Numerical approximation of ergodic BSDEs using non linear Feynman-Kac formulas

Emmanuel GOBET (École Polytechnique)
Adrien RICHOU (Université de Bordeaux)
Lukasz SZPRUCH (University of Edinburgh)

Based on the preprint
https://hal.science/hal-04644887/

International Seminar on SDEs and Related Topics
May 9th, 2025

CMAP

Agenda

1. Introduction
   - Statement of the problem
   - BSDE with large time-horizon
   - State of the art
2. Theoretical results
   - Simplified setting
   - Time-randomized FK representation
   - Contraction properties of the FK representation
3. Numerical scheme
   - Fully implementable numerical scheme
   - Error bounds
   - Numerical experiments

Statement of the problem

Aim: Numerical solution (Y, Z, λ) of the Ergodic Backward Stochastic Differential Equation (EBSDE)

$$Y_t = Y_T + \int_t^T (f(X_s, Z_s) - \lambda)\,\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s, \quad 0 \leqslant t \leqslant T, \qquad (1)$$

1) (Y, Z) take values in some appropriate $\mathbb{L}_2$ space,
2) λ is a scalar (called *ergodic cost*),
3) X is the solution of an ergodic forward SDE.

Literature:

i) introduced first by [Fuhrman et al., 2009]: efficient tool to analyse optimal control problems with ergodic cost functionals.
ii) Alternatively, Hamilton-Jacobi-Bellman equation, see [Arisawa and Lions, 1998] and [Bensoussan and Frehse, 2002].

BSDE with large time-horizon

Consider the BSDE-solution (Y^T, Z^T) parameterized by $T > 0$:

$$Y_t^{T,x} = g(X_T^x) + \int_t^T f(X_s^x, Z_s^{T,x})\mathrm{d}s - \int_t^T Z_s^{T,x}\mathrm{d}W_s, \quad 0 \leqslant t \leqslant T.$$

Adjoint problems for stochastic control problems: [Peng, 1993], [Ma and Yong, 1999], [Zhang, 2017].

Which behavior as $T \to +\infty$?

Theorem 1

Under suitable assumptions [Hu et al., 2015], the following asymptotic expansion result holds: for some constants $L \in \mathbb{R}$ and $C > 0$,

$$\left| Y_0^{T,x} - \lambda T - Y_0^x - L \right| \leqslant C(1 + |x|^3)e^{-T/C} \tag{2}$$

where Y_0^x is the solution of (1) for $X_0 = x$.

$\hookrightarrow$ Solving EBSDE for any x gives explicit approximation of $Y_0^{T,x}$ as $T \to +\infty$.

State of the art

i) Theoretical properties of EBSDEs: [Fuhrman et al., 2009], [Richou, 2009], [Debussche et al., 2011], [Cohen and Hu, 2013], [Madec, 2015], [Guatteri and Tessitore, 2020]

ii) Numerical approximation: [Broux-Quemerais et al., 2024] using random horizon time approximation and a neural network space approximation

Our contributions: our aim is to provide an alternative fully implementable scheme and to study the approximation error.

i) Markov representation of the value function and its gradient

ii) fixed point equation to which the gradient is (the unique) solution

iii) contraction properties of this fixed point equation

ivi) full error controls (w.r.t. the number of Picard iterations, the number of Monte-Carlo samples using subGamma concentration-of-inequalities, the grid mesh)

v) numerical experiments to illustrate theoretical findings

Simplified setting

Forward d-dimensional SDE: $X_t = x + \int_0^t b(X_s)\mathrm{d}s + \Sigma W_t, \quad 0 \leqslant t.$

Assumption 2.1

There exist constants s.t., $\forall x, x' \in \mathbb{R}^d, z, z' \in \mathbb{R}^{1\times d}$,

(A-1) $|f(x,z) - f(x',z')| \leqslant K_{f,x}|x - x'| + K_{f,z}\|z - z'\|$,

(A-2) $|b(x) - b(x')| \leqslant K_{b,x}|x - x'|$,

(A-3) $\langle b(x) - b(x'), x - x' \rangle \leqslant -\eta |x - x'|^2$ *or* $b(x) = -Ax$ *with* $\mathrm{Sp}\, A \subset \{z \in \mathbb{C} | \Re(z) > a > 0\}$,

(A-4) Σ *is invertible.*

Unique strong solution X, with unique invariant measure ν.

⚠ The hypothesis of Hurwitz matrix ($b(x) = -Ax$ with $\mathrm{Sp}\, A \subset \{z \in \mathbb{C} | \Re(z) > a > 0\}$) does not imply the dissipativity assumption: take for instance $A = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$.

Proposition 1

Let us assume that Assumptions 2.1 are in force. Then the ergodic BSDE (1) has a solution (Y, Z, λ) such that

$$Y_t = u(X_t), \qquad Z_t = \bar{u}(X_t) \tag{3}$$

for two measurable functions satisfying the growth

$$|u(x)| \leqslant C(1 + |x|), \qquad |\bar{u}(x)| \leqslant C, \qquad \forall x \in \mathbb{R}^d.$$

Moreover, the solution (Y, Z, λ) is unique (up to a constant for Y) in the class of Markovian solutions with previous growth.

Later, we will also justify $u \in C^1$ and that $\bar{u}(.) = \nabla_x u(.)\Sigma$.

Proposition 2

Under assumptions of Proposition 1, u given by (3) is a viscosity solution of the following elliptic PDE

$$\mathcal{L}u(x) + f(x, \nabla_x u(x)\Sigma) = \lambda,$$

where $\mathcal{L}$ denotes the generator of the semi-group associated to the SDE for X.

Time-randomized Feynman-Kac representation

Informal derivation: start with

$$u(x) = \mathbb{E}\left[u(X_T^x) + \int_0^T \left(f(X_s^x, \bar{u}(X_s^x)) - \lambda\right) \mathrm{d}s\right].$$

Differentiate the above with respect to x and use the Malliavin calculus integration by parts formula [Nualart, 2006]: write $\bar{u} = \nabla_x u \Sigma$

$$v(x) = \nabla u(x) = \mathbb{E}\left[v(X_T^x)\nabla_x X_T^x + \int_0^T U_s^x f(X_s^x, \nabla_x u \Sigma(X_s^x))\mathrm{d}s\right]$$

where U_s^x is the (raw vector valued) Malliavin weight given by

$$U_s^x = \frac{1}{s}\left(\int_0^s (\Sigma^{-1}\nabla_x X_r^x)^\top \mathrm{d}W_r\right)^\top.$$

Time-randomized Feynman-Kac representation

Informal derivation: start with

$$u(x) = \mathbb{E}\left[u(X_T^x) + \int_0^T (f(X_s^x, \bar{u}(X_s^x)) - \lambda)\,\mathrm{d}s\right].$$

Differentiate the above with respect to x and use the Malliavin calculus integration by parts formula [Nualart, 2006]: write $\bar{u} = \nabla_x u \Sigma$

$$v(x) = \nabla u(x) = \mathbb{E}\left[v(X_T^x)\nabla_x X_T^x + \int_0^T U_s^x f(X_s^x, \nabla_x u\Sigma(X_s^x))\mathrm{d}s\right]$$

where U_s^x is the (raw vector valued) Malliavin weight given by

$$U_s^x = \frac{1}{s}\left(\int_0^s (\Sigma^{-1}\nabla_x X_r^x)^\top \mathrm{d}W_r\right)^\top.$$

i) At first sight, $v(.)$ solves a nice fixed-point equation

ii) But the terms inside the above expectation have exploding polynomial moments as T goes to $+\infty$!!

Diving into the details of the time-explosion

Take dimension $d = 1$, with $\Sigma = 1$ and $b(x) = -ax$ for a scalar parameter $a > 0$: then $\nabla_x X_t^x = e^{-at}$ and

$$U_s^x = \frac{1}{s}\left(\int_0^s (\Sigma^{-1}\nabla_x X_r^x)^\top \mathrm{d}W_r\right)^\top = \frac{1}{s}\int_0^s e^{-ar}\mathrm{d}W_r \overset{d}{=} \mathcal{N}\left(0, \frac{1-e^{-2as}}{2as^2}\right).$$

Diving into the details of the time-explosion

Take dimension $d = 1$, with $\Sigma = 1$ and $b(x) = -ax$ for a scalar parameter $a > 0$: then $\nabla_x X_t^x = e^{-at}$ and

$$U_s^x = \frac{1}{s}\left(\int_0^s (\Sigma^{-1}\nabla_x X_r^x)^\top \mathrm{d}W_r\right)^\top = \frac{1}{s}\int_0^s e^{-ar}\mathrm{d}W_r \overset{d}{=} \mathcal{N}\left(0, \frac{1-e^{-2as}}{2as^2}\right).$$

Take a bounded driver f:

$$\left|\int_0^T U_s^x f(X_s^x, \bar{u}(X_s^x))\mathrm{d}s\right|_p \leq \int_0^T |U_s^x f(X_s^x, \bar{u}(X_s^x))|_p \,\mathrm{d}s$$
$$\leq \int_0^T C_p\sqrt{\frac{1-e^{-2as}}{2as^2}}\|f\|_\infty \mathrm{d}s.$$

i) Convergence at $s = 0$

ii) Divergence at $s = +\infty$!

$\hookrightarrow$ **one has to find better Malliavin weights...**

Reminder about the choice of Malliavin weight

Let x be given and let $\mathcal{U}_s^x$ the class of Malliavin weights U_s^x such that

$$\nabla_x \mathbb{E}\left[\varphi(X_s^x)\right] = \mathbb{E}\left[\varphi(X_s^x) U_s^x\right]$$

for any square integrable φ.

Lemma 2 ([Fournié et al., 2001])

The weights must have the same conditional expectation: $\mathbb{E}\left[U_s^x \mid X_s^x\right]$ does not depend on U_s^x. The element with minimal L_2-norm is

$$\mathbb{E}\left[U_s^x \mid X_s^x\right] = \bar{U}_s^x.$$

Reminder about the choice of Malliavin weight

Let x be given and let $\mathcal{U}_s^x$ the class of Malliavin weights U_s^x such that

$$\nabla_x \mathbb{E}\left[\varphi(X_s^x)\right] = \mathbb{E}\left[\varphi(X_s^x) U_s^x\right]$$

for any square integrable φ.

Lemma 2 ([Fournié et al., 2001])

The weights must have the same conditional expectation: $\mathbb{E}\left[U_s^x \mid X_s^x\right]$ does not depend on U_s^x. The element with minimal L_2-norm is

$$\mathbb{E}\left[U_s^x \mid X_s^x\right] = \bar{U}_s^x.$$

When X_s^x has a density: the minimal L_2-norm solution is

$$\nabla_x \mathbb{E}\left[\varphi(X_s^x)\right] = \int_{\mathbb{R}^d} \varphi(x') \nabla_x p(0, x; s, x') \mathrm{d}x' = \mathbb{E}\left[\varphi(X_s^x) \bar{U}_s^x\right]$$

$$\text{where } \bar{U}_s^x = \nabla_x (\log(p(0, x; s, x')))\big|_{x'=X_s^x}.$$

Find a good Malliavin weight without knowing the density?

Restricted framework

Assumption 2.2

Let $b(x) = -Ax$, and for some constants such that

(A-1') $|f(x,z) - f(x',z')| \leqslant K_{f,x}\,|x-x'| + K_{f,z}\|z-z'\|$,

(A-2') $\mathrm{Sp}\,A \subset \{z \in \mathbb{C} | \Re(z) > a > 0\}$,

(A-3') Σ *is invertible.*

$$\hookrightarrow X_t^x = e^{-At}x + e^{-At}\int_0^t e^{As}\Sigma \mathrm{d}W_s, \qquad \left\|e^{-At}\right\| \leqslant C_A e^{-at}, \qquad \forall t \geqslant 0.$$

hence X_t^x is Gaussian, with mean $e^{-At}x$ and covariance

$$\Sigma_t := \int_0^t e^{-Ar}\Sigma\Sigma^\top e^{-A^\top r}\mathrm{d}r.$$

Then

$$\bar{U}_s^x = (X_s^x - e^{-As}x)^\top \Sigma_s^{-1} e^{-As} =: e^{-as}\tilde{U}_s,$$

$$\text{with} \quad |\,|\tilde{U}_s|^2|_2 \leqslant C(1 \vee s^{-1/2}), \quad \forall s > 0.$$

Feynman-Kac representation

Theorem 3

Let us assume that Assumptions 2.2 are in force. Then

i) $u \in C^1(\mathbb{R}^d)$,

ii) $Z_t = v(X_t)\Sigma$ *with* $v := \nabla_x u$, $\|v\|_\infty < +\infty$,

iii) the gradient v *is solution of the four following equations*

$$\begin{aligned}
v(x) &= \mathbb{E}\left[v(X_T^x)e^{-AT} + \int_0^T e^{-as}\tilde{U}_s f(X_s^x, v(X_s^x)\Sigma)\mathrm{d}s\right] \\
&= \mathbb{E}\left[\int_0^{+\infty} e^{-as}\tilde{U}_s f(X_s^x, v(X_s^x)\Sigma)\mathrm{d}s\right] \quad \textit{(take } T = +\infty\textit{)} \\
&= \mathbb{E}\left[v(X_T^x)e^{-AT} + \mathbf{1}_{G\le T\theta}\frac{\sqrt{\pi}}{\theta}\sqrt{G}e^{-(\frac{a}{\theta}-1)G}\tilde{U}_{\frac{G}{\theta}} f\left(X_{\frac{G}{\theta}}^x, v(X_{\frac{G}{\theta}}^x)\Sigma\right)\right] \\
&= \frac{\sqrt{\pi}}{\theta}\mathbb{E}\left[\sqrt{G}e^{-(\frac{a}{\theta}-1)G}\tilde{U}_{\frac{G}{\theta}} f\left(X_{\frac{G}{\theta}}^x, v(X_{\frac{G}{\theta}}^x)\Sigma\right)\right] \quad \textit{(take } T = +\infty\textit{)}
\end{aligned}$$

where $\theta \in (0, a)$ *and* $G \overset{d}{=} \mathcal{G}(1/2, 1)$ *is independent of* W.

Application to BSDE in large horizon

A BSDE with driver independent of Y can be well approximated, as the horizon T is large, by an EBSDE

$$Y_0^{T,x} \approx \lambda T + Y_0^x + L$$

with the error bound (2).

i) Y_0^x is defined up to a constant: L depends on this choice.

ii) Once v is obtained, we get $\lambda = \int_{\mathbb{R}^d} f(x, v(x)\Sigma)\nu(\mathrm{d}x)$.

iii) u is the antiderivative of v up to constant

$$Y_0^x = u(x) = \int_0^1 v(tx)x\mathrm{d}t, \quad \forall x \in \mathbb{R}^d.$$

iv) The tuning of L is delicate. Since $Y_0^{x=0} = 0$, we have

$$L = \lim_{T\to+\infty}(Y_0^{T,x=0} - \lambda T),$$

with an exponential convergence.
$\hookrightarrow$ Naive approach: estimate $Y_0^{T,x=0}$ for a few T, and then get an estimation of L...

Contraction properties of the FK representation

For all $T \in \mathbb{R}^+ \cup \{+\infty\}$, define a map

$$\Phi_T : w \in L_0(\mathbb{R}^d, \mathbb{R}^{1\times d}) \longrightarrow L_0(\mathbb{R}^d, \mathbb{R}^{1\times d})$$

given by, for all $x \in \mathbb{R}^d$,

$$\Phi_T(w)(x) = \mathbb{E}\Big[w(X_T^x) e^{-AT} \mathbf{1}_{T<+\infty} + \mathbf{1}_{G\le T\theta} \frac{\sqrt{\pi}}{\theta} \sqrt{G} e^{-(\frac{a}{\theta}-1)G} \tilde{U}_{\frac{G}{\theta}} f\left(X^x_{\frac{G}{\theta}}, w(X^x_{\frac{G}{\theta}})\Sigma \right) \Big]$$

We seek the solution v such that

$$\Phi_T(v) = v, \quad \forall\ T \in \mathbb{R}^+ \cup \{+\infty\}.$$

In the paper, we have studied the contraction properties for any fixed $T < +\infty$ and for $T = +\infty$: properties are better for $T = +\infty$.

Define weighted norm $\|v\|_\rho = \sup_{x\in\mathbb{R}^d} \frac{\|v(x)\|}{\rho(x)}$ with

$\rho_{\mathrm{pol},\alpha,\beta}(x) = (1+\alpha|x|)^\beta$ (with $\beta \geqslant 1$ and $\alpha > 0$) or $\rho_{\exp,\alpha}(x) = e^{\alpha|x|}$ (with $\alpha > 0$)

Theorem 4

Assume

$$\|\Sigma_s^{-1}\|^{1/2} \leqslant c_{1,(4)} + \frac{c_{2,(4)}}{\sqrt{s}}, \quad \forall s > 0. \tag{4}$$

Then, $\|\Phi_\infty(w_1) - \Phi_\infty(w_2)\|_\rho \leqslant \kappa_\infty \|w_1 - w_2\|_\rho$, *with*

i) If $C_A = 1$ *and* $\rho = \rho_{\exp,\alpha}$, *then*

$$\kappa_\infty \leqslant K_{f,z}\|\Sigma\|\sqrt{d}\left(2e^{\frac{\alpha^2\|\Sigma\Sigma^\top\|}{2a^2}}\mathcal{N}\left(\frac{\alpha\|\Sigma\Sigma^\top\|^{1/2}}{a}\right)\right)^{d/2}\left(\frac{c_{1,(4)}}{a} + \frac{\sqrt{\pi}c_{2,(4)}}{\sqrt{a}}\right).$$

ii) If $C_A > 1$ *and* $\rho = \rho_{\mathrm{pol},\alpha,\beta}$, *then*

$$\kappa_\infty \leqslant C_A K_{f,z}\|\Sigma\|\sqrt{d}\mathbb{E}\left[\left(C_A + \alpha C_A\left(\frac{\|\Sigma\Sigma^\top\|}{2a}\right)^{1/2}|Y|\right)^{2\beta}\right]^{1/2}\left(\frac{c_{1,(4)}}{a} + \frac{\sqrt{\pi}c_{2,(4)}}{\sqrt{a}}\right)$$

where $Y \sim \mathcal{N}(0, I_d)$.

Numerical scheme

$$v(x) = \frac{\sqrt{\pi}}{\theta}\mathbb{E}\left[e^{(\theta-a)G}\sqrt{G}\tilde{U}_G f(X_G^x, v(X_G^x)\Sigma)\right]$$

i) Picard scheme:

$$v^{n+1}(x) = \frac{\sqrt{\pi}}{\theta}\mathbb{E}\left[e^{(\theta-a)G}\sqrt{G}\tilde{U}_G f(X_G^x, v^n(X_G^x)\Sigma)\right]$$

ii) space discretization: $x \in \Pi$,

$$v^{n+1}(x) = \frac{\sqrt{\pi}}{\theta}\mathbb{E}\left[e^{(\theta-a)G}\sqrt{G}\tilde{U}_G f(X_G^x, Pv^n(X_G^x)\Sigma)\right]$$

iii) Monte-Carlo estimation.

Our fully implementable numerical scheme

Definition 5

We construct a sequence of random functions $v_M^n : \Omega \times \Pi \to \mathbb{R}^{1\times d}$, $n \in \mathbb{N}$ such that $v_M^0 = 0$ and, for all $n \in \mathbb{N}$, $z \in \Pi$,

$$v_M^{n+1}(z) = \left\lfloor \frac{1}{M} \sum_{j=1}^{M} R^z_{n+1,j}(P v_M^n) \right\rfloor_B ,$$

where $B \geqslant \|v\|_\infty$, for any $\phi : \mathbb{R}^d \to \mathbb{R}^{1\times d}$, $(R^z_{n,j}(\phi))_{n,j\in\mathbb{N}^*, z\in\Pi}$ are independent random variables and for any $z \in \Pi$, $(R^z_{n,j}(\phi))_{n,j\in\mathbb{N}^*}$ have the same distribution as

$$R^z(\phi) := \frac{\sqrt{\pi}}{\theta}\sqrt{G} e^{-(a-\theta)G}\, \tilde{U}_G f\left(X^z_G, \phi(X^z_G)\Sigma\right),$$

recalling that $\theta \in (0, a)$ and $G \overset{d}{=} \mathcal{G}(1/2, \theta)$ is independent of W.

Note that the random variables to sample have sub-Gamma tails.

Concentration-inequalities of sub-Gamma tails

Orlicz function: $\Psi := \exp(.) - 1$.

Orlicz norm: $|Y|_\Psi := \inf \left\{ c > 0, \mathbb{E}\left[\Psi\left(\frac{|Y|}{c}\right)\right] \leqslant 1 \right\}$.

Proposition 3 ([Talagrand, 1989], [van der Vaart and Wellner, 1996])

i) **[Talagrand inequality]** *There exists a universal constant C_Ψ such that, for all sequence $(Y_k)_{1\leqslant k\leqslant K}$ of independent, mean zero, random variables satifying $|Y_k|_\Psi < +\infty$ for all $0 \leqslant k \leqslant K$, we have*

$$\left|\sum_{k=1}^{K} Y_k\right|_\Psi \leqslant C_\Psi \left(\mathbb{E}\left[\left|\sum_{k=1}^{K} Y_k\right|\right] + \left|\max_{1\leqslant k\leqslant K} |Y_k|\right|_\Psi\right).$$

ii) **[Maximal inequality]** *There exists a universal constant C_Ψ such that, for all sequence $(Y_k)_{1\leqslant k\leqslant K}$ of random variables satisfying $|Y_k|_\Psi < +\infty$ for all $0 \leqslant k \leqslant K$, we have*

$$\left|\max_{1\leqslant k\leqslant K} |Y_k|\right|_\Psi \leqslant C_\Psi \Psi^{-1}(K) \max_{1\leqslant k\leqslant K} |Y_k|_\Psi.$$

A convergence result

We assume that our grid Π is centered in 0, and is given by

$$\left\{(i_1\delta, \dots, i_d\delta) \,|\, i_k \in \{-\widetilde{N}, \dots, \widetilde{N}\}, k \in \{1, \dots, d\}\right\}$$

for a given $\widetilde{N} \in \mathbb{N}$.

Proposition 4

Take $M_z = \tilde{M}(1+|z|)^2\rho^{-2}(z)$.

i) If $C_A = 1$ *and* $\rho = \rho_{\exp,\alpha}$, *then we have*

$$\mathbb{E}\left[\sup_{x\in\mathbb{R}^d}\left\|\frac{Pv^n_M(x) - v(x)}{\rho(x)}\right\|\right] = O\left(\delta^2 + \frac{\ln \tilde{N}}{\sqrt{\tilde{M}}} + e^{-\alpha\tilde{N}\delta} + \kappa^n_\infty\right).$$

ii) If $C_A > 1$ *and* $\rho = \rho_{\mathrm{pol},\alpha,\beta}$, *then we have*

$$\mathbb{E}\left[\sup_{x\in\mathbb{R}^d}\left\|\frac{Pv^n_M(x) - v(x)}{\rho(x)}\right\|\right] = O\left(\delta^2 + \frac{\ln \tilde{N}}{\sqrt{\tilde{M}}} + (1+\alpha\tilde{N}\delta)^{-\beta} + \kappa^n_\infty\right).$$

Some numerical experiments

$$\sigma = 1, \qquad f(x,z) = 1 + \sin(\gamma(|x|+|z|)) + \gamma|z| - \sin(\gamma(|x| + 2|x|e^{-|x|^2})) - (2\gamma|x| + 2|x|^2 - d + 2a|x|^2)e^{-|x|^2}.$$

Solution of the EBSDE: $u(x) = e^{-|x|^2}$, $v(x) = -2x^\top e^{-|x|^2}$ and $\lambda = 1$.

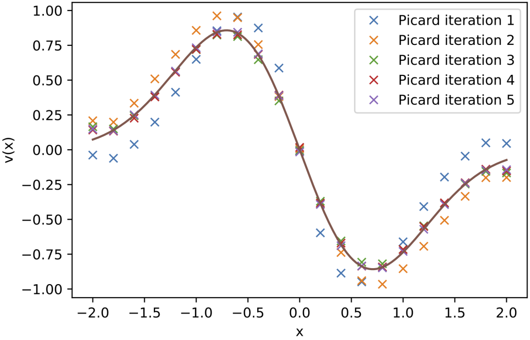

Figure: Solution v at different iterations. Parameters: $d = 1$, $\gamma = 1$, $a = 2$, $\theta = 1.8$, $\widetilde{N} = 10$, $\delta = 0.2$, $M = 10^5$.

Impact of the Lipschitz constant of the driver

$$\mathfrak{E}^{d,r}_{\infty,n} := \sup\Big\{|v^n_M(x) - v(x)| : x = (i_1\delta, \ldots, i_d\delta),$$
$$i_k \in \{-(\widetilde{N} - r), \ldots, (\widetilde{N} - r)\}, k \in \{1, \ldots, d\}\Big\},$$

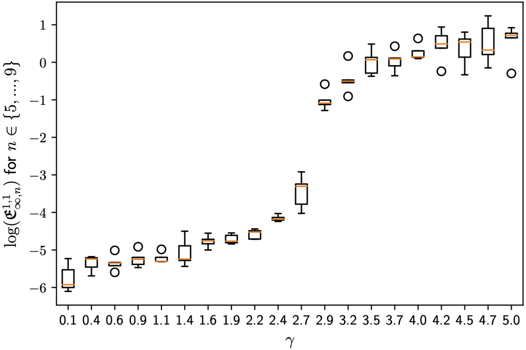

Figure: Box plots of log-sup errors $\mathfrak{E}^{d,r}_{\infty,n}$ (with $d = 1$, $r = 1$) for different n, as a function of γ. Parameters: $a = 2$, $\theta = 1.8$, $\widetilde{N} = 10$, $\delta = 0.2$, $M = 10^5$.

Impact of the parameter θ of the Gamma distribution

Recall that theoretically $\theta < a$.

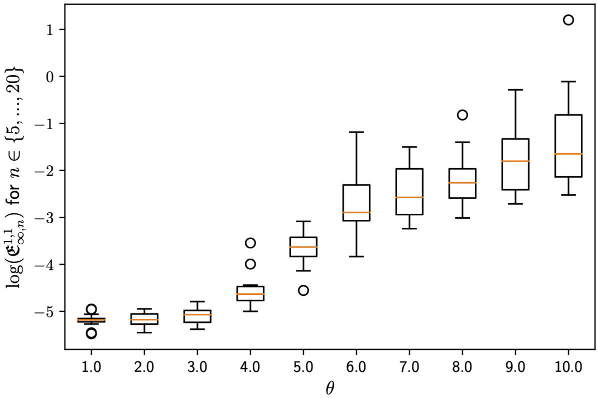

Figure: Box plots of log-sup errors $\mathfrak{E}^{d,r}_{\infty,n}$ (with $d = 1$, $r = 1$) for different n, as a function of θ. Parameters: $a = 2$, $\gamma = 1$, $\widetilde{N} = 10$, $\delta = 0.2$, $M = 10^5$.

Tests in various dimensions

Dimension d	1	2	3	4	5
Sup Error $\mathfrak{E}^{d,1}_{\infty,3}$	$5.49 * 10^{-2}$	$5.69 * 10^{-2}$	$7.69 * 10^{-2}$	$11.9 * 10^{-2}$	$11.3 * 10^{-2}$
Mean Error	$2.31 * 10^{-2}$	$1.94 * 10^{-2}$	$2.16 * 10^{-2}$	$2.76 * 10^{-2}$	$3.49 * 10^{-2}$
Time (s)	4	18	217	4155	86639

Table: Comparison of sup errors and computational times as a function of the dimension d. Parameters: $a = 2$, $\gamma = 1$ $\theta = 1.8$, $\widetilde{N} = 5$, $\delta = 0.4$, $M = 10^4$

Remark. The scheme can be "easily" parallelized. L. Facq and P. Depouilly (Math Institute Bordeaux) used a GPU card with 40Go RAM to do same computations: for $d = 5$, they obtained the result in 90s.

Using neural networks

i) Replacement of the grid approximation by a NN.

ii) Removing Picard iteration:

$$\mathbb{E}\left[\left\|\mathtt{NN}(X_0) - \frac{1}{M}\sum_{i=1}^{M}\frac{\sqrt{\pi}}{\theta}e^{(\theta-a)G^i}\sqrt{G^i}\,\tilde{U}^i_{G^i}f(X^{X_0}_{G^i}, \mathtt{NN}(X^{X_0}_{G^i})\Sigma)\right\|^2\right]$$

$\hookrightarrow$ Ongoing numerical experiments done by S. Chardul for solving (non ergodic) infinite horizon BSDE.

References I

Arisawa, M. and Lions, P.-L. (1998).
On ergodic stochastic control.
Comm. Partial Differential Equations, **23**(11-12):2187–2217.

Bensoussan, A. and Frehse, J. (2002).
Ergodic control Bellman equation with Neumann boundary conditions.
In *Stochastic theory and control (Lawrence, KS, 2001)*, volume **280** of *Lecture Notes in Control and Inform. Sci.*, pages 59–71. Springer, Berlin.

Broux-Quemerais, G., Kaakai, S., Matoussi, A., and Sabbagh, W. (2024).
Deep learning scheme for forward utilities using ergodic bsdes.

Cohen, S. N. and Hu, Y. (2013).
Ergodic BSDEs driven by Markov chains.
SIAM J. Control Optim., 51(5):4138–4168.

Debussche, A., Hu, Y., and Tessitore, G. (2011).
Ergodic BSDEs under weak dissipative assumptions.
Stochastic Process. Appl., 121(3):407–426.

Fournié, E., Lasry, J., Lebuchoux, J., and Lions, P. (2001).
Applications of Malliavin calculus to Monte Carlo methods in finance, II.
Finance and Stochastics, 5(2):201–236.

References II

Fuhrman, M., Hu, Y., and Tessitore, G. (2009).
Ergodic BSDES and optimal ergodic control in Banach spaces.
SIAM J. Control Optim., **48**(3):1542–1566.

Guatteri, G. and Tessitore, G. (2020).
Ergodic BSDEs with multiplicative and degenerate noise.
SIAM J. Control Optim., 58(4):2050–2077.

Hu, Y., Madec, P.-Y., and Richou, A. (2015).
A probabilistic approach to large time behavior of mild solutions of HJB equations in infinite dimension.
SIAM J. Control Optim., 53(1):378–398.

Khasminskii, R. (2012).
Stochastic stability of differential equations, volume 66 of *Stochastic Modelling and Applied Probability*.
Springer, Heidelberg, second edition.
With contributions by G. N. Milstein and M. B. Nevelson.

Ma, J. and Yong, J. (1999).
Forward-Backward Stochastic Differential Equations.
Lecture Notes in Mathematics, 1702, Springer-Verlag.
A course on stochastic processes.

References III

Madec, P.-Y. (2015).
Ergodic BSDEs and related PDEs with Neumann boundary conditions under weak dissipative assumptions.
Stochastic Process. Appl., 125(5):1821–1860.

Nualart, D. (2006).
Malliavin calculus and related topics.
Springer Verlag, second edition.
(with corrections on the webpage of the author).

Peng, S. (1993).
Backward stochastic differential equations and applications to optimal control.
Appl. Math. Optim., **27**(2):125–144.

Richou, A. (2009).
Ergodic BSDEs and related PDEs with Neumann boundary conditions.
Stochastic Process. Appl., **119**(9):2945–2969.

Talagrand, M. (1989).
Isoperimetry and integrability of the sum of independent Banach-space valued random variables.
Ann. Probab., 17(4):1546–1570.

References IV

van der Vaart, A. W. and Wellner, J. A. (1996).
Weak Convergence and Empirical Processes: With Applications to Statistics.
Springer Series in Statistics. Springer-Verlag, New York.

Zhang, J. (2017).
Backward stochastic differential equations, volume 86 of *Probability theory and stochastic modelling*.
Springer, New-York.