# Multi-level stochastic approximation algorithms

Noufel Frikha

Université Paris Diderot, LPMA

28th October, 2013.

Stochastic processes and their statistics in Finance in Okinawa

# Outline of the presentation

# Introduction

▷ Multi-level Monte Carlo paradigm was originally introduced for the computation of :

$$\mathbb{E}_x[f(X_T)]$$

where $f : \mathbb{R}^q \to \mathbb{R}$ and $(X_t)_{t\in[0,T]}$ is a $q$-dimensional process satisfying

$$\forall t \in [0,T],\ X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s. \qquad (SDE_{b,\sigma})$$

When no closed formula is available, one proceeds in two steps :
▷ Step 1 : *Discretization* scheme of $(SDE_{b,\sigma})$ by

$$X_t^n = x + \int_0^t b(X^n_{\phi_n(s)})ds + \int_0^t \sigma(X^n_{\phi_n(s)})dW_s, \quad \phi_n(s) = \sup\{t_i :\ t_i \leq s\}.$$

with time step $\Delta = T/n$ and regular points $t_i = i\Delta$, $i = 0, \cdots, n$.

This step induces a *weak error*

$$\mathcal{E}_D(f,n,T,b,\sigma) = \mathbb{E}_x[f(X_T)] - \mathbb{E}_x[f(X_T^n)] \approx \Delta$$

see Talay & Tubaro (90), Bally & Talay (96), ...

▷ Step 2 : *Estimation* of $\mathbb{E}_x[f(X_T^n)]$ by $M^{-1} \times \sum_{j=1}^{M} f((X_T^n)^j)$ induces a *statistical error* :

$$\mathcal{E}_S(M,f,n,T,b,\sigma) = \mathbb{E}_x[f(X_T^n)] - \frac{1}{M}\sum_{j=1}^{M} f((X_T^n)^j)$$

The global error associated to the computation of $\mathbb{E}_x[f(X_T)]$ writes :

$$\begin{aligned}\mathcal{E}_{Glob}(M,n) &:= \mathbb{E}_x[f(X_T)] - \frac{1}{M}\sum_{j=1}^{M} f((X_T^n)^j)\\ &= \mathcal{E}_D(f,n,T,b,\sigma) + \mathcal{E}_S(M,f,n,T,b,\sigma).\end{aligned}$$

# Complexity analysis

Optimal complexity : How to balance $M$ w.r.t $n$ to achieve a global error of order $\epsilon$ ?

▷ Duffie & Glynn (95) : If the weak discretization error of order $n^{-\alpha}$, i.e.

$$\exists \alpha \in (0,1],\ n^{\alpha}(\mathbb{E}_x[f(X_T)] - \mathbb{E}_x[f(X_T^n)]) \to C(\alpha, f, b, \sigma, T),\quad n \to +\infty$$

then,

$$n^{\alpha}\left(\frac{1}{n^{2\alpha}}\sum_{j=1}^{n^{2\alpha}} f((X_T^n)^j) - \mathbb{E}_x[f(X_T)]\right) \Longrightarrow \mathcal{N}\left(C(\alpha, f, b, \sigma, T), Var(f(X_T))\right).$$

▷ It is optimal to set $M = n^{2\alpha}$ to achieve an error of order $\epsilon = n^{-\alpha}$ :

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

# Statistical Romberg Monte Carlo scheme

▷ To reduce the complexity, Kebaier (05) proposed a two-level Monte Carlo scheme to approximate $\mathbb{E}_x[f(X_T)]$ by :

$$\hat{M}(\gamma_1, \gamma_2, \beta) := \frac{1}{n^{\gamma_1}} \sum_{j=1}^{n^{\gamma_1}} f((\hat{X}_T^{n^\beta})^j) + \frac{1}{n^{\gamma_2} T} \sum_{j=1}^{n^{\gamma_2} T} f((X_T^n)^j) - f((X_T^{n^\beta})^j)$$

- $((\hat{X}_T^{n^\beta})^j)_{j \in [\![1, n^{\gamma_1}]\!]}$ and $(((X_T^n)^j, X_T^{n^\beta})^j)_{j \in [\![1, n^{\gamma_2} T]\!]}$ are independent.
- $(X_T^n, X_T^{n^\beta})$ are computed with the same path but with different time steps.

▷ Main result : If $n^\alpha(\mathbb{E}_x[f(X_T)] - \mathbb{E}_x[f(X_T^n)]) \to C(\alpha, f, b, \sigma, T)$, then

$$n^\alpha \left( \hat{M}(2\alpha, 2\alpha - \beta, \beta) - \mathbb{E}_x[f(X_T)] \right) \Longrightarrow \mathcal{N}\left(C(\alpha, f, b, \sigma, T), Var(f(X_T)) + Var(\nabla f(X_T) U_T\right.$$

▷ Optimal Complexity to achieve an error of order $n^{-\alpha}$ :

$$C_{SR-MC} = C \times (n^\beta n^{2\alpha} + (n^\beta + n) n^{2\alpha - \beta}) \approx C \times n^{2\alpha + \frac{1}{2}}, \ \beta = \frac{1}{2}.$$

# Multi-level Monte Carlo scheme

▷ Generalizing Kebaier's approach, Giles (08) proposed a multi-level Monte Carlo scheme to approximate $\mathbb{E}_x[f(X_T)]$ by :

$$\hat{M}(n) := \frac{1}{N_0}\sum_{j=1}^{N_0} f((X_T^1)^j) + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{j=1}^{N_\ell} f((X_T^{m^\ell})^j) - f((X_T^{m^{\ell-1}})^j)$$

- $L+1$ independent empirical mean sequences.
- Euler schemes with geometric sequence of time steps, $m^L = n$.

$$Var(\hat{M}(n)) = \frac{1}{N_0} Var(f(X_T^1)) + \sum_{\ell=1}^{L}\frac{1}{N_\ell} Var(f(X_T^{m^\ell}) - f(X_T^{m^{\ell-1}})) \leq C\sum_{\ell=0}^{L} N_\ell^{-1} m^{-\ell}$$

▷ Optimal Complexity to achieve an error of order $n^{-\alpha}$ :

$$C_{\text{ML-MC}} = C \times n^{2\alpha}(\log(n))^2, \text{ for } N_\ell := 2c_2 n^{2\alpha}(L+1)T/m^\ell.$$

▷ See also the recent work of Kebaier & Ben Alaya (12).

# Stochastic approximation algorithm

▷ Aim : Extend the scope of the ML-MC method to stochastic optimization by means of stochastic approximation (SA).
▷ Introduced by H.Robbins & S.Monro (1951). It is a recursive simulation-based algorithm to estimate $\theta^*$ solution of

$$h(\theta) := \mathbb{E}[H(\theta, U)] = 0, \;\; H : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d, \;\; U \sim \mu$$

▷ Behind and implicitly assumed : Computation of $h$ is costly compared to the computation of $H$ and to the simulation of $U$.
▷ Devise the following scheme $p \in \mathbb{N}, \; \theta_0 \in \mathbb{R}^d$

$$\theta_{p+1} = \theta_p - \gamma_{p+1} H(\theta_p, U_{p+1}) = \theta_p - \gamma_{p+1} \underbrace{(h(\theta_p) + \Delta M_{p+1})}_{\textit{Corrupted observations of } h(\theta_p)}$$

with $(U_p)_{p\geq 1}$ i.i.d. $\mathbb{R}^q$-valued r.v. with law $\mu$ and

$$\sum_{p\geq 1} \gamma_p = +\infty, \quad \sum_{p\geq 1} \gamma_p^2 < +\infty,$$

to take advantage of an averaging effect along the scheme.

# Asymptotic properties of $(\theta_p)_{p\geq 1}$

## $a.s.$ convergence and convergence rate

▷ $a.s.$ convergence : Robbins-Monro Theorem

- *mean-reverting assumption*

$$\forall \theta \in \mathbb{R}^d, \theta \neq \theta^*, \ \langle \theta - \theta^*, h(\theta) \rangle > 0,$$

- *domination assumption*

$$\forall \theta \in \mathbb{R}^d, \ |h(\theta)|^2 \leq \mathbb{E}|H(\theta, U)|^2 \leq C(1 + |\theta - \theta^*|^2).$$

Then, one has :

$$\theta_p \xrightarrow{a.s.} \theta^*, \ p \to +\infty.$$

▷ Weak convergence rate : under mild assumptions, in “standard cases”, one has

$$\sqrt{\gamma_p^{-1}}(\theta_p - \theta^*) \Longrightarrow \mathcal{N}(0, \Sigma^*), \ p \to +\infty.$$

# Some applications in computational finance

▷ In many applications, notably in computational finance, we are interested in estimating the zero $\theta^*$ of $h(\theta) = \mathbb{E}_x[H(\theta, X_T)]$.

▷ Some examples among others :

- Implied volatility :

$$\sigma \in \mathbb{R}_+ \ s.t. \ \mathbb{E}_x[(X_T(\sigma) - K)_+] = P_{market}.$$

- Implied correlation between $X_T^1$ and $X_T^2$ :

$$\rho \in (-1, 1) \ s.t. \ \mathbb{E}_x[(\max(X_T^1, X_T^2(\rho)) - K)_+] = P_{market}.$$

- VaR and CVaR of a financial portfolio :

$$(\xi, C) \ s.t. \ \mathbb{P}_x(F(X_T) \leq \xi) = \alpha, \ \ C = VaR_\alpha + \frac{1}{1-\alpha}\mathbb{E}_x[(F(X_T) - VaR_\alpha)_+]$$

- Portfolio optimization : $\sup_{\theta \in \mathbb{R}^q} \mathbb{E}_x[U(F(X_T) - \theta.(X_T - x))]$.

- ▷ The function $h$ is generally not known and $X_T$ cannot be simulated.
- ▷ Estimating the zero $\theta^*$ of $h(.) = \mathbb{E}_x[H(., X_T)]$ by a SA is not possible !
- ▷ Therefore, we need to proceed in two steps :
  - Approximate the zero $\theta^*$ by the zero $\theta^{*,n}$ of $h^n(.) := \mathbb{E}_x[H(., X_T^n)]$ :

    Implicit discretization error : $\mathcal{E}_D(n, T, b, \sigma, H) := \theta^* - \theta^{*,n}$.

    Related issues : $\theta^{*,n} \to \theta^*$ ? What about the rate ? Expansion ?
  - Estimate $\theta^{*,n}$ by $M \in \mathbb{N}^*$ steps of the following SA scheme :

$$\theta^n_{p+1} = \theta^n_p - \gamma_{p+1} H(\theta^n_p, (X^n_T)^{p+1}), \quad p \in [\![0, M-1]\!]$$

    Statistical error : $\mathcal{E}_S(n, M, \gamma, T, H) := \theta^{*,n} - \theta^n_M$.
- ▷ Therefore, the global error between $\theta^*$ and its approximation $\theta^n_M$ is :

$$\begin{aligned}\mathcal{E}_{glob}(M, \gamma, H) &= \theta^* - \theta^{*,n} + \theta^{*,n} - \theta^n_M \\ &:= \mathcal{E}_D(n, T, b, \sigma, H) + \mathcal{E}_S(n, M, \gamma, T, H).\end{aligned}$$

# Outline of the presentation

# On the implicit discretization error

## Proposition

$\forall n \in \mathbb{N}^*$, assume that $h$ and $h^n$ satisfy a mean reverting assumption. Moreover, suppose that $(h^n)_{n\geq 1}$ converges loc. unif. towards $h$. Then, one has :

$$\theta^{*,n} \to \theta^* \text{ as } n \to +\infty.$$

## Proposition

Suppose that $h$ and $h^n$, $n \geq 1$, are $\mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ and that $Dh(\theta^*)$ is non-singular. Assume that $(Dh^n)_{n\geq 1}$ conv. loc. unif. to $Dh$. If $\exists \alpha \in [0, 1]$ s.t.

$$\forall \theta \in \mathbb{R}^d, \lim_{n\to+\infty} n^{\alpha}(h^n(\theta) - h(\theta)) = \mathcal{E}(h, \alpha, \theta),$$

then, one has

$$\lim_{n\to+\infty} n^{\alpha}(\theta^{*,n} - \theta^*) = -Dh^{-1}(\theta^*)\mathcal{E}(h, \alpha, \theta^*).$$

# Optimal tradeoff between implicit discretization and statistical errors

Remember that the global error between $\theta^*$ and its estimate $\theta^n_M$ is :

$$\mathcal{E}_{glob}(M,\gamma,H) = \theta^* - \theta^{*,n} + \theta^{*,n} - \theta^n_M$$

Suppose that :

- $\exists\, \underline{\lambda} > 0, \ \ \forall n \geq 1, \ \forall \theta \in \mathbb{R}^d, \ \langle \theta - \theta^{*,n}, h^n(\theta) \rangle \geq \underline{\lambda} |\theta - \theta^{*,n}|^2$.
- $\gamma$ varies regul. with exponent $(-\rho)$, $\rho \in [1/2, 1)$, that is, $\forall x > 0$, $\lim_{t\to+\infty} \gamma(tx)/\gamma(t) = x^{-\rho}$, $\zeta = 0$.
- for $t \geq 1$, $\gamma(t) = \gamma_0/t$ and $\gamma_0$ satisfies $2\underline{\lambda}\gamma_0 > 1$, $\zeta = 1/(2\gamma_0)$

**Theorem**

Under these assumptions, one has

$$n^{\alpha} \left( \theta^n_{\gamma^{-1}(1/n^{2\alpha})} - \theta^* \right) \Longrightarrow -Dh^{-1}(\theta^*)\mathcal{E}(h,\alpha,\theta^*) + \mathcal{N}\left(0, \Sigma^*\right),$$

$$\Sigma^* := \int_0^{\infty} \exp\left(-s(Dh(\theta^*) - \zeta I_d)\right)^T \mathbb{E}_x[H(\theta^*, X_T)H(\theta^*, X_T)^T] \exp\left(-s(Dh(\theta^*) - \zeta I_d)\right) ds$$

# Interpretation

For a global error of order $n^{-\alpha}$, one needs to devise $M = \gamma^{-1}(n^{-2\alpha})$ steps of the SA.

▷ Computational cost of SA is :

$$C_{\text{SA}}(\gamma) = C \times n \times \gamma^{-1}(n^{-2\alpha}),$$

▷ Two basic step sequences :

- if $\gamma(p) = \gamma_0/p$ with $2\underline{\lambda}\gamma_0 > 1$, then $C_{SA} = C \times n^{2\alpha+1}$.
- if $\gamma(p) = \gamma_0/p^{\rho}$, $\frac{1}{2} < \rho < 1$, then $C_{SA} = C \times n^{\frac{2\alpha}{\rho}+1}$.

▷ Optimal complexity is reached for $\gamma(p) = \gamma_0/p$

▷ Main drawback : The constraint on $\gamma_0$ is difficult to handle in practical implementation.

# Outline of the presentation

# The statistical Romberg SA method

It is clearly apparent that : $\theta^{*,n} = \theta^{*,n^\beta} + \theta^{*,n} - \theta^{*,n^\beta}$, $\beta \in (0,1)$. We estimate $\theta^*$ by :

$$\Theta_n^{sr} = \theta_{M_1}^{n^\beta} + \theta_{M_2}^{n} - \theta_{M_2}^{n^\beta}.$$

$\triangleright$ $(\theta_{M_2}^n, \theta_{M_2}^{n^\beta})$ is computed using two Euler approximation schemes with different time steps but with the same Brownian path.
$\triangleright$ $\theta_{M_1}^{n^\beta}$ comes from Brownian paths which are independent to those used for the computation of $(\theta_{M_2}^n, \theta_{M_2}^{n^\beta})$.
To establish a CLT we need the following assumptions :

- $\forall \theta$, $\mathbb{P}(X_T \notin \mathcal{D}_{H,\theta}) = 0$, $\mathcal{D}_{H,\theta} := \{x \in \mathbb{R}^q : x \mapsto H(\theta,x)$ differ. at $x\}$.
- $\forall (\theta, \theta', x) \in (\mathbb{R}^d)^2 \times \mathbb{R}^q,\ |H(\theta,x) - H(\theta',x)| \leq C(1+|x|^r)|\theta - \theta'|$.
- $\forall \theta \in \mathbb{R}^d,\ n^{1/2}\|Dh^n(\theta) - Dh(\theta)\| \to 0,\quad$ as $n \to +\infty$.

# CLT for the two-level SA method

## Theorem

Suppose that $\tilde{\mathbb{E}}(D_xH(\theta^*, X_T)U_T)(D_xH(\theta^*, X_T)U_T)^T$ is positive definite. Assume that $(\gamma(p))_{p\geq 1}$ satisfies one of the following assumptions :

- $\gamma$ varies regul. with expon. $(-\rho)$, $\rho \in (1/2, 1)$, $\zeta = 0$
- for $t \geq 1$, $\gamma(t) = \gamma_0/t$ and $\gamma_0$ satisfies $\underline{\lambda}\gamma_0 > 1$, $\zeta = 1/(2\gamma_0)$.

Then, for $M_1 = \gamma^{-1}(1/n^{2\alpha})$ and $M_2 = \gamma^{-1}(1/(n^{2\alpha-\beta}T))$, one has

$$n^{\alpha}(\Theta_n^{sr} - \theta^*) \Longrightarrow Dh^{-1}(\theta^*)\mathcal{E}(h, \alpha, \theta^*) + \mathcal{N}(0, \Sigma^*), \quad n \to +\infty$$

with

$$\Sigma^* := \int_0^{\infty} \left(e^{-s(Dh(\theta^*)-\zeta I_d)}\right)^T (\mathbb{E}_x[H(\theta^*, X_T)H(\theta^*, X_T)^T] \\ + \tilde{\mathbb{E}}\,(D_xH(\theta^*, X_T)U_T)\,(D_xH(\theta^*, X_T)U_T)^T)e^{-s(Dh(\theta^*)-\zeta I_d)}ds$$

# Sketch of proof 1/3

First use the following decomposition :

$$\Theta_n^{sr} - \theta^* = \theta^{n^\beta}_{\gamma^{-1}(1/n^{2\alpha})} - \theta^{*,n^\beta} + \theta^{n}_{\gamma^{-1}(1/n^{2\alpha-\beta})} - \theta^{n^\beta}_{\gamma^{-1}(1/n^{2\alpha-\beta})} - (\theta^{*,n} - \theta^{*,n^\beta}) + \theta^{*,n} - \theta^*$$

▷ Step 1 : Impl. discret. error : $n^\alpha(\theta^{*,n} - \theta^*) \to -Dh^{-1}(\theta^*)\mathcal{E}(h,\alpha,\theta^*)$.

▷ Step 2 : We also have : $n^\alpha(\theta^{n^\beta}_{\gamma^{-1}(1/n^{2\alpha})} - \theta^{*,n^\beta}) \Longrightarrow \mathcal{N}(0,\Gamma^*)$ with

$$\Gamma^* := \int_0^\infty \exp\left(-s(Dh(\theta^*) - \zeta I_d)\right)^T \mathbb{E}_x[H(\theta^*,X_T)H(\theta^*,X_T)^T] \exp\left(-s(Dh(\theta^*) - \zeta I_d)\right) ds$$

▷ Step 3 : Use the following decomposition :

$$\begin{aligned}
&\theta^{n}_{\gamma^{-1}(1/n^{2\alpha-\beta})} - \theta^{n^\beta}_{\gamma^{-1}(1/n^{2\alpha-\beta})} - (\theta^{*,n} - \theta^{*,n^\beta}) \\
&= \theta^{n}_{\gamma^{-1}(1/n^{2\alpha-\beta})} - \theta_{\gamma^{-1}(1/n^{2\alpha-\beta})} - (\theta^{*,n} - \theta^*) \\
&- (\theta^{n^\beta}_{\gamma^{-1}(1/n^{2\alpha-\beta})} - \theta_{\gamma^{-1}(1/n^{2\alpha-\beta})} - (\theta^{*,n^\beta} - \theta^*))
\end{aligned}$$

where $(\theta_p)_{p\geq 0}$ is the *artificial* SA : $\theta_{p+1} = \theta_p - \gamma_{p+1}H(\theta_p,(X_T)^{p+1})$, $\theta_0 = \theta_0^n$.

# Sketch of proof 2/3

Then, we prove :

$$n^{\alpha}\left(\theta^{n^{\beta}}_{\gamma^{-1}(1/(n^{2\alpha-\beta}T))}-\theta_{\gamma^{-1}(1/(n^{2\alpha-\beta}T))}-(\theta^{*,n^{\beta}}-\theta^{*})\right)\Longrightarrow\mathcal{N}(0,\Theta^{*})$$

with

$$\Theta^{*}:=\int_{0}^{\infty}(e^{-s(Dh(\theta^{*})-\zeta I_d)})^{T}\tilde{\mathbb{E}}\left(D_xH(\theta^{*},X_T)U_T\right)\left(D_xH(\theta^{*},X_T)U_T\right)^{T}e^{-s(Dh(\theta^{*})-\zeta I_d)}ds$$

and $n^{\alpha}\left(\theta^{n}_{\gamma^{-1}(1/(n^{2\alpha-\beta}T))}-\theta_{\gamma^{-1}(1/(n^{2\alpha-\beta}T))}-(\theta^{*,n}-\theta^{*})\right)\xrightarrow{\mathbb{P}}0.$

▷ A Taylor's expansion yields for $p\geq 0$

$$\theta^{n^{\beta}}_{p+1}-\theta^{*,n^{\beta}}=\theta^{n^{\beta}}_{p}-\theta^{*,n^{\beta}}-\gamma_{p+1}Dh^{n^{\beta}}(\theta^{*,n^{\beta}})(\theta^{n^{\beta}}_{p}-\theta^{*,n^{\beta}})+\gamma_{p+1}\Delta M^{n}_{p+1}-\gamma_{p+1}\zeta^{n^{\beta}}_{p}$$
$$\theta_{p+1}-\theta^{*}=\theta_{p}-\theta^{*}-\gamma_{p+1}Dh(\theta^{*})(\theta_{p}-\theta^{*})+\gamma_{p+1}\Delta M_{p+1}-\gamma_{p+1}\zeta_{p},$$

# Sketch of proof 3/3

Simple induction shows that $z_p^{n^\beta} = \theta_p^{n^\beta} - \theta_p - (\theta^{*,n^\beta} - \theta^*)$

$$z_n^{n^\beta} = \Pi_{1,n} z_0^{n^\beta} + \sum_{k=1}^{n} \gamma_k \Pi_{k+1,n} \Delta N_k^{n^\beta} + \sum_{k=1}^{n} \gamma_k \Pi_{k+1,n} \Delta R_k^{n^\beta} + \sum_{k=1}^{n} \gamma_k \Pi_{k+1,n} \Delta S_k^{n}$$

where

- $\Pi_{k,n} := \prod_{j=k}^{n} (I_d - \gamma_j Dh(\theta^*))$,
- non-linear. innov : $\Delta N_k^{n^\beta} = h^{n^\beta}(\theta^*) - h(\theta^*) - (H(\theta^*, (X_T^{n^\beta})^{k+1}) - H(\theta^*, (X_T)^{k+1}))$
- non-linear. in space : $\Delta R_k^{n^\beta} = h^{n^\beta}(\theta_k^{n^\beta}) - h^{n^\beta}(\theta^*) - (H(\theta_k^{n^\beta}, (X_T^{n^\beta})^{k+1}) - H(\theta^*, (X_T^{n^\beta})^{k+1})) + H(\theta_k, (X_T)^{k+1}) - H(\theta^*, (X_T)^{k+1}) - (h(\theta_k) - h(\theta^*))$
- Rest : $\Delta S_k^n := \left(\zeta_{k-1}^n - \zeta_{k-1} + (Dh(\theta^*) - Dh^{n^\beta}(\theta^{*,n^\beta}))(\theta_{k-1}^{n^\beta} - \theta^{*,n^\beta})\right)$

▷ $n^\alpha \Pi_{1,\gamma^{-1}(1/(n^{2\alpha-\beta}T))} z_0^{n^\beta} \xrightarrow{L^1(\mathbb{P})} 0$.

▷ $n^\alpha \sum_{k=1}^{\gamma^{-1}(1/(n^{2\alpha-\beta}T))} \gamma_k \Pi_{k+1,\gamma^{-1}(1/(n^{2\alpha-\beta}T))} \Delta R_k^{n^\beta} \xrightarrow{\mathbb{P}} 0$

▷ $n^\alpha \sum_{k=1}^{\gamma^{-1}(1/(n^{2\alpha-\beta}T))} \gamma_k \Pi_{k+1,\gamma^{-1}(1/(n^{2\alpha-\beta}T))} \Delta S_k^{n} \xrightarrow{\mathbb{P}} 0$.

▷ CLT martingale arrays (see e.g. Hall & Heyde) + Jacod & Protter CLT :

$$n^\alpha \sum_{k=1}^{\gamma^{-1}(1/(n^{2\alpha-\beta}T))} \gamma_k \Pi_{k+1,\gamma^{-1}(1/(n^{2\alpha-\beta}T))} \Delta N_k^{n^\beta} \Longrightarrow \mathcal{N}(0, \Sigma^*)$$

# Multi-level SA scheme

It uses $L$ Euler schemes with time steps given by $T/m^{\ell}$, $\ell \in \{1, \cdots, L\}$ s.t. $m^L = n$ and estimates $\theta^*$ by

$$\Theta_n^{ml} = \theta_{M_0}^1 + \sum_{\ell=1}^{L} \theta_{M_\ell}^{m^\ell} - \theta_{M_\ell}^{m^{\ell-1}}.$$

To establish a CLT we need the following assumptions :

- $\forall \theta$, $\mathbb{P}(X_T \notin \mathcal{D}_{H,\theta}) = 0$, $\mathcal{D}_{H,\theta} := \{x \in \mathbb{R}^q : x \mapsto H(\theta, x)$ differ. at $x\}$.
- $\forall (\theta, \theta', x) \in (\mathbb{R}^d)^2 \times \mathbb{R}^q$, $|H(\theta, x) - H(\theta', x)| \leq C(1 + |x|^r)|\theta - \theta'|$.
- $\exists \beta > 1/2$, $\forall \theta \in \mathbb{R}^d$, $n^\beta \|Dh^n(\theta) - Dh(\theta)\| \to 0$, as $n \to +\infty$.
- Weak error is of order 1 : $\forall \theta \in \mathbb{R}^d, n(h^n(\theta) - h(\theta)) \to \mathcal{E}(h, 1, \theta)$.

# CLT for the Multi-level SA scheme

## Theorem

Suppose that $\tilde{\mathbb{E}}(D_xH(\theta^*,X_T)U_T)(D_xH(\theta^*,X_T)U_T)^T$ is positive definite. Assume that $(\gamma(p))_{p\geq 1}$ satisfies one of the following assumptions :

- $\gamma$ varies regul. with expon. $(-\rho)$, $\rho \in (1/2, 1)$, $\zeta = 0$
- for $t \geq 1$, $\gamma(t) = \gamma_0/t$ and $\gamma_0$ satisfies $\underline{\lambda}\gamma_0 > 1$, $\zeta = 1/(2\gamma_0)$.

Then, for $M_0 = \gamma^{-1}(1/n^2)$, $M_l = \gamma^{-1}(m^\ell \log(m)/(n^2\log(n)(m-1)T))$, $\ell = 1, \cdots, L$, one has

$$n(\Theta_n^{ml} - \theta^*) \Longrightarrow Dh^{-1}(\theta^*)\mathcal{E}(h,1,\theta^*) + \mathcal{N}(0,\Sigma^*), \quad n \to +\infty$$

with

$$\Sigma^* := \int_0^\infty \left(e^{-s(Dh(\theta^*)-\zeta I_d)}\right)^T (\mathbb{E}_x[H(\theta^*,X_T^1)H(\theta^*,X_T^1)^T] \\ + \tilde{\mathbb{E}}\left(D_xH(\theta^*,X_T)U_T\right)\left(D_xH(\theta^*,X_T)U_T\right)^T)e^{-s(Dh(\theta^*)-\zeta I_d)}ds$$

# Complexity Analysis

For a total error of order $n^{-\alpha}$, the complexity of the SR-SA method is

$$C_{\text{SR-SA}}(\gamma) = C \times (n^{\beta}\gamma^{-1}(1/n^{2\alpha}) + (n + n^{\beta})\gamma^{-1}(1/(n^{2\alpha-\beta}T))),$$

For a total error of order $n^{-1}$, the complexity of the ML-SA method is

$$C_{\text{ML-SA}}(\gamma) = C \times \left(\gamma^{-1}(1/n^2) + \sum_{\ell=1}^{L} M_\ell(m^\ell + m^{\ell-1})\right).$$

- If $\gamma(p) = \gamma_0/p$ and $\underline{\lambda}\gamma_0 > 1$ then $\beta^* = 1/2$ is the optimal choice leading to

$$C_{\text{SR-SA}}(\gamma) = C' n^{2\alpha+1/2},$$

$$C_{\text{ML-SA}}(\gamma) = C\left(n^2 + n^2(\log n)^2 \frac{m^2-1}{m(\log m)^2}\right) = \mathcal{O}(n^2(\log(n))^2),$$

- If $\gamma(p) = \gamma_0/p^\rho$, $\frac{1}{2} < \rho < 1$ then $\beta^* = \rho/(1+\rho)$ leading to

$$C_{\text{SR-SA}}(\gamma) = C' n^{\frac{2\alpha}{\rho} + \frac{\rho}{1+\rho}}, \text{ and } C_{\text{ML-SA}}(\gamma) = \mathcal{O}(n^{\frac{2}{\rho}}(\log n)^{\frac{1}{\rho}}).$$

# Outline of the presentation

# Computation of quantiles of a 1-d diffusion process

We consider a GBM : $X_t = x_0 + \int_0^t rX_s ds + \int_0^t \sigma X_s dW_s, \quad t \in [0,T]$. The quantile at level $l \in (0,1)$ of $X_T$ is

$$q_l(X_T) := \inf\{\theta : \mathbb{P}(X_T \leq \theta) \geq l\} = x_0 \exp((r - \sigma^2/2)T + \sigma\sqrt{T}\phi^{-1}(l)).$$

$q_l(X_T)$ is the unique solution to

$$h(\theta) = \mathbb{E}_x[H(\theta, X_T)] = 0, \quad H(\theta, x) = \mathbf{1}_{\{x \leq \theta\}} - l.$$

▷ parameters : $x_0 = 100, \ r = 0.05, \ \sigma = 0.4, \ T = 1, \ l = 0.7$,
▷ reference value : $q_{0.7}(X_T) = 119.69$.
▷ Implicit discretization error : We plot $n \mapsto nh^n(\theta^*)$ (MC estimator) and $n \mapsto n(\theta^{*,n} - \theta^*)$ (SA estimator) for $n = 100, \cdots, 500$, with $M = 10^8$ samples.
▷ optimal tradeoff CLT : Distribution of $n(\theta^n_{\gamma^{-1}(1/n^2)} - \theta^*)$, obtained with $n = 100$ and $N = 1000$ samples,

# Implicit discretization error behavior

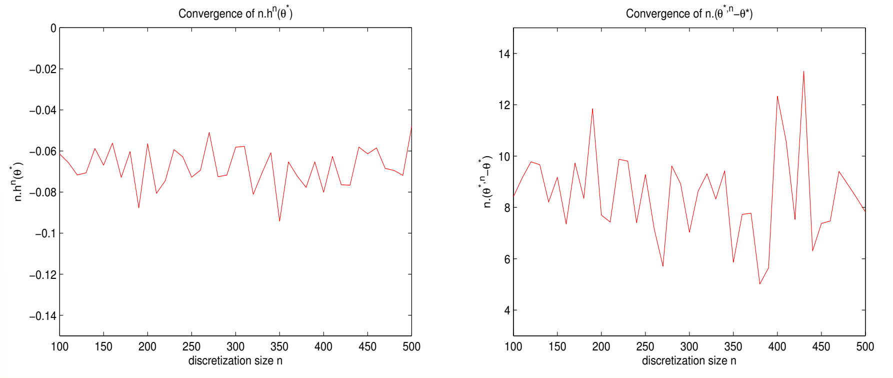



FIGURE : On the left : Weak discretization error $n \mapsto nh^n(\theta^*)$. On the right : Implicit discretization error $n \mapsto n(\theta^{*,n} - \theta^*)$, $n = 100, \cdots, 500$.

# Optimal tradeoff

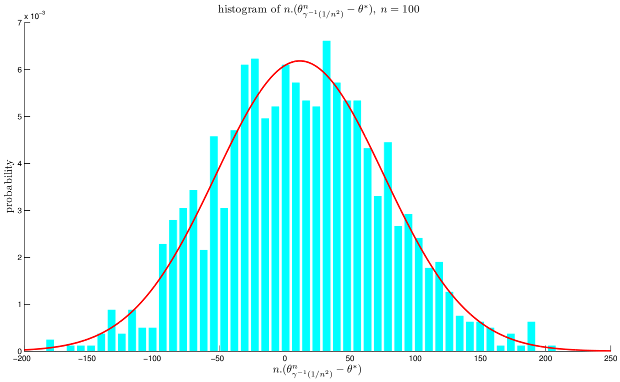



FIGURE : Histogram of $n(\theta^n_{\gamma^{-1}(1/n^2)} - \theta^*)$, $n = 100$, with $N = 1000$ samples.

# Computation of the level of an unknown function

We (still) consider a GBM : $X_t = x_0 + \int_0^t rX_s ds + \int_0^t \sigma X_s dW_s, \quad t \in [0,T]$.
For a fixed $l$, we aim at solving :

$$e^{-rT}\mathbb{E}(X_T - \theta)_+ = l \Longleftrightarrow h(\theta) = 0, \text{ with } H(\theta, x) = l - e^{-rT}(x-\theta)_+$$

We first fix a value $\theta^*$, the BS formula gives $l$ and we estimate $\theta^*$.
▷ parameters : $x_0 = 100,\ r = 0.05,\ \sigma = 0.4,\ T = 1,\ l = 0.7,$
▷ reference value : $\theta^* = 100$.
▷ Implicit discretization error : We plot $n \mapsto nh^n(\theta^*)$ (MC estimator) and $n \mapsto n(\theta^{*,n} - \theta^*)$ (SA estimator) for $n = 100, \cdots, 500$, with $M = 10^8$ samples.
▷ optimal tradeoff CLT : Distributions of $n(\theta^n_{\gamma^{-1}(1/n^2)} - \theta^*)$, $n(\Theta^{sr}_n - \theta^*)$ and $n(\Theta^{ml}_n - \theta^*)$, obtained with $n = 4^4 = 256$ and $N = 1000$ samples

# Implicit discretization error behavior

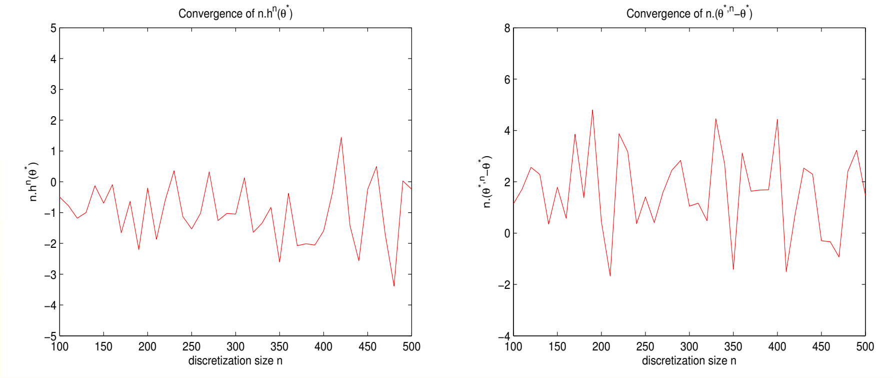



FIGURE : On the left : Weak discretization error $n \mapsto nh^n(\theta^*)$. On the right : Implicit discretization error $n \mapsto n(\theta^{*,n} - \theta^*)$, $n = 100, \cdots, 500$.

# CLT for the different schemes

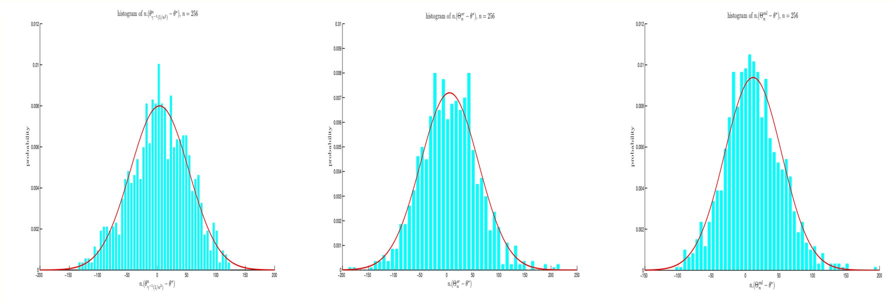



FIGURE : Histograms of $n(\theta_{\gamma^{-1}(1/n^2)} - \theta^*)$, $n(\Theta^{sr}_n - \theta^*)$, $n(\Theta^{ml}_n - \theta^*)$, $n = 100$, $N = 1000$.

# CLT for the different schemes

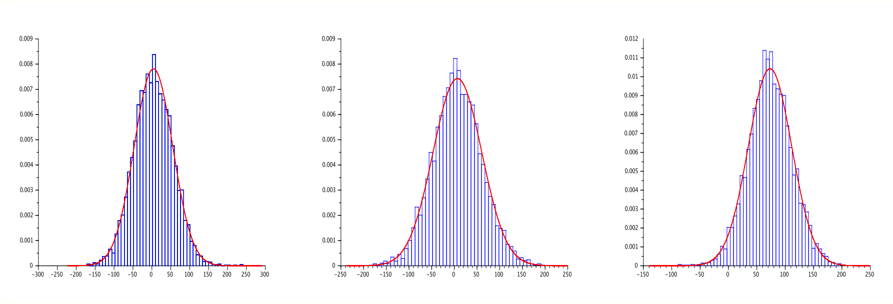


FIGURE : Histograms of $n(\theta_{\gamma^{-1}(1/n^2)} - \theta^*)$, $n(\Theta_n^{sr} - \theta^*)$, $n(\Theta_n^{ml} - \theta^*)$, $n = 625$, $N = 5000$.

# Comparison of the three estimators

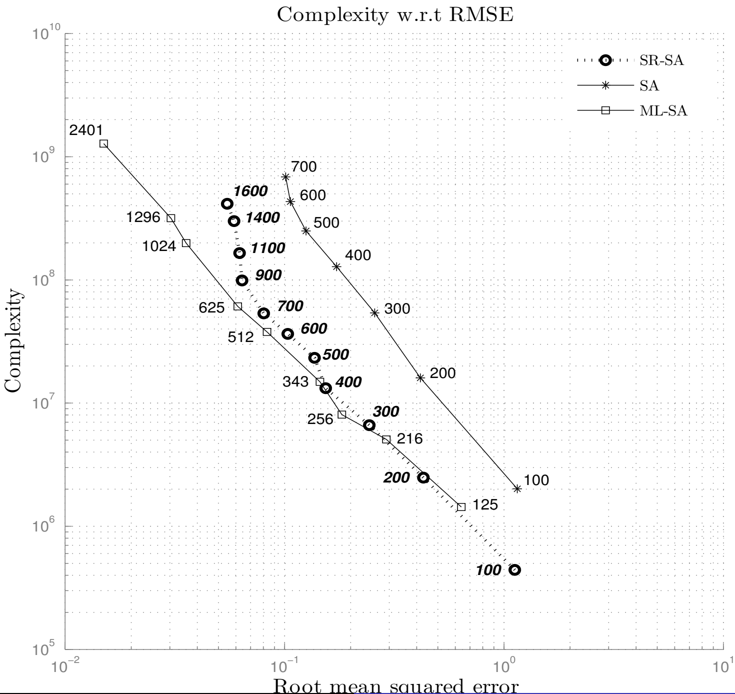
For a set of $N = 200$ different targets $\theta_k^*$ equidistributed on the interval $[90, 110]$ and for different values of $n$, we compute the complexity of each method and its root mean squared error (RMSE) :

$$\text{RMSE} = \left( \frac{1}{N} \sum_{k=1}^{N} (\Theta_k^n - \theta_k^*)^2 \right)^{1/2}$$

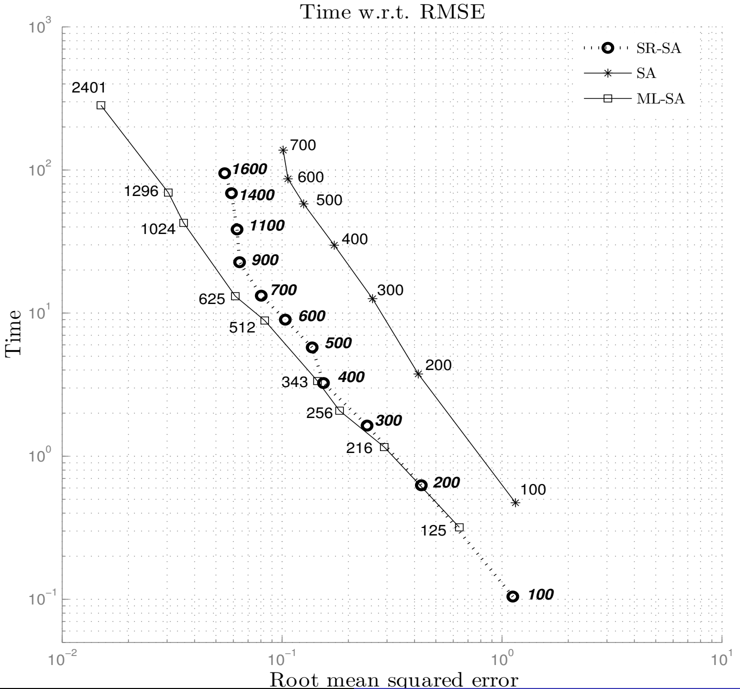
where $\Theta_k^n = \theta^n_{\gamma^{-1}(1/n^2)}$, $\Theta_n^{sr}$ or $\Theta_n^{ml}$ is the considered estimator.

▷ For each given $n$ and for each estimator, we provide a couple $(\text{RMSE}, \text{Complexity})$ (average complexity) and also the couple $(\text{RMSE}, \text{Time})$ (average time).

▷ The multi-level SA estimator has been computed for different values of $m$, $m = 2$ to $m = 7$.



Complexity w.r.t RMSE
Complexity
Root mean squared error
SR-SA
SA
ML-SA
2401
1296
1024
625
512
343
256
216
125
700
600
500
400
300
200
100
1600
1400
1100
900
700
600
500
400
300
200
100


Time w.r.t. RMSE
SR-SA
SA
ML-SA
Time
Root mean squared error
$10^{3}$
$10^{2}$
$10^{1}$
$10^{0}$
$10^{-1}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
2401
1296
1024
625
512
343
256
216
125
1600
1400
1100
900
700
600
500
400
300
200
100
700
600
500
400
300
200
100