself bounded, error estimates of monotone schemes have been obtained first by Krylov [27] for a case where σ is a constant function. These results were developed further in [28, 29, 6, 7, 8] by introducing new tools that allow to consider the case where σ can depend on time, space and also on the control variable. Several other extensions of the theory have been analysed in the literature, let us mention some of these extensions for stopping-game problems [13], for impulsive control systems [14], for integro-partial differential HJB equations [18, 9, 10], and for a general class of coupled HJB systems [16]. Note also that the case of fully uniformly elliptic operators have been also studied in [17] using a different approach than the one introduced by Krylov. Here, we extend the theory of error estimates to an unbounded Lipschitz setting. The proof is still based on "Krylov regularization" and on some refined consistency estimates. To the best of our knowledge, this is the first result in the case where b, σ and the solution to the HJB equation itself are unbounded with respect to the space variable (with linear growth).

The study in this paper is partly motivated by reachability analysis for stochastic systems. Let $\mathcal{C}$ be a non-empty subset of $\mathbb{R}^d$ ("the target"). Let $\rho \in [0,1[$ and $t \leq T$. Consider the backward reachable set under probability of success ρ, that is, the set of initial points x for which the probability that there exists a process $X^u_{t,x}$ solution of (1), associated with an admissible control $u \in \mathcal{U}$ and that reaches $\mathcal{C}$ at time T is higher than ρ:

$$\Omega^\rho_t = \Big\{ x \in \mathbb{R}^d \,\big|\, \exists u \in \mathcal{U},\ \mathbb{P}[X^u_{t,x}(T) \in \mathcal{C}] > \rho \Big\}. \tag{3}$$

The sets Ω^ρ_t can be characterized by using the *level-set approach*. Indeed, it is straithforward to see that Ω^ρ_t is equivalent to:

$$\Omega^\rho_t = \Big\{ x \in \mathbb{R}^d \,\big|\, \exists u \in \mathcal{U},\, \mathbb{E}[\mathbb{1}_{\mathcal{C}}(X^u_{t,x}(T))] > \rho \Big\}.$$

Moreover, by considering the control problem (2) with $\phi(x) := \mathbb{1}_{\mathcal{C}}(x)$, it is possible to show that for every $\rho > 0$ and every $t \in [0, T]$, the backward reachable set Ω^ρ_t is given by the *level-set*:

$$\Omega^\rho_t = \{x \in \mathbb{R}^d, \vartheta(t, x) > \rho\}.$$

The level approach has been introduced in [33] to model front propagation problems. Then, the method has attracted a big interest for studying backward reachable sets of continuous non-linear dynamical systems under general conditions, see [31, 11] and the references therein. The idea of using the level set approach in discrete time stochastic setting has been also considered in [5, 1, 2]. In this case, the value function is obtained by solving the dynamic programming principle. In the present paper, we are interested in the approximation of the probabilistic backward reachable sets for time-continuous stochastic processes. We analyse the approach and we provide error estimates between the exact sets and their numerical approximation.

Let us mention that other numerical methods for reachability analysis have been introduced and analysed in the literature. The most natural numerical algorithm consists in using Monte Carlo [34] simulations to generate a set of trajectories starting from a given initial position $x \in \mathbb{R}^d$. Then the percentage of trajectories reaching the target gives an approximation of the probability of success (for reaching the target) when starting from this position x. On the other hand, for linear stochastic systems, a bound for the probability of hitting a target can be obtained by using the enclosing hulls of the probability density function for time intervals, see [4, 3] for instance. Note that these approaches are used to calculate the probabilities of success but do not allow to define the entire set of points that have the same given probability. In addition, Monte-Carlo-based methods often require a large number of simulations to obtain a good accuracy. We will use such simulations in Section 6 to validate our level-set approach.

The paper is organized as follows: Section 2 introduces the notations and the setting of the control problems (2). In section 3, we derive an error estimate for the value functions when the payoff function Φ is approximated by smooth functions. In section 4, we analyse the error estimates for a semi-Lagrangian scheme for the approximation of the value function. In section 5, we study the characterization and approximation of probabilistic backward reachable sets. Section 6 is devoted to some illustrative numerical examples.

2 Setting of the problem. Basic assumptions

Throughout this paper, $|\cdot|$ denotes the Euclidean norm for any $\mathbb{R}^N$ type space, and $\mathbb{B}_R$ is the closed ball centred at the origin and with radius R.

For a given set $\mathcal{S} \subset \mathbb{R}^N$, the indicator function is given by $\mathbb{1}_{\mathcal{S}}(x) = 1$ if $x \in \mathcal{S}$ and $\mathbb{1}_{\mathcal{S}}(x) = 0$ otherwise. The distance function to $\mathcal{S}$ is $\mathrm{dist}(\mathrm{x}, \mathcal{S}) = \inf\{|\mathrm{x}-\mathrm{y}| : \mathrm{y} \in \mathcal{S}\}$. We also denote by $\mu(\mathcal{S})$ the measure of $\mathcal{S}$ with respect to Lebesgue's measure.

For any real valued function $\varphi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, we say that $\varphi \in C^{k,l}([0,T] \times \mathbb{R}^d)$ (for non-negative integers k, l) iff all the partial derivatives $\partial^i_{t^i} \partial^j_{x^j} \varphi$, for $0 \le i \le k$ and $0 \le j \le l$, exist and are continuous functions. Moreover, we denote by $\|\varphi\|_0$ the norm given by:

$$\|\varphi\|_0 := \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\varphi(t,x)|,$$

and for the matrix $D^k \varphi$ (the k-th derivative with respect to the variable x):

$$\|D^k \varphi\|_0 := \max_{\alpha_i \ge 0,\ \sum \alpha_i = k} \|\frac{\partial^k}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \varphi\|_0.$$

Let $\{\boldsymbol{\Omega}, \mathcal{F}_t, \{\mathcal{F}_t\}_{t \ge 0}, P\}$ be a filtered probability space, $W(.)$ be a given m-dimensional Brownian motion, and $T > 0$. We denote by $\mathcal{U}$ the set of progressively measurable [1] processes valued in $U \subset \mathbb{R}^q$ $(q \ge 1)$ where U is a non empty compact set. Let $(X^u_{t,x}(s))_{0 \le s \le T}$ be a controlled process valued in $\mathbb{R}^d$ solution of the following stochastic differential equation:

$$\begin{cases} dX(s) = b(s, X(s), u(s))ds + \sigma(s, X(s), u(s))dW(s) & \forall s \in [t, T], \\ X(t) = x, \end{cases} \tag{4}$$

where $\sigma : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}^{d \times m}$ and $b : [0,T] \times \mathbb{R}^d \times U \to \mathbb{R}^d$ are two continuous functions satisfying the following standard assumption:

(H1a) there exists $L_0 > 0$ such that for any $(s,t,x,y,u) \in [0,T] \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \times U$, we have:

$$|b(t,x,u) - b(s,y,u)| + |\sigma(t,x,u) - \sigma(s,y,u)| \le L_0(|x-y| + |t-s|^{\frac{1}{2}}).$$

For convenience, we assume also that $|\sigma(0,0,u)| + |b(0,0,u)| \le L_0$ for any $u \in U$.

Assumption (H1a) imposes that b and σ are Lipschitz continuous with respect to x and $\frac{1}{2}$-Hölder continuous with respect to t. Note that the uniform Lipschitz property on b and σ and the compactness of U guarantee the existence of a controlled process on the time interval $[t,T]$ for each given initial data x, and for every admissible control $u \in \mathcal{U}$ (see [23] for more details). A process $X^u_{t,x}$ solution of (4) associated to a control $u \in \mathcal{U}$ will be said *admissible*. Moreover, there exists K_0 depending only on L_0, T, d and m (see [35, page 42] or [23, Appendice D]) such that for any $u \in \mathcal{U}$, $0 \le t \le t' \le T$ and

[1] A precise definition of a progressively measurable function can be found in [23, page 159]

$x, x' \in \mathbb{R}^d$

$$\mathbb{E}\left[\sup_{\theta\in[t,t']}\left|X^u_{t,x}(\theta)-X^u_{t,x'}(\theta)\right|^2\right] \leq K_0^2|x-x'|^2, \tag{5a}$$

$$\mathbb{E}\left[\sup_{\theta\in[t,t']}\left|X^u_{t,x}(\theta)-X^u_{t',x}(\theta)\right|^2\right] \leq K_0^2(1+|x|^2)\ |t-t'|. \tag{5b}$$

Furthermore, for every $p \geq 1$, there exists $K_p > 0$ such that:

$$\mathbb{E}\left[\sup_{\theta\in[t,t']}\left|X^u_{t,x}(\theta)-x\right|^p\right] \leq K_p(1+|x|^p)\ |t-t'|. \tag{5c}$$

For a part of the results that will be presented in this paper, we will also need an ellipticity condition that we state as follows:

(H1b) σ depends only on (t,x) and there exists a real number $\Lambda \geq 1$, such that:

$$\forall (t,x)\in(0,T)\times\mathbb{R}^d, \quad \Lambda I_d \geq \sigma(t,x)\sigma(t,x)^T \geq \Lambda^{-1} I_d \tag{6}$$

where I_d is the identity matrix, and the inequalities (6) are in the sense of symmetric matrices: $\Lambda\|\xi\|^2 \geq \langle \xi, \sigma\sigma^T\xi\rangle \geq \Lambda^{-1}\|\xi\|^2$, $\forall \xi \in \mathbb{R}^d$.

Assumption (H1b) will be used in section 3. It is useful to derive Aronson type estimates [22] on the density of probability associated with the process $X^u_{t,x}$, precise statement is given in Lemma 3.3.

Remark 2.1 Note that more generally, assumption (H1b) can be replaced by a weak Hörmander condition where the diffusion takes part only in some components and the noise propagates through a chain of differential equations. In that context, Aronson type estimates can still be obtained as in [22], and the results of the present paper could be extended.

Throughout the paper, we denote by $\Phi : \mathbb{R}^d \to \mathbb{R}$ a given final cost function satisfying the assumption:

(H2) Φ is measurable, and with linear growth, i.e, there exists $M_0 > 0$ such that:

$$|\Phi(x)| \leq M_0(1+|x|) \qquad \text{a.e. } x \in \mathbb{R}^d.$$

Now, consider the following optimal control problem:

$$\vartheta(t,x) := \sup_{u\in\mathcal{U}} \mathbb{E}\big[\Phi(X^u_{t,x}(T))\big]. \tag{7}$$

Under assumptions (H1a)-(H2), the value function ϑ is well defined but it may be discontinuous. Moreover, according to [25], if Φ is upper semi-continuous (u.s.c) and under some additional convexity assumptions on the drift and the diffusion coefficients, ϑ is u.s.c and satisfies the following HJB equation:

$$-\partial_t\vartheta + \mathcal{H}(t,x,D\vartheta,D^2\vartheta) = 0 \qquad \text{in } (0,T)\times\mathbb{R}^d \tag{8a}$$

$$\vartheta(T,x) = \Phi(x) \quad \text{in } \mathbb{R}^d. \tag{8b}$$

In this paper, we are interested in the error estimates theory for numerical approximations of the value function ϑ. Since ϑ is discontinuous, we shall first introduce a regularized problem with a controlled error with respect to the original problem, and on which further analysis and numerical approximation will be more convenient. For this aim, we consider a family of regularized functions $(\Phi_\epsilon)_{\epsilon>0}$, and denote by $\mathcal{D}_\epsilon$ (for any $\epsilon > 0$) the set where Φ_ϵ and Φ take different values:

$$\mathcal{D}_\epsilon := \{x \in \mathbb{R}^d \mid \Phi_\epsilon(x) \neq \Phi(x)\}.$$

Then we consider the following assumption:

(H3) (i) For every $\epsilon \in]0,1]$, $\Phi_\epsilon : \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz continuous function with Lipschitz constant $L_\epsilon \geq 0$,
(ii) there exists a constants $M_0 > 0$ (independent of ϵ), such that

$$|\Phi_\epsilon(x)| \leq M_0(1 + |x|), \qquad x \in \mathbb{R}^d,$$

(iii) there exists a constants $M_1 > 0$ (independent of ϵ), such that for any $A > 0$

$$\mu(\mathcal{D}_\epsilon \cap \mathbb{B}_A) \leq M_1\, A\, \epsilon, \qquad \epsilon \in]0,1].$$

(The constant M_0 in (H3)-(ii) can be chosen to be the same constant as in (H2) without loss of generality.)

Of course the existence of such approximated Lipschitz continuous functions implicitly imposes some more requirements on the function Φ itself. However, (H3) is satisfied in many cases. For instance, it is possible to construct a family Φ_ϵ satisfying (H3) when the function Φ is piecewise Lipschitz function with discontinuities lying in a union of compact regular sub-manifolds of dimension $d-1$. See also remark 5.2 or Section 5 for construction of such approximations in some particular cases.

Notice also that if $\mathcal{D}_\epsilon \subset \mathbb{B}_A$ for some given $A \geq 1$ and for all $\epsilon \in]0,1]$, then (H3)-(iii) is simply equivalent to assume that there exists $M_1 > 0$ such that $\mu(\mathcal{D}_\epsilon) \leq M_1\epsilon$ for every $\epsilon \in]0,1]$.

Now, consider an approximation of ϑ given by the value function associated to the following control problem:

$$\vartheta_\epsilon(t,x) := \sup_{u\in\mathcal{U}} \mathbb{E}\big[\Phi_\epsilon(X^u_{t,x}(T))\big]. \tag{9}$$

Under $(H3)$ and using the ellipticity condition $(H1b)$, we shall derive an error estimate of $\vartheta - \vartheta_\epsilon$.

The next step consists of deriving error estimates for numerical approximation of ϑ_ϵ which is a Lipschitz continuous function with linear growth in x variable. For every $\epsilon > 0$, this new value function can be characterized as unique Lipschitz viscosity solution of the HJB equation:

$$\begin{aligned} -\partial_t\vartheta_\epsilon + \mathcal{H}(t,x,D\vartheta_\epsilon,D^2\vartheta_\epsilon) &= 0 \qquad \text{in } (0,T)\times\mathbb{R}^d & \text{(10a)}\\ \vartheta_\epsilon(T,x) &= \Phi_\epsilon(x) \quad \text{in } \mathbb{R}^d. & \text{(10b)}\end{aligned}$$

The Lipschitz regularity is suitable for deriving the error estimates when the HJB equation is approximated by a monotone scheme. However, error estimates for second order

HJB equations have been studied so far only for Lipschitz *bounded* viscosity solutions, as well as *bounded* coefficients b, and σ. Here, we are concerned by the case of unbounded viscosity solutions with linear growth and unbounded coefficients with respect to the x variable (which is the case of many real applications such as call options in mathematical finance). We recall here that thouroughout all the paper the control set U is bounded.

3 The regularized problem

Notation. Throughout this sections and the following ones, the constant C will denote a generic positive real number that may depend only on T, d, m, L_0, K_0, K_p.

3.1 Error estimate for the regularization procedure

Here we focus on the error estimate between ϑ and the approximated value function ϑ_ϵ.

Theorem 3.1 *Assume (H1a), (H1b), (H2), and (H3). Let ϑ and ϑ_ϵ be the value functions defined respectively by (7) and (9).*
(i) There exist a constant $C_0 > 0$ (depending only on $T, L_0, \Lambda, M_0, M_1$) and $\epsilon_0 \in]0,1]$, such that for every $0 < \epsilon < \epsilon_0$ the following estimate holds:

$$|\vartheta(t,x) - \vartheta_\epsilon(t,x)| \leq C_0 \frac{1 + |x|^2 + |\log \epsilon|}{(T-t)^{d/2}} \, \epsilon \tag{11}$$

for every $0 \leq t < T$ and $x \in \mathbb{R}^d$.
(ii)Furthermore, if there exists $A > 0$ such that $\mathcal{D}_\epsilon \subset \mathbb{B}_A$ for every $\epsilon \in]0,1]$, then there exist $C_1, C_2 > 0$ (depending only on T, M_0, M_1 and A) such that the following estimate holds:

$$|\vartheta(t,x) - \vartheta_\epsilon(t,x)| \leq \frac{C_1}{(T-t)^{d/2}} \, e^{-C_2 \frac{\mathrm{dist}(\mathrm{x}, \mathcal{D}_\epsilon)^2}{(T-t)}} \, \epsilon \tag{12}$$

for every $\epsilon \in]0,1]$, for every $x \in \mathbb{R}^d$ and every $0 \leq t < T$.

Remark 3.2 In particular, if there exists $A > 0$ such that $\mathcal{D}_\epsilon \subset \mathbb{B}_A$ for every $\epsilon > 0$, then Theorem 3.1(ii) leads directly to the following bound (since $\mathrm{dist}(\mathrm{x}, \mathcal{D}_\epsilon) \geq 0$) :

$$|\vartheta(t,x) - \vartheta_\epsilon(t,x)| \leq \frac{C_1}{(T-t)^{d/2}} \epsilon \tag{13}$$

for every $0 \leq t < T$ and every $x \in \mathbb{R}^d$. Moreover, by using (12) and the fact that $e^{-r} \leq C/r^{d/2}$ for all $r > 0$ (for some constant $C \geq 0$), we conclude that there exists $C_1' \geq 0$ depending on T, M_0, M_1, A such that:

$$|\vartheta(t,x) - \vartheta_\epsilon(t,x)| \leq \frac{C_1'}{[\mathrm{dist}(\mathrm{x}, \mathcal{D}_\epsilon)]^{\mathrm{d}}} \, \epsilon \tag{14}$$

for every $0 \leq t \leq T$ and any $x \in \mathbb{R}^d \setminus \mathcal{D}_\epsilon$.

Before giving the proof of Theorem 3.1, we first recall some known results on the *density of probability* of the process $X^u_{t,x}(\cdot)$, for a given $(t,x) \in [0,T)$ and an admissible control $u \in \mathcal{U}$. We will denote by $y \longmapsto p^u(t,x;s,y)$ the density of probability function associated to the process $X^u_{t,x}(s)$ (for a given admissible control $u \in \mathcal{U}$).

Lemma 3.3 *Assume (H1a) and (H1b). There exist $c_1, c_2, c_3 > 0$ such that for any $(t,s,x,y) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^d$ such that $t < s$, and for any admissible control $u \in \mathcal{U}$, the following estimate holds:*

$$|p^u(t,x;s,y)| \leq \frac{c_1}{(s-t)^{\frac{d}{2}}}\, e^{-c_2\frac{|x-y|^2}{2(s-t)}}\, e^{c_3|x|^2}. \tag{15}$$

Proof. Let $u \in \mathcal{U}$. From [22, Theorem 1.1], there exists $c_1, c_2 > 0$ such that for every $0 \leq t < s \leq T$ and every $x, y \in \mathbb{R}^d$, we have:

$$p^u(t,x;s,y) \leq \frac{c_1}{(s-t)^{\frac{d}{2}}} \exp\left(- c_2 \frac{|\theta^u_{t,x}(s) - y|^2}{s-t}\right) \tag{16}$$

where $\theta(s) := \theta^u_{t,x}(s)$ is the solution of the differential equation:

$$\begin{aligned} &\frac{d}{ds}\theta(s) = b(s,\theta(s),u(s)), \quad s \geq t, \\ &\theta(t) = x. \end{aligned}$$

Note that by [22, Theorem 1.1], the constants c_1 and c_2 depend only on d, Λ and do not depend neither on θ^u nor on the control u. Therefore, the estimate (15) is valid for any control function $u \in \mathcal{U}$. By assumption (H1a) and by some classical estimates we get: $|\theta^u_{t,x}(s) - x| \leq L_0(1+|x|)(s-t)e^{L_0(s-t)}$. On the other hand, a straightforward calculation yields to: $|x-y| \leq |\theta^u_{t,x}(s) - y| + |\theta^u_{t,x}(s) - x|$. Hence

$$-\frac{|\theta^u_{t,x}(s) - y|^2}{s-t} \leq -\frac{|x-y|^2}{2(s-t)} + 2L_0^2(1+|x|^2)(s-t)e^{2L_0(s-t)}.$$

Thus, we obtain

$$\begin{aligned} p^u(t,x;s,y) &\leq \frac{c_1}{(s-t)^{\frac{d}{2}}}\, e^{-c_2\frac{|x-y|^2}{2(s-t)}}\, e^{c_3(1+|x|^2)} \\ &\leq \frac{c_1 e^{c_3}}{(s-t)^{\frac{d}{2}}}\, e^{-c_2\frac{|x-y|^2}{2(s-t)}}\, e^{c_3|x|^2}, \end{aligned}$$

with $c_3 = 2c_2L_0^2Te^{2L_0T}$, which gives the desired upper bound. The lower bound can be derived in the same way. $\square$

Now we turn to the proof of theorem 3.1.

Proof of Theorem 3.1. Let $t \in [0,T]$ and $x \in \mathbb{R}^d$. We have

$$\begin{aligned} |\vartheta(t,x) - \vartheta_\epsilon(t,x)| &= \left| \sup_{u\in\mathcal{U}} \mathbb{E}[\Phi(X^u_{t,x}(T))] - \sup_{u\in\mathcal{U}} \mathbb{E}[\Phi_\epsilon(X^u_{t,x}(T))]\right| \\ &\leq \sup_{u\in\mathcal{U}} \mathbb{E}\big[\big|\Phi(X^u_{t,x}(T)) - \Phi_\epsilon(X^u_{t,x}(T))\big|\big] \\ &\leq \sup_{u\in\mathcal{U}} \int_{\mathbb{R}^d} |\Phi(y) - \Phi_\epsilon(y)|\, p^u(t,x;T,y)dy, \end{aligned}$$

where $y \to p^u(t,x;T,y)$ is the density of probability associated to the process $X^u_{t,x}(T)$ associated to a control function $u \in \mathcal{U}$. Since $\mathrm{supp}(\Phi - \Phi_\epsilon) \subset \mathcal{D}_\epsilon$, it comes:

$$|\vartheta(t,x) - \vartheta^\epsilon(t,x)| \leq \sup_{u\in\mathcal{U}} \int_{\mathcal{D}_\epsilon} |\Phi(y) - \Phi_\epsilon(y)|\, p^u(t,x;T,y)dy. \tag{17}$$

We first consider the proof of (ii). We assume that $\mathcal{D}_\epsilon \subset \mathbb{B}_A$ for some $A > 0$ and for every $\epsilon > 0$. Then by taking into account Lemma 3.3 (using the fact that for every $y \in \mathcal{D}_\epsilon$, we have $|x-y| \geq \mathrm{dist(x,}\mathcal{D}_\epsilon))$, and by assumption (H3) (which implies also that for any $y \in \mathbb{R}^d$, $|\Phi(y) - \Phi_\epsilon(y)| \leq 2M_0(1+|y|))$, we get:

$$\begin{aligned}
&\int_{\mathcal{D}_\epsilon} (\Phi(y) - \Phi_\epsilon(y))p^u(t,x;T,y)dy \\
&\quad \leq \quad 2M_0(1+A)c_1(T-t)^{-\frac{d}{2}}\, e^{-c_2 \frac{\mathrm{dist(x,}\mathcal{D}_\epsilon)^2}{T-t}} e^{c_3 A^2} \mu(\mathcal{D}_\epsilon \cap \mathbb{B}_A) \\
&\quad \leq \quad 2M_0 M_1 c_1 (T-t)^{-\frac{d}{2}}\, e^{-c_2 \frac{\mathrm{dist(x,}\mathcal{D}_\epsilon)^2}{T-t}} e^{c_3 A^2} (1+A)A\,\epsilon,
\end{aligned} \tag{18}$$

for every $u \in \mathcal{U}$, which concludes to the desired bound for (ii).

We come back to the general case (i). For $\epsilon \in]0,1]$, let $X_\epsilon \in \mathbb{R}$ be the unique parameter such that $X_\epsilon \geq \frac{d-1}{2c_2}$ (where c_2 is introduced in Lemma 3.3) and

$$X_\epsilon^{(d-1)/2} e^{-c_2 X_\epsilon} = \epsilon. \tag{19}$$

Then, as $\epsilon \to 0$, it holds $X_\epsilon \sim \frac{1}{c_2}|\log \epsilon|$ and therefore, $X_\epsilon = O(|\log(\epsilon)|)$. Introduce also the positive constant $R_\epsilon := \sqrt{(T-t)X_\epsilon}$.

Let $u \in \mathcal{U}$ be fixed. Using the estimate (16), we obtain a first bound as follows:

$$\begin{aligned}
&\int_{\mathcal{D}_\epsilon \cap (\theta^u_{t,x}(T) + \mathbb{B}_{R_\epsilon})} (\Phi(y) - \Phi_\epsilon(y))p^u(t,x;T,y)dy \\
&\quad \leq \quad 2M_0(1 + |\theta^u_{t,x}(T)| + R_\epsilon)\frac{c_1}{(T-t)^{\frac{d}{2}}} \mu\Big(\mathcal{D}_\epsilon \cap (\theta^u_{t,x}(T) + \mathbb{B}_{R_\epsilon})\Big) \\
&\quad \leq \quad 2M_0 M_1 (1 + |\theta^u_{t,x}(T)| + R_\epsilon)^2 \frac{c_1}{(T-t)^{\frac{d}{2}}}\,\epsilon. \\
&\quad \leq \quad 2M_0 M_1 C(1 + |x| + R_\epsilon)^2 \frac{c_1}{(T-t)^{\frac{d}{2}}}\,\epsilon
\end{aligned} \tag{20}$$

where we have used that $|\theta^u_{t,x}(T)| \leq C(1+|x|)$ for some constant $C > 0$ that only depends on T and L_0 (and does not depend on u).

On the other hand,

$$
\begin{aligned}
\int_{\mathcal{D}_\epsilon \cap\ \mathbb{R}^d \setminus (\theta^u_{t,x}(T) + \mathbb{B}_{R_\epsilon})} & (\Phi(y) - \Phi_\epsilon(y)) p^u(t,x;T,y) dy \\
&\leq \int_{\mathbb{R}^d \setminus (\theta^u_{t,x}(T) + \mathbb{B}_{R_\epsilon})} |\Phi(y) - \Phi_\epsilon(y))| p^u(t,x;T,y) dy \\
&\leq \int_{|y - \theta^u_{t,x}(T)| \geq R_\epsilon} 2M_0(1+|y|) \frac{c_1}{(T-t)^{d/2}}\, e^{-c_2 \frac{|\theta^u_{t,x}(T) - y|^2}{T-t}}\, dy \\
&= \int_{|y| \geq R_\epsilon} 2M_0(1 + |y + \theta^u_{t,x}(T)|) \frac{c_1}{(T-t)^{d/2}}\, e^{-c_2 \frac{|y|^2}{T-t}}\, dy \\
&\leq 2c_1 M_0 \int_{|z| \geq \frac{R_\epsilon}{\sqrt{T-t}}} (1 + |\theta^u_{t,x}(T)| + \sqrt{T-t}|z|) e^{-c_2|z|^2}\, dz
\end{aligned} \tag{21}
$$

On the other hand, we have the following Lemma (see the proof in Appendice B):

Lemma 3.4 *For any $\alpha \geq 0$, there exists a constant $q_\alpha > 0$ (depending also on c_2 and d), such that*

$$
\int_{|z| \geq a, z \in \mathbb{R}^d} |z|^\alpha e^{-c_2|z|^2} dz \leq q_\alpha\ a^{\alpha+d-1} e^{-c_2 a^2}, \quad \textit{as } |a| \geq 1.
$$

Hence, with $a := R_\epsilon / \sqrt{(T-t)} = \sqrt{X_\epsilon}$, it comes:

$$
\begin{aligned}
\int_{\mathcal{D}_\epsilon \cap\ \mathbb{R}^d \setminus (\theta^u_{t,x}(T) + \mathbb{B}_{R_\epsilon})} & (\Phi(y) - \Phi_\epsilon(y)) p^u(t,x;T,y) dy \\
&\leq 2c_1 M_0 \Big(q_0 (1 + C(1+|x|)) X_\epsilon^{\frac{d-1}{2}} + q_1 \sqrt{T-t} X_\epsilon^{\frac{d}{2}} \Big) e^{-c_2 X_\epsilon} \\
&\leq 2c_1 M_0 C (1 + |x| + R_\epsilon) X_\epsilon^{\frac{d-1}{2}} e^{-c_2 X_\epsilon}
\end{aligned} \tag{22}
$$

for some constant $C \geq 0$, under the condition that $X_\epsilon \geq 1$ (which is satisfied whenever ϵ is small enough).

By combining (20) and (22), and taking into account that the two estimates do not depend on the control variable u, we get for every ϵ small enough:

$$
\begin{aligned}
& |\vartheta(t,x) - \vartheta^\epsilon(t,x)| \\
\leq\ & 2(c_1 + M_1) M_0 C (1 + |x| + R_\epsilon)^2 \Big((T-t)^{-d/2}\, \epsilon + X_\epsilon^{\frac{d-1}{2}}\, e^{-c_2 X_\epsilon} \Big).
\end{aligned} \tag{23}
$$

By using the definition of X_ϵ and its properties, we get:

$$
\begin{aligned}
|\vartheta(t,x) - \vartheta^\epsilon(t,x)| &\leq 2(c_1 + M_1) M_0 C (1 + |x| + R_\epsilon)^2 (1 + (T-t)^{-d/2}) \epsilon \\
&\leq 2(c_1 + M_1) M_0 C (1 + |x| + C|\log \epsilon|^{1/2})^2 (1 + (T-t)^{-d/2}) \epsilon \\
&\leq C_0 \frac{1 + |x|^2 + |\log \epsilon|}{(T-t)^{d/2}} \epsilon
\end{aligned}
$$

where the constant $C_0 > 0$ depends only on $T, L_0, M_0, M_1, c_1, c_2$, which concludes the proof of (i). $\square$

3.2 Some regularity results for ϑ_ϵ

In this subsection, we provide an upper bound of the Hölder constant of ϑ^ϵ. By using the fact that the function Φ_ϵ is L_ϵ-Lipschitz continuous, we obtain the following:

Lemma 3.5 *Assume (H1a), (H2) and (H3). There exists a constant $C > 0$ such that for every $\epsilon > 0$, the value function ϑ_ϵ satisfies:*

$$|\vartheta_\epsilon(t,x) - \vartheta_\epsilon(t,y)| \le CL_\epsilon |x-y|, \tag{24}$$

for all $x, y \in \mathbb{R}^d$, $t \in [0,T]$. Moreover,

$$|\vartheta_\epsilon(t,x) - \vartheta_\epsilon(s,x)| \le CL_\epsilon(1+|x|)\,|t-s|^{\frac{1}{2}} \tag{25}$$

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

Proof. (i) By straightforward calculations, we obtain:

$$\begin{aligned}|\vartheta_\epsilon(t,x) - \vartheta_\epsilon(t,y)| &\le \sup_{u\in\mathcal{U}} |\mathbb{E}[\Phi_\epsilon(X^u_{t,x}(T))] - \mathbb{E}[\Phi_\epsilon(X^u_{t,y}(T))]| \\ &\le \sup_{u\in\mathcal{U}} \mathbb{E}[|\Phi_\epsilon(X^u_{t,x}(T)) - \Phi_\epsilon(X^u_{t,y}(T))|].\end{aligned}$$

Then by using the Lipschitz regularity of Φ_ϵ, it follows that:

$$|\vartheta_\epsilon(t,x) - \vartheta_\epsilon(t,y)| \le L_\epsilon \sup_{u\in\mathcal{U}} \mathbb{E}\big[\big|X^u_{t,x}(T) - X^u_{t,y}(T)\big|\big].$$

By using (5), we get the inequality:

$$|\vartheta_\epsilon(t,x) - \vartheta_\epsilon(t,y)| \le K_0 L_\epsilon\,|x-y|.$$

(ii) Without loss of generality, we assume that $s = t+h$ for some $h > 0$. By using the definition of ϑ, we have:

$$\begin{aligned}|\vartheta_\epsilon(t{+}h,x) - \vartheta_\epsilon(t,x)| &\le \sup_{u\in\mathcal{U}} \left(\mathbb{E}\left[\Phi_\epsilon(X^u_{t+h,x}(T))\right] - \mathbb{E}\left[\Phi_\epsilon(X^u_{t,x}(T))\right]\right) \\ &\le \sup_{u\in\mathcal{U}} \mathbb{E}\left[\mathbb{E}\left[\left|\Phi_\epsilon(X^u_{t+h,x}(T)) - \Phi_\epsilon(X^u_{t+h,X^u_{t,x}(t+h)}(T))\right| \mid \mathcal{F}_{t+h}\right] | \mathcal{F}_t\right] \\ &\le \sup_{u\in\mathcal{U}} L_\epsilon\,\mathbb{E}\left[\mathbb{E}\left[\left|X^u_{t+h,x}(T) - X^u_{t+h,X^u_{t,x}(t+h)}(T)\right| \mid \mathcal{F}_{t+h}\right] | \mathcal{F}_t\right]\end{aligned}$$

Finally, taking into account (5), we deduce that:

$$|\vartheta_\epsilon(t{+}h,x) - \vartheta_\epsilon(t,x)| \le K_0^2 L_\epsilon (1+|x|)\,h^{1/2}.$$

Therefore, taking any $C \ge \max(K_0, K_0^2)$, the desired result follows. □

It is also known that ϑ_ϵ satisfies the following dynamic programming principle and the HJB equation:

Proposition 3.6 *Assume (H1a), (H2) and (H3).*
(i) Let $(t,x) \in [0,T] \times \mathbb{R}^d$*, and denote* $\mathcal{T}_{[t,T]}$ *the set of* $(\mathcal{F}_\theta)_{\theta\in[t,T]}$*-adapted stopping times with values a.e. in* $[t,T]$*. Let* $\{\tau^u; u \in \mathcal{U}\}$ *be a subset of* $\mathcal{T}_{[t,T]}$ *(independant of* $\mathcal{F}_t$*). Then*

$$\vartheta_\epsilon(t,x) = \sup_{u\in\mathcal{U}} \mathbb{E}[\vartheta_\epsilon(\tau^u, X^u_{t,x}(\tau^u))] \tag{26}$$

(ii) The function ϑ_ϵ *is the unique continous viscosity solution (see definition 3.7), with linear growth, of the following HJB equation:*

$$\begin{aligned} -\partial_t\vartheta_\epsilon + \mathcal{H}(t,x,D\vartheta_\epsilon,D^2\vartheta_\epsilon) &= 0 \qquad \textit{in } (0,T)\times\mathbb{R}^d && \text{(27a)}\\ \vartheta_\epsilon(T,x) &= \Phi_\epsilon(x) \quad \textit{in } \mathbb{R}^d. && \text{(27b)}\end{aligned}$$

where $\mathcal{H}$ *denotes the Hamiltonian function defined by:*

$$\mathcal{H}(t,x,p,Q) := \inf_{u\in U}\left\{-\frac{1}{2}Tr(\sigma(t,x,u)\sigma^T(t,x,u)\,Q) - b(t,x,u)\cdot p\right\} \tag{28}$$

for every $t \in [0,T], x \in \mathbb{R}^d, p \in \mathbb{R}^d$ *and for every symmetric* $d\times d$*-matrix* Q*.*

Definition 3.7 A usc function $\overline{\vartheta}$ (resp. lsc function $\underline{\vartheta}$) on $[0,T]\times\overline{\mathbb{R}}^d$ is a viscosity sub-solution (resp. super-solution) of (27), if for each function $\varphi \in C^{1,2}([0,T]\times\overline{\mathbb{R}}^d)$, at each maximum (resp. minimum) point (t,x) of $\overline{\vartheta}-\varphi$ (resp. $\underline{\vartheta}-\varphi$) the following inequalities hold

$$\begin{cases} -\partial_t\varphi + \mathcal{H}(t,x,D_x\varphi,D_x^2\varphi) \le 0 & \textit{in } [0,T)\times\mathbb{R}^d \\ \min\left(-\partial_t\varphi + \mathcal{H}(t,x,D_x\varphi,D_x^2\varphi), \overline{\vartheta}-\Phi_\epsilon\right) \le 0 & \textit{on } \{T\}\times\mathbb{R}^d. \end{cases}$$

(resp.

$$\begin{cases} -\partial_t\varphi + \mathcal{H}(t,x,D_x\varphi,D_x^2\varphi) \ge 0 & \textit{in } [0,T)\times\mathbb{R}^d \\ \max\left(-\partial_t\varphi + \mathcal{H}(t,x,D_x\varphi,D_x^2\varphi), \overline{\vartheta}-\Phi_\epsilon\right) \ge 0 & \textit{on } \{T\}\times\mathbb{R}^d. \end{cases}$$

The proof of Proposition 3.6 can be found in [23, Chapter 5]. For the uniqueness of viscosity solution we will use in particular the following comparison principle that holds for unbounded solutions (see [20] for the proof).

Proposition 3.8 *Let* $v_1, v_2 : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ *and assume that* v_1 *is u.s.c and* v_2 *is l.s.c, that there exists a constant* $c \ge 0$ *and* $p \ge 1$ *such that* $v_1 \ge -c(1+|x|^p)$ *and* $v_2 \le c(1+|x|^p)$ *for all* $x \in \mathbb{R}^d$*, and that* v_1 *and* v_2 *are respectively viscosity subsolution and supersolution of* (27)*. Then* $v_1(t,x) \le v_2(t,x), \forall (t,x) \in [0,T]\times\mathbb{R}^d$*.*

4 Error estimate for numerical approximations by a Semi-Lagrangian scheme

4.1 Time semi-discrete scheme

We aim at approximating v, the unique continous viscosity solution, with linear growth, of the following HJB equation:

$$\begin{aligned} -\partial_t v + \mathcal{H}(t,x,Dv,D^2v) &= 0 \qquad \text{in } (0,T)\times\mathbb{R}^d && \text{(29a)}\\ v(T,x) &= \phi(x) \quad \text{in } \mathbb{R}^d. && \text{(29b)}\end{aligned}$$

where $\mathcal{H}$ is the same Hamiltonian function as in (28). This is the same as HJB equation (27) but with a general terminal data ϕ instead of Φ_ϵ. Throughout this section, we assume satisfied the assumption $(H1)$ on the drift and diffusion coefficients b, σ. Also ϕ is assumed to be a Lipschitz continuous function, with Lipschitz constant L_ϕ. The Hölder constant of v will be denoted by L_v.

We aim to give new error estimates for semi-Lagrangian schemes [19], in the case of Lipschitz continuous b and σ yet that can be unbounded (as well as the solution v itself).

For convenience, we will denote by σ_k the column vectors of the matrix σ :

$$\sigma(t,x,a) = \big[\sigma_1, \dots, \sigma_m\big](t,x,a),$$

and let us denote the vectors $(\bar{\sigma}_k)_{k=1,\dots,2m}$ as follows

$$\bar{\sigma}_k(t,x,a) := (-1)^k \sqrt{m}\ \sigma_{\lfloor \frac{k-1}{2} \rfloor}(t,x,a) \tag{30}$$

(where $\lfloor p \rfloor$ denotes the integer part of $p \in \mathbb{R}$).

Let $h = h > 0$ denote a given time step, and consider a semi-discrete scheme defined as (for $x \in \mathbb{R}^d$):

$$V^N(x) = \phi(x) \tag{31a}$$

and, for every $n = N, \dots, 1$,

$$V^{n-1}(x) = \mathcal{S}^h(t_n, x, V^n), \tag{31b}$$

with, for any $t \in [0,T]$, $x \in \mathbb{R}^d$, and any function $w : \mathbb{R}^d \to \mathbb{R}$,

$$\mathcal{S}^h(t,x,w) := \frac{1}{2m} \max_{a \in U} \left\{ \sum_{k=1}^{2m} w(x + hb(t,x,a) + \sqrt{h}\bar{\sigma}_k(t,x,a)) \right\}$$

By V we will denote the linear interpolation of $V^0, \cdots, V^n$ on $t_0, \cdots, t_N$.

The main result of this section is the following:

Theorem 4.1 *Assume that (H1a) is satisfied and that ϕ is Lipschitz continuous function with Lipschitz constant L_ϕ. There exists $C \geq 0$, $\forall n \in [0, \dots, N]$,*

$$|V^n(x) - v(t_n, x)| \leq C L_\phi \,(1 + |x|)^{7/4}\, h^{1/4}.$$

The above theorem is an extension to the error estimates known in the literature for bounded Hölder continuous value functions with bounded and Lipschitz continuous drift b and diffusion σ, see [6, 8, 21]. The proof given here is based on classical shaking and regularization techniques introduced by Krylov [27, 28] combined with a precise consistency estimate and an interpretation of the numerical scheme as value function of a discrete-time control problem.

Remark 4.1 *More precisely if we assume, for some constant* $L_{0,0}, L_{0,1} \geq 0$*:*

$$|b(t,x,a)| + |\sigma(t,x,a)| \leq L_{0,0} + L_{0,1}|x|, \quad |\phi(x)| \leq L_{0,0} + L_{0,1}|x|, \tag{32}$$

for every $(t,x,a) \in [0,T] \times \mathbb{R}^d \times U$*, then there exists a constant* $C \geq 0$*, such that:*

$$|V^n(x) - v(t_n,x)| \leq C L_\phi \left(1 + L_{0,0} + L_{0,1}|x|\right)^{7/4} h^{1/4}.$$

In particular if ϕ, b *and* σ *are bounded functions then the previous estimates hold with* $L_{0,1} := 0$ *and we find the usual error estimate bounded by* $h^{1/4}$ *up to a universal constant (i.e. no growth term in* $|x|^{7/4}$*).*

4.1.1 Properties of (31)

First, we derive the following consistency property:

Lemma 4.2 *For any regular function* $\varphi \in C^{2,4}([0,T] \times \mathbb{R}^d)$*, denoting* $\varphi^n(x) = \varphi(t_n,x)$ *and* $\mathcal{E}_\varphi^n(x)$ *as*

$$\mathcal{E}_\varphi^n(x) := -\partial_t \varphi(t_n,x) + \mathcal{H}(t_n,x,D\varphi,D^2\varphi) - \frac{\varphi^{n-1}(x) - \mathcal{S}^h(t_n,x,\varphi^n)}{h} \tag{33}$$

where $\mathcal{S}^h$ *is defined in* (31)*, it holds*

$$|\mathcal{E}_\varphi^n(x)| \leq C\,(\|\varphi_{tt}\|_0 + \sum_{k=2,3,4} \|D^k\varphi\|_0)\ \sup_{a\in U}\left(|b(t_n,x,a)|^2 + |\sigma(t_n,x,a)|^4\right)\ h,$$

where $C \geq 0$ *is a constant independent of* n*,* h *and* φ*.*

Proof. The result is straightforward by first using a Taylor expansion of fourth order of $\varphi(t-h, y+\sqrt{h}\bar\sigma_k(t,x,a))$ around $\varphi(t,y)$, where $y = x + b(t,x,a)h$, and then by using a second order Taylor expansion of the result around x. □

In particular, by using the Lipschitz regularity of b and σ, and their linear growthness, it also holds that

$$|\mathcal{E}_\varphi^n(x)| \leq C\,(\|\varphi_{tt}\|_0 + \sum_{k=2,3,4} \|D^k\varphi\|_0)\ (1+|x|^4)\ h. \tag{34}$$

Remark 4.3 *More precisely if we assume* (32) *for some constant* $L_{0,0}, L_{0,1} \geq 0$*, then it also holds, for a constant* $C \geq 0$*:*

$$|\mathcal{E}_\varphi^n(x)| \leq C\,(\|\varphi_{tt}\|_0 + \sum_{k=2,3,4} \|D^k\varphi\|_0)\ (1 + L_{0,0} + L_{0,1}|x|)^4\ h. \tag{35}$$

Now, by considering $Q \in \{1,\ldots,2m\}$ a random variable such that $\mathbb{P}[Q=k] = \frac{1}{2m}$, it follows that the scheme (31) is equivalent to:

$$V^{n-1}(x) = \max_{a\in U} \mathbb{E}\Big[V^n(x + hb(t_n,x,a) + \sqrt{h}\,\bar\sigma_Q(t_n,x,a))\Big].$$

For the sequel of the section, it will be useful to define recursively the Markov chain $Z_{n,x}^{k,a}$ as follows. For a given $x \in \mathbb{R}^d$, a given $k \geq n \geq 0$, a sequence of controls $a = (a_n, \ldots, a_k, \ldots)$ with $a_i \in U$, and a sequence $(Q_n, Q_{n+1}, \ldots, Q_k, \ldots)$ of i.i.d. random variables with same law as Q,

- If $k = n$,

$$Z_{n,x}^{n,a} := x.$$

- If $k \geq n$,

$$Z_{n,x}^{k+1,a} := Z_{n,x}^{k,a} + hb(t_k, Z_{n,x}^{k,a}, a_k) + \sqrt{h}\bar{\sigma}_{Q_k}(t_k, Z_{n,x}^{k,a}, a_k).$$

Clearly, $Z_{n,x}^{k,a}$ will depend only of n, x, the first $k - n$ values of $(a_n, \ldots, a_{k-1})$ and of $Q_n, \ldots, Q_{k-1}$.

The scheme can then be written equivalently in the form

$$V^{n-1}(x) = \max_{a \in U} \mathbb{E}\Big[V^n(Z_{n,x}^{n+1,a_0})\Big]. \tag{36}$$

By standard estimates, there exists a constant $C > 0$ (depending on L_0, d, m and T) such that for all $x, y \in \mathbb{R}^d$, for all $0 \leq n \leq k \leq N$ and $0 \leq m \leq N - k$, we have:

$$\max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_{k-1}} \mathbb{E}\big[|Z_{n,x}^{k,a}|^4\big] \cdots\Big]\Big] \leq C(1 + |x|^4), \tag{37a}$$

$$\max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_{k-1}} \mathbb{E}\big[|Z_{n,x}^{k,a} - Z_{n,y}^{k,a}|\big] \cdots\Big]\Big] \leq C|x - y|, \tag{37b}$$

$$\max_{a_n,\ldots,a_k} \mathbb{E}\Big[\max_{a_{k+1}} \mathbb{E}\Big[\cdots \max_{a_{k+m-1}} \mathbb{E}\big[|Z_{n,x}^{k+m,a} - Z_{n,x}^{k,a}|\big] \cdots\Big]\Big] \leq C(1 + |x|)(t_{k+m} - t_k)^{\frac{1}{2}}. \tag{37c}$$

For sake of completeness, a proof of the above estimates is given in Appendix C. Finally, we recall that the scheme is $\frac{1}{2}$-hölder in time and Lispchitz continuous in space:

Lemma 4.4 *There exists $C > 0$ (independent of h), for $0 \leq n \leq n + k \leq N$:*

$$|V^{n+k}(x) - V^n(y)| \leq CL_\phi\Big((1 + |x|)(t_{n+k} - t_n)^{\frac{1}{2}} \ + \ |x - y|\Big),$$

for $x, y \in \mathbb{R}^d$.

Proof. By recursion we have

$$\begin{aligned} V^n(x) &= \max_{a_{n+1}} \mathbb{E}\Big[V^{n+1}(Z_{n+1,x}^{n+2,a_{n+1}})\Big] \\ &= \max_{a_{n+1}} \mathbb{E}\Big[\max_{a_{n+2}} \mathbb{E}\Big[\cdots \max_{a_{n+k}} \mathbb{E}\big[V^{n+k}(Z_{n+1,x}^{n+k+1,a})\big] \cdots\Big]\Big] \end{aligned}$$

with $a = (a_{n+1}, a_{n+2}, \ldots, a_{n+k})$. In particular, with $k = N - n$ and knowing that $V^N(x) = \phi(x)$, it follows that:

$$V^n(x) = \max_{a_{n+1}} \mathbb{E}\Big[\max_{a_{n+2}} \mathbb{E}\Big[\cdots \max_{a_N} \mathbb{E}\big[\phi(Z_{n,x}^{N,a})\big] \cdots\Big]\Big].$$

By using (37b), we deduce that:

$$|V^n(x) - V^n(y)| \leq L_\phi C|x-y|. \tag{38}$$

Similarly, by using (37c) with $k := n$ and $m := k$, we obtain:

$$\begin{aligned}
&|V^n(x) - V^{n+k}(x)| \\
&\quad = \left| \max_{a_{n+1}} \mathbb{E}\Big[\max_{a_{n+2}} \mathbb{E}\Big[\cdots \max_{a_{n+k}} \mathbb{E}\big[V^{n+k}(Z^{n+k+1,a}_{n+1,x}) - V^{n+k}(x) \big] \cdots \Big]\Big] \right| \\
&\quad \leq L_\phi C \max_{a_{n+1}} \mathbb{E}\Big[\max_{a_{n+2}} \mathbb{E}\Big[\cdots \max_{a_{n+k}} \mathbb{E}\big[|Z^{n+k+1,a}_{n+1,x} - x| \big] \cdots \Big]\Big] \\
&\quad \leq L_\phi C^2 (1+|x|)(t_{n+k} - t_n)^{\frac{1}{2}} \qquad (39)
\end{aligned}$$

By combining the inequalities (38) and (39), the desired statement is proved. □

4.1.2 Upper bound

First, we consider a regular super-solution of (29), denoted w, and aim derive an upper bound for

$$e^n := V^n - w^n, \tag{40}$$

where $w^n(x) = w(t_n, x)$.

Lemma 4.5 *Let w be a regular super-solution of (29). For all $0 \leq n \leq N-1$ and every $x \in \mathbb{R}^d$, we have:*

$$\begin{aligned}
e^n(x) \leq &\max_{a_{n+1}} \mathbb{E}\Big[\max_{a_{n+2}} \mathbb{E}\Big[\cdots \max_{a_N} \mathbb{E}\Big[e^N(Z^{N+1,a}_{n+1,x}) \Big] \ldots \Big]\Big] \\
&+ h \sum_{n+1 \leq k \leq N} \max_{a_{n+1}} \mathbb{E}\Big[\max_{a_{n+2}} \mathbb{E}\Big[\cdots \max_{a_{k-1}} \mathbb{E}\Big[\mathcal{E}^k_w(Z^{k,a}_{n+1,x}) \Big] \cdots \Big]\Big].
\end{aligned}$$

Proof. By definition of the scheme,

$$V^{n-1}(x) = \mathcal{S}^h(t_n, x, V^n) = \max_{a_n \in U} \mathbb{E}\Big[V^n(Z^{n+1,a_n}_{n,x}) \Big],$$

and by the consistency estimate of Lemma 4.2 and the super-solution property, it comes that

$$w^{n-1}(x) \geq \mathcal{S}^h(t_n, x, w^n) - h\mathcal{E}^n_w(x) \;=\; \max_{a_n \in U} \mathbb{E}\Big[w^n(Z^{n+1,a_n}_{n,x}) \Big] - h\mathcal{E}^n_w(x).$$

Therefore, for $e_n = V^n - w^n$ we get the estimate

$$e^{n-1}(x) \leq \max_{a_n \in U} \mathbb{E}\Big[e^n(Z^{n+1,a_n}_{n,x}) \Big] + h\mathcal{E}^n_w(x).$$

Hence, by recursion,

$$\begin{aligned} e^{n-1}(x) &\leq \max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[e^{n+1}(Z^{n+2,a_{n+1}}_{n+1,Z^{n+1,a_n}_{n,x}}) \Big] + h\mathcal{E}^{n+1}_w(Z^{n+1,a_n}_{n,x}) \Big] + h\mathcal{E}^n_w(x) \\ &\leq \max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[e^{n+1}(Z^{n+2,a}_{n,x}) \Big]\Big] + h \max_{a_n} \mathbb{E}\Big[\mathcal{E}^{n+1}_w(Z^{n+1,a_n}_{n,x}) \Big] + h\mathcal{E}^n_w(x) \\ &\;\;\vdots \\ &\leq \max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_{n+k}} \mathbb{E}\Big[e^{n+k}(Z^{n+k+1,a}_{n,x}) \Big] \cdots \Big]\Big] \\ &\qquad\qquad + h \sum_{0\leq j\leq k} \max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_{n+j-1}} \mathbb{E}\Big[\mathcal{E}^{n+j}_w(Z^{n+j,a}_{n,x}) \Big] \cdots \Big]\Big] \end{aligned}$$

(where we have denoted $\max\limits_{a_n\in U} \equiv \max\limits_{a_n}$, and the term $j=0$ in the sum corresponds only to $\mathbb{E}[\mathcal{E}^n_w(x)]$). We finally obtain:

$$\begin{aligned} e^{n-1}(x) \leq \max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_N} \mathbb{E}\Big[e^{N}(Z^{N+1,a}_{n,x}) \Big] \cdots \Big]\Big] \\ + h \sum_{n\leq k\leq N} \max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_{k-1}} \mathbb{E}\Big[\mathcal{E}^{k}_w(Z^{k,a}_{n,x}) \Big] \cdots \Big]\Big] \end{aligned}$$

The desired result follows by changing n into $n+1$. □

The previous result holds for any smooth function w that is super-solution of (29). The viscosity solution v is just Hölder continuous. However it is possible to construct a regular function $w \equiv v_\eta$, *close* to v, and which is a classical sub-solution of (29). More precisely, by using the shaking coefficients techniques introduced in [26] combined with a standard regularization by mollification, we have the following result.

Lemma 4.6 *Under assumtion (H1a), for every $\eta > 0$ there exists a C^∞ function v_η such that v_η is a classical super-solution to (29). Moreover, there exists $C > 0$ such that for every $\eta > 0$ the following estimates hold:*

$$|v(t,x) - v_\eta(t,x)| \leq CL_\phi(1+|x|)\eta \tag{41a}$$

$$\Big|\frac{\partial^k v_\eta}{dt^k}(t,x)\Big| \leq \frac{CL_\phi}{\eta^{2k-1}}(1+|x|) \quad \textit{and} \quad \Big\|\frac{\partial^k v_\eta}{dx^k}\Big\|_0 \leq \frac{CL_\phi}{\eta^{k-1}} \tag{41b}$$

for any $k \geq 1$, and for every $(t,x) \in [0,T]\times\mathbb{R}^d$.

The proof of this result can be found in [26] under the additional assumption that b, σ and ϕ are bounded functions. However the arguments used in [26] can be easily extended to the case when (H1a) is satisfied and ϕ is a Lipschitz function (not necessarily bounded). For convenience of the reader, the outline of the proof is given in Appendix A.

Now, we have all the ingredients to conclude the upper bound:

Proof of theorem 4.1: upper bound of $V^n - v(t_n,\cdot)$.

Let $\eta > 0$. By using Lemma 4.6, there exists $C \geq 0$ such that for every $x \in \mathbb{R}^d$, it holds:

$$\mathcal{E}^n_{v_\eta}(x) \leq CL_\phi(1+|x|^4)\ \frac{h}{\eta^3},$$
$$\max_{a_n} \mathbb{E}\Big[\max_{a_{n+1}} \mathbb{E}\Big[\cdots \max_{a_{N-1}} \mathbb{E}\Big[|V^N(Z^{N,a}_{n,x}) - v_\eta(T, Z^{N,a}_{n,x})\Big]\cdots\Big]\Big] \leq CL_\phi(1+|x|)\eta.$$

By applying the result of Lemma 4.5 with $w = v_\eta$, and taking into account estimates (37), we obtain:

$$V^n(x) - v_\eta(t_n, x) \leq CL_\phi(1+|x|)\eta + TCL_\phi(1+|x|^4)\ \frac{h}{\eta^3}.$$

Therefore for $|x| \leq R$, we can choose for η an optimal value of order $\eta \equiv (R^3 h)^{1/4}$ to derive the "upper" bound:

$$\|(V^n - v(t_n, \cdot))_+\|_{L^\infty(\mathbb{B}_R)} \leq CL_\phi R^{7/4} h^{1/4}$$

for any $n = 0, \cdots, N$, with C independant of h, ϕ and R. □

4.1.3 Lower bound

Now, we aim at deriving the lower bound estimate for the semi-discrete scheme (31). For this, we will apply exactly the same techniques as used for the upper bound, reversing the role of the equation and the scheme. The key point is that the solution V of the semi-discrete scheme is also Hölder continuous. We first build a function V^η by considering a scheme with shaking coefficients :

$$V^\eta(t,x) = \max_{\substack{-\eta^2 \leq e_1 \leq 0 \\ |e_2| \leq \eta \\ a \in U}} \mathbb{E}\Big[V^\eta(t+h, x + (hb + \sqrt{h}\bar{\sigma})(t+e_1, x+e_2, a))\Big] \quad \text{in } [-2\eta^2, T) \times \mathbb{R}^d, \tag{42a}$$

$$V^\eta(T,x) = \phi(x) \quad \text{in } \mathbb{R}^d, \tag{42b}$$

(where σ, b are extended in time interval $[-2\eta^2, T]$ in such way (H1a) is still valid).

We define by convolution $V_\eta := V^\eta * \rho_\eta$ where ρ_η is a sequence of mollifiers defined by $\rho_\eta(t,x) := \frac{1}{\eta^{d+2}}\rho(\frac{t}{\eta^2}, \frac{x}{\eta})$ and with ρ such that $\{\rho_\eta\}_\eta$ is the sequence of mollifiers defined by

$$\rho_\eta = \frac{1}{\eta^{d+2}}\rho(\frac{t}{\eta^2}, \frac{x}{\eta})$$
$$\rho \in C^\infty(\mathbb{R}^{d+1}),\ \rho \geq 0, \quad \text{supp}\ \rho \subset [0,1] \times B_1, \quad \int_{\mathbb{R}}\int_{\mathbb{R}^d} \rho(s,x)\, dx ds = 1. \tag{43}$$

Then, by sing the same arguments as in Appendix A, we get the following Lemma.

Lemma 4.7 *Under assumtion (H1a), for every $\eta > 0$, V_η is a C^∞ function such that*

$$V_\eta(t,x) - \mathcal{S}^h(t+h, x, V_\eta(t+h,\cdot)) \geq 0 \qquad \forall (t,x) \in [0, T-h] \times \mathbb{R}^d.$$

Moreover, there exists $C > 0$ depending on T and L_0 such that for every $\eta > 0$ the following estimates hold for every $t \in [0,T]$ and every $x \in \mathbb{R}^d$:

$$|V(t,x) - V_\eta(t,x)| \leq CL_\phi \ (1+|x|)\eta \tag{44a}$$

$$|\frac{\partial^k V_\eta}{dt^k}(t,x)| \leq \frac{CL_\phi}{\eta^{2k-1}}(1+|x|) \quad \textit{and} \quad \|\frac{\partial^k V_\eta}{dx^k}\|_0 \leq \frac{CL_\phi}{\eta^{k-1}} \tag{44b}$$

for any $k \geq 1$.

By straithforward calculations, one can check that:

$$\mathcal{E}^n_{V_\eta}(x) \geq -\frac{CL_\phi}{\eta^3}(1+|x|^4)\, h. \tag{45}$$

By using the consistency estimate of Lemma 4.2, and the Hölder estimates on V_η (that can be inferred from the one of the scheme), we then deduce:

$$\begin{aligned} -\partial_t V_\eta + \mathcal{H}(t_n, x, DV_\eta, D^2V_\eta) &= \frac{V_\eta^{n-1}(x) - \mathcal{S}^h(t_n, x, V_\eta^n)}{h} + \mathcal{E}^n_{V_\eta}(x) \\ &\geq \mathcal{E}^n_{V_\eta}(x) \\ &\geq -\frac{CL_\phi}{\eta^3}(1+|x|^4)\, h \end{aligned}$$

(for some constant $C \geq 0$). In the same way we establish the same estimate for any $t \in [t_{n-1}, t_n]$. Hence

$$-\partial_t V_\eta + \mathcal{H}(t, x, DV_\eta, D^2V_\eta) \geq -\frac{CL_\phi}{\eta^3}(1+|x|^4)\, h, \quad \forall (t,x) \in [0,T] \times \mathbb{R}^d. \tag{46}$$

Let ζ be the following function:

$$\zeta(t,x) := \bar{C} L_\phi e^{\lambda(T-t)} \Big((1+|x|^4)\frac{h}{\eta^3} + \sqrt{1+|x|^2}\eta \Big)$$

where $\lambda > 0$, $\bar{C} > 0$ will be fixed later on. The definitions of $\mathcal{H}$ and of ζ, and the linear growth of b and σ with respect to $|x|$, yield the following bounds:

$$\begin{aligned} -\partial_t \zeta \ = \ \lambda \zeta &\geq \lambda \bar{C} e^{\lambda(T-t)} \Big((1+|x|^4)\frac{h}{\eta^3} + |x|\eta \Big) \\ \mathcal{H}(t, x, D\zeta, D^2\zeta) &\geq -K\bar{C} e^{\lambda(T-t)} \big(|x|^4 \frac{h}{\eta^3} + |x|\eta \big) \end{aligned}$$

for some constant $K \geq 0$ tht depends on L_0. In particular, choosing $\lambda := K+1$, it holds

$$-\partial_t \zeta + \mathcal{H}(t, x, D\zeta, D^2\zeta) \ \geq \ \bar{C}\Big((1+|x|^4)\frac{h}{\eta^3} + \sqrt{1+|x|^2}\eta \Big). \tag{47}$$

Combining (46) and (47), it comes for any $t \in [0,T]$ and $x \in \mathbb{R}^d$:

$$\begin{aligned}
-\partial_t(V_\eta+\zeta) &+ \mathcal{H}(t,x,D(V_\eta+\zeta),D^2(V_\eta+\zeta)) \\
&\geq -\partial_t V_\eta + \mathcal{H}(t,x,DV_\eta,D^2V_\eta) - \partial_t\zeta + \mathcal{H}(t,x,D\zeta,D^2\zeta) \\
&\geq -CL_\phi(1+|x|^4)\frac{h}{\eta^3} + \bar{C}\left((1+|x|^4)\frac{h}{\eta^3} + \sqrt{1+|x|^2}\eta\right) \\
&\geq 0, \qquad (48)
\end{aligned}$$

for any $\bar{C} \geq CL_\phi$. On the other hand it also holds

$$\begin{aligned}
V_\eta(T,x)+\zeta(T,x) &\geq \phi(x) - CL_\phi(1+|x|)\eta + \bar{C}\sqrt{1+|x|^2}\eta \\
&\geq 0, \qquad (49)
\end{aligned}$$

for any $\bar{C} \geq \sqrt{2}CL_\phi$. Hence by choosing $\bar{C} := \sqrt{2}CL_\phi$, (48) and (49) hold and $V_\eta+\zeta$ is a viscosity super-solution of (29). Furthermore, $V_\eta+\zeta$ has a quadratic upper bound growth: $V_\eta(t,x)+\zeta(t,x) \leq c_a\ (1+|x|^4)$ for some constant $c_a > 0$. The exact solution v is also a viscosity sub-solution of (29), with a linear growth (so a linear bound from below of the form $v(t,x) \geq -c_b(1+|x|$ for some constant $c_b > 0$). Therefore, according to the comparison principle stated in Lemma 3.8, it comes that $V_\eta+\zeta \geq v$ on $[0,T]\times\mathbb{R}^d$. By consequence,

$$v(t_n,x) - V^n(x) \leq \zeta(t_n,x) \leq \sqrt{2}CL_\phi e^{\lambda T}\big((1+|x|^4)\,\frac{h}{\eta^3} + \sqrt{1+|x|^2}\eta\big).$$

Finally, by choosing η such that $\eta^4 \equiv R^3h$ and for $|x| \leq R$ we obtain the following reverse estimate, for some constant $C \geq 0$:

$$\|(v(t_n,.)-V^n)_+\|_{L^\infty(\mathbb{B}_R)} \leq CL_\phi\, R^{7/4}\, h^{1/4}.$$

This concludes the proof of theorem 4.1.

4.2 Fully discrete scheme

Now, Consider a spatial discretisation of $\mathbb{R}^d$ (which can be assumed uniform for simplicity): for some given mesh steps $\Delta x_i > 0$, $x_i = i\Delta x \equiv (i_1\Delta x_1,\ldots,i_d\Delta x_d)$ with $i \in \mathbb{Z}^d$. We will denote $|\Delta x|$ the Euclidean norm of Δx, $\mathcal{G} := \{i\Delta x,\ i \in \mathbb{Z}^d\}$, and $\mathcal{G}_h := \{t_0,\ldots,t_N\}\times\mathcal{G}$.

Fully discrete scheme: for $n = N,\ldots,1$, for all $x_i \in \mathcal{G}$:

$$V_i^{n-1} = V^{n-1}(x_i) = \frac{1}{2m}\max_{a\in U}\left\{\sum_{k=1}^{2m}[V^n](x_i + h\,b(t_n,x_i,a) + \sqrt{h}\,\bar{\sigma}_k(t_n,x_i,a))\right\}, \qquad (50)$$

where $[V^n]$ denotes the bilinear interpolation of (V_i^n) on (x_i), and with

$$V_i^N = V^N(x_i) = \phi(x_i),\ \forall x_i \in \mathcal{G}. \qquad (51)$$

Theorem 4.2 *Assume (H1a) and assume ϕ is a L_ϕ-Lipschitz continuous function. Let v be the continuous solution of (29) , and let V^Δ the numerical solution satisfying the scheme (50), with $\Delta = (h, \Delta x)$ the time and space steps. There exists $C > 0$ depending only on T, L_0 such that for every $R > 0$, we have:*

$$\|v - V^\Delta\|_{L^\infty(\mathbb{B}_R)} \leq C L_\phi \left(R^{7/4} h^{1/4} + \frac{|\Delta x|}{h} \right).$$

Proof. By theorem 4.1, we have already an error estimate between v and the solution V of the semi-discrete scheme (31). Now, the error between V and V^Δ is a classical result (see [21, 12] for details). $\square$

4.3 Comments

The convergence and error estimates results of this section are still valid for more general schemes in the form of Markov chain approximations as (36) with general probability density (see also [30, 32]). Indeed, all the arguments developed in this section are mainly based on the formulation (36) and do not depend on the probability distribution of the random variable Q neither on the formulation as semi-Lagrangian scheme (31).

For the numerical simulations in section 6, we will consider the scheme (related to Milstein's approximation):

$$\mathcal{S}^h(t, x, w) := \frac{1}{2m} \max_{a \in U} \left\{ \sum_{k=1}^{2m} w(x + h\widetilde{b}(t, x, a) + \sqrt{h}\bar{\sigma}_k(t, x, a)) \right\} \tag{52}$$

where $\widetilde{b}(t, x, a) := \frac{1}{2}(b(t, x, a) + b(t, x + hb(t, x, a), a))$. This scheme is a little bit more precise for the approximation of the deterministic part of the processes when $b(\cdot, \cdot, a)$ is non-constant. The error estimates of the present section can be easily extended to such an approximation.

5 Application: Probabilistic reachability analysis

Let $\mathcal{C}$ be a nonempty subset of $\mathbb{R}^d$ with non-zero measure ("the target"). Let $\rho \in [0, 1[$ and $t \leq T$. Consider Ω_t^ρ the backward reachable set under probability of success ρ, that is, the set of initial points x for which the probability that there exists trajectory $X_{t,x}^u$ solution of (4), associated with an admissible control $u \in \mathcal{U}$ and that reaches $\mathcal{C}$ at time T is at least ρ:

$$\Omega_t^\rho = \left\{ x \in \mathbb{R}^d \,\middle|\, \exists u \in \mathcal{U},\ \mathbb{P}[X_{t,x}^u(T) \in \mathcal{C}] > \rho \right\}. \tag{53}$$

Such backward reachable sets play an important role in many applications. For instance the set Ω_t^ρ can be interpreted as a "safety region" for reaching $\mathcal{C}$, with confidence ρ.

For time discrete stochastic systems, stochastic backward reachable sets of the form of (53) have been analysed and characterized via an adequate stochastic optimal control problem in [1] and [2]. In this case, the control problem is solved via the dynamic programming approach.

In the context of financial mathematics, the problem of characterizing the backward reachable set with a given probability was first introduced by Föllmer and Leukert [24]. This problem was also studied and converted into the class of stochastic target problems by Touzi, Bouchard and Elie in [15].

In order to characterize the domain Ω_t^ρ for different values of ρ, we consider the level-set approach and introduce the following optimal control problem:

$$\vartheta(t,x) := \sup_{u\in\mathcal{U}} \mathbb{E}\big[\mathbb{1}_{\mathcal{C}}(X^u_{t,x}(T))\big] \;\equiv\; \sup_{u\in\mathcal{U}} \mathbb{P}\big[X^u_{t,x}(T)\in\mathcal{C}\big]. \tag{54}$$

Therefore, it is straightforward to show the following:

Proposition 5.1 *Assume (H1a), and let ϑ defined in (54). Then, $\forall t\in[0,T]$:*

$$\Omega_t^\rho = \{x\in\mathbb{R}^d,\ \vartheta(t,x) > \rho\}. \tag{55}$$

Following the results of section 3, we first regularize the function $\mathbb{1}_{\mathcal{C}}(\cdot)$ by functions Φ^ϵ (for $\epsilon>0$), defined as follows:

$$\Phi^\epsilon(x) = \min(1, \max(0, -\frac{1}{\epsilon}\mathrm{dist}(\mathrm{x},\mathcal{C}))). \tag{56}$$

Notice that the Φ^ϵ is $\frac{1}{\epsilon}$-Lipschitz continuous (see Figure 1).

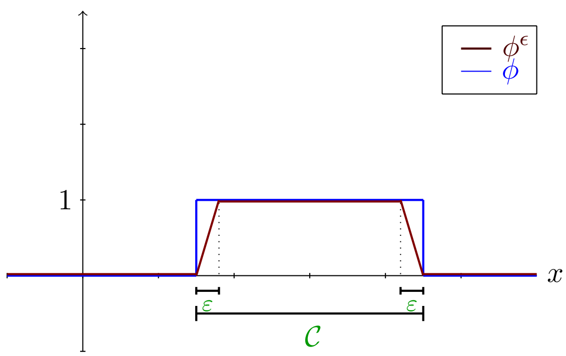

Figure 1: Regularization Φ^ϵ of the indicator function $\mathbb{1}_{\mathcal{C}}$ for a given set $\mathcal{C}$.

Then, we consider the following "regularized" control problem:

$$\vartheta^\epsilon(t,x) := \sup_{u\in\mathcal{U}} \mathbb{E}[\Phi^\epsilon(X^u_{t,x}(T))],$$

and we denote by $\vartheta^{\epsilon,\Delta}$ a numerical approximation of ϑ^ϵ obtained by solving the fully discretized scheme (50). In order to obtain an error estimates of $\vartheta-\vartheta^{\epsilon,\Delta}$, we shall need to assume the following hypothesis on the target set $\mathcal{C}$:

(H4) $\mathcal{C}$ is a non-empty Borelean subset of $\mathbb{R}^d$. Moreover, if we denote by $\mathcal{C}_\epsilon$ the set defined by:

$$\mathcal{C}_\epsilon := \{x\in\mathcal{C},\ \mathrm{dist}(\mathrm{x},\partial\mathcal{C}) \le -\epsilon\},$$

then there exists a constant $M_1>0$ such that, for every $A>0$,

$$\mu((\mathcal{C}\setminus\mathcal{C}_\epsilon)\cap\mathbb{B}_A) \le M_1 A\,\epsilon.$$

Remark 5.2 The above assumption is satisfied in many cases, for example when $\mathcal{C}$ is a half space or when $\mathcal{C}$ is a finite union of bounded, convex polytopes $\mathcal{O}_i \subset \mathbb{R}^d$ with non empty interiors:

$$\mathcal{C} := \bigcup_{i=1,\dots,p} \mathcal{O}_i.$$

Theorem 5.3 *Assume that (H1a), (H1b) and (H4) are satisfied.*
(i) There exists $C > 0$ such that for every $\epsilon > 0$ and every $\Delta = (h, \Delta x)$ mesh steps, for every $\forall t \in [0, T)$, $\forall x \in \mathbb{R}^d$, the following holds:

$$|\vartheta^{\epsilon,\Delta}(t,x) - \vartheta(t,x)| \leq C\frac{1 + |x|^2 + |\log \epsilon|}{(T-t)^{\frac{d}{2}}}\epsilon + \frac{C}{\epsilon}\left(|x|^{\frac{7}{4}}h^{\frac{1}{4}} + \frac{\Delta x}{h}\right) \tag{57}$$

(ii) If there exists $A > 0$ such that $\mathcal{C} \setminus \mathcal{C}_\epsilon \subset \mathbb{B}_A$ for every $\epsilon \in]0,1]$, then there exists $C > 0$ such that

$$|\vartheta^{\epsilon,\Delta}(t,x) - \vartheta(t,x)| \leq C\frac{1}{(T-t)^{\frac{d}{2}}}e^{-C_2\frac{\mathrm{dist}(x,\mathcal{C}\setminus\mathcal{C}_\epsilon)^2}{T-t}}\epsilon + \frac{C}{\epsilon}\left(|x|^{\frac{7}{4}}h^{\frac{1}{4}} + \frac{\Delta x}{h}\right) \tag{58}$$

for every $\forall t \in [0, T)$, $\forall x \in \mathbb{R}^d$, $\forall \epsilon \in]0, 1]$.

Proof. Under assumption (H4), all the requirements of assumptions (H2)-(H3) are satisfied for $\Phi = \mathbb{1}_{\mathcal{C}}$ and the regularized function Φ_ϵ defined in (56). Thus the result of theorems 3.1 and 5.3 can be applied and lead to the result. □

To get the optimal rates in (57) and (58), one can choose ϵ, h and Δx in such way to minimize the error in the right hand side of the estimates. For instance in case there exists $A > 0$ such that $\mathcal{C} \setminus \mathcal{C}_\epsilon \subset \mathbb{B}_A$, and for every $0 \leq T - \delta$ (with $\delta > 0$), and for every $x \in \mathbb{B}_R$ with $R > 1$, the error estimate in (58) becomes:

$$|\vartheta^{\epsilon,\Delta}(t,x) - \vartheta(t,x)| \leq C\frac{\epsilon}{\delta^{\frac{d}{2}}} + \frac{C}{\epsilon}R^{\frac{7}{4}}\left(h^{\frac{1}{4}} + \frac{\Delta x}{h}\right).$$

The optimal estimate is then obtained by choosing ϵ, $\frac{h^{1/4}}{\epsilon}$ and $\frac{1}{\epsilon}\frac{\Delta x}{h}$ to be of the same order. This leads to $\epsilon \sim h^{1/8} \sim \Delta x^{1/10}$ and to the following estimate:

$$|\vartheta^{\epsilon,\Delta}(t,x) - \vartheta(t,x)| \leq C\frac{R^{\frac{7}{4}}}{\delta^{\frac{d}{4}}}\Delta x^{1/10}. \tag{59}$$

Therefore, we obtain the following approximation of Ω_t^ρ, for $0 \leq t \leq T - \delta$:

$$\left\{x,\ \vartheta^{\epsilon,\Delta}(x,t) > \rho + C\frac{R^{\frac{7}{4}}}{\delta^{\frac{d}{4}}}\Delta x^{\frac{1}{10}}\right\} \subset \Omega_t^\rho \cap \mathbb{B}_R. \tag{60a}$$

and

$$\Omega_t^\rho \cap \mathbb{B}_R \subset \left\{x,\ \vartheta^{\epsilon,\Delta}(x,t) > \rho - C\frac{R^{\frac{7}{4}}}{\delta^{\frac{d}{4}}}\Delta x^{\frac{1}{10}}\right\}, \tag{60b}$$

Hence we can approximate the region Ω_t^ρ by levels sets of the numerical approximation of $\vartheta^{\epsilon,\Delta}$.

The above approximation, of order $O(\Delta x^{1/10})$, is rough. However, in practice, we have observed numerically that it is sufficient to take $h \equiv \Delta x$ and, in that case, the error behaves like $O(\Delta x)$, so the errors in (59) or (60) are also of the order of Δx (see section 6 for more details).

To conclude, we have given a simple numerical approximation procedure for the characterisation of probabilistic backward reachable sets and how to control rigorously the error made in the approximation.

6 Numerical simulations

In all this section the numerical scheme considered is the fully discrete Semi-Lagrangian scheme (52) where the maximization operation is performed on a subset of control values $\{a_1, \ldots, a_{N_u}\}$ that represents a discretization of U with a mesh size Δu. In all the simulations, the regularization parameter ϵ will be chosen as $\epsilon = \frac{1}{\Delta x}$.

Example 1. We consider the following stochastic differential equation with no drift term and no control:

$$\begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} = \sigma \begin{pmatrix} c & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} \tag{61}$$

where $c = 0.2$ and $\sigma = 0.2$. The time horizon is $T = 1.0$. The target $\mathcal{C}$ is the diamond of summits $(-0.8, 1), (1.2, 1), (0.8, 1), (-1.2, 1)$ (see Figure 2(up-left)). If we consider the initial data

$$\phi(x, y) = \begin{cases} 1 & \text{if } x \in \mathcal{C} \\ 0 & \text{otherwise} \end{cases} \tag{62}$$

then the exact solution to the HJB equation is known and the level-set function is given by:

$$\vartheta(t, x, y) = v_1(t, x + cy)\, v_1(t, y)$$

where $v_1(t, r) := \frac{1}{\sqrt{2\pi\ \sigma^2\ t}} \int_{-1-r}^{1-r} e^{-\frac{s^2}{2\sigma^2\ t}} ds$.

First, Figure 2(up-right) shows the backward reachable set Ω_t^ρ for $\rho = 0.05$ at time $t = 0$. For this simulation, we have considered the computational domain $D = [-4, 4]^2$ with a uniform grid and zero condition outside the domain D:

$$\vartheta(t, x, y) = 0, \quad \forall t \in [0, T],\ \forall (x, y) \notin D \tag{63}$$

(which amounts to take homogenous Dirichlet boundary condition on ∂D). One can observe a good matching between the numerical front (computed using the scheme approximation), and the exact front (computed by using the exact value function).

In Figure 2(down), we have also plotted different sets corresponding to different level set values (when using $\Delta x = \Delta y = 0.016$). This corresponds to different level of confidence for reaching the target.

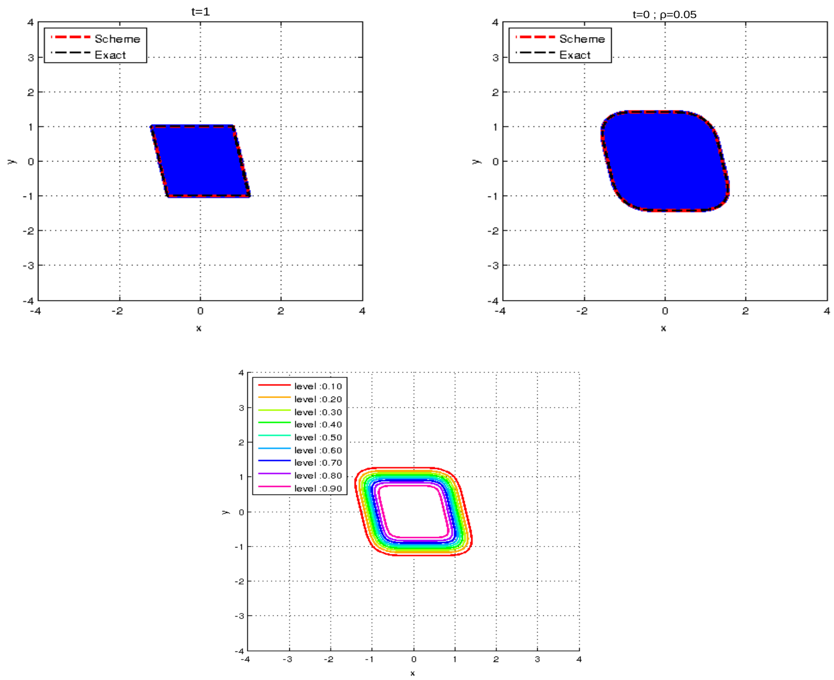

Figure 2: Example 1 with $\underline{\sigma = 0.2}$: Target (up-left); backward reachable set Ω_0^ρ for $\rho = 0.05$ (up-right); backward reachable sets Ω_0^ρ for different values of ρ (down).

Remark 6.1 *The reason behind the use of a diamond as a target set is to validate the behaviour of the numerical scheme for non standard target shape.*

We consider also a case with smaller target, and set $\sigma = 0.5$ (σ is taken large in order to see the impact of the diffusion). Similar simulations as in before are performed in this case and the results are given in Figure 3, where we can observe again a good approximation of the backward reachable sets.

In Table 1, we summarize the error estimates between the exact solution and the numerical approximation, for $\sigma = 0.25$, showing L^∞, L^1 and L^2 errors. The discretization parameters h and Δx are chosen of the same order ($h \equiv \Delta x$), and the numerical simulations show (roughly) a convergence of first order.

Table 1: Example 1 with $\sigma = 0.25$: Error estimates $\vartheta - V$ at time $t = 0$, using $h = T/N$ and $\Delta x = h$.

N	error L^∞	error L^1	error L^2	CPU time (s)
20	3.31 e-2	1.02 e-1	3.88 e-2	3.33 $\times 10^{-1}$
40	1.56 e-2	4.65 e-2	1.82 e-2	2.43 $\times 10^{0}$
80	6.97 e-3	1.99 e-2	8.01 e-3	2.05 $\times 10^{1}$
160	3.78 e-3	1.14 e-2	4.34 e-3	1.55 $\times 10^{2}$
320	2.01 e-3	6.13 e-3	2.36 e-3	1.27 $\times 10^{3}$

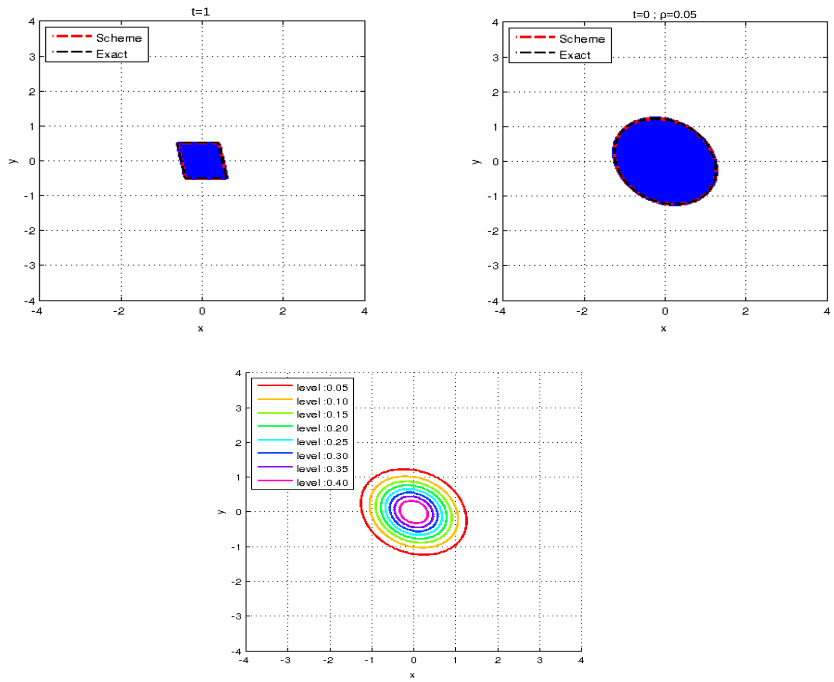

Figure 3: Example 1 with $\underline{\sigma = 0.5}$: the target set (up-left); backward reachable set Ω_0^ρ for $\rho = 0.05$ (up-right); backward reachable sets for different values of ρ.

Remark 6.2 *In this simple case, we can observe numerically that the error estimate decreases with order 1 which is better than what one can prove theoretically (order of* $1/4$*). Also, the choice* $h = \Delta x$ *seems to give better numerical approximations than what we would get if we choose the optimal ratio between* h *and* Δx *established in section 6. The first-order numerical behavior can be justified in this example by the fact that the exact solution is very smooth on* $[0, T) \times \mathbb{R}^2$.

Example 2. Now we deal with the following controlled stochastic system:

$$\begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} = u\sigma \begin{pmatrix} c & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} \tag{64}$$

where u is a control taking values in $[0, 1]$, $c = 0.2$, $\sigma = 0.25$ and $T = 1.0$. The initial data and the boundary conditions are the same as the ones used in the first example. For this kind of problem the exact solution is not known. The solution obtained with $N = 160$ is taken as the reference solution. The error estimates computed at time $t = 0$ are summarized in Table 2. As for Example 1, we observe again a convergence of order 1.

Example 3. In this example, we consider a controlled stochastic system with a drift:

$$dx(t) = \begin{pmatrix} -1 & -4 \\ 4 & -1 \end{pmatrix} x(t)dt + u(t)dt + \begin{pmatrix} 0.7 & 0 \\ 0.7 & 0.7 \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} \tag{65}$$

where $u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}$ and $u_i \in [-0.1, 0.1]$, for $i = 1, 2$.

Table 2: (Example 2) Error table, using $\Delta x = h = T/N$.

N	error L^∞	error L^1	error L^2	CPU time (s)
10	1.68 e-1	1.16 e-3	8.41 e-3	6.42×10^{-1}
20	8.80 e-2	7.49 e-4	4.34 e-3	5.06×10^{0}
40	4.43 e-2	2.55 e-4	1.79 e-3	4.01×10^{1}
80	1.62 e-2	1.12 e-4	7.13 e-4	3.20×10^{2}

The linear system (65) is used in [4] to illustrate an approximation of the probability of reaching a target by using enclosing hulls of probability density functions. Here, we set $T = 1.75$ and consider a target set represented by the green small square in Figure 4. We compute for different times $t \in \{0.75; 0.25; 0\}$, the set Ω_t^ρ for $\rho = 0.4$, see Figure 4. The numerical simulation is performed on a computational domain $D = [-8, 8]^2$ with a uniform grid and boundary conditions as (63).

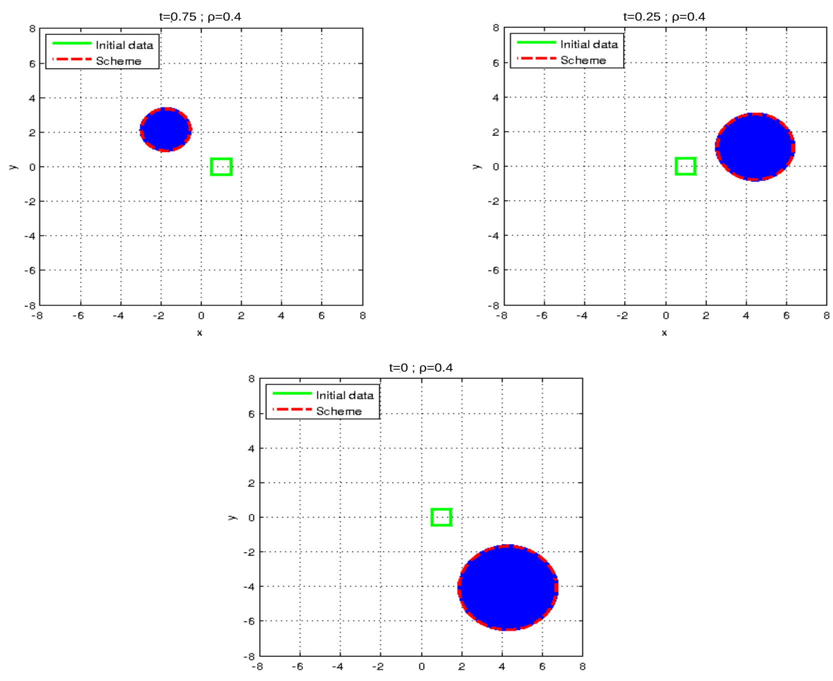

Figure 4: (Example 3) Reachable sets at different times $t \in \{0.75,\ 0.25,\ 0\}$ for a time horizon $T = 1.75$. The target set is represented by the green square.

Once the numerical approximation V of the value function and the backward reachable set $\Omega_{t_n}^\rho$ are computed, and in order to validate the numerical simulations, we generate different trajectories starting from the backward reachable set using the algorithm described below. Let $\bar{x}$ a given initial position, the following algorithm aims to reconstruct a trajectory on starting at time t_n from the position $\bar{x}$:

<u>**Algorithm (trajectory reconstruction)**</u> Initialization: Set $X_n = \bar{x}$.

For $k = n$ to $N-1$:

Step 1 Compute optimal control at $t = t_k$:

$$u_k = \underset{a \in \{a_1, \dots, a_{N_u}\}}{\text{argmax}} \ \mathbb{E}[V(X^a_{k+1}, t_{k+1})].$$

Step 2 Compute the next point at iteration k:

$$X_{k+1} := X_k + b(t_k, X_k, u_k)dt + \sigma(t_k, X_k, u_k)\sqrt{h}B \tag{66}$$

where B is a random variable with the normal law $\mathcal{N}(0, 1)$.

Figure 5 shows some controlled process issued from a starting point located in the backward reachable sets Ω^ρ_t for $t \in \{0.75, 0.25, 0\}$.

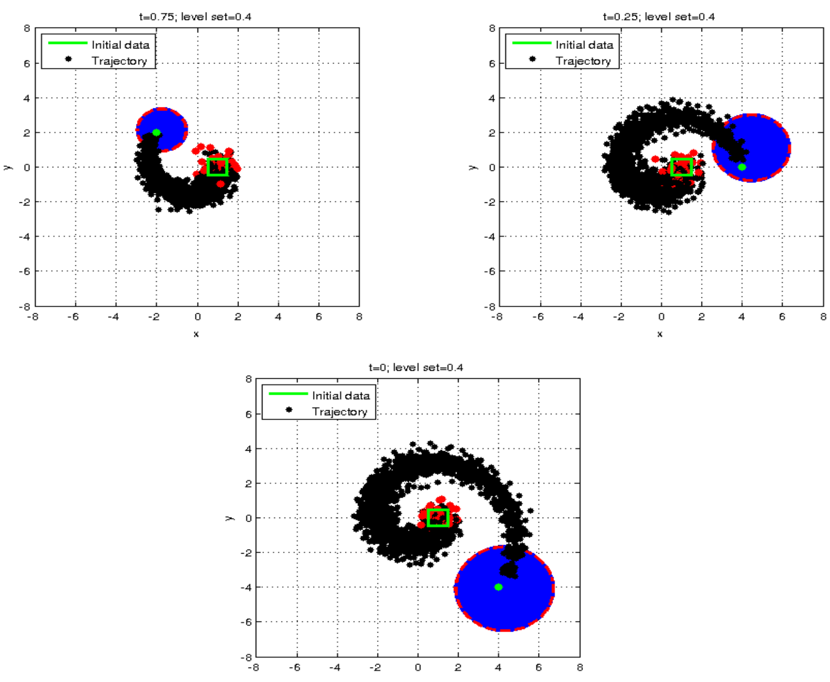

Figure 5: Example 3: Behaviour of controlled processes starting from the backward reachable sets at times $t \in \{0.75,\ 0.25,\ 0\}$ for a final time horizon $T = 1.75$.

Now consider $\bar{x} := (-1.0, 2.0)^\mathsf{T}$ and set $t_n = 0.75$. We compute an approximation $V(t_n, \bar{x})$ of the level-set function (by numerically solving the corresponding HJB equation). On a grid with the discretization parameters $\Delta u = \Delta x = h = \frac{T}{N}$, we obtain the values given in the second column of Table 3. The third column of this table gives the differences between two values of $V(t_n, \bar{x})$ computed on two successive grids. When $N = 80$, we obtain that $V(t_n, \bar{x})$ is approximately 0.503, and the difference between the last two computed values is of order 0.014. Hence we can extrapolate numerically to say that a process starting from $\bar{x}$ reaches the target set with a probability 0.503 ± 0.007.

Table 3: Example 3: The value of $V(t_n, \bar{x})$ for different mesh parameters, and difference between two successive values (here $t_n = 0.75$ and $\bar{x} = (-1.0, 2.0)^\mathsf{T}$).

N	$V(t_n, \bar{x})$ estimate	differences
10	0.41768	-
20	0.46359	0.046
40	0.48926	0.026
80	0.50326	0.014

Now, let $N = 20$ and call the trajectory reconstruction algorithm described above to generate some trajectories starting from $\bar{x}$ (by Monte Carlo simulations). The results are reported in Table 4 with M is the number of the simulated trajectories, p is the percentage of trajectories reaching the target set and C.I denotes the confidence interval at 95%. Notice that the percentage p is an approximation of the probability to reach the target. The results of Table 4 show that the value $V(t_n, \bar{x})$ at point $\bar{x}$ is inside the confidence interval.

Table 4: Example 3: Percentage p of simulated trajectories that reach the target set, corresponding confidence interval (C.I.), and a Monte Carlo error estimate (MC-error)

M	p	C.I.	MC-error
3000	0.51233	(0.4944, 0.5302)	0.0179
6000	0.51317	(0.5005, 0.5258)	0.0127
12000	0.51575	(0.5068, 0.5247)	0.0090
25000	0.50912	(0.5029, 0.5153)	0.0062
50000	0.50876	(0.5044, 0.5131)	0.0044
100000	0.50969	(0.5066, 0.5128)	0.0031

A Proof of Lemma 4.5

First, we notice that the functions σ and b are defined for times $t \in [0, T]$, but they can be extended to times $[-2\eta^2, T + 2\eta^2]$ in such a way that assumption (H1a) still holds.

For any $\eta > 0$, let $\mathcal{E}$ be the set of progressively measurable processes (α, χ) valued in $[-\eta^2, 0] \times \mathbb{B}_\eta \subset \mathbb{R} \times \mathbb{R}^d$ that is,

$$\mathcal{E} := \big\{\text{prog. meas. process } (\alpha, \chi) \text{ valued in } [-\eta^2, 0] \times \mathbb{B}_\eta\big\}.$$

Now, consider the function v^η associated to the perturbed control problem (with $\eta > 0$):

$$v^\eta(t, x) := \inf_{\substack{u \in \mathcal{U}, \\ (\alpha,\chi) \in \mathcal{E}}} \mathbb{E}\left[\phi\left(X_{t,x}^{u,(\alpha,\chi)}(T)\right)\right],$$

where $X_{t,x}^{u,(\alpha,\chi)}$ is the solution of the perturbed system of SDEs

$$\begin{cases} dX(s) = b(s + \alpha(s), X(s) + \chi(s), u(s))ds + \sigma(s + \alpha(s), X(s) + \chi(s), u(s))d\mathcal{B}(s) \\ X(t) = x. \end{cases}$$

By classical arguments, we can show that v^η is the unique viscosity solution of the perturbed HJB equation:

$$\begin{cases} -\partial_t v_t^\eta + \inf_{-\eta^2 \le s \le 0, |e| \le \eta} \mathcal{H}(t+s, x+e, Dv^\eta, D^2 v^\eta) = 0 \quad \text{in } Q_{\eta^2} \\ v^\eta(T,x) = \phi(x) \quad \text{in } \mathbb{R}^d \end{cases} \tag{67}$$

where $Q_{\eta^2} := (-\eta^2, T] \times \mathbb{R}^d$. Using similar argument to those used in Lemma 3.5, the function v^η satisfies the following relations:

$$|v^\eta(t,x) - v^\eta(t,y)| \le CL_\phi |x-y| \tag{68a}$$

$$|v^\eta(t,x) - v^\eta(s,x)| \le CL_\phi (1+|x|)|t-s|^{\frac{1}{2}} \tag{68b}$$

for all $x, y \in \mathbb{R}^d$, $t, s \in [0,T]$.

The bound on the difference between the pertubed function v^η and the value function v at every point (t,x) follows by the Lipschitz property of ϕ and (H1a). Indeed, we have:

$$\begin{aligned} |v(t,x) - v^\eta(t,x)| &= \Big| \sup_{\substack{u\in\mathcal{U} \\ (\alpha,\chi)\in\mathcal{E}}} \mathbb{E}\big[\phi(X_{t,x}^u(T))\big] - \sup_{\substack{u\in\mathcal{U} \\ (\alpha,\chi)\in\mathcal{E}}} \mathbb{E}\big[\phi(X_{t,x}^{u,(\alpha,\chi)}(T))\big] \Big| \\ &\le \Big| \sup_{\substack{u\in\mathcal{U} \\ (\alpha,\chi)\in\mathcal{E}}} \mathbb{E}\big[\phi(X_{t,x}^u(T)) - \phi(X_{t,x}^{u,(\alpha,\chi)}(T))\big] \Big| \\ &\le L_\phi \sup_{\substack{u\in\mathcal{U} \\ (\alpha,\chi)\in\mathcal{E}}} \mathbb{E}\big[\big|X_{t,x}^u(T) - X_{t,x}^{u,(\alpha,\chi)}(T)\big|\big] \end{aligned} \tag{69}$$

On the other hand, for every $\tau \in [t,T]$, we have:

$$\begin{aligned} &\mathbb{E}\big[\big|X_{t,x}^u(\tau) - X_{t,x}^{u,(\alpha,\chi)}(\tau)\big|^2\big] \\ \le\ &\mathbb{E}\Big[\Big| \int_t^\tau \big[b(s+\alpha(s), X_{t,x}^{u,(\alpha,\chi)}(s) + \chi(s), u(s)) - b(s, X(s), u(s))\big] ds \\ +\ &\int_t^\tau \big[\sigma(s+\alpha(s), X_{t,x}^{u,(\alpha,\chi)}(s) + \chi(s), u(s)) - \sigma(s, X(s), u(s))\big] dW(s) \Big|^2\Big]. \end{aligned}$$

With assumption (H1a), Cauchy-Schwartz inequality and Gronwall Lemma, we obtain that:

$$\mathbb{E}\big[\big|X_{t,x}^u(T) - X_{t,x}^{u,(\alpha,e)}(T)\big|^2\big] \le C\eta^2. \tag{70}$$

By combining (69) and (70), we finally get:

$$|v(t,x) - v^\eta(t,x)| \le L_\phi C\eta. \tag{71}$$

On the other hand, by a change of variables, we see that for $-\eta^2 \le s \le 0, |e| \le \eta$, the function $v^\mu(\cdot - s, \cdot - e)$ is a supersolution of the following equation:

$$-\partial_t \varphi + \mathcal{H}(t, x, D\varphi, D^2\varphi) = 0 \quad \text{in } (-\eta^2, T+s) \times \mathbb{R}^d \tag{72}$$

In order to regularize v^η, we construct the following sequence $v_\eta = v^\eta * \rho_\eta$ where $\{\rho_\eta\}_\eta$ is the sequence of mollifiers defined by

$$\rho_\eta = \frac{1}{\eta^{d+2}}\rho(\frac{t}{\eta^2},\frac{x}{\eta})$$
$$\rho \in C^\infty(\mathbb{R}^{d+1}),\ \rho \geq 0, \quad \text{supp } \rho \subset [0,1]\times B_1, \quad \int_{\mathbb{R}}\int_{\mathbb{R}^d}\rho(s,x)\,dxds = 1. \tag{73}$$

A Riemmann-sum approximation shows that $v_\eta(t,x)$ can be viewed as the limit of convex combinaisons of $v^\eta(t-s,x-e)$. By the stability result for viscosity supersolutions, and by using same arguments as in [6, Appendix A], we can conclude that v_η is itself a supersolution of (72).

Moreover, for a small $\eta << 1$ and using (68a)-(68b), we have

$$\begin{aligned} |v_\eta(t,x) - v^\eta(t,x)| &= \int_0^1\int_{\mathbb{B}_1}|v^\eta(t-\eta^2\tau, x-\eta z) - v^\eta(t,x)|\rho(\tau,z)\,dzd\tau \\ &\leq CL_\phi\eta\int_0^1\int_{\mathbb{B}_1}\big((1+|x|+\eta|z|)\sqrt{\tau}+|z|\big)\rho(\tau,z)\,dzd\tau \end{aligned}$$

Thus, we obtain for $\eta\in(0,1)$ that $|v_\eta(t,x) - v^\eta(t,x)| \leq CL_\phi(1+|x|)\eta$, which together with (71) yield to the desired estimate:

$$|v(t,x) - v_\eta(t,x)| \leq CL_\phi\eta(1+|x|) \quad \text{For all } t,x\in[0,T]\times\mathbb{R}^d \tag{74}$$

for a positive constant $C>0$.

Bound estimates (41b) for the derivatives of v_η can be derived in straightforward way by using the definition of mollification and the Hölder estimates of v^η (see [14, section] for instance).

B Proof of Lemma 3.4

By using spherical coordinates in $\mathbb{R}^d$ it first holds that

$$I(a) := \int_{|z|\geq a, z\in\mathbb{R}^d}|z|^\alpha e^{-c_2|z|^2}dz = |S_d|\int_a^\infty r^\beta e^{-c_2r^2}dr, \quad \text{with } \beta := \alpha+d-1,$$

where $|S_d|$ denotes the surface of the unit sphere of $\mathbb{R}^d$. Then the following identity holds:

$$2(1+\frac{1-\beta}{2c_2}\frac{1}{r^2})r^\beta e^{-c_2r^2} = -\frac{1}{c_2}\frac{d}{dr}(r^{\beta-1}e^{-c_2r^2}).$$

In the case when $a^2 \geq 2\frac{|\beta-1|}{2c_2} = \frac{|\beta-1|}{c_2}$, and for $r\geq a$, we observe that $1\leq 2(1+\frac{1-\beta}{2c_2}\frac{1}{r^2})$, hence

$$r^\beta e^{-c_2r^2} \leq -\frac{1}{c_2}\frac{d}{dr}(r^{\beta-1}e^{-c_2r^2}).$$

By integration over $r\in[a,\infty[$, we obtain $I(a)\leq\frac{|S_d|}{c_2}a^{\beta-1}e^{-c_2a^2}$. On the other hand, if $a\in J := [1,\sqrt{\frac{|\beta-1|}{c_2}}]$, then $F(a) := I(a)/\big(a^{\beta-1}e^{-c_2a^2}\big)$ is a continuous function on the interval J so it is bounded by some $q_\alpha > 0$. We can furthermore chose $q_\alpha \geq \frac{|S_d|}{c_2}$. In all cases, for $a \geq 1$, it holds $F(a) \leq q_\alpha$. Hence $I(a) \leq q_\alpha\, a^{\alpha+d-2} e^{-c_2 a^2}$. Using that $a^{\alpha+d-2} \leq a^{\alpha+d-1}$ for $a \geq 1$, we obtain the desired result.

C Proof of estimates (37)

Let us first prove, by recursion, the following estimate: there exists $C \geq 0$ such that, for any $p \in \{2,4\}$, for $0 \leq n \leq k \leq N$, $\forall x$ and $h \leq 1$:

$$\begin{aligned} \max_{a_n} \mathbb{E}\bigg[\max_{a_{n+1}} \mathbb{E}\bigg[\cdots \max_{a_{k-1}} \mathbb{E}\Big[|Z^{k,a}_{n,x}|^p \Big] \cdots \bigg]\bigg] & \\ \leq \ \leq e^{C(k-n)h} \bigg(|x|^p + C(k-n)h \bigg) & \qquad (75) \end{aligned}$$

For $p = 4$ and using that $(k-n)h \leq Nh = T$, this will give the desired estimate (37a).

We first start with the case $p = 2$. Using conditional expectations, let us first show that, for some constant $C_1, C_2 \geq 0$:

$$\max_{a_{k-1}} \mathbb{E}\big[|Z^{k,a}_{n,x}|^2 \,|\, Z^{k-1,a}_{n,x} = y \big] \leq |y|^2 e^{C_1 h} + C_2 h. \qquad (76)$$

Denoting $b(y) = b(t_k, y, a_k)$ as well as $\bar\sigma_q(y) = \bar\sigma_q(t_k, y, a_k)$, it holds

$$\begin{aligned} \mathbb{E}\big[|Z^{k,a}_{n,x}|^2 \,|\, Z^{k-1,a}_{n,x} = y \big] &= \mathbb{E}\big[|y + hb(y) + \sqrt{h}\bar\sigma_Q(y)|^2 \big] \\ &= |y + hb(y)|^2 + h\mathbb{E}[|\bar\sigma_Q(y)|^2], \end{aligned}$$

where we have used that $\mathbb{E}[\bar\sigma_Q(y)] = 0$ by the definition of the random variable $\bar\sigma_Q$. Hence it holds, since $\bar\sigma_Q(y)$ has a linear growth in y:

$$\mathbb{E}\big[|Z^{k,a}_{n,x}|^2 \,|\, Z^{k-1,a}_{n,x} = y \big] \leq |y + hb(y)|^2 + Ch(1 + |y|^2).$$

Notice that for $h \leq 1$ it holds, for the Euclidean norm, and for any vectors A and B of $\mathbb{R}^d$,

$$|A + hB|^2 \leq |A|^2(1 + h) + 2h|B|^2 \qquad (77)$$

(using Cauchy-Schwarz inequality and $h^2 \leq h$). Hence we obtain a bound of the form

$$\max_{a_{k-1}} \mathbb{E}\big[|Z^{k,a}_{n,x}|^2 \,|\, Z^{k-1,a}_{n,x} = y \big] \leq |y|^2(1 + C_1 h) + C_2 h.$$

We conclude to (76) by using $1 + C_1 h \leq e^{C_1 h}$.

Then we can iterate the previous bound, to obtain

$$\begin{aligned} & \max_{a_{k-2}} \mathbb{E}\bigg[\max_{a_{k-1}} \mathbb{E}\bigg[|Z^{k,a}_{n,x}|^2 \,|\, Z^{k-2,a}_{n,x} = y \bigg]\bigg] \\ & \quad \leq \max_{a_{k-2}} \mathbb{E}\bigg[|Z^{k-1,a}_{n,x}|^2 e^{C_1 h} + C_2 h |\, Z^{k-2,a}_{n,x} = y \bigg] \\ & \quad \leq \bigg(|y|^2 e^{C_1 h} + C_2 h \bigg) e^{C_1 h} + C_2 h \\ & \quad \leq |y|^2 e^{2C_1 h} + C_2 h(1 + e^{C_1 h}). \end{aligned}$$

By a recursion argument and since $Z^{n,a}_{n,x} = x$, it holds

$$\begin{aligned}
&\max_{a_n} \mathbb{E}\left[\max_{a_{n+1}} \mathbb{E}\left[\cdots \max_{a_{k-1}} \mathbb{E}\left[|Z^{k,a}_{n,x}|^2 \right] \cdots \right]\right] \\
&\qquad \leq \quad e^{C_1(k-n)h}|x|^2 + C_2 h \sum_{j=0,\ldots,k-n-1} e^{C_1 jh} \\
&\qquad \leq \quad e^{C_1(k-n)h}\Big(|x|^2 + C_2(k-n)h\Big).
\end{aligned}$$

Now we turn to the case $p = 4$. Let us show that a similar estimate to (76) holds, mainly:

$$\max_{a_{k-1}} \mathbb{E}\Big[|Z^{k,a}_{n,x}|^4 \,|\, Z^{k-1,a}_{n,x} = y \Big] \leq |y|^4 e^{C_1 h} + C_2 h. \tag{78}$$

The rest of the proof of (75) then follows the same idea as for the case $p = 2$, and is left to the reader.

To prove (78), assuming first that $d = 1$ to simplify the argument, denoting $A = y + hb(y)$ and $B = \bar{\sigma}_Q(y)$, we have

$$\begin{aligned}
\mathbb{E}\big[|Z^{k,a}_{n,x}|^4 \,|\, Z^{k-1,a}_{n,x} = y \big] &= \mathbb{E}[|A + \sqrt{h}B|^4] \\
&= |A|^4 + 16hA^2\mathbb{E}[B^2] + h^2\mathbb{E}[B^4],
\end{aligned}$$

where we have used that $\mathbb{E}[B] = \mathbb{E}[B^3] = 0$. Then $\mathbb{E}[B^2] \leq C(1 + |y|^2)$ and $\mathbb{E}[B^4] \leq C(1 + |y|^4)$, it can be shown that $|y + hb(y)|^4 \leq |y|^4(1 + Ch) + Ch$ for some constant $C \geq 0$ (for intance by using twice (77)), and (78) is deduced from these estimates. The case $d \geq 1$ can be treated in a similar way.

The proof of (37b) can be obtained in a similar way as for the proof of (37a) for $p = 2$. It is first established that

$$\max_{a_n} \mathbb{E}\left[\max_{a_{n+1}} \mathbb{E}\left[\cdots \max_{a_{k-1}} \mathbb{E}\left[|Z^{k,a}_{n,x} - Z^{k,a}_{n,y}|^2 \right] \cdots \right]\right] \leq C|x - y|^2. \tag{79}$$

Then, using that $E[|X|] \leq (\mathbb{E}[|X|^2])^{1/2}$, the desired estimate is obtained.

Finaly we consider the proof of (37c). In a complete similar manier as for the proof of (75), we can establish that for any given $x_0 \in \mathbb{R}^d$,

$$\begin{aligned}
&\max_{a_n} \mathbb{E}\left[\max_{a_{n+1}} \mathbb{E}\left[\cdots \max_{a_{k-1}} \mathbb{E}\left[|Z^{k,a}_{n,x} - x_0|^2 \right] \cdots \right]\right] \\
&\qquad \leq \quad e^{C(k-n)h}\Big(|x - x_0|^2 + C(k-n)h(1 + |x|^2) \Big).
\end{aligned} \tag{80}$$

In particular for $x_0 = x$, for some other constant $C \geq 0$, we obtain:

$$\begin{aligned}
&\max_{a_n} \mathbb{E}\left[\max_{a_{n+1}} \mathbb{E}\left[\cdots \max_{a_{k-1}} \mathbb{E}\left[|Z^{k,a}_{n,x} - x|^2 \right] \cdots \right]\right] \\
&\qquad \leq C(1 + |x|^2)\ (k-n)h
\end{aligned} \tag{81}$$

By using (81) and the fact that for $y := Z^{k,a}_{n,x}$, $Z^{k+m,a}_{n,x} = Z^{m,a'}_{k,y}$ (with controls $a' = (a_k, a_{k+1}, \dots, a_{k+m-1})$), we have

$$\max_{a_k} \mathbb{E}\left[\max_{a_{k+1}} \mathbb{E}\left[\cdots \max_{a_{k+m-1}} \mathbb{E}\left[|Z^{k+m,a}_{n,x} - Z^{k,a}_{n,x}|^2 \,|\, Z^{k,a}_{n,x} = y\right] \cdots\right]\right] \le C(1+|y|^2)\, mh$$

Then

$$\max_{a_k} \mathbb{E}\left[\max_{a_{k+1}} \mathbb{E}\left[\cdots \max_{a_{k+m-1}} \mathbb{E}\left[|Z^{k+m,a}_{n,x} - Z^{k,a}_{n,x}|^2\right] \cdots\right]\right] \le C\mathbb{E}\left[1 + |Z^{k,a}_{n,x}|^2\right] mh,$$

By using (37a), the right hand side term is bounded by $C(1+|x|^2)mh = C(1+|x|^2)(t_{m+k} - t_k)$. Using again inequalities of the type $\mathbb{E}[|x|] \le \mathbb{E}[|X|^2]^{1/2}$, we obtain the desired bound (37c).

References

[1] A. Abate, S. Amin, M. Prandini, J. Lygeros and S. Sastry, Computational approaches to reachability analysis of stochastic hybrid systems, *Hybrid Systems*, **4416** (2007), 4–17.

[2] A. Abate, M. Prandini, J. Lygeros and S. Sastry, Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems, *Automatica*, **44** (2008), 2724–2734.

[3] M. Althoff, O. Stursberg and M. Buss, Safety assessement of autonomous cars using verification techniques, *American Control Conference*, (2007), 4154–4159.

[4] M. Althoff, O. Stursberg and M. Buss, Safety assessement for stochastic linear systems using enclosing hulls of probability density functions, *European Control Conference*, 625–630.

[5] S. Amin, A. Abate, M. Prandini, S. Sastry and J. Lygeros, Reachability analysis for controlled discrete time stochastic hybrid systems, in *Lecture Notes in Computer Science LNCS*, **3927** (2006), 49–63.

[6] G. Barles and E. R. Jakobsen, On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations, *ESAIM:M2AN*, **36** (2002), 33–54.

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

[8] G. Barles and E. R. Jakobsen, Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations, *Mathematics of Computations*, **76** (2007), 1861–1893.

[9] I. H. Biswas, E. R. Jakobsen and K. H. Karlsen, Difference quadrature schemes for nonlinear degenerate parabolic integro-pde, *SIAM J. Numer. Anal.*, **48** (2010), 1110–1135.

[10] I. H. Biswas, E. R. Jakobsen and K. H. Karlsen, Viscosity solutions for a system of integro-pdes and connections to optimal switching and control of jump-diffusion processes, *Applied mathematics and optimization*, **62** (2010), 47–80.

[11] O. Bokanowski, N. Forcadel and H. Zidani, Reachability and minimal times for state constrained nonlinear problems without any controllability assumption, *SIAM J. Control and Optimization. Doi: 10.1137/090762075*, **48** (2010), 4292–4316.

[12] O. Bokanowski, A. Picarelli and H. Zidani, Dynamic programming and error estimates for stochastic control problems with maximum cost, *Applied Math. and Optimization*, **71** (2015), 125–163.

[13] J. Bonnans, S. Maroso and H. Zidani, Error bounds for stochastic differential games: The adverse stopping case, *IMA, J. Numer. Anal.*, **26** (2006), 188–212.

[14] J. Bonnans, S. Maroso and H. Zidani, Error estimates for a stochastic impulse control problem, *Applied. Math. and Optimisation*, **55** (2007), 327–357.

[15] B. Bouchard, R. Elie and N. Touzi, Stochastic target problems with controlled loss, *SIAM*, **48** (2008), 3123–3150.

[16] A. Briani, F. Camilli and H. Zidani, Approximation schemes for monotone systems of nonlinear second order partial differential equations: convergence result and error estimate, *Differential Equations and Applications*, **4** (2012), 297–317.

[17] L. Caffarelli and P. E. Souganidis, A rate of convergence for monotone finite difference approximations to fully nonlinear uniformly elliptic pde, *Comm. Pure Appl. Math.*, **61** (2008), 1–17.

[18] F. Camilli and E. R. Jakobsen, A finite element like scheme for integro-partial differential Hamilton-Jacobi-Bellman equations, *SIAM J. Numer. Anal.*, **47** (2009), 2407–2431.

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

[20] F. Da Lio and O. Ley, Uniqueness results for convex Hamilton-Jacobi equations under $p > 1$ growth conditions on data, *Applied Math. and Optimization*, **63** (2011), 309–339.

[21] K. Debrabant and E. R. Jakobsen, Semi-Lagrangian schemes for linear and fully non-linear diffusion equations, *Mathematics of Computations*, **82** (2013), 1433–1462.

[22] F. Delarue and S. Menozzi, Density estimates for a random noise propagating through a chain of differential equations, *J. Funct. Anal.*, **259** (2010), 1577–1630.

[23] W. H. Fleming and M. H. Soner, *Controlled Markov Processes and Viscosity Solutions*, vol. 25 of Stochastic Modelling and Applied Probability, 2nd edition, Springer, New York, 2006.

[24] H. Föllmer and P. Leukert, Quantile hedging, *SIAM*, **3** (1999), 251–273.

[25] D. Goreac and O.-S. Serea, Mayer and optimal stopping stochastic control problems with discontinuous cost, *J. Math. Anal. Appl.*, **380** (2011), 327–342.

[26] N. V. Krylov, Mean value theorems for stochastic integrals, *Ann. Probab.*, **29** (2001), 385–410.

[27] N. Krylov, On the rate of convergence of finite difference approximation for Bellman's equation, *St. Petersburg Math. J.*, **9** (1998), 639–650.

[28] N. Krylov, On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients, *Probability Theory and Related Fields*, **117** (2000), 1–16.

[29] N. Krylov, On the rate of convergence for finite-difference approximations for bellman equations with lipschitz coefficients, *Applied Mathematics and Optimization*, **52** (2005), 365–399.

[30] H. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, vol. 24 of Applications of mathematics, Springer, New York, 2001, Second edition.

[31] I. Mitchell, A. Bayen and C. Tomlin, A time-dependent Hamiliton-Jacobi formulation of reachable sets for continuous dynamic games, *IEEE Transactions on automatic control*, **50** (2005), 947–957.

[32] R. Munos and H. Zidani, Consistency of a simple multidimensional scheme for hjb equations, *C. R. Acad. Sci. Paris, Ser. I*, **340** (2005), 499–502.

[33] (MR965860) [10.1016/0021-9991(88)90002-2] S. Osher and J. A. Sethian, Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations, *J. Comput. Phys.*, **79** (1988), 12–49.

[34] R. Rubinstein and D. Kroese, *Simulation and the Monte Carlo Method*, Wiley, 2008, Second edition.

[35] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Stochastic Modelling and Applied Probability, Springer, Paris, 1999.