A Symmetrized Milstein scheme with strong rate of convergence for some CEV-like SDEs

Mireille Bossy and Héctor Olivero

Inria Sophia Antipolis   Universidad de Chile

Closing international conference of Thematic Cycle on Monte-Carlo techniques

Paris July 5-8th 2016

On the discretization of SDEs with non-Lipschitz diffusion

Focus on

- one dimensional case
- special form of non-Lipschitz diffusion : constant elasticity of variance (CEV) type
- strong error

Let $(X_t)_{t\geq 0}$ be the $\mathbb{R}$-valued process solution to

$$X_t = x_0 + \int_0^t b(X_s)ds + \sigma \int_0^t |X_s|^\alpha dW_s, \tag{1}$$

$x_0 > 0$ and $\sigma > 0$ are given.
$(W_t)_{t\geq 0}$ is a $\mathbb{R}$-valued Brownian motion on a given $(\Omega, \mathscr{F}, \mathbb{P})$.

(1) has a unique strong solution when

Hypotheses

(H0) $\alpha \in [1/2, 1)$.

(H1) *drift function* b *such that* $b(0) > 0$, and Lipschitz
$(|b(x) - b(y)| \leq K|x-y|, \ \forall (x,y) \in \mathbb{R}^2)$.

Then for any $q \geq 0$,

$$\mathbb{E}\Big[\sup_{0\leq t\leq T} X_t^q\Big] \leq C.$$

Some popular examples in finance

- Cox, Ingersoll and Ross (1985) proposed to model the dynamics of the short term interest rate r_t as the solution of

$$\begin{cases} dr_t^x = (a - br_t^x)dt + \sigma\sqrt{r_t^x}dW_t, \\ r_0^x = x \geq 0. \end{cases}$$

- Hull and White (1990) proposed the following mean-reverting diffusion process

$$dr_t = (a(t) - b(t)r_t)dt + \sigma(t)r_t^\alpha dW_t$$

  with $0 \leq \alpha \leq 1$.

- Basic volatility models :

$$dS_t = \mu S_t dt + \sqrt{\alpha_t} S_t dW_t,$$

$$\begin{aligned} \text{Heston model} \quad &: \quad d\alpha_t = \Theta(\omega - \alpha_t)dt + \zeta\sqrt{\alpha_t}dB_t, \\ \text{3/2 model} \quad &: \quad d\alpha_t = \Theta(\omega - \alpha_t)dt + \zeta(\alpha_t)^{3/2}dB_t. \end{aligned}$$

Other stochastic volatility models

- Hagan et al. (2002) proposed the $SABR$–model (stochastic–$\alpha\beta\rho$ model) as stochastic volatility model for forward prices :

$$dF_t = \sigma_t F_t^{\beta} dW_t^1$$
$$d\sigma_t = \alpha \sigma_t dB_t,$$

(W^1, W^2) is a 2d–Brownian motion, $B_t = \rho W_t^1 + \sqrt{(1-\rho^2)} W_t^2$, $\rho \in [-1, 1]$.

- General models for stochastic volatility (Lions Musiela 2007) :

$$dF_t = \sigma_t^{\delta} F_t^{\beta} dW_t^1$$
$$d\sigma_t = \alpha \sigma_t^{\gamma} dB_t + b(\sigma_t) dt,$$

with $\alpha, \delta > 0$, $\gamma, \beta \in (0, 1]$ + other restrictions.

Other application domain : turbulent flow modelling

Lagrangian models for turbulent particle transport or direct turbulent (subgrid) modelling

$$\begin{cases} d\mathrm{X}_t = \mathrm{U}_t dt, \\ d\mathrm{U}_t^{(i)} = \left[-\dfrac{\partial \langle \mathscr{P} \rangle}{\partial x_i}(t,\mathrm{X}_t) - C \langle \mathscr{W} \rangle (t,\mathrm{X}_t) \left(\mathrm{U}_t^{(i)} - \left\langle \mathscr{U}^{(i)} \right\rangle (t,\mathrm{X}_t) \right) \right] dt \\ \qquad\qquad + \sqrt{C_0 k(t,\mathrm{X}_t) \langle \mathscr{W} \rangle (t,\mathrm{X}_t)} dW_t^{(i)}, \ \forall\, i \in \{1,2,3\} \end{cases} \tag{2}$$

with

$$d\omega_t = -C_3 \langle \mathscr{W} \rangle (t,\mathrm{X}_t) \left(\omega_t - \langle \mathscr{W} \rangle (t,\mathrm{X}_t) \right) dt - S_\omega(t,\mathrm{X}_t) \langle \mathscr{W} \rangle (t,\mathrm{X}_t) \omega_t dt \\ + \sqrt{C_4 \langle \mathscr{W} \rangle^2 (t,\mathrm{X}_t) \omega_t} dW_t^{(4)}$$

Kinetic turbulent energy : $k(t,x) = \dfrac{1}{2} \sum_{i=1}^{3} \langle \mathrm{U}^{(i)} \mathrm{U}^{(i)} \rangle (t,x) - \langle \mathrm{U}^{(i)} \rangle^2 (t,x)$

Pressure $\langle \mathscr{P} \rangle (t,x)$ must be recovered by $\nabla^2 \langle \mathscr{P} \rangle = -\dfrac{\partial^2 \left\langle \mathscr{U}^{(i)} \mathscr{U}^{(j)} \right\rangle}{\partial x_i \partial x_j}$.

$$\left\langle g(\mathscr{U}, \mathscr{W}) \right\rangle (t,x) = \mathbb{E} \left[g(\mathrm{U}_t, \omega_t) / \mathrm{X}_t = x \right].$$

(non exhaustive) Bibliography (I)

- CIR case $b(x) = a - bx$, $b > 0$ and $\alpha = \frac{1}{2}$

A. Alfonsi.
High order discretization schemes for the CIR process: application to affine term structure and Heston models.
Mathematics of Computation, 79(269):209–237, 2010.

Martin Hutzenthaler, Arnulf Jentzen, and Marco Noll.
Strong convergence rates and temporal regularity for Cox-Ingersoll-Ross processes and Bessel processes with accessible boundaries.
arXiv preprint arXiv:1403.6385, 2014.

- Lamperti transform based / implicit drift based approach $(\frac{1}{2} \leq \alpha \leq 1$, b Lipschitz)

A. Alfonsi.
Strong order one convergence of a drift implicit Euler Scheme: Application to CIR process.
Statistic and Probability Letters, (83):602–607, 2013.

J.-F. Chassagneux, A. Jacquier, and I. Mihaylov.
An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients.
arXiv preprint arXiv:1405.3561t, 2014.

A. Neuenkirch and L. Szpruch.
First order strong approximations of scalar SDEs defined in a domain.
Numerische Mathematik, 128(1):103–136, 2014.

(non exhaustive) Bibliography (II)

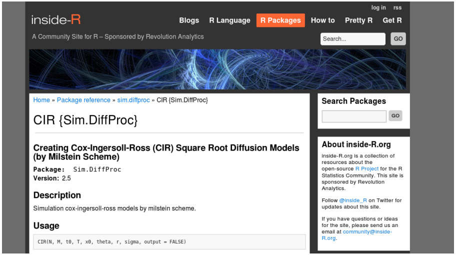
- Direct approach - (explicit/implicit drift) schemes ($\frac{1}{2} \leq \alpha \leq 1$, b Lipschitz)

C. Kahl and H. Schurz.
Balanced Milstein methods for ordinary SDEs.
Monte Carlo Methods and Applications, 12(2):143–170, 2006.

A. Berkaoui, M. Bossy, and A. Diop.
Euler sheme for SDEs with non-Lipschitz diffusion coeffcient : strong convergence.
ESAIM Probability and Statistics, 12:1–11, 2008.

D. J Higham, X. Mao, and L. Szpruch.
Convergence, non-negativity and stability of a new Milstein scheme with applications to finance.
Discrete Contin. Dyn. Syst. Ser. B, 18(8):2083–2100, 2013.

Many references about the use of Milstein scheme for CIR on the web:



Milstein scheme for CEV type SDEs is a good idea

$x_0 > 0$, $T > 0$, and $N \in \mathbb{N}$. $\Delta t = T/N$ and $t_k = k\Delta t$. Define $(\overline{X}_{t_k}, k = 0, \ldots, N)$ by

$$\overline{X}_{t_k} = \begin{cases} x_0, \text{ for } k = 0, \\ \left| \overline{X}_{t_{k-1}} + b(\overline{X}_{t_{k-1}})\Delta t + \sigma \overline{X}^{\alpha}_{t_{k-1}} (W_{t_k} - W_{t_{k-1}}) + \frac{\alpha\sigma^2}{2} \overline{X}^{2\alpha-1}_{t_{k-1}} \left[(W_{t_k} - W_{t_{k-1}})^2 - \Delta t \right] \right|, \\ \qquad \text{for } k = 1, \ldots, N \quad \text{is the Symmetrized Milstein Scheme (SMS)} \\ \max\left(0, \overline{X}_{t_{k-1}} + b(\overline{X}_{t_{k-1}})\Delta t + \sigma \overline{X}^{\alpha}_{t_{k-1}} (W_{t_k} - W_{t_{k-1}}) \right. \\ \qquad \left. + \frac{\alpha\sigma^2}{2} \overline{X}^{2\alpha-1}_{t_{k-1}} \left[(W_{t_k} - W_{t_{k-1}})^2 - \Delta t \right] \right), \\ \qquad \text{for } k = 1, \ldots, N \quad \text{is the projected Milstein Scheme (PMS)} \end{cases}$$

Milstein increment :

$$\overline{X}_{t_{k-1}} + b(\overline{X}_{t_{k-1}})\Delta t + \sigma \overline{X}^{\alpha}_{t_{k-1}} (W_{t_k} - W_{t_{k-1}}) + \underbrace{\frac{\alpha\sigma^2}{2} \overline{X}^{2\alpha-1}_{t_{k-1}} \left[(W_{t_k} - W_{t_{k-1}})^2 - \Delta t \right]}_{\text{pushes the process in } (0, +\infty)}$$

MLMC application : a motivation for studding the strong error

Theorem – (Bossy Diop 2004)

For the symmetrized Euler scheme. Let b and f be a $\mathbb{R}$-valued C^4 bounded function, with bounded spatial derivatives up to the order 4.
When $\alpha = \frac{1}{2}$, if $b(0) \geq 2\sigma^2$, then

$$\left|\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T)\right| \leq C\left(\Delta t + \left(\frac{\Delta t}{x_0}\right)^{\frac{b(0)}{\sigma^2}}\right) \quad \text{for } \Delta t \leq \frac{1}{2K} \wedge x_0.$$

When $\frac{1}{2} < \alpha < 1$, $\quad |\mathbb{E}f(X_T) - \mathbb{E}f(\overline{X}_T)| \leq C\left(1 + \frac{1}{x_0^{q(\alpha)}}\right)\Delta t, \quad$ for $\Delta t \leq \frac{1}{2K} \wedge \frac{2x_0}{b(0)}$

Let $\mathbb{E}\widehat{B}(\Delta t_l)$ a discrete-time weak approximation of $\mathbb{E}f((X_t, 0 \leq t \leq T))$.
Multilevel Monte Carlo estimator :

$$\widehat{Y}_T = \frac{1}{N_0}\sum_{i=1}^{N_0} \widehat{B}^{(i)}(\Delta t_0) + \sum_{l=1}^{L} \frac{1}{N_l}\sum_{i=1}^{N_l}\left(\widehat{B}^{(i)}(\Delta t_l) - \widehat{B}^{(i)}(\Delta t_{l-1})\right).$$

Optimal Choice for $L, \Delta t_t = \frac{T}{M^l}, N_l, l = 1, \ldots, L$ in order to minimize the computational time (complexity) for a targeted RMS error ε (Giles '08)

$$\mathbb{E}[(\widehat{Y}_T - \mathbb{E}f((Xt, 0 \leq t \leq T)))^2] = \mathscr{O}(\varepsilon^2)$$

MLMC experiment on the Zero Coupon Bound pricing with CIR model

ZCB pricing of maturity T, $B(0,T) = \mathbb{E}[\exp(-\int_0^T r_s ds)]$, that admits a wellknown closed-form solution, easy to compute with precision

$$B(0,T) = A(T)e^{-B(T)x_0}, \quad A(T) = \left[\frac{2\lambda e^{(b+\lambda)T/2}}{(\lambda+b)(e^{\lambda T}-1)+2\lambda}\right]^{\frac{2a}{\sigma^2}}, \quad B(T) = \frac{2(e^{\lambda T}-1)}{(\lambda+b)(e^{\lambda T}-1)+2\lambda}.$$

where $\lambda = \sqrt{b^2+2\sigma^2}$.

Given ε, choose the following a priori parametrization of the MLMC method (Giles '08) :

- $\Delta t_l = \frac{1}{2^{(l+1)}}$,
- $L = \dfrac{\log \varepsilon^{-1}}{\log 2}$
- $N_l = \frac{2}{\varepsilon^2}\sqrt{V_l \Delta t_l}\left(\sum_{l=0}^{L}\sqrt{V_l/dt_l}\right)$
- with $V_l = \overline{\mathbb{V}ar}\left(\widehat{B}^{(1)}(\Delta t_l) - \widehat{B}^{(1)}(\Delta t_{l-1})\right)$

MLMC experiment on the ZCB pricing with CIR model

Comparison with

- Alfonsi Implicit Scheme (AIS) for CIR (Alfonsi 2005)
- Balanced Milstein Scheme (BMS) (Kahl and Schurz 2006)

ε = 1.0e-03 ($L = 9$, $\Delta t_L = 1/2^{10}$)	SMS	PMS	AIS	BMS
CPU time ($N_0 + \cdots + N_L$) (effective error)	0.2304 (792 651) (1.970e-05)	0.2657 (950 838) (3.347e-04)	0.264 (990 769) (3.132e-04)	0.274 (992 432) (3.292e-04)

ε = 1.0e-04 ($L = 13$, $\Delta t_L = 1/2^{14}$)	SMS	PMS	AIS	BMS
CPU time ($N_0 + \cdots + N_L$) (effective error)	16.871 (56 229 224) (4.870e-05)	20.843 (70 876 600) (1.091e-04)	17.311 (73 824 621) (9.538e-06)	16.95 (73 668 115) (2.203e-06)

Theoretical strong rate of convergence for the SMS

Theorem

Assume that

- (i) Let $p \geq 1$. If $\alpha > \frac{1}{2}$ we assume $b(0) > 2\alpha(1-\alpha)^2\sigma^2$. Whereas for $\alpha = \frac{1}{2}$ we assume $b(0) > 3\left(2[p \vee 2]+1\right)\sigma^2/2$.
- (ii) The drift coefficient b is of class $\mathscr{C}^2(\mathbb{R})$, and b'' has polynomial growth.

Define a maximum step size $\Delta_{\max}(\alpha)$ as

$$\Delta_{\max}(\alpha) = \frac{x_0}{(1-\sqrt{\alpha})b_\sigma(\alpha)} \wedge \begin{cases} \dfrac{1}{4\alpha K(\alpha)}, & \text{for } \alpha \in (\frac{1}{2}, 1) \\ \dfrac{1}{4K} \wedge x_0, & \text{for } \alpha = \frac{1}{2}. \end{cases} \tag{3}$$

Then for any $p \geq 1$ that allows (i), there exists a constant C depending on p, T, $b(0)$, α, σ, K, and x_0, but not on Δt, such that for all $\Delta t \leq \Delta_{\max}(\alpha)$,

$$\sup_{0 \leq t \leq T} \left(\mathbb{E}\left[|X_t - \overline{X}_t|^p\right]\right)^{\frac{1}{p}} \leq C\Delta t. \tag{4}$$

Summary of convergence results for $\alpha = 1/2$

Condition over the parameters for the strong convergence for various schemes

Scheme	Norm	Drift	Convergence's Condition	Theoretical rate
SMS	$L^p, p \geq 1$	$b \in \mathcal{C}^2$, b Lipschitz b'' with polynomial growth	$b(0) > 3\left(2[p \vee 2] + 1\right)\frac{\sigma^2}{2}$	1
AIS [2]	$L^p, p \in [1, \frac{4b(0)}{3\sigma^2})$	$b(x) = a - bx$	$b(0) > (1 \vee \frac{3}{4}p)\sigma^2$	1
BMS				undetermined
MES [6]	L^1	$b(x) = \mu_1(x) - \mu_2(x)x$ $\mu_i \in \mathcal{C}_b^2 \cap \mathcal{C}_b^0,\ \mu_1 \geq 0$ $\mu_1' \leq 0,\ \mu_2' \geq 0$	$b(0) > \frac{5\sigma^2}{2}$	1
			$b(0) > \frac{3\sigma^2}{2}$	$1/2$
			$b(0) > \sigma^2$	$\left(\frac{1}{6}, \frac{1}{2} - \frac{\sigma^2}{2b(0)+\sigma^2}\right)$
SES[3]	$L^p, p \geq 1$	b Lipschitz	$b(0) > \left[\sqrt{\frac{8}{\sigma^2}\mathcal{K}(\frac{p}{2} \vee 1)} + 1\right]\frac{\sigma^2}{2}$, $\mathcal{K}(q) = K(16q - 1)$ $\vee 4\sigma^2(8p-1)^2$	$1/2$

- Alfonsi Implicit Scheme (AIS) for CIR (Alfonsi 2005)
- Balanced Milstein Scheme (BMS) (Kahl and Schurz 2006)
- Modified Euler Scheme (MES) (Chassagneux et al 2015),
- Symmetryzed Euler Scheme (SES) (Berkaoui et al 2008)

Numerical experiments for $\alpha = 1/2$

$dX_t = 10 - 10X_t + \sigma\sqrt{X_t}dW_t, 0 \leq t \leq 1$ and $x_0 = 1$. $\Delta t \in \{\frac{\Delta_{\max}(\alpha)}{2^n}, n = 1, \ldots 9\}$,

σ^2	Observed $L^1(\Omega)$ convergence rate $\hat{\rho}$ (and its R–squared value)									
	SMS		AIS		BMS		MES		SES	
	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)
1	1.0674	(99.8%)	0.9575	(99.9%)	0.9576	(99.9%)	1.0586	(99.8%)	0.5951	(99.2%)
4	1.0632	(99.8%)	0.9565	(99.9%)	0.9578	(99.9%)	1.0544	(99.9%)	0.5355	(99.8%)
6.25	1.0647	(99.8%)	0.9567	(99.9%)	0.9579	(99.9%)	1.0479	(99.9%)	0.5251	(99.9%)
9	1.0669	(99.8%)	0.9536	(9.97%)	0.9493	(99.9%)	1.0399	(99.9%)	0.5166	(99.9%)
36	0.6500	(9.98%)	0.5719	(99.9%)	0.4447	(99.4%)	0.0593	(95.2%)	0.4738	(99.9%)

- $\sigma^2 = 1$, $b(0) > 6\sigma^2$: theoretical rate of order 1 for SMS AIS MES
- $\sigma^2 = 4$, $b(0) \in (5\frac{\sigma^2}{2}, 6\sigma^2)$: theoretical rate of order 1 for AIS MES only
- $\sigma^2 = 6.5$, $b(0) \in (3\sigma^2/2, 5\sigma^2/2)$: theoretical rate of order 1 for AIS only
- $\sigma^2 = 9$, $b(0) \in (\sigma^2, 3\sigma^2/2)$: theoretical rate of order 1 for AIS only
- $\sigma^2 = 36$, $b(0) < \sigma^2$: only sublinear convergence are expected

Summary of convergence results for $1/2 < \alpha \leq 1$

Condition over the parameters for the strong convergence for various schemes

Scheme	Norm	Drift	Convergence's Condition	Theoretical rate
SMS	$L^p, p \geq 1$	$b \in \mathcal{C}^2$, b Lipschitz b'' with polynomial growth	$b(0) > 2\alpha(1-\alpha)^2\sigma^2$	1
AIS	$L^p, p \in [1, \frac{4b(0)}{3\sigma^2})$	$b(x) = a - bx$	$b(0) > 0$	1
BMS				undetermined
MES	L^1	$b(x) = \mu_1(x) - \mu_2(x)x$ $\mu_i \in \mathcal{C}^2_b \cap \mathcal{C}^0_b,\ \mu_1 \geq 0$ $\mu'_1 \leq 0,\ \mu'_2 \geq 0$	$b(0) > 0$	1
SES	$L^p, p \geq 1$	b Lipschitz	$b(0) > 0$	$1/2$

Numerical experiments for $\alpha = 1/2$

$dX_t = 10 - 10X_t + \sigma X_t{}^{\alpha} dW_t, 0 \leq t \leq 1$ and $x_0 = 1$. $\Delta t \in \{\frac{\Delta_{\max}(\alpha)}{2^n}, n = 1, \ldots 9\}$,

Parameters		Observed $L^1(\Omega)$ convergence rate $\hat{\rho}$ (and its R^2 value)					
		SMS		BMS		SES	
α	σ^2	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)	$\hat{\rho}$	(R^2)
0.6	49	1.0068	(99.8%)	0.7306	(99%)	0.5250	(99.8%)
	53.29	1.0015	(99.8%)	0.7785	(99.8%)	0.5135	(99.9%)
	144	0.6720	(98.6%)	0.4314	(9.73%)	0.5079	(99.9%)
0.7	64	1.0300	(99.9%)	0.9027	(99.7%)	0.5233	(99.8%)
	81	1.0203	(99.9%)	0.8811	(99.6%)	0.5324	(99.7%)
	225	0.9294	(99.5%)	0.6534	(97.8%)	0.6405	(99.3%)

- First and fourth rows : $b(0) > 2\alpha(1-\alpha)^2\sigma^2$.
- Second and fifth rows : the parameters do not satisfy $b(0) \geq 2\alpha(1-\alpha)^2\sigma^2$
- Third and six rows : $\sigma \gg b(0)$, SMS performs better than the BMS, specially when σ^2 grows.

Some elements from the proof
(i) basic facts about $(X_t, 0 \leq t \leq T)$

When $\frac{1}{2} < \alpha < 1$, for any $q > 0$,

$$\sup_{0 \leq t \leq T} \mathbb{E}\left[X_t^{-q}\right] \leq C(1 + x_0^{-q}).$$

When $\alpha = \frac{1}{2}$, for any q such that $1 < q < 2\frac{b(0)}{\sigma^2} - 1$,

$$\sup_{0 \leq t \leq T} \mathbb{E}\left[X_t^{-q}\right] \leq C x_0^{-q}.$$

For all $\mu \geq 0$, there exist a positive constant $C(T,\mu)$, increasing in μ and T, depending also on $b,\ \sigma,\ \alpha$ and x_0 such that

$$\mathbb{E}\exp\left(\mu \int_0^T \frac{ds}{X_s^{2(1-\alpha)}}\right) \leq C(T,\mu).$$

The inequality holds if $b(0) > \frac{\sigma^2}{2}$ and $\mu \leq \frac{\sigma^2}{8}(2\frac{b(0)}{\sigma^2} - 1)^2$.

(i) basic facts about $(\overline{X}_t, 0 \leq t \leq T)$

Lemma Local error

For any $p \geq 1$, $\sup_{0 \leq t \leq T} \mathbb{E}\left[|\overline{X}_t - \overline{X}_{\eta(t)}|^{2p}\right] \leq C\Delta t^p$.

For $\alpha > \frac{1}{2}$, assume $b(0) > 2\alpha(1-\alpha)^2\sigma^2$, whereas for $\alpha = \frac{1}{2}$, assume $b(0) > 3(2p+1)\sigma^2/2$. Then the Corrected Local Error satisfies

$$\sup_{0 \leq t \leq T} \mathbb{E}\left[\left|\sigma\overline{X}_t^{\alpha} - \sigma\overline{X}_{\eta(t)}^{\alpha} - \alpha\sigma^2\overline{X}_{\eta(t)}^{2\alpha-1}(W_t - W_{\eta(t)})\right|^{2p}\right] \leq C\Delta t^{2p}.$$

For any $\alpha \in [\frac{1}{2}, 1)$, if $b(0) > 2\alpha(1-\alpha)^2\sigma^2$, and $\Delta t \leq 1/(2K(\alpha))$,

$$\sup_{0 \leq t \leq T} \mathbb{P}\left(\overline{Z}_t \leq 0\right) \leq C\exp\left(-\frac{\gamma}{\Delta t}\right).$$

Fix $\rho \in (0, 1]$, and set $\bar{x}(\alpha) = b_\sigma(\alpha)/K(\alpha)$.

$$\mathbb{P}\left[\overline{Z}_t \leq (1-\rho)b_\sigma(\alpha)\Delta t,\ \overline{X}_{\eta(s)} < \rho\bar{x}(\alpha)\right] = 0.$$

For $\Delta t \leq \Delta_{\max}$ then there exist a positive $\gamma > 0$ such that

$$\mathbb{E}\left(L_T^0(\overline{X})^2\right) \leq C\frac{1}{\sqrt{\Delta t}}\exp\left(\frac{-\gamma}{2\Delta t}\right).$$

A direct proof

$$\mathcal{E}_t := \overline{X}_t - X_t$$

Itô Lemma + $\forall x \geq 0,\ y \geq 0, \quad |x^\alpha - y^\alpha|(x^{1-\alpha} + y^{1-\alpha}) \leq 2\alpha|x-y|$

$$\begin{aligned}\mathbb{E}\left[\mathcal{E}_t^{2p}\right] \leq C \int_0^t \sup_{u \leq s} \mathbb{E}\left(\mathcal{E}_u^{2p}\right) ds + 2p \int_0^t \mathbb{E}\left(\mathcal{E}_s^{2p-1}\left[b(X_{\eta(s)}) - b(X_s)\right]\right) ds \\ + 8p(2p-1) \int_0^t \mathbb{E}\left(\mathcal{E}_s^{2p-2}\left\{\sigma \overline{X}_s^\alpha - \sigma X_s^\alpha\right\}^2\right) ds + C\Delta t^{2p}.\end{aligned} \tag{5}$$

Trick: $\mathbb{E}\left(\mathcal{E}_s^{2p-2}\left\{\sigma \overline{X}_s^\alpha - \sigma X_s^\alpha\right\}^2\right) \leq C\mathbb{E}\left(\mathcal{E}_s^{2p} X_s^{-2(1-\alpha)}\right) = C\mathbb{E}\left[\Gamma_s \mathcal{E}_s^{2p} X_s^{-2(1-\alpha)} \Gamma_s^{-1}\right]$,

Applying Cauchy-Schwartz inequality:

$$\mathbb{E}\left[\Gamma_s \mathcal{E}_s^{2p} \frac{1}{X_s^{2(1-\alpha)}} \Gamma_s^{-1}\right] \leq \left[\mathbb{E}\left(\Gamma_s^2 \mathcal{E}_s^{4p}\right)\right]^{\frac{1}{2}} \times \left[\mathbb{E}\left(\frac{1}{X_s^{4(1-\alpha)}} \Gamma_s^{-2}\right)\right]^{\frac{1}{2}}.$$

Choice of a good weight process $(\Gamma_t, 0 \leq t \leq T)$.

The weighted error

Introduce

$$\beta_t = 2p\|b'\|_\infty + p(4p-1) + \frac{4\alpha^2(1+\delta^2)p(4p-1)\sigma^2}{X_t^{2(1-\alpha)}},$$

and the Weight Process $(\Gamma_t, 0 \le t \le T)$ defined by $\Gamma_t = \exp\left(-\int_0^t \beta_s ds\right)$.

Lemma Weighted Error

For $p \ge 1$ and $\alpha \in [\frac{1}{2}, 1)$, for all $\Delta t \le \Delta_{\max}(\alpha)$

$$\sup_{0 \le t \le T} \mathbb{E}\left(\Gamma_t^2 \mathscr{E}_t^{4p}\right) \le C\Delta t^{4p}. \tag{6}$$

$$\begin{aligned}
\mathbb{E}\left(\Gamma_t^2 \mathscr{E}_t^{4p}\right) \le\ & 4p\int_0^t \mathbb{E}\left(\Gamma_s^2 \mathscr{E}_s^{4p-1}\left[b(X_{\eta(s)}) - b(X_s)\right]\right) ds \\
& + 4p\|b'\|_\infty \int_0^t \mathbb{E}\left(\Gamma_{\eta(s)}^2 \mathscr{E}_{\eta(s)}^{4p}\right) ds + 4p\|b'\|_\infty \int_0^t \mathbb{E}\left(\Gamma_s^2 \mathscr{E}_s^{4p}\right) ds \\
& + 2(1+\delta^2)p(4p-1)\mathbb{E}\left(\int_0^t \Gamma_s^2 \mathscr{E}_s^{4p-2}\left[\sigma X_s^\alpha - \sigma \overline{X}_s^\alpha\right]^2 ds\right) + 2p(4p-1)\mathbb{E}\left(\int_0^t \Gamma_s^2 \mathscr{E}_s^{4p} ds\right) \\
& + 2p(4p-1)\int_0^t \mathbb{E}\left(D_s(\overline{X})^{4p}\right) ds - \mathbb{E}\left(\int_0^t 2\beta_s \Gamma_s^2 \mathscr{E}_s^{4p} ds\right) + \int_0^t \mathbb{E}\left(R_s \mathbb{1}_{\{\overline{Z}_s < 0\}}\right) ds + C\Delta t^{4p}.
\end{aligned}$$

Two concluding remarks

- A proof with 1D - arguments only in the exact process control of (negative) moments.

- A too strong constraint on $b(0)$, but generic smooth drift are allowed.