Importance Sampling and Statistical Romberg method

Ahmed Kebaier
(joint work with Mohamed Ben Alaya and Kaouther Hajji)
University Paris 13, France

Stochastic Processes and their Statistics in Finance

October 28th 2013

Outline of The Talk

Outline

The Model

- $X \in \mathbb{R}^d$ be solution to

$$dX_t = b(X_t)dt + \sum_{j=1}^{q} \sigma_j(X_t)dW_t^j, \quad X_0 = x \in \mathbb{R}^d$$

where $W = (W^1, \ldots, W^q)$ is a q-dimensional Brownian motion.

- Functions $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ and $\sigma_j : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, $1 \leq j \leq q$, satisfy condition

$$(\mathcal{H}_{b,\sigma})\ \forall x, y \in \mathbb{R}^d\ |b(x) - b(y)| + \sum_{j=1}^{q} |\sigma_j(x) - \sigma_j(y)| \leq C_{b,\sigma}|x - y|,$$

Discretization error

Let X^n be the Euler scheme with time step $\delta = T/n$

$$dX_t^n = b(X_{\eta_n(t)})dt + \sum_{j=1}^{q} \sigma_j(X_{\eta_n(t)})dW_t^j, \quad \eta_n(t) = [t/\delta]\delta.$$

under condition $(\mathcal{H}_{b,\sigma})$ we have property

$$(\mathcal{P}) \ \ \forall p \geq 1, \ X, X^n \in L^p \quad \text{ and } \mathbb{E}[\sup_{0\leq t\leq T} |X_t - X_t^n|^p] \leq \frac{K_p(T)}{n^{p/2}}.$$

Discretization error

In the context of possibly degenerate diffusions

- For a given function ψ, we set

$$\varepsilon_n := \mathbb{E}\psi(X_T) - \mathbb{E}\psi(X_T^n)$$

- If ψ b and σ_j are $\mathcal{C}_P^4$ then $\varepsilon_n \simeq 1/n$

- However, if ψ is only of class $\mathcal{C}^1$, then we have $\varepsilon_n \simeq 1/n^\alpha$ for any $\alpha \in [1/2, 1]$

From now on, we suppose

$$(\mathcal{H}_{\varepsilon_n}^\alpha)\ n^\alpha \varepsilon_n := n^\alpha \mathbb{E}\psi(X_T^n) - \mathbb{E}\psi(X_T) \to C_\psi(T, \alpha) \text{ for } \alpha \in [1/2, 1].$$

CLT for Monte Carlo method

Theorem

Let $\psi \in \mathcal{C}^1$ *s.t. we have*

$$(\mathcal{H}^{\alpha}_{\varepsilon_{\mathbf{n}}}) \qquad \lim_{n\to\infty} n^{\alpha}\varepsilon_n = C_\psi(T,\alpha)$$

Then,

$$n^{\alpha}\Big(\frac{1}{n^{2\alpha}}\sum_{i=1}^{n^{2\alpha}}\psi(X^n_{T,i}) - \mathbb{E}\,\psi(X_T)\Big) \Rightarrow \sigma G + C_\psi(T,\alpha),$$

with $\sigma^2 = Var(\psi(X_T))$.

CLT for Monte Carlo method

Theorem

Let $\psi \in \mathcal{C}^1$ s.t. we have

$$(\mathcal{H}^{\alpha}_{\varepsilon_{\mathbf{n}}}) \qquad \lim_{n\to\infty} n^{\alpha}\varepsilon_n = C_{\psi}(T,\alpha)$$

Then,

$$n^{\alpha}\Big(\frac{1}{n^{2\alpha}}\sum_{i=1}^{n^{2\alpha}} \psi(X^n_{T,i}) - \mathbb{E}\,\psi(X_T)\Big) \Rightarrow \sigma G + C_{\psi}(T,\alpha),$$

with $\sigma^2 = Var(\psi(X_T))$.

Optimal time complexity

$$C_{MC} = C \times n^{2\alpha+1}$$

Statistical Romberg algorithm

- We construct two Euler schemes X_T^n and $X_T^{\sqrt{n}}$ with time step T/n and $T/\sqrt{n}$.
- Let
$$E = \mathbb{E}\psi\left(X_T^{\sqrt{n}}\right).$$
- We set
$$Q = \psi\left(X_T^n\right) - \psi\left(X_T^{\sqrt{n}}\right) + E$$
- Note that
$$\mathbb{E}(Q) = \mathbb{E}\psi(X_T^n) \text{ and } Var(Q) = O\left(\frac{1}{\sqrt{n}}\right)$$

Statistical Romberg method

The statistical Romberg routine that approximates $\mathbb{E}\psi(X_T)$ using only two empirical means

$$V_n := \frac{1}{N_1}\sum_{i=1}^{N_1}\psi(\hat{X}^{\sqrt{n}}_{T,i}) + \frac{1}{N_2}\sum_{i=1}^{N_2}\psi(X^{n}_{T,i}) - \psi(X^{\sqrt{n}}_{T,i}).$$

Under assumption $(\mathcal{H}^{\alpha}_{\varepsilon_n})$, this method is tamed by a central limit theorem with a rate of convergence equal to n^{α} (Kebaier 2005). More precisely, for $N_1 = n^{2\alpha}$, $N_2 = n^{2\alpha-1/2}$ we have

$$n^{\alpha}(V_n - \mathbb{E}\psi(X_T)) \to \mathcal{N}(C_{\psi}(T,\alpha), \sigma^2), \text{ with}$$

$$\sigma^2 := \mathrm{Var}\left(\psi(X_T)\right) + \tilde{\mathrm{Var}}\left(\nabla\psi(X_T)\cdot U_T\right),$$

Statistical Romberg method

- The process U is the weak limit process of the error $\sqrt{n}(X^n - X)$ and is solution to

$$dU_t = \dot{b}(X_t)U_t dt + \sum_{j=1}^{q} \dot{\sigma}_j(X_t)U_t dW_t^j - \frac{1}{\sqrt{2}} \sum_{j,\ell=1}^{q} \dot{\sigma}_j(X_t)\sigma_\ell(X_t) d\tilde{W}_t^{\ell j},$$

  where $\tilde{W}$ is a q^2-dimensional standard Brownian motion, independent of W, and $\dot{b}$ (respectively $(\dot{\sigma}_j)_{1\leq j\leq q}$) is the Jacobian matrix of b (respectively $(\sigma_j)_{1\leq j\leq q}$).
- This result is due to Jacod-Kurtz-Protter (91-98) provided that b and σ are $\mathcal{C}^1$.

Importance Sampling

We define the family of $\mathbb{P}_\theta$, as all the equivalent probability measures with respect to $\mathbb{P}$ such that

$$L_t^\theta = \frac{d\mathbb{P}_\theta}{d\mathbb{P}}|_{\mathcal{F}_t} = \exp\left(\theta \cdot W_t - \frac{1}{2}|\theta|^2 t\right).$$

Hence, $B_t^\theta := W_t - \theta t$ is a Brownian motion under $\mathbb{P}_\theta$. This leads to

$$\mathbb{E}\psi(X_T) = \mathbb{E}_\theta\left[\psi(X_T)e^{-\theta\cdot B_T^\theta - \frac{1}{2}|\theta|^2 T}\right].$$

The optimal θ parameter is chosen so that it reduces

$$Var_\theta\left[\psi(X_T)e^{-\theta\cdot B_T^\theta - \frac{1}{2}|\theta|^2 T}\right]$$

Let us introduce the process X_t^θ solution, under $\mathbb{P}$, to

$$dX_t^\theta = \left(b(X_t^\theta) + \sum_{j=1}^{q} \theta_j \sigma_j(X_t^\theta) \right) dt + \sum_{j=1}^{q} \sigma_j(X_t^\theta) dW_t^j,$$

$(B_t^\theta, X_t)_{t\geq 0}$ under $\mathbb{P}_\theta$ has the same law as $(W_t, X_t^\theta)_{t\geq 0}$ under $\mathbb{P}$ we get

$$\mathbb{E}\psi(X_T) = \mathbb{E}g(\theta, X_T^\theta, W_T), \text{ with } g(\theta, x, y) = \psi(x) e^{-\theta\cdot y - \frac{1}{2}|\theta|^2 T}.$$

We also introduce the Euler continuous approximation $X^{n,\theta}$ of the process X^θ solution, under $\mathbb{P}$, to

$$dX_t^{n,\theta} = \left(b(X_{\eta_n(t)}^{n,\theta}) + \sum_{j=1}^{q} \theta_j \sigma_j(X_{\eta_n(t)}^{n,\theta}) \right) dt + \sum_{j=1}^{q} \sigma_j(X_{\eta_n(t)}^\theta) dW_t^j,$$

Our target now is to approximate $\mathbb{E}\psi(X_T) = \mathbb{E}g(\theta, X_T^\theta, W_T)$ by

$$\frac{1}{N_1}\sum_{i=1}^{N_1} g(\theta, \hat{X}_{T,i}^{\sqrt{n},\theta}, \hat{W}_{T,i}) + \frac{1}{N_2}\sum_{i=1}^{N_2} g(\theta, X_{T,i}^{\sqrt{n},\theta}, W_{T,i}) - g(\theta, X_{T,i}^{\sqrt{n},\theta}, W_{T,i}).$$

According to Kebaier (2005), we have a CLT with limit variance

$$\mathrm{Var}\left(g(\theta, X_T^\theta, W_T)\right) + \tilde{\mathrm{Var}}\left(\nabla_x g(\theta, X_T^\theta, W_T)\cdot U_T^\theta\right)$$

where U^θ is the weak limit process of the error $\sqrt{n}(X^{n,\theta} - X^\theta)$, solution to

$$dU_t^\theta = \left(\dot{b}(X_t^\theta) + \sum_{j=1}^{q}\theta_j\dot{\sigma}_j(X_t^\theta)\right)U_t^\theta dt + \sum_{j=1}^{q}\dot{\sigma}_j(X_t^\theta)U_t^\theta dW_t^j$$
$$- \frac{1}{\sqrt{2}}\sum_{j,\ell=1}^{q}\dot{\sigma}_j(X_t^\theta)\sigma_\ell(X_t^\theta)d\tilde{W}_t^{\ell j}.$$

it boils down to choose $\theta^* = \underset{\theta \in \mathbb{R}^q}{\operatorname{argmin}}\, v(\theta)$

$$v(\theta) := \tilde{\mathbb{E}}\left(\left[\psi(X_T^\theta)^2 + (\nabla\psi(X_T^\theta)\cdot U_T^\theta)^2\right] e^{-2\theta.W_T - |\theta|^2 T}\right)$$

Note that $v(\theta)$ is not explicit, we introduce $\theta_n^* := \underset{\theta \in \mathbb{R}^q}{\operatorname{argmin}}\, v_n(\theta)$

$$v_n(\theta) := \tilde{\mathbb{E}}\left(\left[\psi(X_T^{n,\theta})^2 + (\nabla\psi(X_T^{n,\theta})\cdot U_T^{n,\theta})^2\right] e^{-2\theta.W_T - |\theta|^2 T}\right)$$

with $U^{n,\theta}$ is the Euler discretization scheme of U^θ, solution to

$$dU_t^{n,\theta} = \left(\dot{b}(X_{\eta_n(t)}^{n,\theta}) + \sum_{j=1}^{q} \theta_j \dot{\sigma}_j(X_{\eta_n(t)}^{n,\theta})\right) U_{\eta_n(t)}^{n,\theta} dt$$

$$+ \sum_{j=1}^{q} \dot{\sigma}_j(X_{\eta_n(t)}^{n,\theta}) U_{\eta_n(t)}^{n,\theta} dW_t^j - \frac{1}{\sqrt{2}} \sum_{j,\ell=1}^{q} \dot{\sigma}_j(X_{\eta_n(t)}^{n,\theta}) \sigma_\ell(X_{\eta_n(t)}^{n,\theta}) d\tilde{W}_t^{\ell j}.$$

Theorem

Suppose σ and b are in $\mathcal{C}^2$ with bounded first and second derivatives. Then for any $\theta \in \mathbb{R}$ the following property holds

$$(\tilde{\mathcal{P}})\ \forall p \geq 1,\ U^{\theta}, U^{n,\theta} \in L^p \quad \text{and} \quad \tilde{\mathbb{E}}\left[\sup_{0\leq t\leq T} |U_t^{\theta} - U_t^{n,\theta}|^p\right] \leq \frac{K_p(T)}{n^{p/2}}.$$

In particular, for $\theta = 0$ the above property holds for the processes U and U^n.

existence and uniqueness of θ^*

Theorem

Suppose σ and b are in $\mathcal{C}^2$ with bounded first and second derivatives and let ψ in $\mathcal{C}^1$ such that $\mathbb{P}(\psi(X_T) \neq 0) > 0$.

- *If there exists $a > 1$ such that $\mathbb{E}\left[\psi^{2a}(X_T)\right]$ and $\mathbb{E}\left[|\nabla\psi(X_T)|^{2a}\right]$ are finite,*

Then the function $\theta \mapsto v(\theta)$ is $\mathcal{C}^2$ and strictly convex with $\nabla v(\theta) = \tilde{\mathbb{E}} H(\theta, X_T, U_T, W_T)$ where

$$H(\theta, X_T, U_T, W_T) := (\theta T - W_T)\left[\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2\right] e^{-\theta\cdot W_T + \frac{1}{2}|\theta|^2 T}.$$

Moreover, there exists a unique $\theta^* \in \mathbb{R}^q$ such that $min_{\theta\in\mathbb{R}^q} v(\theta) = v(\theta^*)$.

Proof

First of all, note the process (B, X, U) under $\tilde{\mathbb{P}}_\theta$ has the same law as (W, X^θ, U^θ) under $\tilde{\mathbb{P}}$. So, using a change of probability, we get

$$v(\theta) := \tilde{\mathbb{E}}\left(\left[\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2\right] e^{-\theta.W_T + \frac{1}{2}|\theta|^2 T}\right).$$

It follows that

- The map $\theta \mapsto \left[\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2\right] e^{-\theta.W_T + \frac{1}{2}|\theta|^2 T}$ is $\mathcal{C}^1$
- $\nabla v(\theta) = H(\theta, X_T, U_T, W_T)$

For $c > 0$ we have,

$$\sup_{|\theta|\le c} |H(\theta, X_T, U_T, W_T)| \le (cT + |W_T|)\left[\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2\right] e^{c|W_T| + \frac{1}{2}c^2 T}.$$

Proof

- Using Holder's inequality, $\tilde{\mathbb{E}} \sup_{|\theta| \leq c} |H(\theta, X_T, U_T, W_T)|$ is bounded by

$$e^{\frac{1}{2}c^2 T} \Big(\|\psi^2(X_T)\|_a \|e^{c|W_T|}(cT + |W_T|)\|_{\frac{a}{a-1}} \\ + \||\nabla\psi(X_T)|^2\|_a \||U_T|^2\|_{\frac{2a}{a-1}} \|e^{c|W_T|}(cT + |W_T|)\|_{\frac{2a}{a-1}} \Big).$$

- Using property $(\tilde{\mathcal{P}})$ and $\mathbb{E}\psi^{2a}(X_T)$ and $\mathbb{E}|\nabla\psi(X_T)|^{2a}$ are finite we conclude the boundedness of $\tilde{\mathbb{E}} \sup_{|\theta| \leq c} |H(\theta, X_T, U_T, W_T)|$.

- In the same way, we prove that v is of class $\mathcal{C}^2$ in $\mathbb{R}^q$

$$\mathrm{Hess}(v(\theta)) = \tilde{\mathbb{E}}\Big[((\theta T - W_T)(\theta T - W_T)^* + T I_q) \times(\psi^2(X_T) + (\nabla\psi(X_T)\cdot U_T)^2)e^{-\theta.W_T+\frac{1}{2}|\theta|^2 T}\Big].$$

Since $\mathbb{P}(\psi(X_T) \neq 0) > 0$, we get for all $u \in \mathbb{R}^q \backslash \{0\}$

$$u^* \ \mathrm{Hess}(v(\theta)) \ u = \tilde{\mathbb{E}}\Big[T|u|^2 + (u.(\theta T - W_T))^2(\psi^2(X_T) + (\nabla\psi(X_T)\cdot U_T)^2)e^{-\theta.W_T+\frac{1}{2}|\theta|^2 T}\Big] > 0.$$

- Now it will be sufficient to prove that $\lim_{|\theta|\to\infty} \ v(\theta) = +\infty$

$$v(\theta) = \tilde{\mathbb{E}}\left[(\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2)e^{-\theta.W_T+\frac{1}{2}|\theta|^2 T}\right].$$

$$+\infty = \tilde{\mathbb{E}}\left[\liminf_{|\theta|\to\infty}(\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2)e^{-\theta.W_T+\frac{1}{2}|\theta|^2 T}\right]$$

$$\leq \liminf_{|\theta|\to+\infty} \tilde{\mathbb{E}}\left[(\psi(X_T)^2 + (\nabla\psi(X_T)\cdot U_T)^2)e^{-\theta.W_T+\frac{1}{2}|\theta|^2 T}\right].$$

The same results can be obtained for the Euler scheme X^n.

Theorem

Suppose σ and b are in $\mathcal{C}^2$ with bounded first and second derivatives.Let ψ be $\mathcal{C}^1$ such that $\mathbb{P}(\psi(X_T^n) \neq 0) > 0$.

- *If there exists $a > 1$ such that $\mathbb{E}\left[\psi^{2a}(X_T^n)\right]$ and $\mathbb{E}\left[|\nabla\psi(X_T^n)|^{2a}\right]$ are finite*

Then the function $\theta \mapsto v_n(\theta)$ is $\mathcal{C}^2$ and strictly convex with

$$\nabla v_n(\theta) = \tilde{\mathbb{E}} H(\theta, X_T^n, U_T^n, W_T).$$

- *Moreover, there exists a unique $\theta_n^* \in \mathbb{R}^q$ such that $min_{\theta\in\mathbb{R}^q} v_n(\theta) = v_n(\theta_n^*)$.*

Further, we prove the convergence of θ_n^* towards θ^* as n tends to infinity.

Theorem

Suppose σ and b are in $\mathcal{C}^2$ with bounded first and second derivatives. Let ψ be $\mathcal{C}^1$ such that $\mathbb{P}(\psi(X_T) \neq 0) > 0$ and for all $n \in \mathbb{N}$, $\mathbb{P}(\psi(X_T^n) \neq 0) > 0$.

- *If there exists $a > 1$ such that $\mathbb{E}\left[\psi^{2a}(X_T)\right]$, $\sup_{n\in\mathbb{N}} \mathbb{E}\left[\psi^{2a}(X_T^n)\right]$, $\mathbb{E}\left[|\nabla\psi(X_T)|^{2a}\right]$ and $\sup_{n\in\mathbb{N}} \mathbb{E}\left[|\nabla\psi(X_T^n)|^{2a}\right]$ are finite.*

Then,

$$\theta_n^* \longrightarrow \theta^*, \qquad \text{as } n \to \infty.$$

Outline

- The aim now is to construct for fixed n some sequences $(\theta_i^n)_{i\in\mathbb{N}}$ such that $\lim_{i\to\infty}\theta_i^n = \theta_n^* \arg\min_{\theta\in\mathbb{R}} v_n(\theta) =$ almost surely.

- Indeed, using the Robbins-Monro algorithm, we construct recursively the sequence of random variables $(\theta_i^n)_{i\in\mathbb{N}}$ in $\mathbb{R}^q$ given by

$$\theta_{i+1}^n = \theta_i^n - \gamma_{i+1} H(\theta_i^n, X_{T,i+1}^n, U_{T,i+1}^n, W_{T,i+1}),\ i \geq 0,\ \theta_0^n \in \mathbb{R}^q,$$

$(\gamma_i)_{i\geq 1}$ is a decreasing sequence of positive real numbers satisfying

$$\sum_{i=1}^{\infty} \gamma_i = \infty \quad \text{and} \quad \sum_{i=1}^{\infty} \gamma_i^2 < \infty$$

- To obtain the a.s. convergence of θ_i^n to θ_n^*, we need to check
  - $\forall \theta \neq \theta_n^*,\ \langle \nabla v_n(\theta), \theta - \theta_n^* \rangle > 0,$
  - (NEC) $\tilde{\mathbb{E}}\left[|H(\theta, X_T^n, U_T^n, W_T)|^2\right] \leq C(1+|\theta|^2),$ for all $\theta \in \mathbb{R}^q$.

Unfortunately, this condition is not satisfied in our context.

Constrained stochastic algorithm

Let $(\mathcal{K}_i)_{i\in\mathbb{N}}$ denote an increasing sequence of compact sets satisfying $\cup_{i=0}^{\infty} \mathcal{K}_i = \mathbb{R}^d$ and $\mathcal{K}_i \subsetneq \mathring{\mathcal{K}}_{i+1}, \forall i \in \mathbb{N}$. For $\theta_0^n \in \mathcal{K}_0$, $\alpha_0^n = 0$ and a gain sequence $(\gamma_i)_{i\in\mathbb{N}}$ satisfying (**??**), we define the sequence $(\theta_i^n, \alpha_i^n)_{i\in\mathbb{N}}$ recursively by

$$\begin{cases} \textbf{if} & \theta_i^n - \gamma_{i+1} H(\theta_i^n, X_{T,i+1}^n, U_{T,i+1}^n, W_{T,i+1}) \in \mathcal{K}_{\alpha_i^n}, \ \textbf{then} \\ & \theta_{i+1}^n = \theta_i^n - \gamma_{i+1} H(\theta_i^n, X_{T,i+1}^n, U_{T,i+1}^n, W_{T,i+1}), \text{and } \alpha_{i+1}^n = \alpha_i^n \\ \textbf{else} & \theta_{i+1}^n = \theta_0^n \text{ and } \alpha_{i+1}^n = \alpha_i^n + 1, \end{cases}$$

Constrained stochastic algorithm

Theorem

Suppose σ *and* b *are* $\mathcal{C}^2$ *with bounded first and second derivatives and* ψ *is* $\mathcal{C}^1$*. Assume that for all* $n \in \mathbb{N}$, $\mathbb{P}(\psi(X_T^n) \neq 0) > 0$.

- *there exists* $a > 1$ *s.t.* $\mathbb{E}\left[\psi^{4a}(X_T^n)\right]$ *and* $\mathbb{E}\left[|\nabla\psi(X_T^n)|^{4a}\right] < \infty$

Then the sequence $(\theta_i^n)_{i\geq 0}$ *satisfies*

1. *For all* $n \in \mathbb{N}$*, we have* $\theta_i^n \underset{i\to\infty}{\longrightarrow} \theta_n^*$*, a.s.*
2. *Reversely, for all* $i \in \mathbb{N}$*, we have* $\theta_i^n \underset{n\to\infty}{\longrightarrow} \theta_i$*, a.s.,*

$$\begin{cases} \textbf{if} & \theta_i - \gamma_{i+1} H(\theta_i, X_{T,i+1}, U_{T,i+1}, W_{T,i+1}) \in \mathcal{K}_{\alpha_i}, \ \textbf{then} \\ & \theta_{i+1} = \theta_i - \gamma_{i+1} H(\theta_i, X_{T,i+1}, U_{T,i+1}, W_{T,i+1}), \text{and } \alpha_{i+1} = \alpha_i \\ \textbf{else} & \theta_{i+1} = \theta_0 \text{ and } \alpha_{i+1} = \alpha_i, \end{cases}$$

The following corollary follows immediately

Corollary

Under above assumptions the constrained algorithm given satisfies

$$\lim_{i,n\to\infty} \theta_i^n = \lim_{i\to\infty}\big(\lim_{n\to\infty} \theta_i^n\big) = \lim_{n\to\infty}\big(\lim_{i\to\infty} \theta_i^n\big) = \theta^*, \quad \tilde{\mathbb{P}}\text{-}a.s.,$$

where $\theta^* = \underset{\theta\in\mathbb{R}^q}{\operatorname{argmin}}\, v(\theta)$

$$v(\theta) := \tilde{\mathbb{E}}\left(\left[\psi(X_T^\theta)^2 + (\nabla\psi(X_T^\theta)\cdot U_T^\theta)^2\right] e^{-2\theta.W_T - |\theta|^2 T}\right)$$

Unconstrained stochastic algorithm

- We use the idea proposed by Lemaire and Pagès (2009), a new algorithm that satisfies (NEC). In our context we have

$$\nabla v_n(\theta) = \tilde{\mathbb{E}}\left((\theta T - W_T)\left[\psi(X_T^n)^2 + (\nabla\psi(X_T^n)\cdot U_T^n)^2\right] e^{-\theta\cdot W_T + \frac{1}{2}|\theta|^2 T}\right).$$

- To do so, we apply Girsanov theorem, with shift parameter $-\theta$.

$$B_t^{(-\theta)} := W_t + \theta t \;\text{ and }\; L_t^{(-\theta)} := \frac{d\mathbb{P}_{(-\theta)}}{d\mathbb{P}}|_{\mathcal{F}_t} = e^{-\theta . W_t - \frac{1}{2}|\theta|^2 t}$$

$$\begin{aligned}\nabla v_n(\theta) &= \tilde{\mathbb{E}}_{(-\theta)}\left[(2\theta T - B_T^{(-\theta)})\left[\psi(X_T^n)^2 + (\nabla\psi(X_T^n)\cdot U_T^n)^2\right] e^{|\theta|^2 T}\right] \\ &= \tilde{\mathbb{E}}\left[(2\theta T - W_T)\left[\psi(X_T^{n,(-\theta)})^2 + (\nabla\psi(X_T^{n,(-\theta)})\cdot U_T^{n,(-\theta)})^2\right] e^{|\theta|^2 T}\right],\end{aligned}$$

since $(B^{(-\theta)}, X^n, U^n, \tilde{\mathbb{P}}_{(-\theta)}) \stackrel{law}{=} (W, X^{n,(-\theta)}, U^{n,(-\theta)}, \tilde{\mathbb{P}})$

- We need in this context to strengthen our assumptions on ψ and suppose that $\partial_\alpha \psi$ are with polynomial growth for $|\alpha| \leq 1$

- we introduce for a given $\eta > 0$, a new function

$$\tilde{H}_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T) = e^{-\eta|\theta|^2 T}(2\theta T - W_T) \times \left[\psi(X_T^{n,(-\theta)})^2 + (\nabla\psi(X_T^{n,(-\theta)}) \cdot U_T^{n,(-\theta)})^2\right].$$

Then, the algorithm is given by

$$\theta_{i+1}^n = \theta_i^n - \gamma_{i+1} H_\eta(\theta_i^n, X_{T,i+1}^{n,(-\theta_i^n)}, U_{T,i+1}^{n,(-\theta_i^n)}, W_{T,i+1}), \quad \theta_0 \in \mathbb{R}. \quad (1)$$

This algorithm would behave like a classical Robbins-Monro one and does not suffer from the violation of (NEC).

Theorem

Suppose σ and b are $\mathcal{C}^2$ with bounded first and second derivatives. Let ψ in $\mathcal{C}^1$ such that for and for all $n \in \mathbb{N}$, $\mathbb{P}(\psi(X_T^n) \neq 0) > 0$. In addition, assume that for $\lambda > 0$ we have

$$|\nabla\psi(x)| \leq C_\psi(1 + |x|^\lambda) \quad \text{for all } x \in \mathbb{R}^d \text{ and } C_\psi > 0.$$

Then, the sequence $(\theta_i^n)_{i\geq 0}$ given by routine (1), satisfies

$$\forall n \in \mathbb{N}, \quad \theta_i^n \underset{i\to\infty}{\longrightarrow} \theta_n^*, \quad a.s.$$

$$\theta_n^* := \underset{\theta\in\mathbb{R}^q}{\operatorname{argmin}} \tilde{\mathbb{E}}\left(\left[\psi(X_T^{n,\theta})^2 + (\nabla\psi(X_T^{n,\theta}) \cdot U_T^{n,\theta})^2\right] e^{-2\theta.W_T - |\theta|^2 T}\right)$$

Proof

- We have to check first that $\forall\theta \neq \theta_n^*$

$\langle h_n(\theta), \theta - \theta_n^*\rangle > 0$, where $h_n(\theta) = \tilde{\mathbb{E}}\, H_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T)$.

This is immediate since $h_n(\theta) = K_\eta(\theta)\nabla v_n(\theta)$ with $K_\eta > 0$.

- It remains to prove

$\sup_{\theta\in\mathbb{R}^q} \tilde{\mathbb{E}}\left[|H_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T)|^2\right] < \infty$,

By Cauchy-Schwartz inequality we obtain for $\lambda_1 = 4\lambda \vee 2(\lambda+1)$,

$$\tilde{\mathbb{E}}\left[|H_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T)|^2\right] \leq e^{-2\eta|\theta|^2 T}\left\||2\theta T - W_T|^2\right\|_2 \times\left(\left\|\psi(X_T^{n,(-\theta)})^2\right\|_2 + \left\|(\nabla\psi(X_T^{n,(-\theta)})\cdot U_T^{n,(-\theta)})^2\right\|_2\right).$$

$$\tilde{\mathbb{E}}\left[|H_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T)|^2\right] \leq Ce^{-2\eta|\theta|^2 T}(1+|\theta|^2)\left(1 + \left\||X_T^{n,(-\theta)}|^{\lambda_1}\right\|_2 + \left\||U_T^{n,(-\theta)}|^4\right\|_2\right).$$

- Using properties $(\mathcal{P})$ and $(\tilde{\mathcal{P}})$, we get

$$\tilde{\mathbb{E}}\left[|H_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T)|^2\right] \leq Ce^{-2\eta|\theta|^2 T}(1+|\theta|^2) \times \left(1 + \left\| |X_T^{n,(-\theta)} - X_T^n|^{\lambda_1} \right\|_2 + \left\| |U_T^{n,(-\theta)} - U_T^n|^4 \right\|_2 \right)$$

- Using Gronwall inequality, we obtain that

$$\tilde{\mathbb{E}}\left|X_T^{n,(-\theta)} - X_T^n\right|^{2\lambda_1} \leq C|\theta|^{2\lambda_1} \sum_{j=1}^q \tilde{\mathbb{E}}\left|\int_0^T |\sigma_j(X_s^{n,(-\theta)})|ds\right|^{2\lambda_1},$$

$$\tilde{\mathbb{E}}\left|U_T^{n,(-\theta)} - U_T^n\right|^8 \leq C|\theta|^8 \tilde{\mathbb{E}}\left|\int_0^T |U_s^{n,(-\theta)}|ds\right|^8.$$

- As $(B^{(-\theta)}, X^n, U^n, \tilde{\mathbb{P}}_{(-\theta)}) \overset{law}{=} (W, X^{n,(-\theta)}, U^{n,(-\theta)}, \tilde{\mathbb{P}})$,

$$\tilde{\mathbb{E}}\left|\int_0^T |\sigma_j(X_s^{n,(-\theta)})|ds\right|^{2\lambda_1} = \tilde{\mathbb{E}}\left(\left|\int_0^T |\sigma_j(X_s^n)|ds\right|^{2\lambda_1} e^{-\theta\cdot W_T - \frac{1}{2}|\theta|^2 T}\right)$$

In the same way

$$\tilde{\mathbb{E}}\left|\int_0^T |U_s^{n,(-\theta)}|ds\right|^8 = \tilde{\mathbb{E}}\left(\left|\int_0^T |U_s^n|ds\right|^8 e^{-\theta\cdot W_T - \frac{1}{2}|\theta|^2 T}\right).$$

Now using Holder's inequality, with $\frac{1}{r}+\frac{1}{r'}=1$, the linear growth of $(\sigma_j)_{1\leq j\leq q}$, properties $(\mathcal{P})$ and $(\tilde{\mathcal{P}})$ we obtain

$$\begin{aligned}
&\tilde{\mathbb{E}}\left[|H_\eta(\theta, X_T^{n,(-\theta)}, U_T^{n,(-\theta)}, W_T)|^2\right] \\
&\qquad \leq Ce^{-2\eta|\theta|^2 T}(1+|\theta|^2)\left(1+(|\theta|^{\lambda_1}+|\theta|^4)e^{\frac{r-1}{4}|\theta|^2 T}\right).
\end{aligned}$$

We complete the proof by choosing $r\in]1,1+8\eta[$. $\square$

Theorem

Suppose b and σ are $\mathcal{C}^2$ with bounded first and second derivatives and s.t. all the derivatives of order 2 are Lipschitz continuous. Let $\nabla\psi$ in $\mathcal{C}^1$ s.t. for all $n \in \mathbb{N}$, $\mathbb{P}(\psi(X_T^n) \neq 0) > 0$.

- *Assume also that $\partial_\alpha \psi$, for $|\alpha| \leq 2$, are with polynomial growth*

Then, for all $\forall i \in \mathbb{N}$, $p \geq 1$, there exists $C_i > 0$ such that

$$\forall n \in \mathbb{N}^*, \quad \tilde{\mathbb{E}}|\theta_{i+1}^n - \theta_{i+1}|^{2p} \leq \frac{C_i}{n^p}.$$

Moreover,

$$\forall i \in \mathbb{N}, \quad \theta_i^n \underset{n\to\infty}{\longrightarrow} \theta_i, \quad a.s.$$

where the sequence $(\theta_i)_{i\geq 0}$ is introduced in the above corollary.

Proof

We give the proof in the case of dimension one and proceed by induction on $i \in \mathbb{N}$ For all $p \geq 1$ relation (1) yields

$$\begin{aligned}\tilde{\mathbb{E}}(\theta_{i+1}^n - \theta_{i+1})^{2p} \leq C\tilde{\mathbb{E}}(\theta_i^n - \theta_i)^{2p} \\ + C\gamma_{i+1}^{2p}\tilde{\mathbb{E}}\Big(H_\eta(\theta_i^n, X_{T,i+1}^{n,(-\theta_i^n)}, U_{T,i+1}^{n,(-\theta_i^n)}, W_{T,i+1}) \\ -H_\eta(\theta_i, X_{T,i+1}^{(-\theta_i)}, U_{T,i+1}^{(-\theta_i)}, W_{T,i+1})\Big)^{2p}.\end{aligned}$$

Using the induction assumption we only need to control the second term bounded by $C \times (\tilde{\mathbb{E}}H_1^{2p} + \tilde{\mathbb{E}}H_2^{2p})$

$$\begin{aligned}H_1 := e^{-\eta|\theta_i^n|^2 T}(2\theta_i^n T - W_{T,i+1}) \\ \times \Big[\psi(X_{T,i+1}^{n,(-\theta_i^n)})^2 + \Big(\psi'(X_{T,i+1}^{n,(-\theta_i^n)})U_{T,i+1}^{n,(-\theta_i^n)}\Big)^2 \\ -\psi(X_{T,i+1}^{(-\theta_i^n)})^2 - \Big(\psi'(X_{T,i+1}^{(-\theta_i^n)})U_{T,i+1}^{(-\theta_i^n)}\Big)^2\Big]\end{aligned}$$

$$H_2 := e^{-\eta|\theta_i^n|^2 T}(2\theta_i^n T - W_{T,i+1})\left[\psi(X_{T,i+1}^{(-\theta_i^n)})^2 + \left(\psi'(X_{T,i+1}^{(-\theta_i^n)})U_{T,i+1}^{(-\theta_i^n)}\right)^2\right]$$

$$-e^{-\eta|\theta_i|^2 T}(2\theta_i T - W_{T,i+1})\left[\psi(X_{T,i+1}^{(-\theta_i)})^2 + \left(\psi'(X_{T,i+1}^{(-\theta_i)})U_{T,i+1}^{(-\theta_i)}\right)^2\right].$$

- We give the proof for H_2. Here, we need first to introduce, for all $u \in \mathbb{R}$, the couple of u-sensitivity processes $(Y_t^{(-u)}, Z_t^{(-u)})_{t\in[0,T]}$

given by $Y_t^{(-u)} := \frac{\partial X_t^{(-u)}}{\partial u}$ and $Z_t^{(-u)} := \dfrac{\partial U_t^{(-u)}}{\partial u}, \quad t \in [0, T]$

Therefore, we write $\tilde{\mathbb{E}} H_2^{2p} = \tilde{\mathbb{E}} B(\theta_i^n, \theta_i)$

$$B(\theta,\theta') = \tilde{\mathbb{E}}\left[\int_{(\theta',\theta)} \frac{\partial}{\partial u}\left\{e^{-\eta|u|^2 T}(2uT - W_T)\right.\right.$$
$$\left.\left.\times\left[\psi(X_T^{(-u)})^2 + \left(\psi'(X_T^{(-u)})U_T^{(-u)}\right)^2\right]\right\} du\right]^{2p}.$$

- Now, $B(\theta,\theta') \leq C\sum_{i=1}^{4} B_i(\theta,\theta')$. These four terms are of the same type, so we only treat one of them let's say B_3

$$B_3(\theta,\theta') \leq C|\theta-\theta'|^{2p-1}$$
$$\times \int_{(\theta',\theta)} \tilde{\mathbb{E}}\left[(2uT - W_T)^{2p} e^{-2p\eta|u|^2 T} (U_T^{(-u)})^{2p} \psi'(X_T^{(-u)})^{4p} (Z_T^{(-u)})^{2p}\right] du$$

- Note that the same probability change leading to cancel the u-term in the drift part of $X^{(-u)}$ operates in the same way for the other processes $U^{(-u)}$, $Y^{(-u)}$ and $Z^{(-u)}$. So for all $u \in \mathbb{R}$, we get $(B^{(-u)}, X, U, Y, Z, \tilde{\mathbb{P}}_{(-u)}) \overset{law}{=} (W, X^{(-u)}, U^{(-u)}, Y^{(-u)}, Z^{(-u)}, \tilde{\mathbb{P}})$

$$B_3(\theta,\theta') \leq C|\theta-\theta'|^{2p-1} \int_{(\theta',\theta)} \tilde{\mathbb{E}}\Big[(uT - W_T)^{2p} e^{-2p\eta|u|^2 T}$$
$$\times (U_T)^{2p} \psi'(X_T)^{4p} (Z_T)^{2p} e^{-uW_T - \frac{1}{2}|u|^2 T}\Big] du.$$

$$B_3(\theta,\theta') \leq C|\theta-\theta'|^{2p},$$

- This is immediate, since X and U satisfy properties $(\mathcal{P})$ and $(\tilde{\mathcal{P}})$ and Z is a diffusion process with enough smooth coefficients satisfying likewise the same type of properties.
- So that, we obtain for all $p>1$

$$B(\theta,\theta') \leq C|\theta-\theta'|^{2p}.$$

Now, since $\tilde{\mathbb{E}}H_2^{2p} = \tilde{\mathbb{E}}B(\theta_i^n,\theta_i)$, it follows for all $p>1$

$$\tilde{\mathbb{E}}H_2^{2p} \leq C\mathbb{E}|\theta_i^n-\theta_i|^{2p} \leq \frac{C_i}{n^p}.$$

Corollary

Under above assumptions if $\mathbb{P}(\psi(X_T) \neq 0) > 0$, *then the unconstrained algorithm satisfies*

$$\lim_{i,n\to\infty} \theta_i^n = \lim_{i\to\infty} \big(\lim_{n\to\infty} \theta_i^n \big) = \lim_{n\to\infty} \big(\lim_{i\to\infty} \theta_i^n \big) = \theta^*, \quad \tilde{\mathbb{P}}\text{-}a.s.,$$

where $\theta^* = \operatorname*{argmin}_{\theta\in\mathbb{R}^q} v(\theta)$

$$v(\theta) := \tilde{\mathbb{E}}\left(\left[\psi(X_T^\theta)^2 + (\nabla\psi(X_T^\theta)\cdot U_T^\theta)^2 \right] e^{-2\theta.W_T - |\theta|^2 T} \right)$$

Outline

Adaptative Statistical Romberg method

• The adaptative importance sampling algorithm for the statistical Romberg method approximates our initial quantity of interest $\mathbb{E}\psi(X_T) = \mathbb{E}\left[\psi(X_T^\theta)e^{-\theta\cdot W_T - \frac{1}{2}|\theta|^2 T}\right]$ by

$$\frac{1}{N_1}\sum_{i=1}^{N_1} g(\hat{\theta}_i^m, \hat{X}_{T,i+1}^{m,\hat{\theta}_i^m}, \hat{W}_{T,i+1})$$
$$+ \frac{1}{N_2}\sum_{i=1}^{N_2}\left(g(\theta_i^n, X_{T,i+1}^{n,\theta_i^n}, W_{T,i+1}) - g(\theta_i^n, X_{T,i+1}^{m,\theta_i^n}, W_{T,i+1})\right),$$

where for all $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^q$, $g(\theta, x, y) = \psi(x)e^{-\theta\cdot y - \frac{1}{2}|\theta|^2 T}$.

• Here the paths generated by W and $\hat{W}$ are of course independent.

Lindeberg Feller CLT

Theorem

Suppose that $(\Omega, \mathbb{F}, \mathbb{P})$ *is a probability space and that for each* n, *we have a filtration* $\mathbb{F}_n = (\mathcal{F}_k^n)_{k\geq 0}$, *a sequence* $k_n \to \infty$ *as* $n \to \infty$ *and a real square integrable vector martingale* $M^n = (M_k^n)_{k\geq 0}$ *which is adapted to* $\mathbb{F}_n$ *such that*

- *There exists a deterministic symmetric positive semi-definite matrix* Γ, *such that* $\langle M\rangle_{k_n}^n = \sum_{k=1}^{k_n} \mathbb{E}\left[|M_k^n - M_{k-1}^n|^2 | \mathcal{F}_{k-1}^n\right] \xrightarrow[n\to\infty]{\mathbb{P}} \Gamma.$
- *There exists a real number* $a > 1$, *such that* $\sum_{k=1}^{k_n} \mathbb{E}\left[|M_k^n - M_{k-1}^n|^{2a} | \mathcal{F}_{k-1}^n\right] \xrightarrow[n\to\infty]{\mathbb{P}} 0.$

Then

$$M_{k_n}^n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma) \quad \text{as} \;\; n \to \infty.$$

Toeplitz Lemma

Lemma

Let $(a_i)_{1\leq i\leq k_n}$ a sequence of real positive numbers, where $k_n \uparrow \infty$ as n tends to infinity, and $(x_i^n)_{i\geq 1, n\geq 1}$ a double indexed sequence such that

(i) $\lim\limits_{n\to\infty} \sum_{1\leq i\leq k_n} a_i = \infty$

(ii) $\lim\limits_{i,n\to\infty} x_i^n = \lim\limits_{i\to\infty}\big(\lim\limits_{n\to\infty} x_i^n\big) = \lim\limits_{n\to\infty}\big(\lim\limits_{i\to\infty} x_i^n\big) = x < \infty$

Then

$$\lim_{n\to+\infty} \frac{\sum_{i=1}^{k_n} a_i x_i^n}{\sum_{i=1}^{k_n} a_i} = x.$$

The adaptative Monte Carlo method

Theorem

- *Let* $(\theta_i^n)_{i\geq 0}$, $n \in \mathbb{N}$ *and* $(\theta_i)_{i\geq 0}$ *satisfying*

$(\mathcal{H}_\theta)$ $\lim_{i,n\to\infty} \theta_i^n = \lim_{i\to\infty}\big(\lim_{n\to\infty} \theta_i^n\big) = \lim_{n\to\infty}\big(\lim_{i\to\infty} \theta_i^n\big) = \theta^*, \quad \tilde{\mathbb{P}}$*-a.s.*,

- *Assume that* b *and* σ *satisfy* $(\mathcal{H}_{b,\sigma})$ *and the function* ψ *is a real valued function satisfying assumption* $(\mathcal{H}_{\varepsilon_n})$, *with* $\alpha \in [1/2, 1]$ *and* $C_\psi \in \mathbb{R}$, *s.t.* $|\psi(x) - \psi(y)| \leq C(1 + |x|^p + |y|^p)|x - y|$, *then the following convergence holds*

$$n^\alpha \left(\frac{1}{n^{2\alpha}} \sum_{i=1}^{n^{2\alpha}} g(\theta_i^n, X_{T,i+1}^{n,\theta_i^n}, W_{T,i+1}) - \mathbb{E}\psi(X_T) \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(C_\psi, \sigma^2\right).$$

$$\sigma^2 := \mathbb{E}\left(\psi(X_T)^2 e^{-\theta^*.W_T - \frac{1}{2}|\theta^*|^2 T} \right) - [\mathbb{E}\psi(X_T)]^2$$

The adaptative statistical Romberg method

Theorem

- *Let* $(\theta_i^n)_{i\geq 0}$, $n \in \mathbb{N}$ *and* $(\theta_i)_{i\geq 0}$ *satisfying* $(\mathcal{H}_\theta)$.
- *Assume that* b *and* σ *are* $\mathcal{C}^1$ *satisfying* $(\mathcal{H}_{b,\sigma})$ *and* ψ *is* $\mathcal{C}^1$, *satisfying* $(\mathcal{H}_{\varepsilon_n})$, *with constants* $\alpha \in (1/2, 1]$ *and* $C_\psi \in \mathbb{R}$, *s.t.*

$$|\psi(x) - \psi(y)| \leq C(1 + |x|^p + |y|^p)|x - y|, \quad \text{for some } C, p > 0.$$

If we choose $N_1 = n^{2\alpha}$, $N_2 = n^{2\alpha-\beta}$ *then*

$$n^\alpha \left(V_n - \mathbb{E}\psi(X_T)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(C_\psi, \sigma^2 + \tilde{\sigma}^2\right) \qquad \text{as } n \to \infty,$$

where $\sigma^2 = \mathbb{E}\left[\psi(X_T)^2 e^{-\theta^*\cdot W_T - \frac{1}{2}|\theta^*|^2 T}\right] - \left[\mathbb{E}\psi(X_T)\right]^2$ *and*

$$\tilde{\sigma}^2 := \tilde{\mathbb{E}}\left[\left[\nabla\psi(X_T)\cdot U_T\right]^2 e^{-\theta^*\cdot W_T - \frac{1}{2}|\theta^*|^2 T}\right].$$

Proof

- We prove only the convergence of the secod empirical mean in the Statistical Romberg method. To do so, we introduce the martingale arrays $(M_k^n)_{k\geq 1}$. For $\beta = 1/2$

$$M_k^n := \frac{1}{n^{\alpha-\beta}} \sum_{i=1}^{k} \Big(g(\theta_i^n, X_{T,i+1}^{n,\theta_i^n}, W_{T,i+1}) - g(\theta_i^n, X_{T,i+1}^{n^\beta,\theta_i^n}, W_{T,i+1}) - \mathbb{E}[\psi(X_T^n) - \psi(X_T^{n^\beta})] \Big),$$

- The quadratic variation of M evaluated at $n^{2\alpha-\beta}$ is given by

$$\langle M\rangle^n_{n^{2\alpha-\beta}} = \frac{1}{n^{2\alpha-\beta}} \sum_{i=1}^{n^{2\alpha-\beta}} n^\beta \xi_n(\theta_i^n) - \left(n^{\frac{\beta}{2}} [\mathbb{E}\psi(X_T^n) - \mathbb{E}\psi(X_T^{n^\beta})] \right)^2,$$

where $\forall \theta \in \mathbb{R}^q$, $\xi_n(\theta) := \mathbb{E}\left([\psi(X_T^n) - \psi(X_T^{n^\beta})]^2 e^{-\theta\cdot W_T + \frac{1}{2}|\theta|^2 T} \right).$

- We focus now on the asymptotic behavior of $n^{\beta}\xi_n(\theta)$. Applying Taylor expansion theorem twice we get for all $\theta \in \mathbb{R}^q$

$$n^{\frac{\beta}{2}}[\psi(X_T^n) - \psi(X_T^{n^\beta})]e^{-\frac{1}{2}\theta\cdot W_T + \frac{1}{4}|\theta|^2 T} = n^{\frac{\beta}{2}}\nabla\psi(X_T)\cdot[X_T^n - X_T^{n^\beta}]e^{-\frac{1}{2}\theta\cdot W_T + \frac{1}{4}|\theta|^2 T} + R_n,$$

$$R_n := n^{\frac{\beta}{2}}(X_T^n - X_T)\varepsilon(X_T, X_T^n - X_T) - n^{\frac{\beta}{2}}(X_T^{n^\beta} - X_T)\varepsilon(X_T, X_T^{n^\beta} - X_T)$$

with $\varepsilon(X_T, X_T^n - X_T) \xrightarrow{\mathbb{P}\text{-}a.s.} 0$ and $\varepsilon(X_T, X_T^{n^\beta} - X_T) \xrightarrow{\mathbb{P}\text{-}a.s.} 0$

- Further, as b and σ are $\mathcal{C}^1$ functions then we have the tightness of $n^{\frac{\beta}{2}}(X_T^n - X_T)$ and $n^{\frac{\beta}{2}}(X_T^{n^\beta} - X_T)$ and we deduce that $R_n \to 0$.

- It follows

$$n^{\frac{\beta}{2}}[\psi(X_T^n)-\psi(X_T^{n^\beta})]e^{-\frac{1}{2}\theta\cdot W_T+\frac{1}{4}|\theta|^2 T} \overset{stably}{\Longrightarrow} \nabla\psi(X_T)\cdot U_T e^{-\frac{1}{2}\theta\cdot W_T+\frac{1}{4}|\theta|^2 T}.$$

- Otherwise, $\forall\theta \in \mathbb{R}^q$ and $a' > 1$ we have thanks to the assumption on ψ together with property $(\mathcal{P})$, we obtain

$$\sup_n \mathbb{E}\left|n^{\frac{\beta}{2}}[\psi(X_T^n) - \psi(X_T^{n^\beta})]e^{-\frac{1}{2}\theta\cdot W_T+\frac{1}{4}|\theta|^2 T}\right|^{2a'} < \infty$$

So, $\lim_{n\to\infty} n^\beta \xi_n(\theta) = \tilde{\mathbb{E}}\left([\nabla\psi(X_T)\cdot U_T]^2\, e^{-\theta\cdot W_T+\frac{1}{2}|\theta|^2 T}\right) := \xi(\theta)$

- Using property $(\mathcal{P})$ with assumption on ψ, we check the equicontinuity of the family functions $(n^\beta \xi_n)_{n\geq 1}$.
- So under assumption $(\mathcal{H}_\theta)$, we get

$$\lim_{i,n\to\infty} n^\beta \xi_n(\theta_i^n) = \xi(\theta^*) \quad \tilde{\mathbb{P}}\text{-}a.s.$$

Then, Toeplitz Lemma yields $\lim_{n\to\infty}\langle M\rangle^n_{n^{2\alpha-\beta}} = \xi(\theta^*)$, $\tilde{\mathbb{P}}$-$a.s.$

Outline

The popular stochastic volatility model in finance is the Heston model solution to

$$\begin{cases} dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^1 \\ dV_t = \kappa(\bar{v} - V_t)dt + \sigma\sqrt{V_t}\rho dW_t^1 + \sigma\sqrt{V_t}\sqrt{1-\rho^2}dW_t^2, \end{cases}$$

where W^1 and W^2 are two independent Brownian motions. Parameters κ, σ, $\bar{v}$, $r > 0$ and $|\rho| \leq 1$.

• Our aim is to use the importance sampling method in order to reduce the variance when computing the price is

$$e^{-rT}\mathbb{E}\psi(S_T) = e^{-rT}\mathbb{E}\left[g(\theta, S_T^\theta)\right] = e^{-rT}\mathbb{E}\left[\psi(S_T^\theta) \quad e^{-\theta.W_T - \frac{1}{2}|\theta|^2 T}\right],$$

To approximate S_T^θ, we consider the step T/n and we discretize the stochastic process using the Euler scheme. For $i \in [0, n-1]$ and $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$,

$$\begin{cases} S^{n,\theta}_{t_{i+1}} = S^{n,\theta}_{t_i}\left(1+(r+\theta_1\sqrt{V^{n,\theta}_{t_i}})\dfrac{T}{n}+\sqrt{V^{n,\theta}_{t_i}\dfrac{T}{n}}Z_{1,i+1}\right), \\ V^{n,\theta}_{t_{i+1}} = \Bigg|V^{n,\theta}_{t_i}+\Bigg(\kappa(\bar{v}-V^{n,\theta}_{t_i}) \\ \qquad +\sigma\sqrt{V^{n,\theta}_{t_i}}(\rho\theta_1+\sqrt{1-\rho^2}\theta_2)\Bigg)\dfrac{T}{n}+\sigma\sqrt{V^{n,\theta}_{t_i}\dfrac{T}{n}}Z_{2,i+1}\Bigg|, \end{cases}$$

Hence, the price is firstly approximated by

$$e^{-rT}\mathbb{E}\left[g(\theta,S^{n,\theta}_T)\right]=e^{-rT}\mathbb{E}\left[\psi(S^{n,\theta}_T)\;\; e^{-\theta.W_T-\frac{1}{2}|\theta|^2T}\right],\quad \theta\in\mathbb{R}^2.$$

- The optimal θ for a Monte Carlo method

$$\theta^*_n=\underset{\theta\in\mathbb{R}^2}{\arg\min}\,\mathbb{E}\left[\psi^2(S^{n,\theta}_T)\;\; e^{-2\theta.W_T-|\theta|^2T}\right].$$

- The optimal θ for the second one is

$$\tilde{\theta}_n^* = \arg\min_{\theta \in \mathbb{R}^2} \mathbb{E}\left[\left(\psi^2(S_T^{n,\theta}) + (\nabla\psi(S_T^{n,\theta}).U_T^{n,\theta})^2 \right) e^{-2\theta.W_T - |\theta|^2 T} \right],$$

- Here, we have also the choice of the algorithm approximating both θ_n^* and $\tilde{\theta}_n^*$ by the constrained algorithm or by the unconstrained algorithm
- We fix $S_0 = 100$, $V_0 = 0.01$, $K = 100$, the free interest rate $r = \log(1.1)$, $\sigma = 0.2$, $k = 2$, $\bar{v} = 0.01$, $\rho = 0.5$ and maturity time $T = 1$.

	Constrained algorithm	**Unconstrained algorithm**
θ_n^*	$(0.7906, 0.0516)$	$(0.7904, 0.0532)$
$\tilde{\theta}_n^*$	$(0.7884, 0.0587)$	$(0.7898, 0.0576)$

Table: *Estimation of θ_n^* and $\tilde{\theta}_n^*$*

1. First, we choose $\gamma_i = \gamma_0/i^\alpha$, for $\alpha \in (\frac{1}{2}, 1)$ and $\gamma_0 > 0$.
2. Then, we compute $\bar{\theta}^n_{i+1} := \frac{1}{i+1}\sum_{k=0}^{i} \tilde{\theta}^n_k$.

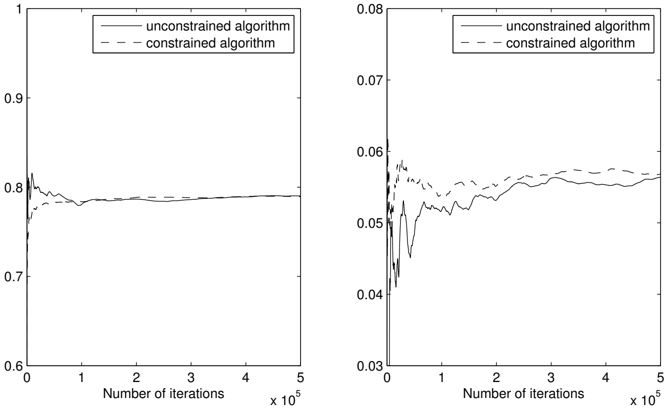

Figure: *Values obtained with* $n = 100$, $\gamma_0 = 0.01$ *and* $\alpha = 0.75$.

Our aim now, is to compare

- MC+IS method: European call option price approximation with $N = n^2$

$$\frac{e^{-rT}}{N}\sum_{i=1}^{N} g(\theta_M^n, S_{T,i+1}^{n,\theta_M^n})$$

- SR+IS method: European call option price approximation method with $N_1 = n^2$ and $N_2 = n^{\frac{3}{2}}$

$$\frac{e^{-rT}}{N_1}\sum_{i=1}^{N_1} g(\tilde{\theta}_M^n, \hat{S}_{T,i+1}^{\sqrt{n},\tilde{\theta}_M^n}) + \frac{e^{-rT}}{N_2}\sum_{i=1}^{N_2}\left(g(\tilde{\theta}_M^n, S_{T,i+1}^{n,\tilde{\theta}_M^n}) - g(\tilde{\theta}_M^n, S_{T,i+1}^{\sqrt{n},\tilde{\theta}_M^n})\right).$$

Method	n	**Price**	**CI length**	**time**
MC+IS	400	9.641444	0.060094	10.38
	900	9.661192	0.029409	91.5
	1600	9.656892	0.016538	512.29
SR+IS	600	9.659409	0.057454	3.36
	1600	9.660062	0.019933	26.79
	3600	9.65673	0.008584	194.6

Table: *Call Price for the Heston model*

Method	n	Price	CI length	time
MC+IS	400	0.863968	0.00721	9.39
	900	0.863291	0.003151	91.58
	1600	0.863766	0.001774	515.31
SR+IS	600	0.867441	$0,007249$	3.27
	1600	0.864213	0.002541	27.02
	3600	0.862589	0.001095	202.2

Table: *Delta call price for the Heston model*

The first method (MC+IS) is already implemented and available in the free online version of Premia platform (`https://www.rocq.inria.fr/mathfi/Premia/index.html`) and our method (SR+IS) is now added in the latest premium version.

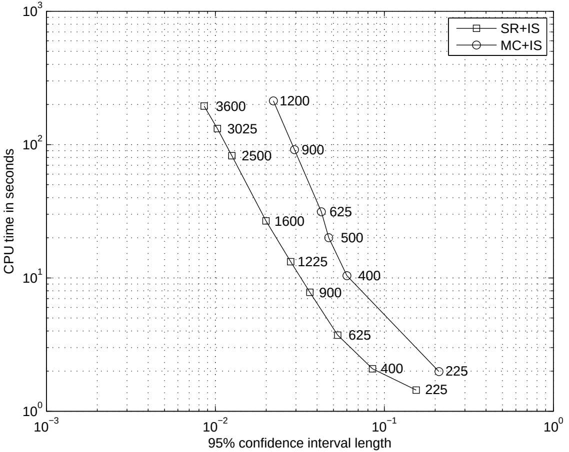

Figure: *CPU time versus the 95%-confidence interval length*

Thank you!

Multilevel Monte Carlo and the Euler scheme

- We use $L+1$ Euler schemes with time steps $\dfrac{T}{m^\ell}$ for $\ell \in \{0, 1, \cdots, L\}$ such that $m^L = n$.
- We can write

$$\mathbb{E}(f(X_T^n)) = \mathbb{E}\left(f(X_T^{m^0})\right) + \sum_{\ell=1}^{L} \mathbb{E}\left(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})\right).$$

- The Multilevel method consits on estimating independently each of the expectations above.

$$Q_n = \frac{1}{N_0}\sum_{k=1}^{N_0} f(X_{T,k}^{m^0}) + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{k=1}^{N_\ell}\left(f(X_{T,k}^{m^\ell}) - f(X_{T,k}^{m^{\ell-1}})\right).$$

we have

$$Var(Q_n) = O\left(\sum_{\ell=1}^{L} N_\ell^{-1} m^{-\ell}\right).$$