Adaptive importance sampling for multilevel Monte Carlo

Ahmed Kebaier (joint work with J. Lelong)

LAGA, Université Paris 13

Conference in honor of Vlad Bally October 6–9, 2015 – Le Mans

Monte-Carlo framework

- Consider a diffusion process X, a time horizon T and a function ψ, we are interested in the computation of $\mathbb{E}[\psi(X_T)]$.
- Usually, one discretizes X on a time grid with n time steps and considers the Monte Carlo approximation of $\mathbb{E}[\psi(X_T^n)]$.
- How to improve things, ie reduce the mean squared error?
  - Use importance sampling to reduce the variance.
  - Use multilevel Monte Carlo to reduce both the bias and the variance.
- In this work, we aim at mixing both techniques.

Outline

1. General framework
   - Importance sampling for SDEs
   - Multilevel Monte Carlo
2. Coupling Importance sampling and multilevel MC
3. Numerical examples

General framework

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x \in \mathbb{R}^d$$

where W is a Brownian motion in $\mathbb{R}^q$.
Consider the continuous time Euler approximation X^n with time step $\delta = T/n$

$$dX_t^n = b(X^n_{\eta_n(t)})dt + \sigma(X^n_{\eta_n(t)})dW_t, \quad \eta_n(t) = \lfloor t/\delta \rfloor \delta.$$

Assume b and σ are globally Lipschitz.

- Strong error

$$\forall p \geq 1, \quad \text{and} \quad \mathbb{E}\left[\sup_{0\leq t\leq T} |X_t - X_t^n|^p\right] \leq \frac{K_p(T)}{n^{p/2}}, \quad \text{with } K_p(T) > 0.$$

- Weak error: if ψ is only $\mathcal{C}^1$

$$\text{for some } \gamma \in [1/2, 1], \quad n^\gamma(\mathbb{E}[\psi(X_T^n)] - \mathbb{E}[\psi(X_T)]) \to C_\psi(T,\gamma), \quad (1)$$

Monte Carlo approach

Approximate $\mathbb{E}[\psi(X_T)]$ by $M^{MC}_{n,N} = \frac{1}{N}\sum_{i=1}^{N}\psi(X^n_{T,i})$

$$\begin{aligned} MSE =& \mathbb{E}\left[\left|M^{MC}_{n,N} - \mathbb{E}[\psi(X_T)]\right|^2\right] \\ &= \underbrace{\mathbb{E}\left[\left|M^{MC}_{n,N} - \mathbb{E}[\psi(X^n_T)]\right|^2\right]}_{\text{Variance}} + \underbrace{\left(\mathbb{E}[\psi(X^n_T)] - \mathbb{E}[\psi(X_T)]\right)^2}_{\text{Bias}^2} \\ &= O(N^{-1}) + O(n^{-2\gamma}). \end{aligned}$$

Choose $N \approx n^{2\gamma}$! This yields $MSE = O(n^{-2\gamma})$ for a complexity of order $O(n^{2\gamma+1})$.

- Use some variance reduction technique to cut down on the constants appearing in the variance for instance using *Importance Sampling*.
- Use multilevel Monte Carlo to reduce the complexity.

Importance Sampling for SDEs

Assume we work on $(\Omega, (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$.

- Introduce for $\theta \in \mathbb{R}^q$, $\mathbb{P}_\theta \sim \mathbb{P}$ s.t. $\frac{d\mathbb{P}_\theta}{d\mathbb{P}}_{|\mathcal{F}_t} = L_t^\theta = \exp\left(-\theta \cdot W_t - \frac{1}{2}|\theta|^2 t\right) = \mathcal{E}^-(W,\theta)$
- Define $(B^\theta)_{t\leq T}$ by $B_t^\theta = W_t + \theta t$, which is a BM under $\mathbb{P}_\theta$.
- Introduce $X(\theta)$ solution of

$$dX(\theta)_t = b(X_t(\theta))dt + \sigma(X_t(\theta))dB_t^\theta,$$
$$dX(\theta)_t = \{b(X_t(\theta)) + \sigma(X_t(\theta)\}\, dt + \sigma(X_t(\theta))dW_t$$

- Under $\mathbb{P}$, X has the same distribution as $X(\theta)$ under $\mathbb{P}_\theta$.

$$\mathbb{E}[\psi(X_T)] = \mathbb{E}_{\mathbb{P}_\theta}[\psi(X_T(\theta))] = \mathbb{E}[\psi(X_T(\theta))\mathcal{E}^-(W,\theta)].$$

- Widely studied: Arouna (2004), Lemaire and Pagès (2009), Lapeyre and Lelong (2011)

Importance Sampling for SDEs

- Then, find θ minimizing

$$v(\theta) = \mathbb{E}\left[\psi(X_T(\theta))^2 e^{-2\theta\cdot W_T - |\theta|^2 T}\right] = \mathbb{E}\left[\psi(X_T)^2 \mathcal{E}^+(W,\theta)\right] \text{ or}$$
$$v_n(\theta) = \mathbb{E}\left[\psi(X_T^n(\theta))^2 e^{-2\theta\cdot W_T - |\theta|^2 T}\right] = \mathbb{E}\left[\psi(X_T^n)^2 \mathcal{E}^+(W,\theta)\right]$$

for $\mathcal{E}^+(W,\theta) = e^{-\theta\cdot W_T + \frac{1}{2}|\theta|^2 T}$ or its Monte Carlo counter part

$$v_{n,N}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \psi(X_{T,i}^n)^2 \mathcal{E}^+(W_i,\theta).$$

Set $\theta^* = \arg\min_\theta v(\theta)$ and $\theta_{n,N} = \arg\min_\theta v_{n,N}(\theta)$.

- The sample average approximation $v_{n,N}$ is infinitely differentiable and strongly convex without smoothness assumptions on ψ. See Jourdain and Lelong (2009).
- Importance sampling focuses only on the variance.

Importance Sampling for SDEs

Proposition 1

Assume for all θ, $\mathbb{E}[\psi(X_T)^2 \, \mathrm{e}^{-\theta \cdot W_T}] < \infty$.
Let $N = N_n$ *be an increasing function of* n, $\lim_{n\to\infty} N_n = \infty$.

- *For* n *large enough, the function* v_{n,N_n} *is strongly convex and smooth.*
- *If* ψ *is locally Hölder, for all* $K > 0$, $\sup_{|\theta|\leq K} |v_{n,N_n}(\theta) - v(\theta)| \xrightarrow[n\to\infty]{} 0$ *a.s.*
- $\theta_{n,N_n} \xrightarrow[n\to+\infty]{a.s.} \theta^*$ *and* $\sqrt{N_n}(\theta_{n,N_n} - \theta^*) \xrightarrow[n\to+\infty]{\mathcal{D}} N(0, \Gamma(\theta^*))$ *where* $\theta^* = \arg\min \mathbb{E}\left[\psi(X_T)^2 \mathcal{E}^+(W,\theta)\right]$.

The proof of the Proposition is based on the following technical result.

Importance Sampling for SDEs

Theorem 1

Let $(X_{n,m})_{n,m}$ *be a doubly indexed sequence of vector valued random variables such that for all n,* $\mathbb{E}[X_{n,m}] = x_m$ *with* $\lim_{m\to+\infty} x_m = x$ *. We define* $\overline{X}_{n,m} = \frac{1}{n}\sum_{i=1}^{n} X_{i,m}$*. Assume that the two following assumptions are satisfied*

1. $\sup_n \sup_m n \operatorname{Var}\left(\overline{X}_{n,m}\right) < +\infty$.
2. $\sup_n \sup_m \operatorname{Var}\left(X_{n,m}\right) < +\infty$.

Then, for all functions $\rho : \mathbb{N} \to \mathbb{N}, \ \nearrow$,

$$\overline{X}_{n,\rho(n)} \xrightarrow[n\to+\infty]{} x \text{ a.s. and in } \mathbb{L}^2.$$

Importance Sampling for SDEs

Define the adaptive Monte Carlo estimator

$$M_{n,N_n} = \frac{1}{N_n} \sum_{i=1}^{N_n} \psi(\tilde{X}^n_{T,i}(\theta_{n,N_n}))\mathcal{E}^-(\tilde{W}_i, \theta_{n,N_n}).$$

Proposition 2

If ψ is locally Holder $\alpha \geq 1$ and a weak error holds with rate n^γ then

- $M_{n,N_n} \longrightarrow \mathbb{E}[\psi(X_T)]$ *a.s. as* $n \to \infty$.
- $\sqrt{N_n}(M_{n,N_n} - \mathbb{E}[f(X_T)]) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(C_\psi(T,\alpha), \sigma^2)$ *where* $\sigma^2 = \mathbb{E}\left[\psi(X_T)^2\mathcal{E}^+(W,\theta^*)\right] - \mathbb{E}[\psi(X_T)]^2$.

We achieve the same variance as if we could directly sample X_T without any bias.

1. General framework
   - Importance sampling for SDEs
   - Multilevel Monte Carlo
2. Coupling Importance sampling and multilevel MC
3. Numerical examples

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

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

- For $N_1 = n^{2\gamma}$, $N_2 = n^{2\gamma-1/2}$ we have

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

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

$$C_{SR} = C \times n^{2\gamma+\frac{1}{2}}.$$

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 (1991-1998)provided that b and σ are $\mathcal{C}^1$.

Multilevel Monte Carlo

Choose n of the form m^L.

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

Define the Monte Carlo approximation as

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

- The samples used in two different blocks $\ell \neq \ell'$ are independent.
- In block ℓ, we use m^ℓ time steps and $N_\ell = \frac{m^{2\gamma L}(m-1)TL}{m^\ell}$ samples.
- Inside block ℓ, the discretizations $X_{T,\ell}^{m^\ell}$ and $X_{T,\ell}^{m^{\ell-1}}$ are obtained using the same Brownian paths.

Multilevel Monte Carlo

We know from Ben Alaya and K. (2015) that

$$m^{\gamma L}(Q_L - \mathbb{E}[\psi(X_T)]) \xrightarrow[L\to\infty]{\mathcal{L}} \mathcal{N}(C_\psi(T,\gamma), \mathbb{E}[(\nabla\psi(X_T)\cdot U_T)^2])$$

- If $\gamma = 1$ then, the optimal sample sizes

$$N_\ell = m^{2L-\ell}(m-1)TL.$$

- The optimal complexity is then

$$C_{MMC} = C \times m^{2L}L^2 = n^2(\log n)^2$$

- However the complexity for a crude Monte Carlo is

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

1 General framework
- Importance sampling for SDEs
- Multilevel Monte Carlo

2 Coupling Importance sampling and multilevel MC

3 Numerical examples

Coupling Importance sampling and MLMC: a first idea

From the previous CLT

$$m^{\gamma L}(Q_L - \mathbb{E}[\psi(X_T)]) \xrightarrow[L\to\infty]{\mathcal{L}} \mathcal{N}(C_\psi(T,\gamma), \mathbb{E}[(\nabla\psi(X_T)\cdot U_T)^2])$$

- Apply MLMC to $\psi(X_T(\theta))\mathcal{E}^-(W,\theta)$ instead of $\psi(X_T)$.
- Compute $\theta^* = \arg\min \mathbb{E}[(\nabla\psi(X_T)\cdot U_T)^2\mathcal{E}^+(W,\theta)]$.
- We would obtain the optimal limiting variance but the computational price would be far too high:
  - $\nabla\psi$ requires extra implementation and smoothness (in practice, not only for the theory)
  - Sample both X and U.
    See Ben Alaya, Hajji and K. (2015) in the setting of Statistical Romberg method.

Coupling IS and MLMC: a better approach

- Introduce an importance sampling parameter per level.
- Define for any $\Lambda_L = (\lambda_0, \dots, \lambda_L) \in (\mathbb{R}^q)^L$,

$$Q_L(\lambda_0, \dots, \lambda_L) = \frac{1}{N_0} \sum_{k=1}^{N_0} \psi(X^{m^0}_{T,0,k}(\lambda_0)) \mathcal{E}^-(W_{0,k}, \lambda_0) + \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(\psi(X^{m^\ell}_{T,\ell,k}(\lambda_\ell)) - \psi(X^{m^{\ell-1}}_{T,\ell,k}(\lambda_\ell)) \right) \mathcal{E}^-(W_{\ell,k}, \lambda_\ell).$$

Coupling IS and MLMC: a better approach

- Compute Λ_L^* minimizing $\mathrm{Var}(Q_L(\Lambda_L))$.
- As all the levels are independent, it is clear that λ_ℓ^* minimizes the variance of level ℓ

$$v_\ell(\lambda) = \mathbb{E}\left[\left(\left(\psi(X_T^{m^\ell}(\lambda)) - \psi(X_{T,}^{m^{\ell-1}}(\lambda))\right)\mathcal{E}^-(W,\lambda)\right)^2\right]$$
$$= \mathbb{E}\left[\left(\psi(X_T^{m^\ell}) - \psi(X_{T,}^{m^{\ell-1}})\right)^2\mathcal{E}^+(W,\lambda)\right].$$

- Approximate v_ℓ using a Monte Carlo method with N'_ℓ samples

$$v_{\ell,N'_\ell}(\lambda) = \frac{1}{N'_\ell}\sum_{k=1}^{N'_\ell}\frac{m^\ell}{(m-1)T}\left|\psi(X_{T,\ell,k}^{m^\ell}) - \psi(X_{T,\ell,k}^{m^{\ell-1}})\right|^2\mathcal{E}^+(W_{\ell,k},\lambda)$$

with $N_\ell \neq N'_\ell$.

The algorithm

for $\ell = 0 : L$ **do**

1. Generate $(X^{m^\ell}_{T,\ell,1}, X^{m^{\ell-1}}_{T,\ell,1}), \ldots, (X^{m^\ell}_{T,\ell,N'_\ell}, X^{m^{\ell-1}}_{T,\ell,N'_\ell}) \overset{i.i.d.}{\sim} (X^{m^\ell}_T, X^{m^{\ell-1}}_T)$
2. Compute the minimizer $\hat{\lambda}_\ell$ of of v_{ℓ,N'_ℓ} by solving $\nabla v_{\ell,N'_\ell}(\hat{\lambda}_\ell) = 0$.

end for

for $\ell = 1 : L$ **do**

3. Conditionally on $\hat{\lambda}_\ell$, generate $(\tilde{X}^{m^\ell}_{T,\ell,1}(\hat{\lambda}_\ell), \tilde{X}^{m^{\ell-1}}_{T,\ell,1}(\hat{\lambda}_\ell)), \ldots, (\tilde{X}^{m^\ell}_{T,\ell,N_\ell}(\hat{\lambda}_\ell), \tilde{X}^{m^{\ell-1}}_{T,\ell,N_\ell}(\hat{\lambda}_\ell)) \overset{i.i.d.}{\sim}$ $(X^{m^\ell}_T(\hat{\lambda}_\ell), X^{m^{\ell-1}}_T(\hat{\lambda}_\ell))$ independently of Step 1.

end for

4. Compute the multilevel importance sampling estimator

$$Q_L(\hat{\lambda}_0, \ldots, \hat{\lambda}_L) = \frac{1}{N_0} \sum_{k=1}^{N_0} \psi(\tilde{X}^{m^0}_{T,0,k}(\hat{\lambda}_0)) \mathcal{E}^-(\tilde{W}_{0,k}, \hat{\lambda}_0)$$
$$+ \sum_{\ell=1}^{L} \frac{1}{N_\ell} \sum_{k=1}^{N_\ell} \left(\psi(\tilde{X}^{m^\ell}_{T,\ell,k}(\hat{\lambda}_\ell)) - \psi(\tilde{X}^{m^{\ell-1}}_{T,\ell,k}(\hat{\lambda}_\ell)) \right) \mathcal{E}^-(\tilde{W}_{\ell,k}, \hat{\lambda}_\ell).$$

Some remarks on our approach

- Each component of $\hat{\Lambda}_L$ is the solution of a strictly convex minimization problem (can be solved in parallel)
- We directly minimize the true variance of the estimator.
- We do not rely on the asymptotic variance of Q_L, no need of $\nabla\psi$ nor of the process U.
- In the expression v_{ℓ,N'_ℓ}, the parameter λ_ℓ is not involved in the function ψ. Hence, the quantities $\psi(X^{m^\ell}_{T,\ell,k}) - \psi(X^{m^{\ell-1}}_{T,\ell,k})$ must only be computed once. Huge computational time savings.

Complexity of the algorithm

- Complexity of the standard ML estimator

$$C_{ML} = \sum_{\ell=0}^{L} N_\ell m^\ell.$$

- Complexity of the ML+IS estimator

$$C_{MLIS} = \sum_{\ell=0}^{L} N'_\ell (m^\ell + 3K_\ell) + \sum_{\ell=0}^{L} N_\ell m^\ell$$

  - K_ℓ: number of Newton's iterations to approximate $\hat{\lambda}_\ell$
  - the factor 3: building $\nabla u_{\ell,N'_\ell}$ and $\nabla^2 u_{\ell,N'_\ell}$ basically boils down to three Monte Carlo summations.

In practice, $K_\ell \leq 5$. So, if we choose $N'_\ell = \frac{N_\ell m^\ell}{m^\ell + 15}$, $C_{MLIS} = 2C_{ML}$.

SLLN for Multilevel Monte Carlo

- Let us introduce a sequence $(a_\ell)_{\ell\in\mathbb{N}}$ of positive real numbers such that $\lim_{L\to\infty}\sum_{\ell=1}^{L} a_\ell = \infty$ and let

$$N_{\ell,L}^{\rho} = \frac{\rho(L)}{m^\ell a_\ell}\sum_{k=1}^{L} a_k,\ \ell \in \{0,\cdots,L\} \text{ for some } \rho : \mathbb{N} \to \mathbb{R} \nearrow . \qquad (2)$$

- Moreover, we prove a SLLN for Multilevel Monte Carlo method

Theorem 2

Assume that $\sup_L \sup_\ell \frac{L^2 a_\ell}{\rho(L)\sum_{k=1}^{L} a_k} < +\infty$. *Then, under our assumptions*

$$Q_L(\widehat{\lambda}_0,\ldots,\widehat{\lambda}_L) \longrightarrow \mathbb{E}[\psi(X_T)] \text{ a.s. when } L \to +\infty.$$

Lindeberg Feller CLT

Theorem 3

Under ou assumptions, for $N^{\rho}_{\ell,L}$ given by (2) with $\rho(L) = m^{2\gamma L}(m-1)T$ and the sequence $(a_\ell)_\ell$ satisfying

$$\lim_{L\to\infty} \frac{1}{\left(\sum_{\ell=1}^{L} a_\ell\right)^{p/2}} \sum_{\ell=1}^{L} a_\ell^{p/2} = 0, \quad \textit{for } p > 2$$

we have

$$m^{\gamma L}\left(Q_L(\widehat{\lambda}_0, \ldots, \widehat{\lambda}_L) - \mathbb{E}[\psi(X_T)]\right) \xrightarrow[L\to+\infty]{\mathcal{D}} \mathcal{N}(C_\psi(T,\gamma), v(\lambda^*))$$

where $v(\lambda) = \mathbb{E}\left[(\nabla\psi(X_T)\cdot U_T)^2\, \mathcal{E}^+(W,\lambda)\right]$ *and* $\lambda^* = \arg\min_\lambda v(\lambda)$.

Some remarks

- We recover the optimal rate : $v(\lambda^*)$ is the limiting variance obtained when directly minimizing the asymptotic variance.
- To avoid minimizing (and computing) the limiting variance involving U, we have to solve L optimization problems but it makes the method fully automatic. Worth it.
- Even though the limiting variance is given by the blocks $\ell > 0$, and the block $\ell = 0$ plays no role in the variance, it is not advisable in practice (for a finite value of L) not to perform importance sampling for the block $\ell = 0$.

1. General framework
   - Importance sampling for SDEs
   - Multilevel Monte Carlo
2. Coupling Importance sampling and multilevel MC
3. Numerical examples

Measuring the efficiency

We compare the mean squared errors for

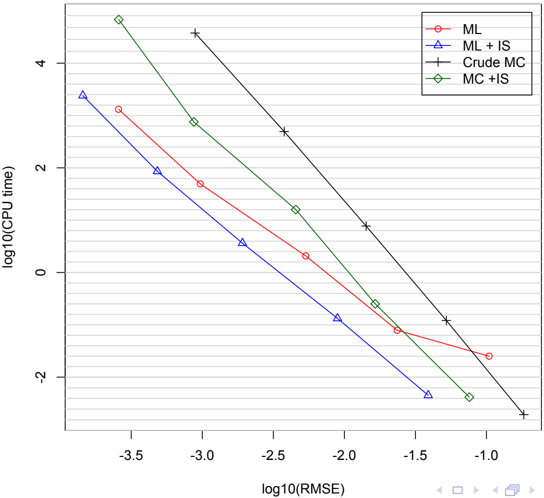
- crude Monte Carlo (MC)
- Monte Carlo with Importance Sampling (MCIS) Jourdain, L. (2009)
- Multi Level Monte Carlo (ML)
- Multi Level Monte Carlo with importance sampling (MLIS)

A "true" value is computed using a ML with a large number of levels ($L = 7$). In this case, for $m = 4$, $N_0 = 5.6\ 10^9$. Computing the benchmark price takes ages, use parallel computing to split the resolution of one level.

Basket option in a local volatility model

- $dS_t^i = S_t^i(rdt + \sigma_i(t, S_t^i)dW_t^i)$ for $i = 1, \cdots, I$
  with $\sigma_i(t,x) = 0.6(1.2 - \mathrm{e}^{-0.1t}\,\mathrm{e}^{-0.001(x\,\mathrm{e}^{rt} - S_0^i)^2})\,\mathrm{e}^{-0.05\sqrt{t}}$
  and $(W^1, \ldots, W^I)$ correlated Brownian motions Cov $(W_t^i, W_t^j) = \rho\, t$ if $i \neq j$.
- payoff : $\left(K - \frac{1}{I}\sum_{i=1}^{I} S_T^i\right)_+$
- Parameters: $I = 5$, $r = 0.05$, $T = 1$, $S_0 = 100$, $K = 100$, $m = 4$.

Basket option in a local volatility model



Best of option in multidimensional Heston model

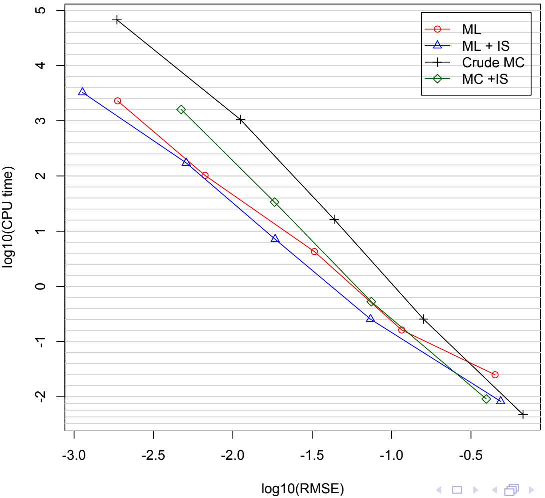
- $$\begin{aligned} dS_t^i &= rS_t^i dt + \sqrt{\sigma_t^i} S_t^i dB_t^i \\ d\sigma_t^i &= \kappa^i (a^i - \sigma_t^i) dt + \nu_t^i \sqrt{\sigma_t^i} (\gamma^i dB_t^i + \sqrt{1 - (\gamma^i)^2} d\tilde{B}_t^i) \end{aligned}$$

  with

  $$d\langle B \rangle_t = \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \dots & \rho & 1 \end{pmatrix} dt \quad \text{and} \quad d\langle \tilde{B} \rangle_t = I_d \, dt$$

- payoff : $\left(\max_{1 \leq i \leq I} S_T^i - K \right)_+$
- Parameters: $I = 5$, $r = 0.03$, $T = 1$, $S_0 = 100$, $K = 140$, $m = 4$, $\nu = 0.25$, $\kappa = 2$, $a = 0.04$, $\gamma = 0.2$, $\rho = 0.5$.

Best of option in multidimensional Heston model



Thank you ! Thanks to Vlad !