A fully backward representation of semilinear PDEs applied to the control of thermostatic loads in power systems

LUCAS IZYDORCZYK *, NADIA OUDJANE † AND FRANCESCO RUSSO ‡

September 10th 2021

Abstract

We propose a fully backward representation of semilinear PDEs with application to stochastic control. Based on this, we develop a fully backward Monte-Carlo scheme allowing to generate the regression grid, backwardly in time, as the value function is computed. This offers two key advantages in terms of computational efficiency and memory. First, the grid is generated adaptively in the areas of interest and second, there is no need to store the entire grid. The performances of this technique are compared in simulations to the traditional Monte-Carlo forward-backward approach on a control problem of thermostatic loads.

Key words and phrases: Ornstein-Uhlenbeck processes; probabilistic representation of PDEs; time-reversal of diffusion; stochastic control; HJB equation; regression Monte-Carlo scheme; demand-side management.

2020 AMS-classification: 60H10; 60H30; 60J60; 65C05; 49L25; 35K58.

1 Introduction

The numerical resolution of non-linear PDEs is a crucial issue in many applications. In particular, stochastic control problems can be formulated by mean of the Hamilton-Jacobi-Bellman (HJB) equations with terminal condition. In this paper, we focus more particularly on control problems raised by demand-side management in power systems. The difficulties come especially from the high dimensionality of the state space, which motivates the use of probabilistic representations. The main issue of numerical schemes is then to concentrate the computing effort in specific regions of interest in the state space. In classical regression Monte-Carlo approaches, the solution

*ENSTA Paris, Institut Polytechnique de Paris. Unité de Mathématiques Appliquées (UMA). E-mail: lucas.izydorczyk@ensta-paris.fr

†EDF R&D, and FiME (Laboratoire de Finance des Marchés de l'Energie (Dauphine, CREST, EDF R&D) www.fime-lab.org). E-mail:nadia.oudjane@edf.fr

‡ENSTA Paris, Institut Polytechnique de Paris. Unité de Mathématiques Appliquées (UMA). E-mail:francesco.russo@ensta-paris.fr.

is evaluated backwardly in time from the final time to the initial time, while the regression grid is generated forwardly from the initial time to the final one. In this paper, we propose a fully backward probabilistic approach which allows to generate adaptively the regression grid, as the solution is evaluated, taking advantage of the calculations already performed. Besides, there is no need to store the entire grid, since the points are generated as they are used for calculations. Our grid will be indeed simulated according to the time-reversal of some diffusion starting from a judicious terminal distribution.

We are interested in semilinear PDEs of the type

$$\begin{cases} \partial_t v(t,x) + H(t,x,v(t,x),\nabla_x v(t,x)) + \frac{1}{2} Tr[\sigma\sigma^\top(t)\nabla_x^2 v(t,x)] = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d \\ v(T,x) = g(x), \end{cases} \tag{1.1}$$

where in particular σ is a deterministic non-degenerate matrix-valued function. Under suitable conditions, there exists a unique viscosity solution v of (1.1) in the class of continuous functions with polynomial growth. One classical probabilistic representation of v is provided by Forward-Backward SDEs (FBSDEs), see e.g. [30]. First a forward diffusion is fixed, with an arbitrary drift $\tilde{b}$

$$dX_t = \widetilde{b}(t,X_t)dt + \sigma(t)dW_t. \tag{1.2}$$

Then the solution of (1.1) is represented by $v(s,x) = Y_s^{s,x}$, where $(Y,Z) = (Y^{s,x}, Z^{s,x})$ is the unique solution of the BSDE

$$Y_t = g(X_T) + \int_t^T F(r,X_r,Y_r,Z_r)dr - \int_t^T Z_r dW_r, \tag{1.3}$$

with $X = X^{s,x}$ being the solution of (1.2) starting at time s with value x and F being related to H by

$$F(t,x,y,z) := H(t,x,y,\left(\sigma^{-1}(t)\right)^\top z) - \left\langle \widetilde{b}(t,x), \left(\sigma^{-1}(t)\right)^\top z \right\rangle. \tag{1.4}$$

Considering a time discretization mesh $t_k = k\delta t$, with $\delta t = \frac{T}{n}$ and $k = 0,\cdots,n$, for a given positive integer n, [13] proved that one can approximate (Y_{t_k}, Z_{t_k}) by $(\hat{Y}_k, \hat{Z}_k)$ such that $\hat{Y}_n = g(X_T)$ and for $k = 0,\cdots,n-1$

$$\begin{cases} \hat{Y}_k &= \mathbb{E}\left(\sum_{\ell=k+1}^{n} F(t_\ell, X_{t_\ell}, \hat{Y}_\ell, \hat{Z}_{\ell-1})\delta t + g(X_T) \middle| X_{t_k}\right) \\ \hat{Z}_k &= \frac{1}{\delta t}\mathbb{E}\left(\hat{Y}_{k+1}(W_{t_{k+1}} - W_{t_k}) \middle| X_{t_k}\right). \end{cases} \tag{1.5}$$

Most of probabilistic numerical schemes (see e.g. regression Monte-Carlo [15, 3], Kernel Monte-Carlo [4], Quantization [8]) rely on that representation. The common idea is then articulated in two steps. First, one generates a grid discretizing the forward process (1.2) in space and time on $[0,T]$, (by Monte-Carlo simulations or Quantization, etc.). Then, one calculates the conditional expectations (1.5) on the grid points in order to estimate $(\hat{Y}, \hat{Z})$. These techniques have generally two limitations.

1. The degree of freedom in the choice of the forward diffusion X is difficult to exploit although it has a major impact on the numerical scheme efficiency: how to chose a reasonable drift $\tilde{b}$ without a priori information on v ?

2. The entire grid discretizing the forward process has to be stored in memory to be revisited backwardly in time in order to compute the solution process (Y, Z). This approach naturally raises some huge memory issues which in general limit drastically the number of Monte-Carlo runs and time steps, hence the accuracy of the procedure.

To overcome such limitations some approaches were proposed in the domain of mathematical finance, in particular for the evaluation of American style options. One technique, intended to deal with the memory problem, relies on bridge simulation, see e.g. [34, 35]. However, this approach requires specific developments for each price model (based for instance on the Brownian bridge for Brownian prices or on the gamma bridge for variance gamma prices) and remains difficult to generalize to a wide class of models. To address the efficiency issue, [1] developed a scheme based on Picard's type iterations that avoids the use of nested conditional expectations backwardly in time, which are replaced by nested conditional expectations along the iterations. In the same line, [14] proposes an adaptive variance reduction technique which combines Picard's iterations and control variate to solve the BSDE. A parallel version of that algorithm was proposed in [27]. However, those approaches require, at each iteration, to approximate the solution on the whole time horizon. Similarly, importance sampling and Girsanov's theorem, were considered to force the exploration of the space towards areas of interest [2]. In particular, this type of approach was derived in the case of stochastic control in [10] providing an iterative scheme that is capable of learning the optimally controlled drift. Here again, that method requires several estimations of the value function on the whole time horizon. Besides [17] proposed an adaptive importance sampling scheme for FBSDEs allowing to select the drift adaptively, as the calculations are performed backwardly. Unfortunately, that approach is limited to situations where the driver F does not depend on Z. In the present paper, we introduce a new adaptive approach to address both the memory problem and the efficiency issue (related to the drift selection) in the general case where the driver may depend on X, Y and Z.

We propose to choose adaptively the drift $\tilde{b}$ at the same time as we discover the function v such that

$$v(t, X_t) = \mathbb{E}\left(\int_t^T H\left(s, X_s, v(s, X_s), \nabla_x v(s, X_s)\right) - \left\langle \tilde{b}(s, X_s), \nabla_x v(s, X_s)\right\rangle ds + g(X_T) \middle| X_t\right), \tag{1.6}$$

by simulating the time-reversal of a solution X of (1.2) starting from the distribution of X_T. More specifically, to take advantage of the Ornstein-Uhlenbeck setting, we choose the drift $\tilde{b}$ to be affine w.r.t. the space variable. We fix a Gaussian distribution ν and look for solutions ξ of the McKean

SDE

$$
\begin{cases}
\xi_0 \sim \nu, \\
\xi_t = \xi_0 - \displaystyle\int_0^t \widetilde{b}\,(T-s,\xi_s) + \sigma\sigma^\top\,(T-s)\,Q\,(T-s)^{-1}\,(\xi_s - m\,(T-s))\,ds + \int_0^t \sigma\,(T-s)\,d\beta_s, \\
m(T-t) = \mathbb{E}\,(\xi_t)\,, \\
Q(T-t) = \mathrm{Cov}\,(\xi_t) \quad \text{for } t \in]0,T].
\end{cases}
\tag{1.7}
$$

Rev-SDEIntro

By Proposition 3.7, (1.7) admits exactly one solution ξ, provided Assumption 1 in Section 3.1 is verified. That assumption depends on the covariance matrix of ν, the drift $\tilde{b}$ and the volatility σ. Indeed, one important limitation is that the covariance matrix should be chosen carefully to ensure that the process is well-defined until T. Point 2. of Proposition 3.7 and Lemma 3.6 say that the time-reversal process $\hat{\xi}$, i.e. $\hat{\xi}_t := \xi_{T-t}$, is an Ornstein-Uhlenbeck process solution of (1.2) such that the law of X_0 is Gaussian with mean $m(0)$ and covariance $Q(0)$. This leads to the first result of this paper which consists of the fully backward representation stated in Theorem 3.10. The proof is based on Feynman-Kac type formula instead of BSDEs and it does not require explicitly the uniqueness of viscosity solution of the PDE (1.1). The second contribution of the paper is Corollary 4.4 which is the "instantiation" of Theorem 3.10 in the framework of stochastic control, i.e. the representation of its value function (solution of a Hamilton-Jacobi-Bellman equation). This holds when the running and terminal cost have polynomial growth with respect to the state space variable. We also suppose that the value function is of class $\mathcal{C}^{0,1}$ whose gradient has polynomial growth. In particular, we derive in Corollary 4.6, a representation involving the gap between the optimally controlled drift and the instrumental drift $\tilde{b}$. In Section 5, we present a fully backward Monte-Carlo regression scheme, where the instrumental drift is adaptively updated in order to mimic the optimally controlled dynamics, see Algorithm 1. We expect that this approach is particularly well-suited when the final cost has a strong impact on the global cost and when the terminal cost function is localized in a small region of the space, so that the initial distribution ν can be chosen in an appropriate way. Finally, in Section 6 we illustrate the interest of this new algorithm applied to the problem of controlling the consumption of a large number of thermostatic loads in order to minimize an aggregative cost. We compare our approach to the classical regression Monte-Carlo scheme based on a forward grid.

2 Notations

SNotat

Let us fix $T > 0$, $d, k \in \mathbb{N}^*$. For a given $p \in \mathbb{N}^*$, $[\![1,p]\!]$ denotes the set of all integers between 1 and p included. $\langle \cdot,\cdot \rangle$ denotes the usual scalar product on $\mathbb{R}^d$ and $|\cdot|$ the associated norm. Elements of $\mathbb{R}^d$ are supposed to be column vectors. $M_d\,(\mathbb{R})$ stands for the set of $d \times d$ matrices, $S_d\,(\mathbb{R})$ for the subset of symmetric matrices, $S_d^+\,(\mathbb{R})$ the subset of symmetric positive semi-definite matrices (in particular with non-negative eigenvalues) and $S_d^{++}\,(\mathbb{R})$ for the subset of strictly positive definite symmetric matrices. For a given $A \in M_d(\mathbb{R})$, $A^\top$ will denote its transpose, $Tr(A)$ its trace, $Sp(A)$ its spectrum, i.e. the set of its eigenvalues, $e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$ its exponential and $||A|| := \sup_{x \in \mathbb{R}^d, |x|=1} |Ax|$. For a given $A \in S_d^+(\mathbb{R})$, $\sqrt{A}$ denotes the unique element of $S_d^+(\mathbb{R})$ such that $(\sqrt{A})^2 = A$.

For a given continuous function $f : [0,T] \mapsto \mathbb{R}^d$ (resp. $g : [0,T] \mapsto M_d(\mathbb{R})$), we set $||f||_\infty := \sup_{t \in [0,T]} |f(t)|$ (resp. $||g||_\infty := \sup_{t \in [0,T]} ||g(t)||$). $\mathcal{C}^{1,2}([0,T], \mathbb{R}^d)$ (resp. $\mathcal{C}^{0,1}([0,T], \mathbb{R}^d)$) denotes the set of real-valued functions defined on $[0,T] \times \mathbb{R}^d$ being continuously differentiable in time and twice continuously differentiable in space (resp. continuous in time and continuously differentiable in space). $\mathcal{C}^0([0,T] \times \mathbb{R}^d)$ (resp $\mathcal{C}^1(\mathbb{R}^d)$) denotes the set of continuous (resp continuously differentiable) real-valued functions defined on $[0,T] \times \mathbb{R}^d$ (resp. $\mathbb{R}^d$). ∇_x will denote the gradient operator and ∇_x^2 the Hessian matrix. For each $p \in \mathbb{N}$, $P_p(\mathbb{R}^d)$ denotes the set of polynomial functions on $\mathbb{R}^d$ with degree p.

In the whole paper, we say that a function $v : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}$ has *polynomial growth* if there exists $q, K > 0$ such that for all $(t,x) \in [0,T] \times \mathbb{R}^d$

$$|v(t,x)| \leq K(1 + |x|^q).$$

When v verifies previous property with $q = 1$, we say that it has *linear growth*.

For a given random vector X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{E}_{\mathbb{P}}(X)$ (resp. $\mathrm{Cov}_{\mathbb{P}}(X) := \mathbb{E}_{\mathbb{P}}\left((X - \mathbb{E}_{\mathbb{P}}(X))(X - \mathbb{E}_{\mathbb{P}}(X))^\top\right)$) will denote its expectation (resp. its covariance matrix) under $\mathbb{P}$. When self-explanatory, the subscript will be omitted in the sequel. For a given $(m, Q) \in \mathbb{R}^d \times S_d^+(\mathbb{R})$, $\mathcal{N}(m, Q)$ denotes the Gaussian probability on $\mathbb{R}^d$ with mean m and covariance matrix Q.

For any stochastic process X, $\mathcal{F}^X$ will denote its canonical filtration. $\widehat{X}$ will denote the time-reversal process $X_{T-\cdot}$.

3 Representation of semilinear PDEs

oSection

3.1 Around two backward ODEs

R41

Let a (resp. c) be Borel bounded functions from $[0,T]$ to $M_d(\mathbb{R})$ (resp. $\mathbb{R}^d$).

In the sequel we will fix a Gaussian Borel probability ν on $\mathbb{R}^d$ with mean $\bar{m}^\nu$ and covariance matrix $\bar{Q}^\nu$. We consider the functions $m^\nu : [0,T] \mapsto \mathbb{R}^d$ and $Q^\nu : [0,T] \mapsto S_d(\mathbb{R})$ denoting respectively the unique solutions of the backward ODEs

$$\begin{cases} \frac{d}{dt} m^\nu(t) = a(t) m^\nu(t) + c(t), \quad t \in [0,T] \\ m^\nu(T) = \bar{m}^\nu, \end{cases} \tag{3.1}$$

ODE_m

$$\begin{cases} \frac{d}{dt} Q^\nu(t) = Q^\nu(t) a(t)^\top + a(t) Q^\nu(t) + \Sigma(t), t \in [0,T] \\ Q^\nu(T) = \bar{Q}^\nu, \end{cases} \tag{3.2}$$

ODE_Q

for which existence and uniqueness hold since they are linear.

We introduce an hypothesis on ν which will be used in the sequel.

Assumption 1. $Q^{\nu}(0) \in \mathcal{S}_d^{+}(\mathbb{R})$.

Easy computations imply for all $t \in [0,T]$

$$m^{\nu}(t) = \mathcal{A}(t)\left(\mathcal{A}(T)^{-1}\bar{m}^{\nu} - \int_t^T \mathcal{A}(s)^{-1} c(s)\, ds\right), \tag{3.3}$$

$$Q^{\nu}(t) = \mathcal{A}(t)\left(\mathcal{A}(T)^{-1}\bar{Q}^{\nu}\left(\mathcal{A}(T)^{-1}\right)^{\top} - \int_t^T \mathcal{A}(s)^{-1}\Sigma(s)\left(\mathcal{A}(s)^{-1}\right)^{\top} ds\right)\mathcal{A}(t)^{\top}, \tag{3.4}$$

where $\mathcal{A}(t), t \in [0,T]$ is the unique solution of the matrix ODE

$$\begin{cases} \frac{d}{dt}\mathcal{A}(t) = a(t)\mathcal{A}(t), t \in [0,T] \\ \mathcal{A}(0) = I_d. \end{cases} \tag{3.5}$$

We recall that for all $t \in [0,T]$, $\mathcal{A}(t)$ is invertible and the matrix valued function $t \mapsto \mathcal{A}(t)^{-1}$ solves the ODE

$$\begin{cases} \frac{d}{dt}\mathcal{A}(t)^{-1} = -\mathcal{A}(t)^{-1} a(t), t \in [0,T] \\ \mathcal{A}(0)^{-1} = I_d, \end{cases} \tag{3.6}$$

see Chapter 8 in [5] for similar and further properties.

Note that in the case $a(t) = a,\ t \in [0,T]$ for a given $a \in M_d(\mathbb{R})$, then $\mathcal{A} : t \to e^{at}$ and identities (3.3), (3.4) simplify as follows:

$$m^{\nu}(t) = e^{-a(T-t)}\bar{m}^{\nu} - \int_t^T e^{-a(s-t)} c(s)\, ds, \tag{3.7}$$

$$Q^{\nu}(t) = e^{-a(T-t)}\bar{Q}^{\nu} e^{-a^{\top}(T-t)} - \int_t^T e^{-a(s-t)}\Sigma(s) e^{-a^{\top}(s-t)} ds, \tag{3.8}$$

for all $t \in [0,T]$.

Remark 3.1. *Suppose that $Q^{\nu}(0)$ belongs to $S_d^{+}(\mathbb{R})$. Identity (3.4) gives in particular*

$$Q^{\nu}(t) = \mathcal{A}(t)\left(Q^{\nu}(0) + \int_0^t \mathcal{A}(s)^{-1}\Sigma(s)\left(\mathcal{A}(s)^{-1}\right)^{\top} ds\right)\mathcal{A}(t)^{\top},\ t \in [0,T]. \tag{3.9}$$

Combining (3.9) and the fact $\sigma(t)$ is invertible for all $t \in [0,T]$, we remark that $Q^{\nu}(t)$ belongs to $S_d^{++}(\mathbb{R})$ for all $t \in]0,T]$.

Finally we give a condition depending on $\mathcal{A}, \sigma, \bar{Q}^{\nu}$ and T to ensure the measure ν fulfills Assumption 1.

Proposition 3.2. *Suppose that*

$$\min Sp\left(\bar{Q}^{\nu}\right) \geq \int_0^T ||\sigma(s)||^2 \left|\left|\left(\mathcal{A}(T)\mathcal{A}(s)^{-1}\right)^{\top}\right|\right|^2 ds. \tag{3.10}$$

Then,

$$Q^{\nu}(0) \in S_d^{+}(\mathbb{R}). \tag{3.11}$$

Proof. Since $\mathcal{A}(T)$ is invertible and $Q^{\nu}(0)$ belongs to $S_d(\mathbb{R})$, (3.11) is equivalent to

$$\mathcal{A}(T)\,Q^{\nu}(0)\,\mathcal{A}(T)^{\top} \in S_d^{+}(\mathbb{R}). \tag{3.12}$$

To prove (3.12), taking into account (3.4), it suffices to show that the matrix

$$\bar{Q}^{\nu} - \int_0^T \mathcal{A}(T)\,\mathcal{A}(s)^{-1}\,\Sigma(s)\left(\mathcal{A}(T)\,\mathcal{A}(s)^{-1}\right)^{\top} ds \in S_d^{+}(\mathbb{R}),$$

or, equivalently, that for all $x \in \mathbb{R}^d$

$$\lambda := x^{\top}\bar{Q}^{\nu}x - \int_0^T x^{\top}\mathcal{A}(T)\,\mathcal{A}(s)^{-1}\,\Sigma(s)\left(\mathcal{A}(T)\,\mathcal{A}(s)^{-1}\right)^{\top} x ds \geq 0. \tag{3.13}$$

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

$$\begin{aligned}\lambda &\geq \min Sp\left(\bar{Q}^{\nu}\right)|x|^2 - \int_0^T \left|\sigma(s)^{\top}\left(\mathcal{A}(T)\,\mathcal{A}(s)^{-1}\right)^{\top} x\right|^2 ds,\\ &\geq \left(\min Sp\left(\bar{Q}^{\nu}\right) - \int_0^T \left\|\sigma(s)^{\top}\right\|^2 \left\|\left(\mathcal{A}(T)\,\mathcal{A}(s)^{-1}\right)^{\top}\right\|^2 ds\right)|x|^2,\\ &\geq 0,\end{aligned}$$

since (3.10) holds. This ends the proof. □

Remark 3.3. *In the case $a(t) = a,\ t \in [0,T]$ for a given $a \in M_d(\mathbb{R})$, Condition (3.10) is satisfied in particular if*

$$\min Sp\left(\bar{Q}^{\nu}\right) \geq \left\|\sigma^{\top}\right\|_{\infty}^2 \int_0^T \left\|e^{a^{\top}s}\right\|^2 ds \tag{3.14}$$

is verified.

Remark 3.4. *Let X be a solution of*

$$X_t = X_0 + \int_0^t \widetilde{b}(s, X_s)\,ds + \int_0^t \sigma(s)\,dW_s,\ t \in [0,T[, \tag{3.15}$$

where σ is a deterministic matrix-valued function and $\widetilde{b}$ the piecewise affine function

$$\widetilde{b}(t,x) = a(t)x + c(t),\ t \in [0,T],$$

and X_0 be a square integrable r.v. It is well-known that X is a square integrable process. Let, for every $t \in [0,T]$, $m(t) = \mathbb{E}(X_t)$ and $Q(t)$ the covariance matrix of X_t. Setting $\bar{m}^{\nu} = \mathbb{E}(X_T)$ and $\bar{Q}^{\nu}$ the covariance matrix of X_T. Then

$$m = m^{\nu}, \quad Q = Q^{\nu}. \tag{3.16}$$

Indeed, by Problem 6.1 in Chapter 5 in [25] m (resp. Q) is solution of (3.1) (resp. (3.2)). (3.16) follows by uniqueness of previous ODEs.

3.2 The representation formula for a general semilinear PDE

In the whole paper σ will be a continuous function defined on $[0,T]$ with values in $M_d(\mathbb{R})$ such that for all $t \in [0,T]$, $\sigma(t)$ is invertible. We will set $\Sigma := \sigma\sigma^\top$.

Let $\widetilde{b} : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and $b_c : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \times S_d^{++}(\mathbb{R}) \mapsto \mathbb{R}^d$ defined by

$$b_c : (t,x,m,Q) \mapsto \Sigma(t) Q^{-1}(x-m), \; \widetilde{b} : (t,x) \mapsto a(t)x + c(t), \tag{3.17}$$

where a, c were defined at Section 3.1. Let $H : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$. The goal of this subsection is to provide a probabilistic representation of viscosity solutions, being continuous in time and continuously differentiable in space, of the semilinear PDE

$$\begin{cases} \partial_t v(t,x) + \frac{1}{2} Tr\left(\Sigma(t) \nabla_x^2 v(t,x)\right) + H\left(t,x,v(t,x), \nabla_x v(t,x)\right) = 0, \; (t,x) \in [0,T[\times \mathbb{R}^d \\ v(T,\cdot) = g. \end{cases} \tag{3.18}$$

To formulate the result we consider the following assumption.

Assumption 2. *g is continuous and has polynomial growth.*

Let ν be a Gaussian Borel probability on $\mathbb{R}^d$ with mean $\bar{m}^\nu$ and covariance $\bar{Q}^\nu$. Let $t \mapsto m^\nu(t)$ defined in (3.3), $t \mapsto Q^\nu(t)$ be given by (3.4) and suppose that ν fulfills Assumption 1.

We fix a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$ on which are defined a d-dimensional Brownian motion β and a random vector ξ_0 distributed according to ν and independent of β.

Let ξ be the unique strong solution of

$$\xi_t = \xi_0 - \int_0^t \widetilde{b}(T-s,\xi_s) + b_c\left(T-s,\xi_s,m^\nu(T-s),Q^\nu(T-s)\right) ds + \int_0^t \sigma(T-s)\, d\beta_s, \; t \in [0,T[. \tag{3.19}$$

Remark 3.5. *(3.19) admits a unique strong solution on $[0,T[$ since its drift is affine with time-dependent continuous coefficients.*

Lemma 3.6. *1. The process $\widehat{\xi} := \xi_{T-\cdot}$ solves the SDE*

$$X_t = X_0 + \int_0^t \widetilde{b}(s,X_s)\, ds + \int_0^t \sigma(s)\, dW_s, \; t \in [0,T[, \tag{3.20}$$

where W is an $\mathcal{F}^{\widehat{\xi}}$-Brownian motion independent of $X_0 \sim \mathcal{N}(m^\nu(0), Q^\nu(0))$.

2. $\widehat{\xi}$ extends continuously to $[0,T]$.

Proof. i) The SDE (3.20) admits in particular existence in law. Let X be a solution of (3.20). To prove the first statement, we first show that the laws of $\widehat{\xi}$ and X coincide.

For this it is enough to prove that $\widehat{X} = X_{T-\cdot}$ and the solution ξ of (3.19) are identically distributed. By Problem 6.1 in Chapter 5 in [25] and by uniqueness of the ODE (3.1) (resp.

(3.2)) with initial condition $m^\nu(0)$ (resp. $Q^\nu(0)$), we get $\mathbb{E}(X_t) = m^\nu(t)$ and $\mathrm{Cov}(X_t) = Q^\nu(t)$ for all $t \in [0,T]$. By Problem 6.2, Chapter 5 in [25]) X is a Gaussian process so

$$\widehat{\xi}_t \sim \mathcal{N}\left(m^\nu(t), Q^\nu(t)\right),\ t \in [0,T]. \tag{3.21}$$

ENormXi

By (3.21) and Theorem 2.1 in [20], X is a solution (in law) of (3.19) on $[0,T[$. Pathwise uniqueness for (3.19) implies uniqueness in law on $[0,T[$ and the first statement of Lemma 3.6 is established.

ii) We proceed now with the proof of the first statement. Let X be a solution of (3.20), so that we know that W is a Brownian motion independent of X_0. On the other hand the process

$$M_t^X := X_t - X_0 - \int_0^t \tilde{b}(u, X_u)du,\ t \in [0,T].$$

is an $\mathcal{F}^X$-martingale with quadratic variation

$$[M^X, (M^X)^\top] = \int_0^\cdot \Sigma(u)du.$$

We have

$$W \equiv \int_0^\cdot \sigma^{-1}(u) dM_u^X. \tag{3.22}$$

EM11

Since $[W, W^\top]_t \equiv tI_d$, by Lévy's characterization theorem, W is a standard $(\mathcal{F}_t^X)$-Brownian motion. We set

$$M_t^{\hat{\xi}} := \hat{\xi}_t - \hat{\xi}_0 - \int_0^t \tilde{b}(u, \hat{\xi}_u)du,\ t \in [0,T],$$

and we denote $W^{\hat{\xi}} := \int_0^\cdot \sigma^{-1}(u)dM_u^{\hat{\xi}}$. Taking i) into account and the fact that $\hat{\xi}$ and X have the same law, then W and $W^{\hat{\xi}}$ are identically distributed and so $W^{\hat{\xi}}$ is an $\mathcal{F}^{\hat{\xi}}$-Brownian motion. Moreover the couple (X_0, W) has the same distribution as $(\hat{\xi}_0, W^{\hat{\xi}})$. Consequently $W^{\hat{\xi}}$ is an $\mathcal{F}^{\hat{\xi}}$-standard Brownian motion (independent of $\hat{\xi}_0$) and the statement 1. follows.

iii) It remains to prove the second statement. For this we show

$$\mathbb{E}\left(\int_0^T \left| b_c\left(s, \widehat{\xi}_s, m^\nu(s), Q^\nu(s)\right)\right| ds\right) < \infty. \tag{3.23}$$

finiteMean

On the one hand, for all $t \in]0,T]$,

$$\begin{aligned}\left|b_c\left(t, \widehat{\xi}_t, m^\nu(t), Q^\nu(t)\right)\right| &= \left|\Sigma(t)\sqrt{Q^\nu(t)^{-1}}\sqrt{Q^\nu(t)^{-1}}\left(\widehat{\xi}_t - m^\nu(t)\right)\right| \\ &\le ||\Sigma||_\infty \sqrt{\left|\left|Q^\nu(t)^{-1}\right|\right|}\left|\sqrt{Q^\nu(t)^{-1}}\left(\widehat{\xi}_t - m^\nu(t)\right)\right| \\ &= \frac{||\Sigma||_\infty}{\sqrt{||Q^\nu(t)||}}\left|\sqrt{Q^\nu(t)^{-1}}\left(\widehat{\xi}_t - m^\nu(t)\right)\right|,\end{aligned}$$

remembering that $Q^\nu(t)$ belongs to $S_d^{++}(\mathbb{R})$.

On the other hand, by (3.21) $\left|\sqrt{Q^{\nu}(t)^{-1}}\left(\widehat{\xi}_t - m^{\nu}(t)\right)\right| \sim |Z|$ where $Z \sim \mathcal{N}(0, I_d)$. Then, (3.23) is verified if we show

$$\int_0^T \frac{1}{\sqrt{||Q^{\nu}(t)||}} dt < \infty. \tag{3.24}$$

If $Q^{\nu}(0) = 0$, then for all $t \in]0,T]$, for all $t \in]0,T]$, Remark 3.1 implies

$$\frac{Q^{\nu}(t)}{t} = \mathcal{A}(t)\left(\frac{1}{t}\int_0^t \mathcal{A}(s)^{-1}\Sigma(s)\left(\mathcal{A}(s)^{-1}\right)^{\top} ds\right)\mathcal{A}(t)^{\top} \underset{t\to 0}{\longrightarrow} \Sigma(0).$$

If $Q^{\nu}(0) \neq 0$, then for all $]0,T]$, again Remark 3.1 yields

$$\frac{||Q^{\nu}(t)||}{t} \geq \left|\left|\left|\mathcal{A}(t)\frac{Q^{\nu}(0)}{t}\mathcal{A}(t)^{\top}\right|\right| - \left|\left|\mathcal{A}(t)\left(\frac{1}{t}\int_0^t \mathcal{A}(s)^{-1}\Sigma(s)\left(\mathcal{A}(s)^{-1}\right)^{\top} ds\right)\mathcal{A}(t)^{\top}\right|\right|\right| \underset{t\to 0}{\longrightarrow} +\infty,$$

where we have also used the fact $\mathcal{A}(0) = I_d$ and the fact $\frac{1}{t}\int_0^t \mathcal{A}(s)^{-1}\Sigma(s)\left(\mathcal{A}(s)^{-1}\right)^T ds$ tends to $\Sigma(0)$ as t tends to 0 thanks to the continuity of $\Sigma, \mathcal{A}^{-1}$ on $[0,T]$.

Hence, for all $t \in]0,T]$,

$$\lim_{t\to 0}\frac{\sqrt{t}}{\sqrt{||Q^{\nu}(t)||}} = \begin{cases} \frac{1}{\sqrt{||\Sigma(0)||}}, \text{ if } Q^{\nu}(0) = 0 \\ 0, \text{ otherwise.} \end{cases} \tag{3.25}$$

This yields (3.24) which implies (3.23); consequently the solution X of (3.19) prolongates to $t = T$ and item 2. is proved.

□

Though, this will not be exploited in the algorithm proposed at Section 5, it is interesting to note that the process ξ introduced in (3.19) can also be seen as the solution of a McKean SDE. Proposition 3.7 below shows that (1.7) admits existence and uniqueness if and only if Assumption 1 is verified. In particular we have the following.

Proposition 3.7. *1. There is at most one solution (ξ, m, Q) of (1.7).*

2. Suppose the validity of Assumption 1. Let ξ be the unique solution of (3.19). Then (ξ, m^{ν}, Q^{ν}) is a solution of (1.7).

Proof. 1. Let (ξ, m, Q) be a solution of (1.7). By definition, ξ solves an SDE of type (3.20) replacing a by $a^{\Sigma} : s \mapsto -a(T-s) - \Sigma(T-s)Q(T-s)^{-1}$ and c by $c^{\Sigma} : s \mapsto -c(T-s) + \Sigma(T-s)Q(T-s)^{-1}m(T-s)$.

By Problem 6.1 Section 5 in [25], the function $t \mapsto \mathbb{E}(\xi_t) (= m(T-t))$ (resp. $t \mapsto \text{Cov}(\xi_t) (= Q(T-t))$) solves the first line of (3.1) (resp. (3.2)) replacing a by a^{Σ} and c by c^{Σ}. Then, the following identities hold for all $t \in]0,T]$:

$$m(T-t) = \mathbb{E}(\xi_0) - \int_0^t a(T-s)m(T-s) + c(T-s)\, ds, \tag{3.26}$$

$$Q(T-t) = \mathrm{Cov}(\xi_0) - \int_0^t Q(T-s)\, a(T-s)^\top + a(T-s)\, Q(T-s)\, ds - \int_0^t \Sigma(T-s)\, ds, \tag{3.27}$$

Id-rev-Q

remarking that

$$a^\Sigma(t)\, m(T-t) + c^\Sigma(t) = -a(T-t)\, m(T-t) - c(T-t),$$

$$Q(T-t)\, a^\Sigma(t)^\top + a^\Sigma(T-t)\, Q(t) = -Q(T-t)\, a(T-t)^\top - a(T-t)\, Q(T-t) - 2\Sigma(T-t).$$

Applying the change of variable $t \mapsto T-t$ in identities (3.26) and (3.27), we show that m (resp. Q) solves the backward ODE (3.1) (resp. (3.2)), which is well-posed. We recall that ξ_0 is distributed according to ν. Then, $m = m^\nu$ and $Q = Q^\nu$, see the beginning of Section 3.1. As a consequence, ξ solves (3.19) and is uniquely determined thanks to Remark 3.5. This shows the validity of item 1.

2. Let ξ be the unique solution of (3.19). Then, the time-reversed process $\widehat{\xi}$ solves (3.20) and $\xi_T \sim \mathcal{N}(m^\nu(0), Q^\nu(0))$, thanks to item 1. of Lemma 3.6. Now, by Remark 3.4, we have $\mathbb{E}\left(\widehat{\xi}_t\right) = m^\nu(t)$, $\mathrm{Cov}\left(\widehat{\xi}_t\right) = Q^\nu(t)$ for all $t \in [0, T[$. This concludes the proof of item 2.

□

RIOR **Remark 3.8.** *1. In [23] we have discussed existence and uniqueness of more general McKean problems involving the densities of the marginal laws instead of expectation and covariance matrix, where the solution is the time-reversal of some (not necessarily Gaussian) diffusion.*

2. In particular, in Section 4.5 of [23] we have investigated existence and uniqueness of

$$\begin{cases} Y_t = Y_0 - \int_0^t \tilde{b}(T-r, Y_r)\, dr + \int_0^t \left\{ \frac{div_y\left(\Sigma_{i.}(T-r)\, p_r(Y_r)\right)}{p_r(Y_r)} \right\}_{i \in [\![1,d]\!]} dr + \int_0^t \sigma(T-r)\, d\beta_r, \\ p_t \text{ density law of } \mathbf{p_t} = \text{law of } Y_t, t \in]0, T[, \\ Y_0 \sim \mathbf{p_T} = \nu, \end{cases} \tag{3.28}$$

MKIntro

where β is a m-dimensional Brownian motion and $\Sigma = \sigma\sigma^\top$, whose solution is the couple $(Y, \mathbf{p})$. Moreover, when the solution exists, there is a probability-valued function $\mathbf{u}$ defined on $[0, T]$ solution of the Fokker-Planck equation

$$\begin{cases} \partial_t \mathbf{u} &= \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}^2 \left((\sigma\sigma^\top)_{i,j}(t) \mathbf{u} \right) - div \left(\tilde{b}(t,x) \mathbf{u} \right) \\ \mathbf{u}(T) &= \nu. \end{cases} \tag{3.29}$$

EDPTerm0Bis

3. Suppose that ν is a Gaussian law on $\mathbb{R}^d$. It is possible to show that Assumption 1 is equivalent to the existence of a probability-valued solution $\mathbf{u}$ of (3.29). In this case the McKean problems (1.7) and (3.28) are equivalent. In particular the component Y of the solution of (3.28) is Gaussian.

We continue with a preliminary lemma. Let W be a Brownian motion. For each $(s,x) \in [0,T] \times \mathbb{R}^d$, $X^{s,x}$ will denote below the process

$$X_t^{s,x} := x + \int_s^t \sigma(r) dW_r,\ t \in [s,T].$$

Lemma 3.9. *Suppose the validity of Assumption 2. Let $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ of class $\mathcal{C}^{0,1}\left([0,T], \mathbb{R}^d\right)$, with polynomial growth and such that the function $H^v : (t,x) \mapsto H\left(t,x,v\left(t,x\right), \nabla_x v\left(t,x\right)\right)$ is continuous with polynomial growth. Then, the following assertions are equivalent.*

1. *v is a viscosity solution of (3.18).*
2. *For each $(s,x) \in [0,T] \times \mathbb{R}^d$,*
$$v\left(s,x\right) = \mathbb{E}\left(\int_s^T H\left(r, X_r^{s,x}, v\left(r, X_r^{s,x}\right), \nabla_x v\left(r, X_r^{s,x}\right)\right) dr + g\left(X_T^{s,x}\right)\right). \tag{3.30}$$
3. *v is of class $\mathcal{C}^{1,2}\left([0,T[, \mathbb{R}^d\right)$ and is a (classical) solution of (3.18).*

Proof. Let v as in the lemma statement.

a) We set

$$w^v(s,x) := \mathbb{E}\left(g(X_T^{s,x}) + \int_s^T H^v(r, X_r^{s,x}) dr\right),\ (s,x) \in [0,T[\times \mathbb{R}^d. \tag{3.31}$$

We show first that w^v is a (classical) solution in $\mathcal{C}^{1,2}\left([0,T[, \mathbb{R}^d\right) \cap \mathcal{C}^0\left([0,T] \times \mathbb{R}^d\right)$ with polynomial growth of the linear PDE

$$\begin{cases} \partial_t w\left(t,x\right) + \frac{1}{2} Tr[\sigma\sigma^\top(t) \nabla_x^2 w\left(t,x\right)] + H^v\left(t,x\right) = 0,\ (t,x) \in [0,T[\times \mathbb{R}^d \\ w(T,\cdot) = g. \end{cases} \tag{3.32}$$

Indeed w^v can be rewritten as

$$w\left(s,x\right) = \int_{\mathbb{R}^d} g\left(z\right) p_T\left(s, z-x\right) dz + \int_s^T \int_{\mathbb{R}^d} H^v\left(r,z\right) p_r\left(s, z-x\right) dz dr,\ (s,x) \in [0,T[\times \mathbb{R}^d, \tag{3.33}$$

where for each $r \in [0,T]$, $s \in [0,r[$, $p_r\left(s,\cdot\right)$ is the density of the r.v. $\int_s^r \sigma(u) dW_u$, i.e. a Gaussian r.v. with mean zero and covariance $\int_s^r \Sigma\left(u\right) du$. Moreover, it is well-known, see e.g. Remark 3.2 in [9], that for each $r \in [0,T]$, $p_r : [0,r[\times \mathbb{R}^d \to \mathbb{R}$ is a smooth solution of

$$\partial_t p_r(t,z) + \frac{1}{2} Tr\left(\Sigma\left(t\right) \nabla_x^2 p_r\left(t,z\right)\right) = 0, (t,z) \in [0,r[\times \mathbb{R}^d. \tag{3.34}$$

Consequently, by usual integration theorems allowing to commute derivation and integrals, one shows (3.32).

b) Consequently w^v is a viscosity solution (3.32).

c) If 1. holds then v is also a viscosity solution of (3.32). By point 1. of Remark 3.12, equation (3.32) admits at most one continuous viscosity solution with polynomial growth. So $v = w^v$ which means 2.

d) If 2. holds then $v = w^v$ and by b) v is a viscosity solution of (3.32) and therefore of (3.18).

e) 3. implies obviously 1. Viceversa, if item 1. holds, a) implies that w^v is a classical solution of (3.32); b) and the uniqueness of viscosity solutions for previous linear equation implies $v = w^v$ and finally item 3.

□

We state now the announced representation result.

Theorem 3.10. *Suppose the validity of Assumption 2. Let ν be a Gaussian probability fulfilling Assumption 1 with associated functions m^ν and Q^ν.*
Let $v \in \mathcal{C}^{0,1}\left([0,T], \mathbb{R}^d; \mathbb{R}\right)$ with polynomial growth and such that $H^v : (t,x) \mapsto H\left(t, x, v\left(t,x\right), \nabla_x v\left(t,x\right)\right)$ is continuous with polynomial growth. Then, v is a viscosity solution of (3.18) if and only if for all $t \in [0,T]$

$$\begin{cases} \xi_t = \xi_0 - \displaystyle\int_0^t \widetilde{b}\left(T-s, \xi_s\right) + b_c\left(T-s, \xi_s, m^\nu\left(T-s\right), Q^\nu\left(T-s\right)\right) ds + \int_0^t \sigma\left(T-s\right) d\beta_s, \\ \xi_0 \sim \nu, \\ v\left(t, \widehat{\xi}_t\right) = \mathbb{E}\left(\displaystyle\int_t^T H\left(s, \widehat{\xi}_s, v\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right) - \left\langle \widetilde{b}\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right\rangle ds + g\left(\widehat{\xi}_T\right) \middle| \widehat{\xi}_t\right). \end{cases} \tag{3.35}$$

Remark 3.11. *The affine drift $\widetilde{b}$ remains a degree of freedom of the representation. In Section 5, in the framework of the Hamilton-Jacobi-Bellman PDEs are given elements to choose rationally $\widetilde{b}$.*

Remark 3.12. *We remark that previous representation (3.35) is valid even if uniqueness does not hold for the semilinear PDE (3.18). In that case even the equation (3.35) does not admit uniqueness. However, we provide below some typical situations for which (3.18) admits at most one viscosity solution, within different classes of solutions.*

1. *Suppose the validity of Assumption 2. Suppose also that H is continuous with polynomial growth in x and linear growth in (y,z). In addition, we suppose that H is Lipschitz in (y,z) uniformly in (t,x) and suppose that for all $R > 0$, there exists $m_R : \mathbb{R} \to \mathbb{R}^+$, tending to 0 at 0^+ such that*

$$\left|H\left(t, x', y, z\right) - H\left(t, x, y, z\right)\right| \leq m_R\left(\left|x' - x\right|\left(1 + |z|\right)\right),$$

for all $t \in [0,T]$, $z \in \mathbb{R}^d$ and $|x|, |x'|, |y| \leq R$. Then, by Theorem 5.1 in [31], implies that (3.18) admits at most one continuous viscosity solution with polynomial growth. In fact that theorem states uniqueness even in a wider class of solutions.

2. *The first theorem in [24] formulates a uniqueness result in a suitable class of bounded uniformly continuous solutions. Alternative assumptions are available to ensure uniqueness in different classes of unbounded functions, for fully non-linear parabolic Cauchy problems. See for instance Corollary 2 in [22], Theorem 3.1 in [29], [7], [21].*

Proof (of Theorem 3.10).

Let v as in the statement.

1. Lemma 3.6 implies that there exists an $\mathcal{F}^{\widehat{\xi}}$-Brownian motion W such that, under $\mathbb{P}$,

$$\widehat{\xi}_t = \widehat{\xi}_0 + \int_0^t \widetilde{b}\left(s, \widehat{\xi}_s\right) ds + \int_0^t \sigma(s)\, dW_s,\ t \in [0,T], \tag{3.36}$$

where $\widehat{\xi}_0 \sim \mathcal{N}\left(m^\nu(0), Q^\nu(0)\right)$. In particular

$$\mathbb{E}\left(\sup_{t\in[0,T]} \left|\widehat{\xi}_s\right|^p\right) < \infty,\ \forall p \geq 1. \tag{3.37}$$

This, together with Assumption 2 and the polynomial growth of H^v also imply that the r.v.

$$\int_0^T H\left(s, \widehat{\xi}_s, v\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right) - \left\langle \widetilde{b}\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right\rangle ds + g(\widehat{\xi}_T)$$

is square integrable.

2. We give now an equivalent formulation of (3.35) using a change of probability measure.

We set $L_s := \sigma(s)^{-1}\widetilde{b}\left(s, \widehat{\xi}_s\right),\ s \in [0,T]$. We denote by $\mathbb{Q}$, the probability equivalent to $\mathbb{P}$ on $\mathcal{F}_T^{\widehat{\xi}}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\left(-\sum_{i=1}^d \int_0^\cdot L_s^i dW_s^i\right)_T$, being well-defined thanks to Lemma 7.1.

The goal is to show that v fulfills (3.35) if and only if it fulfills for all $t \in [0,T]$

$$v\left(t, \widehat{\xi}_t\right) = \mathbb{E}_{\mathbb{Q}}\left(\int_t^T H\left(s, \widehat{\xi}_s, v\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right) ds + g\left(\widehat{\xi}_T\right) \middle| \widehat{\xi}_t\right). \tag{3.38}$$

We remark that,

$$\widehat{\xi}_t = \widehat{\xi}_0 + \int_0^t \sigma(s)\, d\widetilde{W}_s,\ t \in [0,T], \tag{3.39}$$

where

$$\widetilde{W} := W + \int_0^\cdot L_s ds, \tag{3.40}$$

which is a Brownian motion under $\mathbb{Q}$ thanks to Girsanov's Theorem 5.1 in [25]. By item 1.

$$\int_0^T H\left(s, \widehat{\xi}_s, v(s, \widehat{\xi}_s), \nabla_x v\left(s, \widehat{\xi}_s\right)\right) ds + g(\widehat{\xi}_T),$$

is obviously also square integrable under $\mathbb{Q}$.

We set $H_s := H\left(s, \widehat{\xi}_s, v\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right),\ s \in [0,T]$, for the sake of brevity.

We remark first that for each given $s \in [0,T]$,

$$\left\langle \widetilde{b}\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right\rangle = \left\langle \sigma(s)\sigma(s)^{-1}\widetilde{b}\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right\rangle = \left\langle L_s, \sigma(s)^\top \nabla_x v\left(s, \widehat{\xi}_s\right)\right\rangle. \tag{3.41}$$

Then, (3.41) combined with the Markov property of $\widehat{\xi}$ implies that (3.35) is equivalent to

$$v\left(t,\widehat{\xi}_t\right)=\mathbb{E}\left(\int_t^T\left(H_s-\left\langle L_s,\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right\rangle\right)ds+g\left(\widehat{\xi}_T\right)\middle|\mathcal{F}_t^{\widehat{\xi}}\right),$$

which can be rewritten

$$v\left(t,\widehat{\xi}_t\right)=M_t-\int_0^t\left(H_s-\left\langle L_s,\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right\rangle\right)ds,$$

where M is the $\mathbb{P}$-martingale

$$M_t=\mathbb{E}\left(\int_0^T H_s-\left\langle L_s,\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right\rangle ds+g\left(\widehat{\xi}_T\right)\middle|\mathcal{F}_t^{\widehat{\xi}}\right),\ t\in[0,T]. \tag{3.42}$$

Similarly, (3.38) is equivalent to

$$v\left(t,\widehat{\xi}_t\right)=\bar{M}_t-\int_0^t H_s ds,$$

where $\bar{M}$ is the $\mathbb{Q}$-martingale

$$\bar{M}_t=\mathbb{E}_{\mathbb{Q}}\left(\int_0^T H_s ds+g\left(\widehat{\xi}_T\right)\middle|\mathcal{F}_t^{\widehat{\xi}}\right),\ t\in[0,T]. \tag{3.43}$$

To show the aforementioned equivalence, it suffices now to show

$$\bar{M}_t-M_t=\int_0^t\left\langle L_s,\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right\rangle ds,\ t\in[0,T].$$

On the one hand, Theorem 1.7 Chapter 8 in [33] implies that the process $\widetilde{M}:=M+\sum_{i=1}^d[M,\int_0^\cdot L_s^i dW_s^i]$ is a $\mathbb{Q}$-local martingale. On the other hand, for each $i\in[\![1,d]\!]$ by Proposition 3.10 in [19] we have

$$\begin{aligned}[M,\int_0^\cdot L_s^i dW_s^i]&=[v\left(\cdot,\widehat{\xi}\right),\int_0^\cdot L_s^i dW_s^i]\\&=\int_0^\cdot L_s^i\left(\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right)_i ds,\end{aligned}$$

combining (3.43) with the usual properties of covariation for semimartingales. This means that

$$\widetilde{M}=M+\int_0^\cdot\left\langle L_s,\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right\rangle ds$$

is a $\mathbb{Q}$-local martingale. Now,

$$\begin{aligned}\widetilde{M}_T&=M_T+\int_0^T\left\langle L_s,\sigma\left(s\right)^\top\nabla_x v\left(s,\widehat{\xi}_s\right)\right\rangle ds\\&=\int_0^T H_s ds+g\left(\widehat{\xi}_T\right),\end{aligned}$$

thanks to (3.42). Since $\bar{M}$ and $\widetilde{M}$ are $\mathbb{Q}$-local martingales being equal at $t=T$, we have $\bar{M}=\widetilde{M}$. This shows the validity of point 2.

3. For each $(s,x) \in [0,T] \times \mathbb{R}^d$, we set $X^{s,x} := x + \int_s^{\cdot} \sigma(r)\, d\widetilde{W}_r$ where $\widetilde{W}$ is the $\mathbb{Q}$-Brownian motion defined in (3.40). Associated with v, we consider the continuous function

$$w^v(t,x) := \mathbb{E}_{\mathbb{Q}} \left(\int_t^T H\left(r, X_r^{t,x}, v\left(r, X_r^{t,x}\right), \nabla_x v\left(r, X_r^{t,x}\right)\right) dr + g\left(X_T^{t,x}\right) \right),\ (t,x) \in [0,T] \times \mathbb{R}^d.$$

   We observe that v fulfills (3.38) if and only if for all $(t,x) \in [0,T] \times \mathbb{R}^d$

$$v(t,x) = w^v(t,x). \tag{3.44}$$

   Indeed this follows by the freezing lemma of the conditional expectation, the fact that $\widehat{\xi}_t$ is independent of the random field $\left(X_s^{t,x}\right)_{t \le s \le T, x \in \mathbb{R}^d}$ and the flow property

$$X_s^{t,\widehat{\xi}_t} = \widehat{\xi}_s,\ s \in [t,T].$$

4. It remains to show that (3.44) is satisfied if and only if v is a viscosity solution of (3.18). This is the object of Lemma 3.9 applied under the probability $\mathbb{Q}$, in particular to the equivalence between item 1. and item 3.

□

4 Representation of stochastic control problems

Let us briefly recall the link between stochastic control and non-linear PDEs given by the Hamilton-Jacobi-Bellman (HJB) equation. We refer for instance to [11, 32, 37] for more details.

Let $A \subset \mathbb{R}^k$ compact and denote by $\mathcal{A}_0$ the set of all A-valued progressively measurable processes $(\alpha_t)_{t \in [0,T]}$, namely the set of *admissible controls*.

We consider now *state processes* $(X_t^{s,x,\alpha})_{s \le t \le T, \alpha \in \mathcal{A}_0}$ starting at time $s \in [0,T]$ with value $x \in \mathbb{R}^d$, solutions of the controlled SDE

$$dX_t = b(t, X_t, \alpha_t)\, dt + \sigma(t)\, dW_t, \tag{4.1}$$

where W is a d-dimensional Brownian motion and $b : [0,T] \times \mathbb{R}^d \times A \mapsto \mathbb{R}^d$ is supposed to fulfill the following.

Assumption 3. *The function b is continuous and there exists $K \ge 0$ such that*

$$|b(t, x_2, a) - b(t, x_1, a)| \le K |x_2 - x_1|,\ (t, x_1, x_2, a) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \times A.$$

Note that Assumption 3 implies b to have linear growth in space uniformly in time and in the control. Consequently, (4.1) starting at time s with value x admits a unique solution for each $\alpha \in \mathcal{A}_0$, for each $(s,x) \in [0,T] \times \mathbb{R}^d$, by the same arguments as in Theorem 3.1 in [37].

We also introduce the *cost function* $J : [0,T] \times \mathbb{R}^d \times \mathcal{A}_0 \to \mathbb{R}$ defined by

$$J(s,x,\alpha) := \mathbb{E}\left(g(X_T^{s,x,\alpha}) + \int_s^T f\big(r, X_r^{s,x,\alpha}, \alpha_r\big) dr \right), \ (s,x,\alpha) \in [0,T] \times \mathbb{R}^d \times \mathcal{A}_0, \tag{4.2}$$

where the function $f : [0,T] \times \mathbb{R}^d \times A \mapsto \mathbb{R}$ (*running cost*) is supposed to fulfill what follows.

Assumption 4. *The function f is continuous and there exists $m, M \geq 0$ such that*

$$|f(t,x,a)| \leq M(1 + |x|^m), \ (t,x,a) \in [0,T] \times \mathbb{R}^d \times A.$$

Supposing the validity of Assumptions 3 and 4 together with Assumption 2 on the function $g : \mathbb{R}^d \mapsto \mathbb{R}$ (*terminal cost*), we are interested in minimizing, over control processes $\alpha \in \mathcal{A}_0$ the functions $\alpha \mapsto J(0,x,\alpha)$ for every $x \in \mathbb{R}^d$.

To tackle this finite horizon stochastic control problem, the usual approach consists in introducing the associated *value* (or *Bellman*) function $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ representing the minimum expected costs, starting from any time $t \in [0,T]$ at any state $x \in \mathbb{R}^d$, i.e.

$$v(t,x) := \inf_{\alpha \in \mathcal{A}_0} J(t,x,\alpha), \ (t,x) \in [0,T] \times \mathbb{R}^d. \tag{4.3}$$

Note that the terminal condition is known, which fixes $v(T,\cdot) = g$, whereas $v(0,\cdot)$ corresponds to the solution of the original minimization problem.

Remark 4.1. *Suppose the validity of Assumptions 2, 3 and 4.*

1. *The function v is continuous on $[0,T] \times \mathbb{R}^d$ and has polynomial growth, see Theorem 5. Chapter 3. in [26].*

2. *The value function v is a viscosity solution of the Hamilton-Jacobi-Bellman equation*

$$\begin{cases} \partial_t v(t,x) + H(t,x,\nabla_x v(t,x)) + \frac{1}{2} Tr[\sigma\sigma^\top(t)\nabla_x^2 v(t,x)] = 0, \ (t,x) \in [0,T[\times\mathbb{R}^d \\ v(T,\cdot) = g, \end{cases} \tag{4.4}$$

   where H denotes the real-valued function defined on $[0,T] \times \mathbb{R}^d \times \mathbb{R}^d$ by

$$H(t,x,\delta) := \inf_{a \in A} \{ f(t,x,a) + \langle b(t,x,a), \delta \rangle \}, \ (t,x,\delta) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^d, \tag{4.5}$$

   see for example Theorem 7.4 in [37].

3. *By definition, it is obvious that $(x,z) \mapsto H(t,x,z)$ has polynomial growth uniformly with respect to t. It is also clear that H is continuous.*

4. *Under Assumptions 2, 3 and 4, the PDE* (4.4) *admits at most one viscosity solution in the class of continuous solutions with polynomial growth, see Theorem* II.3 *in [28]. Since v has polynomial growth, the value function v is the unique viscosity solution of* (4.4) *in the considered class.*

We formulate below another assumption for the value function v.

_control **Assumption 5.** *v is of class $\mathcal{C}^{0,1}\left([0,T],\mathbb{R}^d\right)$ such that $\nabla_x v$ has polynomial growth.*

RControl **Remark 4.2.** *Under Assumption 5, using Remark 4.1 3. that the function $(t,x) \mapsto H^v(t,x) := H(t,x,\nabla_x v(t,x))$ is continuous with polynomial growth.*

ntrolBis **Remark 4.3.** *1. Assumption 5 is not so restrictive, since whenever g and f are locally Lipschitz with polynomial growth gradient (in space), then v is locally Lipschitz in the space variable. To prove this, it suffices to show that J is locally Lipschitz in x uniformly in t and α. A proof of this fact is given in Lemma 7.2 stated in the Appendix.*

In that context, the value function v is in particular absolutely continuous and for every $t \in [0,T]$, for almost every $x \in \mathbb{R}^d$, $v(t,\cdot)$ is differentiable and $\nabla_x v(t,\cdot)$ exists.

2. Suppose in addition that the functions f, g and b are of class $\mathcal{C}^1$ (in the space variable) and the validity of Assumption 6. Then $\nabla_x v(t,\cdot)$ has polynomial growth as we show below. Indeed, by usual dominated convergence arguments, we can show that for each $(t,\alpha) \in [0,T] \times \mathcal{A}_0$, $x \mapsto J(t,x,\alpha)$ is differentiable with gradient

$$\nabla_x J(t,x,\alpha) = \mathbb{E}\left(Y_T^{t,x,\alpha} \nabla_x g\left(X_T^{t,x,\alpha}\right) + \int_t^T Y_r^{t,x,\alpha} \nabla_x f\left(r, X_r^{t,x,\alpha}, \alpha_r\right) dr\right), \qquad (4.6)$$

obj_func_gra

where $Y^{t,x,\alpha}$ is the unique matrix-valued process fulfilling

$$Y_r^{t,x,\alpha} = I_d + \int_t^r \nabla_x b\left(s, X_s^{t,x,\alpha}, \alpha_s\right) Y_s^{t,x,\alpha} ds, \ r \in [t,T],$$

where $\nabla_x b := \left(\partial_{x_j} b^i\right)_{(i,j) \in [\![1,d]\!]^2}$.

Combining what precedes with Lemma 7.3 stated in the Appendix, we deduce that for all $t \in [0,T]$, for almost every $x \in \mathbb{R}^d$

$$\nabla_x v(t,x) = \nabla_x J(t,x,\alpha^*(t,x)), \qquad (4.7)$$

val_func_gra

where α^* is the Borel function introduced in Assumption 6. In view of (4.6) and (4.7), $\nabla_x v$ has polynomial growth.

aControl **Corollary 4.4.** *Let ν be a Gaussian probability measure fulfilling Assumption 1 with associated functions m^ν and Q^ν. We suppose the validity of Assumptions 2, 3, 4. Among the functions $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ fulfilling Assumption 5, the value function is the unique one which is solution of (3.35). (In this framework H only depends on $\nabla_x v$ and not on v).*

Proof. We recall that H^v has polynomial growth by Remark 4.2 1. Otherwise, on the one hand, by Remark 4.1 and the direct implication in Theorem 3.10, v fulfills (3.35). On the other hand, if a function v fulfills (3.35) then, by the converse implication of Theorem 3.10 v is a viscosity solution of (4.4). By Remark 4.1 3., v can only be the value function.

□

We introduce a supplementary hypothesis on the value function v.

Assumption 6. *There exists a Borel function $\alpha^* : [0,T] \times \mathbb{R}^d \to A$ such that*

$$H\left(t,x,\nabla_x v\left(t,x\right)\right) = \left\langle b\left(t,x,\alpha^*\left(t,x\right)\right), \nabla_x v\left(t,x\right)\right\rangle + f\left(t,x,\alpha^*\left(t,x\right)\right),\ (t,x) \in [0,T] \times \mathbb{R}^d.$$

We state (and show below) a *verification* type result involving α^* without any further regularity assumptions on the value function. That result is somehow classical, but it is not obvious to find it in the literature (see e.g. Chapter 5 of [37] or [18]), with our assumptions. So, for the consistency of the paper we provide a proof. Note to begin that the Borel function $b^* : (t,x) \to b\left(t,x,\alpha^*\left(t,x\right)\right)$ has linear growth thanks to Assumption 3. As a consequence, the *closed loop* equation

$$d\bar{X}_t = b^*\left(t,\bar{X}_t\right)dt + \sigma\left(t\right)dW_t, \tag{4.8}$$

admits a unique strong solution $\bar{X}^x$ starting at time 0 with value x, for each $x \in \mathbb{R}^d$, see Theorem 6 in [38].

Proposition 4.5. *Suppose the validity of Assumptions 2, 3, 4. Let v be the value function defined in (4.3) supposed to be of class $\mathcal{C}^{0,1}$ such that $(t,x) \mapsto H\left(t,x,\nabla_x v\left(t,x\right)\right)$ has polynomial growth.*
Then, the Borel function α^* introduced in Assumption 6 defines an optimal feedback function for the considered control problem in the sense that for each $x \in \mathbb{R}^d$,

$$v\left(0,x\right) = J\left(0,x,\alpha^*\left(\cdot,\bar{X}^x\right)\right). \tag{4.9}$$

Proof (of Proposition 4.5). Let $x \in \mathbb{R}^d$.

1. v is a continuous viscosity solution with polynomial growth of (4.4) and so of (3.18), with $(t,x,y,z) \mapsto H\left(t,x,z\right)$ for the non linearity. By Remark 4.1 we know that H^v is continuous and by assumption it has polynomial growth. So we apply Lemma 3.9 to deduce that v is of class $\mathcal{C}^{1,2}\left([0,T[,\mathbb{R}^d\right)$ and is a classical solution of (4.4).

2. Applying Itô's formula to $v\left(\cdot,\bar{X}^x\right)$ between 0 and $T_0 \in [0,T[$ and using the fact v is a classical solution of (4.4) combined with Assumption 6, we obtain

$$v\left(0,x\right) = v\left(T_0,\bar{X}^x_{T_0}\right) + \int_0^{T_0} f\left(r,\bar{X}^x_r,\alpha^*\left(r,\bar{X}^x_r\right)\right)dr - M_{T_0}, \tag{4.10}$$

where

$$M_t = \int_0^t \nabla_x v\left(r,\bar{X}^x_r\right)^\top \sigma\left(r\right)dW_r,\ t \in [0,T[.$$

By the usual BDG (Burkholder-Davies-Gundy) and Jensen's arguments, $\sup_{t\in[0,T]}|\bar{X}^x_t|$ has all its moments. So, (4.10) implies that the local martingale M extends continuously to a true martingale on $[0,T]$ still denoted by M verifying $\sup_{t\in[0,T]}|M_t| \in L^1$. Indeed v is continuous on $[0,T] \times \mathbb{R}^d$ and v (resp. f) has polynomial growth in space (resp. in the second and third variable). Therefore M is a true martingale. Sending T_0 to T, (4.10) holds with T_0 replaced by T and $v\left(T_0, \bar{X}^x_{T_0}\right)$ replaced by $g\left(\bar{X}^x_T\right)$. Taking the expectation, we obtain

$$v(0,x) = \mathbb{E}\left(g\left(\bar{X}^x_T\right) + \int_0^T f\left(r, \bar{X}^x_r, \alpha^*\left(r, \bar{X}^x_r\right)\right) dr\right). \tag{4.11}$$

3. The process $\alpha^*_t := \alpha^*\left(t, \bar{X}^x_t\right),\ t \in [0,T]$, belongs to the set $\mathcal{A}_0$ of admissible controls and $X = \bar{X}^x$, is a solution of (4.1). Invoking pathwise uniqueness for (4.1), we obtain X^{0,x,α^*} coincides with $\bar{X}^x$. Then, (4.11) implies (4.9).

□

We formulate now a corollary in which is given a representation formula for the value function v involving the optimal feedback function α^*.

Corollary 4.6. *Let ν be a Gaussian probability measure fulfilling Assumption 1 with associated functions m^ν and Q^ν. We suppose the validity of Assumptions 2, 3, 4. Among the functions fulfilling Assumptions 5 and 6, the value function v is the unique one which is solution of*

$$\begin{cases} \xi_t = \xi_0 - \int_0^t \widetilde{b}\left(T-s, \xi_s\right) + b_c\left(T-s, \xi_s, m^\nu\left(T-s\right), Q^\nu\left(T-s\right)\right) ds + \int_0^t \sigma\left(T-s\right) d\beta_s, \\ \xi_0 \sim \nu, \\ v\left(t, \widehat{\xi}_t\right) = \mathbb{E}\left(\int_t^T f\left(s, \widehat{\xi}_s, \alpha^*\left(s, \widehat{\xi}_s\right)\right) - \left\langle \widetilde{b}\left(s, \widehat{\xi}_s\right) - b^*\left(s, \widehat{\xi}_s\right), \nabla_x v\left(s, \widehat{\xi}_s\right)\right\rangle ds + g\left(\widehat{\xi}_T\right) \middle| \widehat{\xi}_t\right), \end{cases} \tag{4.12}$$

for all $t \in [0,T]$.

Proof. The result is a direct consequence of Corollary 4.4, replacing the function H by its expression given in Assumption 6. □

5 A heuristic algorithm

In this section, we propose a heuristic algorithm to solve the control problem described in Section 4. In what follows, the *terminal cost* function g is supposed to belong to $\mathcal{C}^1\left(\mathbb{R}^d\right)$.
Consider a regular time grid with time step $\delta t := \frac{T}{n}$ and grid instants $t_k = k\delta t$ for any $k \in [\![0,n]\!]$. For $k = n-1, n-2, \cdots, 0$, select arbitrarily $\bar{m}_{k+1}, c_{k+1} \in \mathbb{R}^d$ and $\bar{Q}_{k+1} \in S^+_d(\mathbb{R}), a_{k+1} \in M_d(\mathbb{R})$ such that $Q_k(t_k) := e^{-a_{k+1}\delta t}\bar{Q}_{k+1}e^{-a^\top_{k+1}\delta t} - \int_{t_k}^{t_{k+1}} e^{-a_{k+1}(s-t_k)}\Sigma(s)\, e^{-a^\top_{k+1}(s-t_k)} ds \in S^+_d(\mathbb{R})$.
By Corollary 4.6, applied substituting $[0,T]$ with $[t_k, t_{k+1}]$, the solution of (4.4) on $[t_k, t_{k+1}]$, with

terminal condition $v(t_{k+1},\cdot)$, can be represented for $t\in[t_k,t_{k+1}]$ by

$$\begin{cases} \bar\xi_{k+1} & \sim \quad \mathcal{N}(\bar m_{k+1},\bar Q_{k+1}) \\ Y_{k+1} & = \quad v(t_{k+1},\bar\xi_{k+1}) \\ m_k(t) & = \quad e^{-a_{k+1}(t_{k+1}-t)}\bar m_{k+1} - \left(\int_t^{t_{k+1}} e^{-a_{k+1}(s-t)}ds\right) c_{k+1} \\ Q_k(t) & = \quad e^{-a_{k+1}(t_{k+1}-t)}\bar Q_{k+1} e^{-a_{k+1}^\top (t_{k+1}-t)} - \int_t^{t_{k+1}} e^{-a_{k+1}(s-t)}\Sigma(s)\, e^{-a_{k+1}^\top (s-t)}ds \\ \xi_{k,T-t} & = \quad \bar\xi_{k+1} - \int_{t_{n-(k+1)}}^{T-t} \big(a_{k+1}\xi_{k,s} + c_{k+1} + b_c(T-s,\xi_{k,s},m_k(T-s),Q_k(T-s))\big)\,ds \\ & \quad\quad + \int_{t_{n-(k+1)}}^{T-t}\sigma(T-s)d\beta_s \\ \hat\xi_{k,t} & = \quad \xi_{k,T-t} \\ v(t,\hat\xi_{k,t}) & = \quad \mathbb{E}\left(\int_t^{t_{k+1}} F_k\big(s,\hat\xi_{k,s},\nabla_x v(s,\hat\xi_{k,s})\big)ds + Y_{k+1}\middle|\hat\xi_{k,t}\right). \end{cases} \tag{5.1}$$

eq:discrete

In the above recursion, β denotes a d-dimensional Brownian motion on $[0,T]$; for any $k\in[\![0,n-1]\!]$, $(\xi_{k,t})_t$ is a d-dimensional process defined on $[t_{n-(k+1)},t_{n-k}]$ while $(\hat\xi_{k,t})_t$ denotes the associated time reversal defined on $[t_k,t_{k+1}]$; the driver F_k defined on $[t_k,t_{k+1}]\times\mathbb{R}^d\times\mathbb{R}^d$ is such that,

$$F_k(t,x,\delta) := H(t,x,\delta) - \langle a_{k+1}x + c_{k+1},\delta\rangle = \min_{a\in A}\{f(t,x,a) + \langle b(t,x,a),\delta\rangle\} - \langle a_{k+1}x+c_{k+1},\delta\rangle\,. \tag{5.2}$$

eq:Fk

The idea now is to apply a classical numerical method based on linear regressions to approximate the solution to (5.1) recursively in time from $k=n-1$ to $k=0$. For each time instant k, select arbitrarily $\bar m_{k+1}$, $c_{k+1}\in\mathbb{R}^d$ and $\bar Q_{k+1}\in S_d^+(\mathbb{R})$, $a_{k+1}\in M_d(\mathbb{R})$ such that

$$Q_k = e^{-a_{k+1}\delta t}\bar Q_{k+1}e^{-a_{k+1}^\top\delta t} - \Sigma(t_{k+1})\delta t \in S_d^+(\mathbb{R})\,. \tag{5.3}$$

eq:Q

Then we propose to approximate $v(t_k,\cdot)$ by v_k obtained by an explicit time discretization scheme of (5.1) with time step $\delta t = \frac{T}{n}$ as follows.

$$\begin{cases} \bar\xi_{k+1} & \sim \quad \mathcal{N}(\bar m_{k+1},\bar Q_{k+1}) \\ Y_{k+1} & = \quad v_{k+1}(\bar\xi_{k+1}) \\ \xi_k & = \quad \bar\xi_{k+1} - \big(a_{k+1}\bar\xi_{k+1} + c_{k+1} + b_c(t_{k+1},\bar\xi_{k+1},\bar m_{k+1},\bar Q_{k+1})\big)\,\delta t + \sigma(t_{k+1})\sqrt{\delta t}\varepsilon_k \\ v_k(\xi_k) & = \quad \mathbb{E}\left(F_k\big(t_{k+1},\bar\xi_{k+1},\nabla_x v_{k+1}(\bar\xi_{k+1})\big)\delta t + Y_{k+1}\middle|\xi_k\right), \end{cases} \tag{5.4}$$

eq:discreteb

where $(\varepsilon_k)_{0\le k\le n-1}$ are i.i.d. d-dimensional standard Gaussian variables. As in the classical literature, see e.g. [15], we propose to approximate the conditional expectation appearing in (5.4) using Monte-Carlo least squares regression based on a grid constituted by N independent simulations $(\xi_k^i,\bar\xi_{k+1}^i)_{1\le i\le N}$ for $k\in[\![0,n-1]\!]$. In that literature, one generally simulates forwardly that grid. The interest of such *fully backward* representations (5.1)-(5.4), where the grid $(\xi_k^i,\bar\xi_{k+1}^i)_{1\le i\le N}$ is defined backwardly in time, (like the value function), is twofold.

- In terms of computer memory: at each time instant $k+1$, the values of the grid are generated on the fly, $(\xi_k^i, \bar{\xi}_{k+1}^i)_{1 \le i \le N}$. Contrary to the standard approach, there is no need to store the whole grid over the whole set of grid instants $k \in [\![0, n-1]\!]$.

- In terms of the relevance of the grid: at each grid instant, $k+1$ the information acquired on the value function $v(t_{k+1}, \cdot)$ and optimal control strategy $\alpha^*(t_{k+1}, \cdot)$ can be used to adaptively optimize the grid parameter $(a_{k+1}, c_{k+1}, \bar{m}_{k+1}, \bar{Q}_{k+1})$ in order to explore relevant regions of the state space.

We develop some arguments to justify the relevance mentioned above. Indeed, as already announced, the target idea is to generate the grid used for regression computations according to the optimally controlled process dynamics. If this were possible, the sensitivity of the driver F_k w.r.t. the third variable $\nabla_x v$ would vanish. In fact the driver sensitivity w.r.t. $\nabla_x v$ is known to be one major cause of the propagation of numerical errors in approximation schemes, see e.g. [16]. Replacing $\nabla_x v_{k+1}(\bar{\xi}_{k+1})$ by a perturbation $\nabla_x v_{k+1}(\bar{\xi}_{k+1}) + h$ in the last equation of (5.4) we obtain

$$v_k^h(\xi_k) := \mathbb{E}\Big(F_k\big(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1}) + h\big)\delta t + Y_{k+1} \,|\, \xi_k\Big).$$

The impact on $v_k(\xi_k)$ can crudely be evaluated by computing the error $\mathbb{E}[|v_k^h(\xi_k) - v_k(\xi_k)|^2]$. Supposing that no perturbation is impacting Y_{k+1}, fact which will be heuristically justified in Remark 5.2 1., we have

$$\mathbb{E}(|v_k^h(\xi_k) - v_k(\xi_k)|^2) \quad \le \quad \mathbb{E}\Big(\big|F_k\big(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1})\big) - F_k\big(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1}) + h\big)\,\big|^2\Big).$$

Suppose from now on the existence of a Borel function $(t, x, \delta) \mapsto a^*(t, x, \delta)$, such that

$$H(t, x, \delta) := \{f(t, x, a^*(t, x, \delta)) + \langle b(t, x, a^*(t, x, \delta)), \delta\rangle\}\,,\ (t, x, \delta) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d. \tag{5.5}$$

In this case one has $\alpha^*(t, x) = a^*(t, x, \nabla_x v(t, x)),\ (t, x) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$, where α^* was defined in Assumption 6. Coming back to (5.2) we get

$$F_k(t, x, \delta) := H(t, x, \delta) - \langle a_{k+1}x + c_{k+1}, \delta\rangle = \{f(t, x, a^*(t, x, \delta)) + \langle b(t, x, a^*(t, x, \delta)), \delta\rangle\} - \langle a_{k+1}x + c_{k+1}, \delta\rangle. \tag{5.6}$$

A suitable application of the envelope theorem gives

$$\frac{\partial F_k}{\partial \delta}(t, x, \delta) = b(t, x, a^*(t, x, \delta)) - (a_{k+1}x + c_{k+1})\,, \tag{5.7}$$

which yields

$$\begin{aligned}
\mathbb{E}\left(|v_k^h(\xi_k) - v_k(\xi_k)|^2\right) \quad &\le \quad \mathbb{E}\Big|\langle \int_0^1 \frac{\partial F_k}{\partial \delta}(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1}) + \theta h)d\theta\,,\, h\rangle\Big|^2 \\
&= \quad \mathbb{E}\Big|\langle \int_0^1 b\big(t_{k+1}, \bar{\xi}_{k+1}, a^*(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1}) + \theta h)\big)d\theta - (a_{k+1}\bar{\xi}_{k+1} + c_{k+1})\,,\, h\rangle\Big|^2 \\
&\le \quad |h|^2 \mathbb{E}\Big|\int_0^1 b\big(t_{k+1}, \bar{\xi}_{k+1}, a^*(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1}) + \theta h)\big)d\theta - (a_{k+1}\bar{\xi}_{k+1} + c_{k+1})\Big|^2.
\end{aligned}$$

The above relation highlights the fact that the original idea consisting in generating the grid according to a dynamics approaching the optimally controlled process dynamics reduces the propagation of the error induced by the Monte-Carlo regression scheme in terms of least square criteria.

RFBSDEs **Remark 5.1.** *The above relation also shows that previous idea can be read in the more general perspective of the probabilistic representation of a solution v to a semilinear PDE of the type (1.1), via an FBSDE. In that general context, one expects the selected drift of the forward process in the FBSDE to reduce the impact of the sensitivity of the FBSDE driver with respect to $\nabla_x v$.*

Based on that observation, we propose a heuristic algorithm where parameters (a_{k+1}, c_{k+1}) are adaptively chosen as

$$(a_{k+1}, c_{k+1}) \in \arg\min_{a,c} \mathbb{E}\Big|b\big(t_{k+1}, \bar{\xi}_{k+1}, a^*(t_{k+1}, \bar{\xi}_{k+1}, \nabla_x v_{k+1}(\bar{\xi}_{k+1})\big) - (a\bar{\xi}_{k+1} + c)\Big|^2. \tag{5.8}$$ eq:ac

In the above algorithm, the random variables $(\varepsilon_k^i\,,\ k \in [\![0, n-1]\!]\,, i \in [\![1, N]\!])$ are i.i.d. according to $\mathcal{N}(0, I_d)$; $Proj_{S_d^+(\mathbb{R})} : S_d(\mathbb{R}) \mapsto S_d^+(\mathbb{R})$ denotes the Frobenius projection operator on the closed and convex space of semidefinite matrices; for each $p \in \mathbb{N}$, $P_p(\mathbb{R}^d)$ denotes the set of polynomial functions on $\mathbb{R}^d$ with degree p.

em:step5 **Remark 5.2.**

1. *Note that in Step 4, as soon as $Q_k \in S_d^+(\mathbb{R})$ then $(Y_{k+1}^i)_{1\le i\le N}$ results from the update made at previous iteration at Step 8. That updating rule corresponds to the multi-step forward dynamic programming approach [16] which is well-known for not inducing any additional bias error that would propagate backwardly during iterations. However, when $Q_k \notin S_d^+(\mathbb{R})$, in Step 4, then we have to modify $\bar{Q}_{k+1}$, re-generate new variables $(\xi_{k+1}^i)_{1\le i\le N}$ i.i.d. $\sim \mathcal{N}(\bar{m}_{k+1}, \bar{Q}_{k+1})$ and use the update $Y_{k+1}^i = v_{k+1}(\xi_{k+1}^i)$ which adds a bias error. Fortunately, in our numerical simulations it appeared easy to chose a first covariance matrix $\bar{Q}_n$ so that for all $k \in [\![0, n-1]\!]$ we had $Q_k \in S_d^+$. In that situation, the error propagation is only due to the sensitivity of the driver w.r.t. $\nabla_x v$ which is precisely minimized by our heuristics.*

2. *The complexity of Algorithm 1, is comparable to the traditional Monte-Carlo Regression scheme using a **forward grid**. Indeed, Algorithm 1 requires an additional linear regression calculation of order $\mathcal{O}(d^2N)$ at Step 2 which is negligible w.r.t. the polynomial regression computations at Step 7 (operated by both algorithms) inducing $\mathcal{O}(d^4N)$ operations in the specific case considered in simulations where the maximum degree of polynomials is $p = 2$. When $Q_k \notin S_d^+$, Algorithm 1 requires in addition, at Step 4, to implement: a Frobenius projection $Proj_{S_d^+(\mathbb{R})}(Q_k)$ ($\mathcal{O}(d^3)$), N multiplications of matrices $d \times d$ with vectors $d \times 1$ ($\mathcal{O}(d^2N)$); and N independent generations of d-dimensional Gaussian random variables. These additional operations induce a complexity of $\mathcal{O}(d^2N)$ which does not increase the original $\mathcal{O}(d^4N)$ complexity.*

3. *In terms of memory, as already mentioned, we do not have to store the whole regression grid on the whole time horizon constituted of ndN reals but only to consider dN reals at each instant.*

Algorithm 1 Fully Backward Monte-Carlo Regression scheme

Initialization Set $v_n = g$; $k = n-1$; select arbitrarily $(\bar{m}_n, \bar{Q}_n) \in \mathbb{R}^d \times S_d^+(\mathbb{R})$; generate $(\xi_n^i)_{1 \leq i \leq N}$ i.i.d. $\sim \mathcal{N}(\bar{m}_n, \bar{Q}_n)$; set $Y_n^i = g(\xi_n^i)$, for all $i \in [\![1, N]\!]$.

while $k \geq 0$ **do**

1. $\alpha^i{}_{k+1} = \underset{a \in A}{\arg\min} \left\{ f\left(t_{k+1}, \xi_{k+1}^i, a\right) + \left\langle b\left(t_k, \xi_{k+1}^i, a\right), \nabla_x v_{k+1}\left(\xi_{k+1}^i\right) \right\rangle \right\}$, for all $i \in [\![1, N]\!]$.
2. $(a_{k+1}, c_{k+1}) = \underset{(a,c) \in M_d(\mathbb{R}) \times \mathbb{R}^d}{\arg\min} \frac{1}{N} \sum_{i=1}^N \left| a\xi_{k+1}^i + c - b\left(t_{k+1}, \xi_{k+1}^i, \alpha_{k+1}^i\right) \right|^2$.
3. $\bar{m}_k = e^{-a_{k+1}\delta t} \bar{m}_{k+1} - c_{k+1}\delta t$.
4. $Q_k = e^{-a_{k+1}\delta t} \bar{Q}_{k+1} e^{-a_{k+1}^\top \delta t} - \Sigma\left(t_{k+1}\right)\delta t$.
   - **If $\mathbf{Q_k} \in \mathbf{S_d^+}(\mathbb{R})$:** set $\bar{Q}_k = Q_k$,
   - **Else** : set $\bar{Q}_k = Proj_{S_d^+(\mathbb{R})}(Q_k)$; recompute $\bar{Q}_{k+1} = e^{a_{k+1}\delta t}\left(\bar{Q}_k + \Sigma(t_{k+1})\delta t\right) e^{a_{k+1}^\top \delta t}$; regenerate $(\xi_{k+1}^i)_{1 \leq i \leq N}$ i.i.d. $\sim \mathcal{N}(\bar{m}_{k+1}, \bar{Q}_{k+1})$; set $Y_{k+1}^i = v_{k+1}(\xi_{k+1}^i)$, for all $i \in [\![1, N]\!]$.
5. Set $e_{k+1}^i = a_{k+1}\xi_{k+1}^i + c_{k+1} - b\left(t_{k+1}, \xi_{k+1}^i, \alpha_{k+1}^i\right)$, for all $i \in [\![1, N]\!]$.
6. $\xi_k^i = \xi_{k+1}^i - \left(a_{k+1}\xi_{k+1}^i + c_{k+1} + b_c\left(t_{k+1}, \xi_{k+1}^i, \bar{m}_{k+1}, \bar{Q}_{k+1}\right)\right)\delta t + \sigma\left(t_{k+1}\right)\varepsilon_k^i \sqrt{\delta t}$, for all $i \in [\![1, N]\!]$
7. $v_k = \underset{P \in P_p\left(\mathbb{R}^d\right)}{\arg\min} \frac{1}{N} \sum_{i=1}^N \left| Y_{k+1}^i + \left(f\left(t_{k+1}, \xi_{k+1}^i, \alpha_{k+1}^i\right) - \left\langle e_{k+1}^i, \nabla_x v_{k+1}\left(\xi_{k+1}^i\right)\right\rangle\right)\delta t - P\left(\xi_k^i\right) \right|^2$.
8. $Y_k^i = Y_{k+1}^i + \left(f\left(t_{k+1}, \xi_{k+1}^i, \alpha_{k+1}^i\right) - \left\langle e_{k+1}^i, \nabla_x v_{k+1}\left(\xi_{k+1}^i\right)\right\rangle\right)\delta t$, for all $i \in [\![1, N]\!]$
9. $k - 1 \leftarrow k$.

algo **end while**

Remark 5.3. *Suppose that at each time step $k \in [\![0, n-1]\!]$ the matrix Q_k belongs to $S_d^+(\mathbb{R})$. Then, Algorithm 1 is based on the representation formula appearing in Corollary 4.6, on the whole time interval $[0,T]$ with piecewise constant coefficients a, c such that $a(t), c(t) = a_{k+1}, c_{k+1}$ for each $t \in]t_k, t_{k+1}]$, for each $k \in [\![0, n-1]\!]$.*

6 Stochastic control of thermostatically controlled loads

6.1 Model description

With the massive integration of variable renewable energies (like wind farms or solar panels) into power systems, balancing supply and demand in a real time basis requires to develop new leverages. A technical solution is to develop load control schemes in order to automatically adapt consumption to generation. In this section, we propose to apply Algorithm 1 in order to control a large heterogeneous population of air-conditioners on a time horizon $[0,T]$ such that the overall consumption of the population follows a given target profile, while preserving the rooms temperatures within users comfort bounds.

We consider a hierarchical control scheme introduced in [6], where the population is aggregated into d clusters of N^i homogeneous loads (with same air-conditioners and rooms characteristics) for i$\in [\![1, d]\!]$. For each cluster $i \in [\![1, d]\!]$, a *local controller* decides at each time step to turn ON or OFF optimally some air-conditioners of cluster i, in order to satisfy a *prescribed proportion* of devices with status ON in the cluster. The *prescribed proportion* of devices ON in each cluster, at each time step, is computed by a *central controller* controlling the average rooms temperatures in each cluster, $X^i := \frac{1}{N_i}\sum_{j=1}^{N_i} X^{i,j}$, where $X_t^{i,j}$ is the room temperature associated to load $j \in [\![1, N_i]\!]$ of cluster $i \in [\![1, d]\!]$. $(X_t^{i,j})_{0 \le t \le T}$is supposed to follow the usual thermal dynamics (see [12, 36] and references therein)

$$X_t^{i,j} = x_0^{i,j} + \int_0^t \big(-\theta^i (X_s^{i,j} - x_{\text{out}}^i) - \kappa^i P_{\text{max}}^i \alpha_s^{i,j} \big) ds + \sigma^{i,j} W_t^{i,j}, \quad t \in [0,T], \tag{6.1}$$

where for any $j \in [\![1, N_i]\!]$, $\sigma^{i,j} > 0$, $\big(W^{i,j}\big)$ are independent real Brownian motions representing model errors and temperature fluctuations inside the room due to local behavior (window, door opening etc.); $x_0^{i,j}$ is the initial temperature; κ^i is the heat exchange parameter; x_{out}^i denotes the outdoor air temperature; $1/\theta^i > 0$ is the thermal time constant; $P_{\max}^i > 0$ denotes the maximal power consumption; $\alpha_s^{i,j} \in \{0,1\}$ is the status OFF or ON of load (i,j) at time instant $s \in [0,T]$.

We are interested in the problem of the *central controller* who considers the aggregated state process $X := (X^i)_{1 \le i \le d}$, whose dynamics is obtained by averaging dynamics (6.1) over $j \in [\![1, N_i]\!]$, for any $i \in [\![1, d]\!]$,

$$X_t^i = x_0^i + \int_0^t \big(-\theta^i (X_s^i - x_{\text{out}}^i) - \kappa^i P_{\text{max}}^i \alpha_s^i \big) ds + \sigma^i W_t^i, \quad t \in [0,T], \tag{6.2}$$

where the control process $\big(\alpha_s = (\alpha_s^i)_{1 \le i \le d}, s \in [0,T]\big)$ taking values in $[0,1]$ prescribes the proportions of devices ON in each cluster; $x_0^i = \frac{1}{N_i}\sum_{j=1}^{N_i} x_0^{i,j}$; $(\sigma^i)^2 = \frac{1}{N_i^2}\sum_{j=1}^{N_i} (\sigma^{i,j})^2$; $(W^i)_{1 \le i \le d}$ is a d-dimensional Brownian motion.The *central controller* problem can be formulated as a specific instantiation of problem (4.1)-(4.2) with the following:

- the controlled process X driven by a drift coefficient $b := (b^i)_{1\leq i\leq d}$ defined on $[0,T]\times\mathbb{R}^d\times[0,1]^d$ s.t. for any $i\in[\![1,d]\!]$ $b^i(t,x,a) = -\theta^i\left(x^i - x^i_{out}\right) - \kappa^i P^i_{\max}a^i$, with the notation $a := (a^i)_{1\leq i\leq d}$ and $x := (x^i)_{1\leq i\leq d}$;

- the terminal cost $g(x) := \frac{1}{d}\sum_{i=1}^d |x^i - \bar{x}^i|^2$ where $\bar{x}\in\mathbb{R}^d$ denotes given *target* values for the final average temperatures of each cluster;

- the running cost defined on $[0,T]\times\mathbb{R}^d\times[0,1]^d$,

$$f(t,x,a) := \lambda\left(\sum_{i=1}^d \rho^i a^i - r_t\right)^2 + \frac{1}{d}\sum_{i=1}^d\Big(\gamma^i(\rho^i a^i)^2 + \eta^i(x^i - x^i_{\max})_+^2 + \eta^i(x^i_{\min} - x^i)_+^2\Big),$$

  where $\rho^i := \frac{N^i P^i_{\max}}{\sum_{j=1}^d N^j P^j_{\max}}$; $\sum_{i=1}^d \rho^i a^i$ gives the overall current consumption of the population as a proportion of the maximum consumption $\sum_{j=1}^d N^j P^j_{\max}$; $r:[0,T]\mapsto\mathbb{R}_*^+$ denotes the target consumption profile for the overall consumption as a proportion of the maximum consumption $\sum_{j=1}^d N^j P^j_{\max}$; $\lambda>0$ quantifies the incentive for the overall consumption to track the target consumption profile r; $\gamma^i>0$ quantifies the quadratic penalty favoring smooth consumption profiles for cluster i; $\eta^i>0$ is a parameter penalizing excursions outside of the comfort interval $[x^i_{\min}, x^i_{\max}]$ for cluster i average temperature.

Note that b verifies Assumption 3, f verifies Assumption 4 and g Assumption 2.

6.2 Simulation results

Consider the *central controller* problem on a time horizon $T = 3600s$, with a population of heterogeneous air-conditioners composed of $d = 1,2,5,10,15,20$ clusters with $N^i = 20$ identical loads in each cluster. We specify the chosen parameters. In each case, $\kappa = 2.5$°C/J and $\sigma^i = 0.1°\mathrm{C}s^{\frac{1}{2}}$; $x_{\text{out}} = 27°\mathrm{C}$; $\theta^i[s^{-1}]$ is chosen arbitrarily in $[0.1, 0.97]$; $P^i_{max}[W]$ is chosen arbitrarily in $[0.5, 5]$; $x_0 = \bar{x}[°\mathrm{C}]$ is chosen arbitrarily in $[16, 27]$; $x_{\min} = \bar{x} - 1.5°\mathrm{C}$; $x_{\max} = \bar{x} + 1.5°\mathrm{C}$; $\eta = 1(°\mathrm{C})^{-2}$; $\lambda = 20$; γ^i is chosen arbitrarily in $[0.5, 1.5]$. The target profile, r, used in simulations is obtained as the sum of a nominal profile corresponding to the standard (uncontrolled) behavior of air-conditioners and a deviation: $r = r^{\text{nom}} + dev$. The standard dynamics of an (uncontrolled) air-conditioner is driven by a cycling rule of ON/OFF decisions intended to keep the room temperature in $[x^i_{\min}, x^i_{\max}]$. When the air-conditioner is ON, it stays ON at $P^i_{\max}$ until the temperature reaches $x^i_{\min}$ then it switches OFF until the temperature reaches $x^i_{\max}$. Then, the air-conditioner turns ON again and begins a new cycle. The nominal profile r^{nom} has been generated by averaging the consumption of 1000 sets of d clusters of N^i heterogeneous air-conditioners simulated independently according to (6.1), with $(\alpha^{i,j}_t)_{0\leq t\leq T}$ following the cycling rule of ON/OFF decisions and with independent initial conditions for temperature $x^{i,j}_0 \sim \mathcal{N}(x^i_0, 1)$ and ON/OFF status $\alpha^{i,j}_0 \in \{0,1\}$. The deviation profile $dev_t = \frac{20}{100} * \sin(\frac{2\pi t}{T})$ induces a maximal deviation of 20% from the nominal profile and integrates to zero on the time horizon $[0, T]$ so that the target profile corresponds to the same energy consumed on the period $[0, T]$ as the nominal profile.

The time step is $\delta t = 60s$. We have implemented Algorithm 1 with a ***backward grid*** initiated with $\mathcal{N}(m_n = \bar{x}, Q_n = I_d)$. For comparison, we have also implemented the standard Monte-Carlo regression scheme using a ***forward grid*** simulated according to (6.2) with a deterministic control α_s approximating the nominal dynamics (according to the ON/OFF cycling rule) described previously. In both cases, we have used second order polynomials ($p = 2$) as basis functions for regressions. We have considered $N = 10^2$, 10^3, 5×10^3, 10^4, 2×10^4, 5×10^4, 10^5 Monte-Carlo paths for the regression grids. To evaluate the statistical performances of the ***forward*** and ***backward grids***, we have implemented each algorithm independently $N_{\text{grid}} = 100$ times for each value of N. For each run, $i = 1, \cdots, N_{\text{grid}}$, the value functions estimate $(v_k^i)_{0 \le k \le n}$ (and the corresponding gradients) was used to implement the associated strategy $\alpha^i = (\alpha_k^i)_{0 \le k \le n}$ on $M = 1000$ i.i.d. simulations of the Brownian motion W, $\omega^1, \cdots, \omega^j, \cdots, \omega^M$. Then the resulting cost $\mathcal{J}(\alpha^i, \omega^j) := g(X_T^{0,x_0,\alpha^i}(\omega^j)) + \int_0^T f(r, X_r^{0,x_0,\alpha^i}(\omega^j), \alpha_r)dr$ has been computed. The expected cost has been estimated as $\mathbb{E}[\mathcal{J}(\alpha^i, \omega^j)] \approx \hat{J} := \frac{1}{M N_{\text{grid}}} \sum_{i=1}^{N_{\text{grid}}} \sum_{j=1}^{M} \mathcal{J}(\alpha^i, \omega^j)$. The variance of $\hat{J}$ is estimated by $\hat{\sigma}_{\hat{J}}^2$ obtained by replacing, expectations and variances by their empirical approximation based on the sample, $\big(\mathcal{J}(\alpha^i, \omega^j), , i \in [\![1, N_{\text{grid}}]\!]\ j \in [\![1, M]\!]\big)$, in the expression $\hat{\sigma}_{\hat{J}}^2 \approx Var(\hat{J}) = \frac{1}{M N_{\text{grid}}} \mathbb{E}\left[Var\big(\mathcal{J}(\alpha^i, \omega^j)\,|\alpha^i\big)\right] + \frac{1}{N_{\text{grid}}} Var\left(\mathbb{E}\left[\mathcal{J}(\alpha^i, \omega^j)\,|\alpha^i\right]\right)$, for each i and j. We have reported on Table 1 (resp. Table 2) the empirical mean $\hat{J}$ and within parenthesis the empirical standard deviation $\hat{\sigma}_{\hat{J}}$ obtained for each considered pair (d, N) for the ***forward grid*** (resp. ***backward grid***).

One can observe that the ***backward grid*** performs surprisingly well providing with high precision the lowest expected cost achieved by both methods (or almost) with only $N = 5 \times 10^3$ paths whatever the dimension d of the control problem. This is consistent with our intuition based on the idea that localizing the grid around the optimally controlled process paths would bring efficiency and reduce the impact of dimension. The particularity of this problem is that the optimally controlled process is naturally localized in a small region of the state space because, on the one hand a target value, $\bar{x}$, is prescribed for the terminal temperatures (by the terminal cost) and on the other hand a target profile is assigned for the overall power consumption. The ***backward grid*** has the advantage of being initiated around the target state and of following dynamics approaching the optimal strategy. This allows to concentrate the ***backward grid*** in the small region of interest so that restricting the regression basis to polynomials of order $p = 2$ seems already enough to obtain reasonable results. However, one can observe some cases where the ***forward grid*** (for $N = 10^5$ and $d \le 5$) has performed slightly better than the ***backward grid***. This can be interpreted by the fact that the ***forward grid*** knows the initial condition x_0 while the ***backward grid*** has no information about it. To further improve the performances Algorithm 1, an idea would be to find a way to exploit that information on the initial condition. This could constitute the subject of future research.

N	d=1	d=2	d=5	d=10	d=15	d=20
$\mathbf{10^2}$	8.68(0.98)	17.28(1.01)	42.04(1.32)	34.79(0.66)	21.27(0.12)	18.97(0.09)
$\mathbf{10^3}$	7.61($6e^{-4}$)	8.24(0.07)	14.83(0.64)	28.14(0.64)	37.91(0.60)	34.83(0.45)
$\mathbf{5\times 10^3}$	7.60($3e^{-4}$)	7.78($2e^{-3}$)	8.98(0.21)	19.84(0.52)	35.31(0.71)	33.57(0.52)
$\mathbf{10^4}$	7.60($3e^{-4}$)	7.77($1e^{-3}$)	7.69(0.06)	16.06(0.38)	32.20(0.63)	30.66(0.59)
$\mathbf{2\times 10^4}$	7.60($3e^{-4}$)	7.77($2e^{-4}$)	7.37(0.02)	13.58(0.40)	28.97(0.71)	28.17(0.67)
$\mathbf{5\times 10^4}$	7.60($3e^{-4}$)	7.79($2e^{-4}$)	7.28($2e^{-3}$)	7.96(0.25)	26.69(0.65)	26.21(0.69)
$\mathbf{10^5}$	7.61($3e^{-4}$)	7.78($1e^{-4}$)	7.27($8e^{-4}$)	6.12(0.08)	22.54(0.56)	23.26(0.59)

Table 1: Mean, $\hat{J}$ (standard deviation, $\hat{\sigma}_{\hat{J}}$) of the simulated cost with the ***forward grid*** strategy

N	d=1	d=2	d=5	d=10	d=15	d=20
$\mathbf{10^2}$	7.61($3e^{-4}$)	7.78($7e^{-4}$)	7.41($6e^{-3}$)	7.31(0.12)	28.14(0.18)	26.01(0.12)
$\mathbf{10^3}$	7.61($3e^{-4}$)	7.77($2e^{-4}$)	7.39($1e^{-3}$)	6.18($3e^{-3}$)	8.19($6e^{-3}$)	7.87($1e^{-2}$)
$\mathbf{5\times 10^3}$	7.61($3e^{-4}$)	7.77($2e^{-4}$)	7.38($8e^{-4}$)	6.17($1e^{-3}$)	8.15($2e^{-3}$)	7.74($3e^{-3}$)
$\mathbf{10^4}$	7.61($3e^{-4}$)	7.77($2e^{-4}$)	7.38($5e^{-4}$)	6.17($1e^{-3}$)	8.15($2e^{-3}$)	7.73($3e^{-3}$)
$\mathbf{2\times 10^4}$	7.61($3e^{-4}$)	7.77($2e^{-4}$)	7.38($3e^{-4}$)	6.17($8e^{-4}$)	8.15($1e^{-3}$)	7.73($2e^{-3}$)
$\mathbf{5\times 10^4}$	7.60($3e^{-4}$)	7.79($1e^{-4}$)	7.38($2e^{-4}$)	6.16($5e^{-4}$)	8.14($8e^{-4}$)	7.72($1e^{-3}$)
$\mathbf{10^5}$	7.61($3e^{-4}$)	7.79($1e^{-4}$)	7.39($2e^{-4}$)	6.16($4e^{-4}$)	8.14($7e^{-4}$)	7.72($9e^{-4}$)

Table 2: Mean $\hat{J}$ (standard deviation, $\hat{\sigma}_{\hat{J}}$) of the simulated cost with the ***backward grid*** strategy

7 Appendix

7.1 A sufficient condition to obtain an equivalent probability

Lemma 7.1. *We recall that $\widetilde{b}$ was defined in (3.17). Let W be an $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion and X be a solution of*

$$X_t = X_0 + \int_0^t \widetilde{b}\,(s, X_s)\, ds + \int_0^t \sigma\,(s)\, dW_s,\ t \in [0,T], \tag{7.3}$$

where X_0 is a Gaussian random vector independent of W. Set $L_t := \sigma\,(t)^{-1}\,\widetilde{b}\,(t, X_t)\,, t \in [0,T]$. Then, the Doléans exponential $\mathcal{E}\left(-\sum_{i=1}^{d}\int_0^{\cdot} L_s^i dW_s^i\right) := \exp\left(-\int_0^{\cdot}\sum_{i=1}^{d} L_s^i dW_s^i - \frac{1}{2}\int_0^{\cdot} |L_s|^2\, ds\right)$ is an $(\mathcal{F}_t)_{t\in[0,T]}$-martingale.

Proof. Following Corollary 5.14 in [25], it is sufficient to find a constant time step subdivision $(t_n)_{n\in\mathbb{N}}$ of $[0,T]$ such that, for all $n \in \mathbb{N}$,

$$\mathbb{E}\left(\exp\left(\frac{1}{2}\int_{t_n}^{t_{n+1}} |L_s|^2\, ds\right)\right) < \infty.$$

Combining Jensen's inequality and Fubini's theorem, this is fulfilled in particular if for all $n \in \mathbb{N}$,

$$\frac{1}{\delta} \int_{t_n}^{t_{n+1}} \mathbb{E}\left(\exp\left(\frac{\delta \left|L_s\right|^2}{2}\right)\right) ds < \infty,$$

where $\delta := t_{n+1} - t_n$. Let $s \in [0, T]$. Then,

$$\left|L_s\right|^2 \leq 2\delta \left|\left|\sigma^{-1}\right|\right|_\infty^2 \left(\left|\left|a\right|\right|_\infty^2 \left|X_s\right|^2 + \left|\left|c\right|\right|_\infty^2\right), \ \mathbb{P} - \text{a.s},$$

since a, c are bounded and σ^{-1} is also bounded being continuous on $[0, T]$. Furthermore, by item 1. of Lemma 3.6 and (3.21), X is a Gaussian process with mean function m^X (resp. covariance function Q^X) solving the first line of equation (3.1) (resp. (3.2)) with initial condition $\mathbb{E}\left(X_0\right)$ (resp. $\mathrm{Cov}\left(X_0\right)$).

Taking into account the fact that m^X is bounded (since continuous), it suffices to find a subdivision such that

$$\mathbb{E}\left(\exp\left(\frac{1}{2} K \delta \left|Z\right|^2\right)\right) < \infty,$$

where $Z \sim \mathcal{N}\left(0, I_d\right)$ and $K := 4\left|\left|\sigma^{-1}\right|\right|_\infty^2 \left|\left|a\right|\right|_\infty^2 \left|\left|Q^X\right|\right|_\infty > 0$. This is the case in particular if $K\delta < 1$, which ends the proof. □

7.2 Proof of the local Lipschitz property of the cost functional J

Lemma 7.2. *Suppose the validity of Assumption 3. Suppose in addition that the functions g and $x \mapsto f\left(t, x, \alpha\right), \left(t, \alpha\right) \in [0, T] \times \mathcal{A}_0$ are locally Lipschitz with polynomial growth gradient (uniformly in t and α). Then, for each $\left(t, \alpha\right) \in [0, T] \times \mathcal{A}_0$,*

$$x \mapsto J\left(t, x, \alpha\right)$$

is locally Lipschitz, uniformly in t and α.

Proof. We give here a proof of the local Lipschitz property for the term involving the function g since the other term can be treated in the same way.

Let $\left(t, \alpha\right) \in [0, T] \times \mathcal{A}_0$ and x, y in a compact set of $\mathbb{R}^d$. Let K be the Lipschitz constant of b. Using in particular the Cauchy-Schwarz inequality, we get

$$\begin{aligned}\left|\mathbb{E}\left(g\left(X_T^{t,x,\alpha}\right)\right) - \mathbb{E}\left(g\left(X_T^{t,y,\alpha}\right)\right)\right| &\leq \int_0^1 \mathbb{E}\left(\left|\nabla_x g\left(aX_T^{t,x,\alpha} + \left(1-a\right) X_T^{t,y,\alpha}\right)\right| \left|X_T^{t,x,\alpha} - X_T^{t,y,\alpha}\right|\right) da \\ &\leq e^{KT} \int_0^1 \mathbb{E}\left(\left|\nabla_x g\left(aX_T^{t,x,\alpha} + \left(1-a\right) X_T^{t,y,\alpha}\right)\right|\right) da \left|x - y\right| \end{aligned} \tag{7.4}$$

where we have used the estimate $\left|X_T^{t,x,\alpha} - X_T^{t,y,\alpha}\right| \leq e^{KT}\left|x - y\right|$, following from the identity

$$\left|X_r^{t,x,\alpha} - X_r^{t,y,\alpha}\right| \leq \left|x - y\right| + K \int_t^r \left|X_s^{t,x,\alpha} - X_s^{t,y,\alpha}\right| ds, \ r \in [t, T],$$

together with Gronwall's lemma. In view of (7.4), the point is proved if

$$\int_0^1 \mathbb{E}\left(\left|\nabla_x g\left(aX_T^{t,x,\alpha} + (1-a)\, X_T^{t,y,\alpha}\right)\right|\right) da$$

is bounded uniformly in t, x, y, α. This follows from polynomial growth of $\nabla_x g$, classical moment estimates for $\sup_{s\in[t,T]}\left|X_s^{t,z,\alpha}\right|$, $z \in \mathbb{R}^d$ (see for example Corollary 2.5.12 in [26]) and the fact x, y lie in a compact set. □

7.3 A simplified version of the envelope theorem

Lemma 7.3. *Let Λ be an arbitrary set and O be an open subset of $\mathbb{R}^d$. Let $x \in \mathbb{R}^d$. Let $F : O \times \Lambda \mapsto \mathbb{R}$ such that for all $\lambda \in \Lambda$, $F(\cdot, \lambda)$ and $V : x \mapsto \sup_{\lambda\in\Lambda} F(x, \lambda)$ are differentiable at the point x. Suppose also that $\Lambda^*(x) = \{\lambda \in \Lambda, V(x) = F(x, \lambda)\}$ is not empty. Then,*

$$\nabla_x V(x) = \nabla_x F(x, \lambda_x^*),$$

for every $\lambda_x^* \in \Lambda^*(x)$.

Proof. Let x as in the proposition statement and $h \in \mathbb{R}^d$. Let $\lambda_x^* \in \Lambda^*(x)$. Then, using in particular the differentiability of $F(\cdot, \lambda_x^*)$ at the point x, we get

$$\begin{aligned} V(x+h) - V(x) &\geq F(x+h, \lambda_x^*) - F(x, \lambda_x^*) \\ &= \langle \nabla_x F(x, \lambda_x^*), h\rangle + o_0(|h|). \end{aligned} \tag{7.5}$$

By the differentiability of V at the point x, (7.5) implies

$$\langle \nabla_x V(x) - \nabla_x F(x, \lambda_x^*), h\rangle \geq o_0(|h|). \tag{7.6}$$

Setting h to $-h$ in (7.5) and proceeding as before, we obtain

$$\langle \nabla_x V(x) - \nabla_x F(x, \lambda_x^*), h\rangle \leq o_0(|h|). \tag{7.7}$$

Combining (7.6) and (7.7), we get

$$\left\langle \nabla_x V(x) - \nabla_x F(x, \lambda_x^*), \frac{h}{|h|}\right\rangle \underset{h\to 0}{\longrightarrow} 0,$$

which forces $\nabla_x V(x) = \nabla_x F(x, \lambda_x^*)$. This ends the proof.

□

Acknowledgments

The work was supported by a public grant as part of the *Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH*, in a joint call with Gaspard Monge Program for optimization, operations research and their interactions with data sciences.

References

[1] Ch. Bender and R. Denk. A forward scheme for backward SDEs. *Stochastic Processes Appl.*, 117(12):1793–1812, 2007.

[2] Ch. Bender and T. Moseler. Importance sampling for backward SDEs. *Stochastic Anal. Appl.*, 28(2):226–253, 2010.

[3] Ch. Bender and J. Steiner. Least-squares Monte Carlo for backward SDEs. In *Numerical methods in finance. Selected papers based on the presentations at the workshop, Bordeaux, France, June 2010*, pages 257–289. Berlin: Springer, 2012.

[4] B. Bouchard and N. Touzi. Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111:175–206, 2004.

[5] R. Bronson and G. B. Costa. *Matrix Methods: Applied Linear Algebra*. Academic Press, 2008.

[6] D. S. Callaway and I. A. Hiskens. Achieving controllability of electric loads. *Proceedings of the IEEE*, 99(1):184–199, 2010.

[7] M. G. Crandall, H. Ishii, and P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

[8] F. Delarue and S. Menozzi. An interpolated stochastic algorithm for quasi-linear PDEs. *Math. Comp.*, 77(261):125–158 (electronic), 2008.

[9] C. Di Girolami and F. Russo. About classical solutions of the path-dependent heat equation. *Random Oper. Stoch. Equ.*, 28(1):35–62, 2020.

[10] I. Exarchos and E. A. Theodorou. Stochastic optimal control via forward and backward stochastic differential equations and importance sampling. *Automatica*, 87:159–165, 2018.

[11] G. Fabbri, F. Gozzi, and A. Święch. *Stochastic optimal control in infinite dimension*, volume 82 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017. Dynamic programming and HJB equations, With a contribution by Marco Fuhrman and Gianmario Tessitore.

[12] E. Gobet and M. Grangereau. Federated stochastic control of numerous heterogeneous energy storage systems. working paper or preprint, January 2021.

[13] E. Gobet and C. Labart. Error expansion for the discretization of backward stochastic differential equations. *Stochastic Process. Appl.*, 117(7):803–829, 2007.

[14] E. Gobet and C. Labart. Solving BSDE with adaptive control variate. *SIAM J. Numer. Anal.*, 48(1):257–277, 2010.

[15] E. Gobet, J-P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2202, 2005.

[16] E. Gobet and P. Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Math. Comput.*, 85(299):1359–1391, 2016.

[17] E. Gobet and P. Turkedjiev. Adaptive importance sampling in least-squares Monte Carlo algorithms for backward stochastic differential equations. *Stochastic Processes Appl.*, 127(4):1171–1203, 2017.

[18] F. Gozzi and F. Russo. Verification theorems for stochastic optimal control problems via a time dependent Fukushima-Dirichlet decomposition. *Stochastic Process. Appl.*, 116(11):1530–1562, 2006.

[19] F. Gozzi and F. Russo. Weak Dirichlet processes with a stochastic control perspective. *Stochastic Process. Appl.*, 116(11):1563–1583, 2006.

[20] U. G. Haussmann and É. Pardoux. Time reversal of diffusions. *Ann. Probab.*, 14(4):1188–1205, 1986.

[21] H. Ishii. On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic pde's. *Communications on Pure and Applied Mathematics*, 42(1):15–45, 1989.

[22] H. Ishii and K. Kobayasi. On the uniqueness and existence of solutions of fully nonlinear parabolic pdes under the osgood type condition. *Differential Integral Equations*, 7(3-4):909–920, 1994.

[23] L. Izydorczyk, N. Oudjane, F. Russo, and G. Tessitore. Fokker-Planck equations with terminal condition and related McKean probabilistic representation. *Preprint hal-02902615*, 2020.

[24] R. Jensen, P.-L. Lions, and P. E. Souganidis. A uniqueness result for viscosity solutions of second order fully nonlinear partial differential equations. *Proceedings of the American Mathematical Society*, page 975?975, 1988.

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

[26] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.

[27] C. Labart and J. Lelong. A parallel algorithm for solving BSDEs. *Monte Carlo Methods Appl.*, 19(1):11–39, 2013.

[28] P. L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations part 2: viscosity solutions and uniqueness. *Communications in Partial Differential Equations*, 8(11):1229–1276, 1983.

[29] D. Nunziante. Existence and uniqueness of unbounded viscosity solutions of parabolic equations with discontinuous time-dependence. *Nonlinear Analysis: Theory, Methods & Applications*, 18(11):1033 – 1062, 1992.

[30] E. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, pages 79–127. Birkhäuser Boston, Boston, MA, 1998.

[31] É. Pardoux, F. Pradeilles, and Z. Rao. Probabilistic interpretation of a system of semi-linear parabolic partial differential equations. *Annales de l'I.H.P. Probabilités et statistiques*, 33(4):467–490, 1997.

[32] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009.

[33] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[34] C. Ribeiro and N. Webber. Valuing path-dependent options in the variance-gamma model by Monte Carlo with a gamma bridge. *Journal of Computational Finance*, 7(2):81–100, 2004.

[35] P. Sabino. Forward or backward simulation? A comparative study. *Quantitative Finance*, 20(7):1213–1226, 2020.

[36] A. Seguret, C. Alasseur, J. F. Bonnans, A. De Paola, N. Oudjane, and V. Trovato. Decomposition of high dimensional aggregative stochastic control problems. *Preprint arXiv:2008.09827*, 2020.

[37] N. Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2013. With Chapter 13 by Agnès Tourin.

[38] A. Yu. Veretennikov. Parabolic equations and Itô's stochastic equations with coefficients discontinuous in the time variable. *Mathematical Notes of the Academy of Sciences of the USSR*, 31(4):278–283, April 1982.