Forward and backward simulation of Euler scheme

emmanuel.gobet@polytechnique.edu

Centre de Mathématiques Appliquées and FiME,
Ecole Polytechnique and CNRS,
University Paris-Saclay

With the support of La FONDATION du RISQUE

Based on joint works with P. Cohort (Zeliade systems) and M. Mrad (University Paris Nord).

Contents

1 Reversing a Random Number Generator

1.1 Application/motivation

Application: Regression Monte-Carlo with low memory constraints

Dynamic programming equation (DPE) with N times:

$$\mathbf{Y_i} = \mathbb{E}\left[\mathbf{g_i}(\mathbf{Y_{i+1}}, \dots, \mathbf{Y_N}, \mathbf{X_i}, \dots, \mathbf{X_N}) \mid \mathbf{X_i}\right], \quad \mathbf{i} = \mathbf{N} - \mathbf{1}, \dots, \mathbf{0},$$
$$\mathbf{Y_N} = \mathbf{g_N}(\mathbf{X_N}),$$

We want to estimate the function y_i such that $Y_i = y_i(X_i)$.

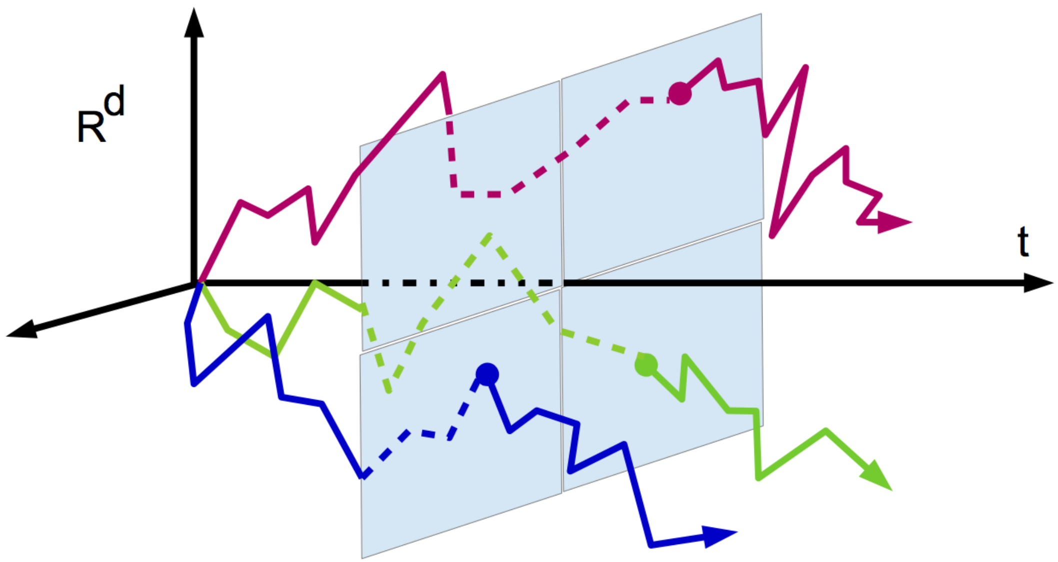
✓ **Optimal Stopping:**

- Value function: $\mathbf{V_i} = \operatorname{ess\,sup}_{\tau \in \mathcal{T}_{\mathbf{i},\mathbf{N}}} \mathbb{E}\left[\mathbf{g}_\tau(\mathbf{X}_\tau) \mid \mathbf{X_i}\right].$
- Continuation value: $\mathbf{Y_i} = \mathbb{E}\left[\mathbf{V_{i+1}} | \mathbf{X_i}\right]$ solves a DPE.

✓ **Semi-linear equations/BSDEs:**

- Non-linear PDE: $\partial_\mathbf{t} \mathbf{u} + \mathcal{L}\mathbf{u} + \mathbf{f}(\mathbf{u}, \sigma \nabla \mathbf{u}) = \mathbf{0}$, $u(1, .) = g(.)$.
- BSDE: $\mathcal{Y}_\mathbf{t} = \mathbb{E}\left[\mathbf{g}(\mathbf{X_1}) + \int_\mathbf{t}^\mathbf{1} \mathbf{f}(\mathbf{s}, \mathcal{Y}_\mathbf{s}, \mathcal{Z}_\mathbf{t}, \mathbf{X_s}) \mathbf{ds} \mid \mathbf{X_t}\right] = \mathbf{u}(\mathbf{t}, \mathbf{X_t})$
- Discrete-time version solves DPE.

Regression Monte-Carlo: **[Longstaff-Schwartz 01, G'-Turkedejiev '16]** value functions are computed using M simulations, and K basis functions



▷ **Convergence of the statistical error:** must have $M \gg K$. More precisely, $\mathbf{M \sim KN^3}$

▷ **Convergence of the approximation error:** $K \sim N^{\alpha d}$

▷ **Memory constraints:**

✓ M paths of length N: MN

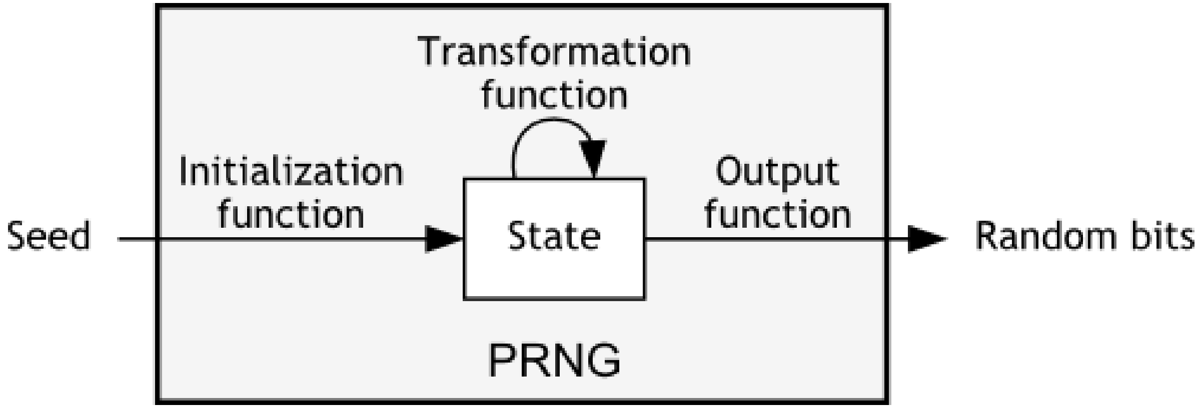
✓ N value functions: KN

In **[Aïd-Campi-Langrené-Pham '14]**, optimal investment in renewable energy. Time discretization: 40 years and 2 decisions per day

➠ $N = 40 \times 2 \times 365 \approx 30000$!!

➠ **Impossible to store the simulations**

1.2 What is a Pseudo-RNG?



Schematic diagram of a pseudo-random number generator

Credits: http://pit-claudel.fr/clement/

Examples.

- ✓ *Linear congruential generator*: $\mathbf{x_{m+1} = a x_m + b}$ MOD $\mathbf{L}$, $\quad \mathbf{u_m = \frac{x_m}{L}}$.
  Ex: $a = 7^5$, $b = 0$ et $L = 2^{31} - 1 = 2147483647$.
- ✓ *Mersenne Twister (1997)*: period $2^{19937} - 1$
- ✓ *Xorshift (2003)*: period $2^{1024} - 1$ (simple and fast)

1.3 Reversing the RNG

PRNG sequence: $\mathbf{u_i = h(G^i(v_0))}$ for some output function h, transformation function G and initialization state v_0.

Is it possible to invert the sequence

$$u_0 \curvearrowright u_1 \curvearrowright \ldots \curvearrowright u_n \qquad \text{to} \qquad u_n \curvearrowright u_1 \curvearrowright \ldots \curvearrowright u_0?$$

▷ A classical example : the L'Ecuyer generator

Classical example among the Multiple Recursive family **[L'Ecuyer '88]**

Transformation function G: defined by

$$G(x,y) := \begin{cases} 40014\,x \text{ MOD } 2147483563 \\ 40692\,y \text{ MOD } 2147483399 \end{cases}$$

Generator output: $h(x,y) = (x-y)/2147483563 \text{ MOD } 1$.

$m_1 := 2147483563$ and $m_2 := 2147483399$ are prime numbers ➠ G is invertible and G^{-1} is of the form of G (same computational complexity).

▷ An up-to-date example : the WELL generators

WELL = *Well Equidistributed Long-period Linear*

Belong to family of Linear Recurrence Modulo 2 generators **[Panneton-L'Ecuyer '06]**.

Transformation function: based on linear recursions $x_{n+1} = Ax_n$, for a binary matrix A.

Output function: $h(x_n) := Bx_n$ for a binary matrix B.

Properties.

- ✓ Generator reversible if A is invertible.
- ✓ G and G^{-1} have similar complexity.
- ✓ Among the 17 WELL generators of **[Panneton-L'Ecuyer '06]**, we found 11 invertible generators (e.g. WELL607a, 19937a, 23209b and 44497a)

2 Euler schemes: forward and backward (v1.0)

IDEA: inverting the RNG allows to backward resimulate the Brownian increments

No need anymore to store them

What about the SDE?

2.1 Forward scheme

$\mathbb{R}^d$-valued diffusion on $[0,1]$: $\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t$.

Forward Euler scheme at times $t_i = i/N$:

$$\overrightarrow{X}_{\mathbf{0}}^{\mathbf{N}} = \mathbf{x}, \quad \overrightarrow{X}_{\mathbf{t_{i+1}}}^{\mathbf{N}} = \overrightarrow{X}_{\mathbf{t_i}}^{\mathbf{N}} + \mu(\mathbf{t_i}, \overrightarrow{X}_{\mathbf{t_i}}^{\mathbf{N}})\frac{\mathbf{1}}{\mathbf{N}} + \sigma(\mathbf{t_i}, \overrightarrow{X}_{\mathbf{t_i}}^{\mathbf{N}})\mathbf{\Delta W_{t_i}}.$$

Theorem. Under Lipschitz assumptions,

$$\sup_{\mathbf{t}\in[\mathbf{0},\mathbf{1}]} \left\| \sup_{|\mathbf{x}|\leq\lambda} |\mathbf{X_t}(\mathbf{x}) - \overrightarrow{X}_{\mathbf{t}}^{\mathbf{N}}(\mathbf{x})| \right\|_{\mathbb{L}_\mathbf{p}} \leq \frac{\mathbf{C}}{\mathbf{N^{1/2}}}\lambda^{\mathbf{2}}, \quad \forall\lambda \geq \mathbf{1}.$$

2.2 Backward scheme

Techniques of stochastic flows [Kunita '97]

✓ **SDE** $X_\cdot(s,x)$ solution from initial condition $x \in \mathbb{R}$ at time $s > 0$.

✓ Inverse flow $\xi_{s,t}(x) := (X_{s,t})^{-1}(x)$.

Theorem ([Kunita '97]). ξ is a semimartingale with respect to t for any (s,x):

$$\begin{cases} d\xi_{s,t}(x) & = & -\text{Jac}(\xi_{s,t}(x))\left[\mu(t,x) - \sum_j \left(\partial_j \sigma(t,x)\right).\sigma^j(t,x)\right]dt + \sigma(t,x)dW_t \\ & + & \frac{1}{2}\sum_{i,j}\partial^2_{ij}\xi_{s,t}(x)\sigma^i(t,x).\sigma^j(t,x)dt, \\ \xi_{s,s}(x) & = & x. \end{cases}$$

ξ is a semimartingale with respect to s for any (t,x):

$$\begin{cases} d\xi_{s,t}(x) & = & -\left[\mu(s,\xi_{s,t}(x)) - \sum_j \left(\partial_j \sigma(s,\xi_{s,t}(x))\right).\sigma^j(s,\xi_{s,t}(x))\right]ds \\ & & -\sigma(s,\xi_{s,t}(x))d\overleftarrow{W}_s, \\ \xi_{t,t}(x) & = & x. \end{cases}$$

Definition (Backward Euler scheme (v1.0)).

$$\overleftarrow{X}^{\mathbf{N}}_{\mathbf{1}} = \overrightarrow{X}^{\mathbf{N}}_{\mathbf{1}},$$

$$\overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} = \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}} - \mu(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\frac{\mathbf{1}}{\mathbf{N}} - \sigma(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\mathbf{\Delta W_{t_i}}$$

$$+ \sum_{\mathbf{j}} \left(\partial_{\mathbf{j}}\sigma(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\right).\sigma^{\mathbf{j}}(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\frac{\mathbf{1}}{\mathbf{N}}.$$

- Easy to simulate using backward regeneration of Brownian increments
- Avoid to invert the forward Euler transform

$$x \mapsto \overrightarrow{X}^{N}_{t_{i+1}} = x + \mu(t_i, x)\frac{1}{N} + \sigma(t_i, x)\Delta W_{t_i}.$$

- ✓ Convergence rate for $\overleftarrow{X}^{N}_{t_i} - X_{t_i}$? or for $\overleftarrow{X}^{N}_{t_i} - \overrightarrow{X}^{N}_{t_i}$?
- Non standard analysis: composition of dependent stochastic maps

2.3 Approximation of compound maps $F(\Theta)$

- ✓ a $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable mapping $(\omega, x) \in (\Omega, \mathbb{R}^d) \mapsto F(\omega, x) \in \mathbb{R}^q$, continuous in x for a.e. ω
- ✓ $\Theta : \Omega \mapsto \mathbb{R}^d$ be a $\mathcal{F}$-random variable.

Our aim: control in $\mathbb{L}_p$ the error $\omega \in \Omega \mapsto F^N(\omega, \Theta^N(\omega)) - F(\omega, \Theta(\omega))$

General result for approximating $F(\Theta)$

(H1) For any $p > 0$, $\exists\ \alpha_p^{(\mathbf{H1})} \geq 0$ s.t.

$$\sup_{\lambda \geq 1} \lambda^{-\alpha_p^{(\mathbf{H1})}} \left\| \sup_{|x| \leq \lambda} |F(\cdot, x)| \right\|_{\mathbb{L}_p} < +\infty.$$

(H2) $\exists\ \kappa \in (0, 1]$ such that $\forall\ p > 0$, $\exists\ \alpha_p^{(\mathbf{H2})} \geq 0$ s.t.

$$\sup_{\lambda \geq 1} \lambda^{-\alpha_p^{(\mathbf{H2})}} \left\| \sup_{x \neq y, |x| \leq \lambda, |y| \leq \lambda} \frac{|F(\cdot, y) - F(\cdot, x)|}{|y - x|^{\kappa}} \right\|_{\mathbb{L}_p} < +\infty.$$

(H3) For any $p > 0$, $\exists\ \alpha_p^{(\mathbf{H3})} \geq 0$ and a non-negative sequence $(\varepsilon_p^{N,(\mathbf{H3})})_{N \geq 1}$ s.t.

$$\sup_{\lambda \geq 1} \lambda^{-\alpha_p^{(\mathbf{H3})}} \left\| \sup_{|x| \leq \lambda} |F^N(\cdot, x) - F(\cdot, x)| \right\|_{\mathbb{L}_p} \leq \varepsilon_{\mathbf{p}}^{\mathbf{N},(\mathbf{H3})}, \quad \forall N \geq 1.$$

(H4) For any $p > 0$, there exists a non-negative sequence $(\varepsilon_p^{N,(\mathbf{H4b})})_{N\geq 1}$ s.t.

$$\sup_{N\geq 1}\left[\|\Theta\|_{\mathbb{L}_p} \vee \|\Theta^N\|_{\mathbb{L}_p}\right] < +\infty,$$

$$\|\Theta^N - \Theta\|_{\mathbb{L}_p} \leq \varepsilon_{\mathbf{p}}^{\mathbf{N},(\mathbf{H4b})}, \quad \forall N \geq 1.$$

Theorem (general result). Assume (**H1-H2-H3-H4**). Then for any $p > 0$ and any $p_2 > p$, there is a constant c independent on N such that

$$\|\mathbf{F}^{\mathbf{N}}(\boldsymbol{\Theta}^{\mathbf{N}}) - \mathbf{F}(\boldsymbol{\Theta})\|_{\mathbb{L}_\mathbf{p}} \leq \mathbf{c}\left(\varepsilon_{\mathbf{2p}}^{\mathbf{N},(\mathbf{H3})} + [\varepsilon_{\kappa\mathbf{p_2}}^{\mathbf{N},(\mathbf{H4b})}]^{\kappa}\right).$$

Corollary (rule of thumb). If

✓ $F_N - F$ "=" $O(N^{-\gamma_F})$ in any L_p,

✓ $\Theta^N - \Theta = O(N^{-\gamma_\theta})$ in any L_p,

the order of L_p-convergence of $F^N(\Theta^N) - F(\Theta)$ is $\gamma_\mathbf{F} \wedge (\kappa\gamma_\theta)$.

Application to backward Euler scheme

Corollary (Backward Euler vs forward diffusion). For any $p \geq 1$,

$$\sup_{\mathbf{t_i} \leq \mathbf{1}} \left\| \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} - \mathbf{X}_{\mathbf{t_i}} \right\|_{\mathbb{L}_\mathbf{p}} = \mathbf{O}(\mathbf{N}^{-\mathbf{1/2}}).$$

Proof. Recall $\xi_{s,t}(x) := (X_{s,t})^{-1}(x)$. Then

$$\overleftarrow{X}^N_{t_i} - X_{t_i} = \xi^N_{t_i,1}\Big(X^N_{0,1}(x_0)\Big) - \xi_{t_i,1}\Big(X_{0,1}(x_0)\Big),$$

$$\sup_{\lambda \geq 1} \lambda^{-2} \left\| \sup_{|x| \leq \lambda} |\xi^N_{t_i,1}(x) - \xi_{t_i,1}(x)| \right\|_{\mathbb{L}_p} \leq \frac{C}{N^{1/2}},$$

$$\sup_{\lambda \geq 1} \lambda^{-1} \left\| \sup_{x \neq y, |x| \leq \lambda, |y| \leq \lambda} \frac{|\xi_{t_i,1}(y) - \xi_{t_i,1}(x)|}{|y - x|} \right\|_{\mathbb{L}_p} < +\infty.$$

Corollary (Backward Euler vs forward Euler). For any $p \geq 1$,

$$\sup_{\mathbf{t_i} \leq \mathbf{1}} \left\| \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} - \overrightarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} \right\|_{\mathbb{L}_\mathbf{p}} = \mathbf{O}(\mathbf{N}^{-1/2}).$$

Next?

✓ Can we improve the rate of retrieving $\overrightarrow{X}^{N}_{t_i}$ by backward generation?

✓ Rate N^{-1}?

☺ It would mean that the backward generation brings a smaller error than the forward generation.

3 Backward Euler scheme (v2.0)

3.1 Definition

$$\overleftarrow{X}^{\mathbf{N}}_{\mathbf{1}} = \overrightarrow{X}^{\mathbf{N}}_{\mathbf{1}},$$

$$\overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} = \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}} - \mu(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\frac{\mathbf{1}}{\mathbf{N}} - \sigma(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\mathbf{\Delta W_{t_i}}$$
$$+ \mathrm{Jac}\big(\sigma(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\mathbf{\Delta W_{t_i}}\big)\sigma(\mathbf{t_{i+1}}, \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_{i+1}}})\mathbf{\Delta W_{t_i}}.$$

In (v1.0), the blue term was $\sum_j \big(\partial_j \sigma(t_{i+1}, \overleftarrow{X}^N_{t_{i+1}})\big).\sigma^j(t_{i+1}, \overleftarrow{X}^N_{t_{i+1}})\frac{1}{N}$.

3.2 Convergence rate and CLT

Theorem (L_p-error). For any $p \geq 1$,

$$\sup_{\mathbf{t_i} \leq \mathbf{1}} \left\| \overleftarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} - \overrightarrow{X}^{\mathbf{N}}_{\mathbf{t_i}} \right\|_{\mathbb{L}_\mathbf{p}} = \mathbf{O}(\mathbf{N}^{-1}).$$

Theorem (CLT). For some processes a and b and an extra Brownian motion B, the error $\mathcal{E}_t^N = \overleftarrow{X}_t^N - \overrightarrow{X}_t^N$ satisfies

$$\left\{N\mathcal{E}_t^N : 0 \leq t \leq 1\right\} \stackrel{d}{\Longrightarrow} \int_.^1 a_s ds + \int_.^1 b_s dB_s.$$

Proof (of L_p error and CLT).

▷ First prove that $\sup_{t_i \leq 1} \left\|\mathcal{E}_{t_i}^N\right\|_{\mathbb{L}_p} = O(N^{-1/2})$.

▷ Then expand $\mathcal{E}_t^N$ so that, for some adapted processes, we have

$$\sup_{0<N,\, 0\leq k \leq N} N^{3/2} \left\|\mathcal{E}_{t_k}^N - L_{t_k}^N\right\|_{\mathbb{L}_p} < +\infty$$

where

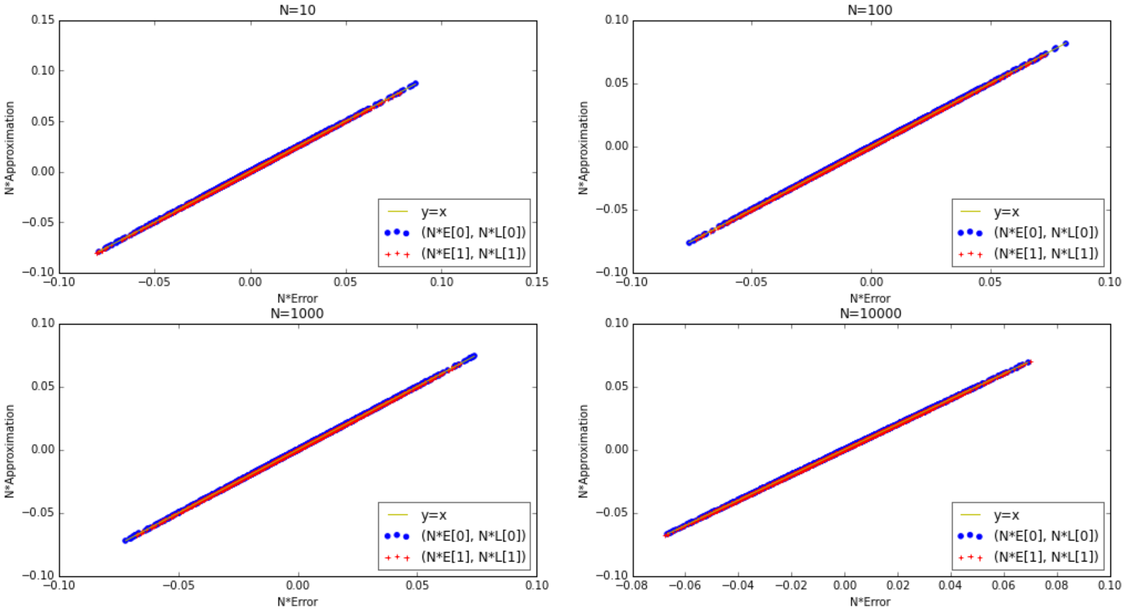
$$L_{t_k}^N := \sum_{j=k}^{N-1} a_k^{(j)} \left(d_j(\Delta W_j^3) + e_j(\Delta W_j^1)\Delta t + f_j(\Delta W_j^4) + g_j(\Delta W_j^2)\Delta t + h_j \Delta t^2\right).$$

3.3 Numerical results

- ✓ 10000 paths
- ✓ empirical joint distribution of $(N\mathcal{E}_0^N, NL_0^N)$
- ✓ examples in dimension 2

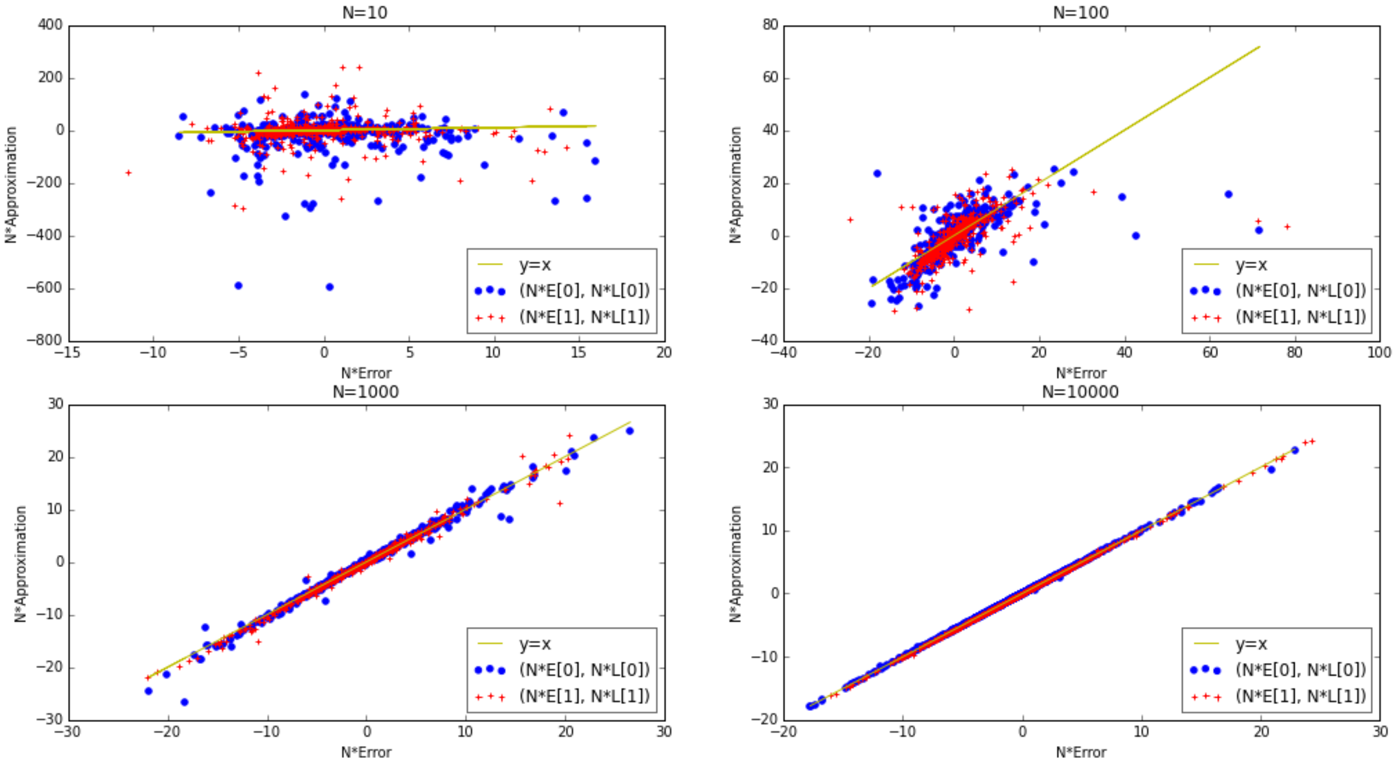
Linear Vol: standard non correlated two-dimensional Black&Scholes case

$$\mu(x) = (0,0), \qquad \sigma(x) = \begin{pmatrix} 0.2\,x_0 & 0 \\ 0 & 0.2\,x_1 \end{pmatrix}, \qquad x_0 = (1,1).$$



Non linear Vol:

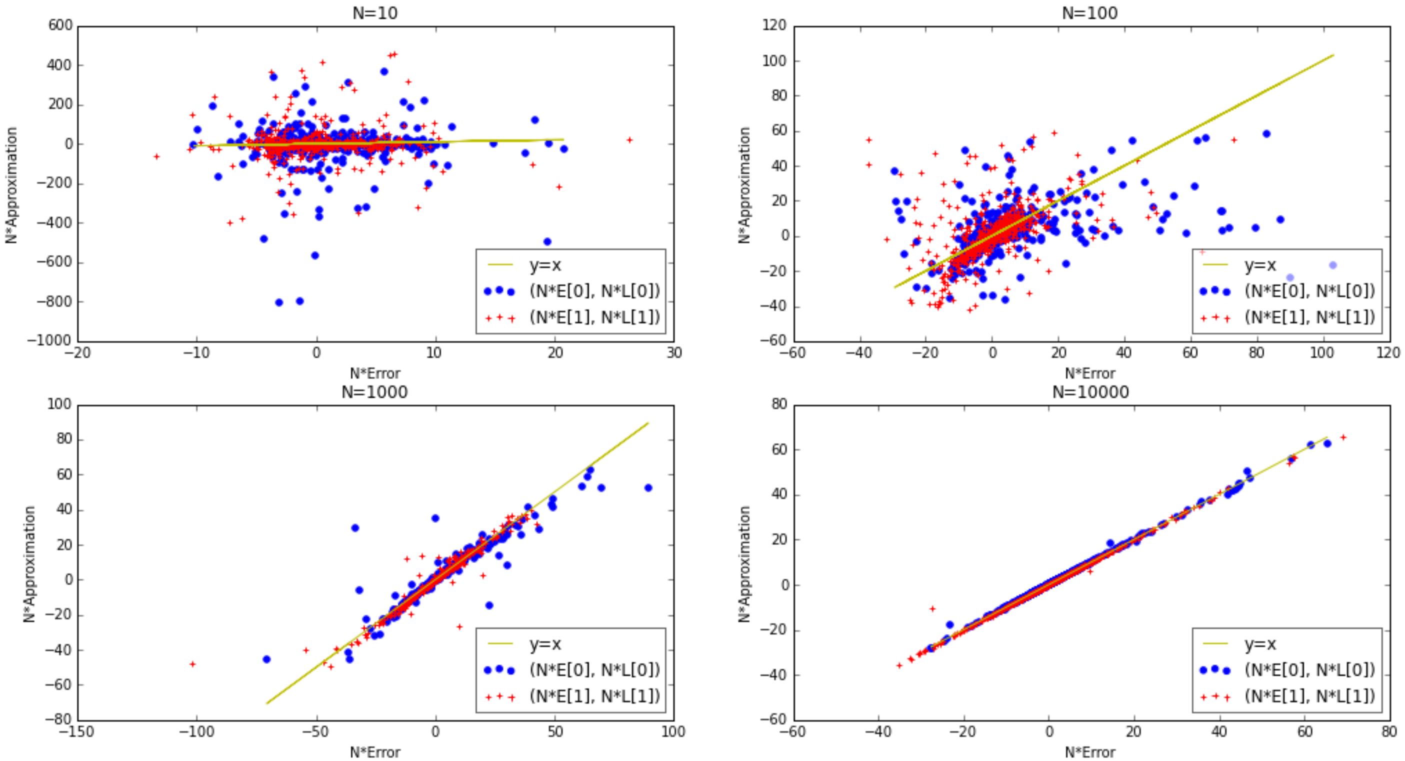
$$\mu(x) = (0,0), \quad \sigma(x) = \begin{pmatrix} 0.2\,x_0(1+\sin(2\pi x_0)) & 0.2\,x_0(1+\cos(2\pi x_1)) \\ 0.2\,x_1(1+\cos(2\pi x_0)) & 0.2\,x_1(1+\sin(2\pi x_1)) \end{pmatrix}, \quad Y_0 = (1,1).$$



Non linear drift and vol:

$$\mu(x) = (0.2\,x_0(1+\sin(2\pi x_0)), 0.2\,x_0(1+\sin(2\pi x_1)))$$

$$\sigma(x) = \begin{pmatrix} 0.2\,x_0(1+\sin(2\pi x_0)) & 0.2\,x_0(1+\cos(2\pi x_1)) \\ 0.2\,x_1(1+\cos(2\pi x_0)) & 0.2\,x_1(1+\sin(2\pi x_1)) \end{pmatrix}, \qquad Y_0 = (1,1).$$



4 Conclusion

✓ Accurate resampling of Euler schemes using backward regeneration of Brownian increments

✓ Useful for Regression Monte Carlo algorithms with strong memory constraints

✓ New tools for analyzing the approximation of compound stochastic maps.

✓ Other on-going applications:

- approximation of processes at random times (fBM, Diffusion at Brownian times,. . .)
- approximation of SPDEs by composition of stochastic flows
- approximation of Progressive Stochastic Utilities **[Musiela, Zariphopoulou '10; El Karoui, Mrad '13]**.

References

[ACLP14] R. Aïd, L. Campi, N. Langrené, and H. Pham. A probabilistic numerical method for optimal multiple switching problem in high dimension. *SIAM J. on Financial Mathematics*, 5(1):191–231, 2014.

[CGM16] P. Cohort, E. Gobet, and M. M'Rad. Forward and backward Euler scheme and Random Number Generator inversion. *Preprint*, 2016.

[EM13] N. El Karoui and M. M'Rad. An exact connection between two solvable SDEs and a nonlinear utility stochastic PDE. *SIAM J. Financial Math.*, 4(1):697–736, 2013.

[GM15] E. Gobet and M. M'Rad. Convergence rate of strong approximations of compound random maps. *Preprint hal-01141320*, 2015.

[GM17] E. Gobet and M. M'Rad. Strong approximation of stochastic processes at random times and application to their exact simulation. *Stochastics*, 2017.

[Kun97] H. Kunita. *Stochastic flows and stochastic differential equations.* Cambridge Studies in Advanced Mathematics. 24. Cambridge: Cambridge University Press, 1997.

[L'E88] P. L'Ecuyer. Efficient and portable combined random number generators. *Communications of the ACM*, 31(6):742–774, 1988.

[LS01] F. Longstaff and E.S. Schwartz. Valuing American options by simulation: A simple least squares approach. *The Review of Financial Studies*, 14:113–147, 2001.

[PL06] F. Panneton and P. L'Ecuyer. Improved long-period generators based on linear recurrences modulo 2. *ACM Transactions on Mathematical Software*, 32(1):1–16, 2006.