An adaptive splitting method for the Cox-Ingersoll-Ross process

Cónall Kelly[1] and Gabriel Lord[2]

[1]University College Cork, Ireland.

[2]Radboud University, The Netherlands.

Irish Numerical Analysis Forum,
Online Seminar,
Feb 9, 2023

Outline

Outline

Motivating example

ODE with non-globally Lipschitz coefficient:

$$x'(t) = -x^3(t), \quad t \geq 0.$$

$x(t) \equiv 0$ is globally asymptotically stable.

Explicit Euler discretisation:

$$X_{n+1} = X_n - hX_n^3, \quad n \in \mathbb{N}.$$

- $X_n \equiv 0$ locally asy. stable: $x_0 \in \left(-\sqrt{2/h}, \sqrt{2/h}\right)$,
- $\left\{-\sqrt{2/h}, \sqrt{2/h}\right\}$ an unstable 2-cycle.
- $\lim_{\to\infty} |X_n| = \infty$ iff $X_0 \in \left(-\infty, -\sqrt{2/h}\right) \cup \left(\sqrt{2/h}, \infty\right)$;
- Scheme converges on $[0, T]$ as $h \to 0$, but new dynamics for fixed $h > 0$.

Motivating example

ODE with non-globally Lipschitz coefficient:

$$x'(t) = -x^3(t), \quad t \geq 0.$$

$x(t) \equiv 0$ is globally asymptotically stable.

Explicit Euler discretisation:

$$X_{n+1} = X_n - hX_n^3, \quad n \in \mathbb{N}.$$

- $X_n \equiv 0$ locally asy. stable: $x_0 \in \left(-\sqrt{2/h}, \sqrt{2/h}\right)$,
- $\left\{-\sqrt{2/h}, \sqrt{2/h}\right\}$ an unstable 2-cycle.
- $\lim_{\to\infty} |X_n| = \infty$ iff $X_0 \in \left(-\infty, -\sqrt{2/h}\right) \cup \left(\sqrt{2/h}, \infty\right)$;
- Scheme converges on $[0, T]$ as $h \to 0$, but new dynamics for fixed $h > 0$.

Motivating example

ODE with non-globally Lipschitz coefficient:

$$x'(t) = -x^3(t), \quad t \geq 0.$$

$x(t) \equiv 0$ is globally asymptotically stable.

Explicit Euler discretisation:

$$X_{n+1} = X_n - hX_n^3, \quad n \in \mathbb{N}.$$

- $X_n \equiv 0$ locally asy. stable: $x_0 \in \left(-\sqrt{2/h}, \sqrt{2/h}\right)$,
- $\left\{-\sqrt{2/h}, \sqrt{2/h}\right\}$ an unstable 2-cycle.
- $\lim_{\to\infty} |X_n| = \infty$ iff $X_0 \in \left(-\infty, -\sqrt{2/h}\right) \cup \left(\sqrt{2/h}, \infty\right)$;
- Scheme converges on $[0, T]$ as $h \to 0$, but new dynamics for fixed $h > 0$.

Introducing a stochastic perturbation

Let B be a standard Brownian motion. The stochastic differential equation

$$dX(t) = -X(t)^3 + dB(t), \quad t \geq 0.$$

has Euler-Maruyama approximation

$$X_{n+1} = X_n - hX_n^3 + \underbrace{(B((n+1)h) - B(nh))}_{\mathcal{N}(0,h)}, \quad n \in \mathbb{N}.$$

The perturbation may push trajectories out of the basin of attraction $\left(-\sqrt{2/h}, \sqrt{2/h}\right)$. Problem!

Introducing a stochastic perturbation

Let B be a standard Brownian motion. The stochastic differential equation

$$dX(t) = -X(t)^3 + dB(t), \quad t \geq 0.$$

has Euler-Maruyama approximation

$$X_{n+1} = X_n - hX_n^3 + \underbrace{(B((n+1)h) - B(nh))}_{\mathcal{N}(0,h)}, \quad n \in \mathbb{N}.$$

The perturbation may push trajectories out of the basin of attraction $\left(-\sqrt{2/h}, \sqrt{2/h}\right)$. Problem!

Euler-Maruyama blows up asymptotically

On $[0, \infty)$, with fixed step size $h > 0$:

Mattingly, Stuart, Higham, 2002:

- Pathwise instability with positive probability:

$$\mathbb{P}[|X_n| \geq 2^n/\sqrt{h}, \text{ for all } n \in \mathbb{N}] > 0;$$

- Second moment instability:

$$\lim_{n\to\infty} \mathbb{E}[|X_n|^2] = \infty.$$

Milstein & Tretyakov, 2005:

- Modified scheme: discard trajectories that leave a sufficiently large sphere;
- Weak convergence on $[0, T]$ for non-globally Lipschitz coefficient equations.

Euler-Maruyama blows up asymptotically

On $[0, \infty)$, with fixed step size $h > 0$:

Mattingly, Stuart, Higham, 2002:

- Pathwise instability with positive probability:

$$\mathbb{P}[|X_n| \geq 2^n/\sqrt{h}, \text{ for all } n \in \mathbb{N}] > 0;$$

- Second moment instability:

$$\lim_{n\to\infty} \mathbb{E}[|X_n|^2] = \infty.$$

Milstein & Tretyakov, 2005:

- Modified scheme: discard trajectories that leave a sufficiently large sphere;
- Weak convergence on $[0, T]$ for non-globally Lipschitz coefficient equations.

Weak and strong convergence

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad t \geq 0.$$

Euler-Maruyama approximation:

$$X_{n+1} = X_n + hf(X_n) + g(X_n)(W_{(n+1)h} - W_{nh}), \quad n \in N.$$

- Weak convergence with order γ if there exists $C \in \mathbb{R}$ such that

$$\|\mathbb{E}[\bar{p}(X(T))] - \mathbb{E}[\bar{p}(X_N)]\| \leq Ch^{\gamma}$$

  for any sufficiently smooth $\bar{p}$.
- Strong convergence at time T with order β in $\mathcal{L}_p$ if there exists $p \in [1, \infty)$ and constants $C_p, \beta > 0$ such that

$$\left(\mathbb{E}\left[\|X(T) - X_N\|^p\right]\right)^{1/p} \leq C_p h^{\beta}.$$

Euler-Maruyama does not converge

On $[0, T]$, with decreasing step size $h = T/N$:

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad t \geq 0.$$

Euler-Maruyama approximation:

$$X_{n+1} = X_n + hf(X_n) + g(X_n)(W_{(n+1)h} - W_{nh}), \quad n \in N.$$

If either f or g

- are not globally Lipschitz continuous;
- and satisfy a polynomial growth condition,

No weak or strong-convergence (Hutzenthaler, Jentzen, Kloeden, 2011):

Euler-Maruyama does not converge

On $[0, T]$, with decreasing step size $h = T/N$:

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad t \geq 0.$$

Euler-Maruyama approximation:

$$X_{n+1} = X_n + hf(X_n) + g(X_n)(W_{(n+1)h} - W_{nh}), \quad n \in N.$$

If either f or g

- are not globally Lipschitz continuous;
- and satisfy a polynomial growth condition,

No weak or strong-convergence (Hutzenthaler, Jentzen, Kloeden, 2011):

References

J. C. Mattingly, A. M. Stuart, and D. J. Higham
Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise.
Stochastic Process. Appl., **101** (2002), pp. 185–232.

G. N. Milstein and M. N. Tretyakov
Numerical integration of stochastic differential equations with non globally Lipschitz coefficients.
SIAM J. Numer. Anal., **43** (2005), pp. 1139–1154.

M. Hutzenthaler, A. Jentzen, and P. E. Kloeden
Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients.
Proc. R. Soc. A, **467**:2130 (2011), pp. 1563–1576.

Outline

Example: control stability via local dynamics

Explicit Euler-Maruyama for $dX(t) = -X(t)^3 + dB(t)$:

$$X_{n+1} = X_n - hX_n^3 + \Delta W_{n+1}, \quad n \in \mathbb{N}.$$

Unperturbed equation: basin of attraction $(-\sqrt{2/h}, \sqrt{2/h})$.

Strategy: when trajectory escapes the basin of attraction, increase it by choosing h sufficiently small.

Suggests

$$h_{n+1} = \frac{c}{2|X_n|^2}, \quad c \in (0,1).$$

Example: control stability via local dynamics

Explicit Euler-Maruyama for $dX(t) = -X(t)^3 + dB(t)$:

$$X_{n+1} = X_n - hX_n^3 + \Delta W_{n+1}, \quad n \in \mathbb{N}.$$

Unperturbed equation: basin of attraction $(-\sqrt{2/h}, \sqrt{2/h})$.

Strategy: when trajectory escapes the basin of attraction, increase it by choosing h sufficiently small.

Suggests

$$h_{n+1} = \frac{c}{2|X_n|^2}, \quad c \in (0,1).$$

A square-root diffusion SDE

Cox, Ingersoll, and Ross, 1985

Linear drift and square-root diffusion:

$$dX(t) = \kappa(\theta - X(t))\,dt + \sigma\sqrt{X(t)}dW(t),\ \ t \in [0, T].$$

- Long run mean: $\theta > 0$;
- Speed of reversion: $\kappa > 0$;
- Stochastic intensity: $\sigma > 0$;
- Solutions a.s. nonnegative.

Theorem (Feller's condition)

Solutions are a.s. positive when $X(0) = X_0 > 0$ and

$$2\kappa\theta \geq \sigma^2.$$

A square-root diffusion SDE

Cox, Ingersoll, and Ross, 1985

Linear drift and square-root diffusion:

$$dX(t) = \kappa\,(\theta - X(t))\,dt + \sigma\sqrt{X(t)}dW(t), \ \ t \in [0, T].$$

- Long run mean: $\theta > 0$;
- Speed of reversion: $\kappa > 0$;
- Stochastic intensity: $\sigma > 0$;
- Solutions a.s. nonnegative.

Theorem (Feller's condition)

Solutions are a.s. positive when $X(0) = X_0 > 0$ and

$$2\kappa\theta \geq \sigma^2.$$

Why discretise the CIR model?

Applications in finance:

- Feller condition: interest rate modelling;
- Outside Feller condition: stochastic volatility modelling;

Exact sampling difficult:

- No analytic solution available in terms of W;
- $X(t)|X(s)$, $0 \leq s < t$: non-central chi-square distribution;
- Exact sampling for Monte Carlo estimation possible if W is uncorrelated, but numerically inefficient;
- Stochastic volatility models (e.g. Heston) use CIR in a vector SDE system with correlated noises.

The challenge for strong numerical approximation:

- Diffusion coefficeint $\sigma\sqrt{X}$: non-Lipschitz, even locally.

Why discretise the CIR model?

Applications in finance:

- Feller condition: interest rate modelling;
- Outside Feller condition: stochastic volatility modelling;

Exact sampling difficult:

- No analytic solution available in terms of W;
- $X(t)|X(s)$, $0 \leq s < t$: non-central chi-square distribution;
- Exact sampling for Monte Carlo estimation possible if W is uncorrelated, but numerically inefficient;
- Stochastic volatility models (e.g. Heston) use CIR in a vector SDE system with correlated noises.

The challenge for strong numerical approximation:

- Diffusion coefficeint $\sigma\sqrt{X}$: non-Lipschitz, even locally.

Why discretise the CIR model?

Applications in finance:

- Feller condition: interest rate modelling;
- Outside Feller condition: stochastic volatility modelling;

Exact sampling difficult:

- No analytic solution available in terms of W;
- $X(t)|X(s)$, $0 \leq s < t$: non-central chi-square distribution;
- Exact sampling for Monte Carlo estimation possible if W is uncorrelated, but numerically inefficient;
- Stochastic volatility models (e.g. Heston) use CIR in a vector SDE system with correlated noises.

The challenge for strong numerical approximation:

- Diffusion coefficeint $\sigma\sqrt{X}$: non-Lipschitz, even locally.

The Heston stochastic volatility model

The spot price of an asset satisfies

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)(\sqrt{1-\rho^2}dW^{(1)}(t) + \rho dW^{(2)}(t)),$$

with volatility process

$$dV(t) = \lambda V(t)(\mu - V(t))dt + \sigma\sqrt{V(t)}dW^{(2)}(t).$$

- $f(v) = \lambda\mu v - \lambda v^2$, $g(v) = b\sigma v^{1/2}$: non-Lipschitz;
- Explicit Euler-Maruyama method does not converge (strongly/weakly);
- Strong convergence required for multi-level Monte Carlo approximation of $\mathbb{E}[\Lambda(S)]$.

The Heston stochastic volatility model

The spot price of an asset satisfies

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)(\sqrt{1-\rho^2}dW^{(1)}(t) + \rho dW^{(2)}(t)),$$

with volatility process

$$dV(t) = \lambda V(t)(\mu - V(t))dt + \sigma\sqrt{V(t)}dW^{(2)}(t).$$

- $f(v) = \lambda\mu v - \lambda v^2$, $g(v) = b\sigma v^{1/2}$: non-Lipschitz;
- Explicit Euler-Maruyama method does not converge (strongly/weakly);
- Strong convergence required for multi-level Monte Carlo approximation of $\mathbb{E}[\Lambda(S)]$.

Sources of error

Estimate value of derivative with payoff Λ_T written on S with risk-free rate r via Monte Carlo approximation of

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\int_0^t r(s)ds}\Lambda_T\right]$$

The overall error associated with a sample size of M is

Sampling error $+$ Numerical approximation error.

- Sampling error depends on M (and is $O(1/\sqrt{M})$);
- Numerical approximation error depends on N. Order of strong/weak convergence.
- Variance reduction (e.g. MLMC) may require strong convergence.

Sources of error

Estimate value of derivative with payoff Λ_T written on S with risk-free rate r via Monte Carlo approximation of

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\int_0^t r(s)ds}\Lambda_T\right]$$

The overall error associated with a sample size of M is

Sampling error $+$ Numerical approximation error.

- Sampling error depends on M (and is $O(1/\sqrt{M})$);
- Numerical approximation error depends on N. Order of strong/weak convergence.
- Variance reduction (e.g. MLMC) may require strong convergence.

Approach 1: Direct approximation of CIR

Fully truncated Euler method

$$\widetilde{X}_{n+1} = \widetilde{X}_n + h\kappa(\theta - \max\{\widetilde{X}_n, 0\}) + \sigma\sqrt{\max\{\widetilde{X}_n, 0\}}\Delta W_{n+1}$$
$$X_{n+1} = \max\{\widetilde{X}_{n+1}, 0\}.$$

- Maintains non-negativity by truncating at zero.
- Widely used in practice.
- Lord, Koekkoek, and Van Dijk (2010): Strong L_1 convergence, no rate.
- Cozma & Reisinger (2020): Strong L_p convergence, rate $1/2$ when

$$2\kappa\theta > 3\sigma^2 \quad \text{for} \quad 2 \leq p < \frac{2\kappa\theta}{\sigma^2} - 1.$$

Approach 1: Direct approximation of CIR

Fully truncated Euler method

$$\widetilde{X}_{n+1} = \widetilde{X}_n + h\kappa(\theta - \max\{\widetilde{X}_n, 0\}) + \sigma\sqrt{\max\{\widetilde{X}_n, 0\}}\Delta W_{n+1}$$
$$X_{n+1} = \max\{\widetilde{X}_{n+1}, 0\}.$$

- Maintains non-negativity by truncating at zero.
- Widely used in practice.
- Lord, Koekkoek, and Van Dijk (2010): Strong L_1 convergence, no rate.
- Cozma & Reisinger (2020): Strong L_p convergence, rate $1/2$ when

$$2\kappa\theta > 3\sigma^2 \quad \text{for} \quad 2 \leq p < \frac{2\kappa\theta}{\sigma^2} - 1.$$

Approach 2: Approximation of transformed CIR

Drift-implicit square-root Euler method

$$dX(t) = \kappa\left(\theta - X(t)\right)dt + \sigma\sqrt{X(t)}dW(t),\ t \in [0, T].$$

- Apply the transform $Y = \sqrt{X}$ to get

$$dY(t) = \left(\frac{\alpha}{Y(t)} + \beta Y(t)\right)dt + \gamma dW_t,\ t \in [0, T],$$

where $\alpha = (4\kappa\theta - \sigma^2)/8$, $\beta = (-4\kappa/8)$, $\gamma = \sigma/2$.

- Drift $f(y) = \alpha/y + \beta y$ is not globally Lipschitz continuous. Satisfies a one-sided Lipschitz condition:

$$\langle f(x) - f(y), x - y\rangle \le C|x - y|^2,\ \text{for all } x, y \in \mathbb{R}^+.$$

- Diffusion $g(y) = \gamma$ is constant.

Approach 2: Approximation of transformed CIR

Drift-implicit square-root Euler method

$$dX(t) = \kappa(\theta - X(t))\,dt + \sigma\sqrt{X(t)}dW(t), \ \ t \in [0, T].$$

- Apply the transform $Y = \sqrt{X}$ to get

$$dY(t) = \left(\frac{\alpha}{Y(t)} + \beta Y(t)\right) dt + \gamma dW_t, \ \ t \in [0, T],$$

  where $\alpha = (4\kappa\theta - \sigma^2)/8$, $\beta = (-4\kappa/8)$, $\gamma = \sigma/2$.
- Drift $f(y) = \alpha/y + \beta y$ is not globally Lipschitz continuous. Satisfies a one-sided Lipschitz condition:

$$\langle f(x) - f(y), x - y\rangle \leq C|x - y|^2, \ \ \text{for all } x, y \in \mathbb{R}^+.$$

- Diffusion $g(y) = \gamma$ is constant.

Approach 2: Approximation of transformed CIR

Drift-implicit square-root Euler method

With a uniform step size $h > 0$,

$$Y_{k+1} = Y_k + \left(\frac{\alpha}{Y_{k+1}} + \beta Y_{k+1} \right) h + \gamma \triangle W_{k+1}, \quad k \in \mathbb{N}$$

has unique positive solution

$$Y_{k+1} = \frac{Y_k + \gamma \triangle W_{n+1}}{2(1-\beta h)} + \sqrt{\frac{(Y_k + \gamma \triangle W_{n+1})^2}{4(1-\beta h)^2} + \frac{\alpha h}{1-\beta h}}, \quad k \in \mathbb{N}.$$

- Initial value $Y_0 = \sqrt{X(0)}$;
- Invert transform: $X_k = Y_k^2$;

Approach 2: Approximation of transformed CIR

Drift-implicit square-root Euler method

With a uniform step size $h > 0$,

$$Y_{k+1} = Y_k + \left(\frac{\alpha}{Y_{k+1}} + \beta Y_{k+1} \right) h + \gamma \triangle W_{k+1}, \quad k \in \mathbb{N}$$

has unique positive solution

$$Y_{k+1} = \frac{Y_k + \gamma \triangle W_{n+1}}{2(1 - \beta h)} + \sqrt{\frac{(Y_k + \gamma \triangle W_{n+1})^2}{4(1 - \beta h)^2} + \frac{\alpha h}{1 - \beta h}}, \quad k \in \mathbb{N}.$$

- Initial value $Y_0 = \sqrt{X(0)}$;
- Invert transform: $X_k = Y_k^2$;

Approach 2: Approximation of transformed CIR

Drift-implicit square-root Euler method

Theorem

Let $2\kappa\lambda > \sigma^2$*. For all*

$$1 \leq p < \frac{4\kappa\lambda}{3\sigma^2}$$

there exists a constant $K_p > 0$ *such that*

$$\left(\mathbb{E} \left[\max_{t \in [0,T]} |X(t) - \widetilde{X}_t|^p \right] \right)^{1/p} \leq K_p \cdot h$$

where $\widetilde{X}_t$ *is continuous-time extension of* X_k*.*

- $\mathcal{L}_2$-convergence if $2\kappa\lambda > 3\sigma^2$;
- Alphonsi (2005, 2013);
- See also Dereich, Neuenkirch, and Szpruch (2012)

Approach 2: Approximation of transformed CIR

Drift-implicit square-root Euler method

Theorem

Let $2\kappa\lambda > \sigma^2$. For all

$$1 \leq p < \frac{4\kappa\lambda}{3\sigma^2}$$

there exists a constant $K_p > 0$ such that

$$\left(\mathbb{E}\left[\max_{t\in[0,T]} |X(t) - \widetilde{X}_t|^p\right]\right)^{1/p} \leq K_p \cdot h$$

where $\widetilde{X}_t$ is continuous-time extension of X_k.

- $\mathcal{L}_2$-convergence if $2\kappa\lambda > 3\sigma^2$;
- Alphonsi (2005, 2013);
- See also Dereich, Neuenkirch, and Szpruch (2012)

Approach 2: Approximation of transformed CIR

Projected square-root Euler method

$$
\begin{aligned}
Y_{n+1} &= \hat{Y}_n + \left(\frac{\alpha}{\hat{Y}_n} - \beta \hat{Y}_n\right) h + \gamma \Delta W_{n+1}; \\
\hat{Y}_n &= \max\left\{h^{1/4}, Y_n\right\}
\end{aligned}
$$

Explicit scheme. Small values projected back up.

- $\kappa\theta > 5\sigma^2/2$: L_1 convergence of order 1;
- $\kappa\theta > 3\sigma^2/2$: L_1 convergence of order $1/2$;
- $\kappa\theta > \sigma^2$: L_1 convergence of order

$$
\max\left\{\frac{1}{6}, \frac{1}{2} - \frac{\sigma^2}{2\kappa\theta + \sigma^2}\right\}.
$$

- Chassagneux, Jacquier, and Mihaylov (2016).

Approach 2: Approximation of transformed CIR

Projected square-root Euler method

$$
\begin{aligned}
Y_{n+1} &= \hat{Y}_n + \left(\frac{\alpha}{\hat{Y}_n} - \beta \hat{Y}_n \right) h + \gamma \Delta W_{n+1}; \\
\hat{Y}_n &= \max \left\{ h^{1/4}, Y_n \right\}
\end{aligned}
$$

Explicit scheme. Small values projected back up.

- $\kappa\theta > 5\sigma^2/2$: L_1 convergence of order 1;
- $\kappa\theta > 3\sigma^2/2$: L_1 convergence of order $1/2$;
- $\kappa\theta > \sigma^2$: L_1 convergence of order

$$
\max \left\{ \frac{1}{6}, \frac{1}{2} - \frac{\sigma^2}{2\kappa\theta + \sigma^2} \right\}.
$$

- Chassagneux, Jacquier, and Mihaylov (2016).

Approach 3: Truncated Milstein-type methods

Hefter & Hurzwurm (2018)

$$
\begin{aligned}
R_1 &= \max\left\{\sigma\sqrt{\Delta t}/2, \sqrt{\max\{\sigma^2\Delta t/4, X_n\} + \frac{\sigma}{2}\Delta W_{n+1}}\right\} \\
X_{n+1} &= \max\left\{R_1^2 + \Delta t(\kappa\theta - \frac{\sigma^2}{4} - \kappa X_2), 0\right\}
\end{aligned}
$$

- L_p-convergence across all parameter values with rate

$$
\min\left\{\frac{1}{2p}, \frac{2\kappa\theta}{p\sigma^2}\right\} - \varepsilon.
$$

- For $p = 2$, this gives rate (any fixed $\varepsilon > 0$)
  - $1/4 - \varepsilon$ when $\kappa\theta \geq 4\sigma^2$.
  - $\kappa\theta/\sigma^2 - \varepsilon$ when $\kappa\theta < 4\sigma^2$.

Some schemes with known convergence rates

Scheme	Norm	Parameter Range	Rate
Truncated Milstein (2018)	L_p	no restriction	$\frac{1}{2p} \wedge \frac{2\kappa\theta}{p\sigma^2} - \epsilon$
Drift Implicit Square-Root Euler (2013)	L_p $p \in [1, \frac{4\kappa\theta}{3\sigma^2})$	$\kappa\theta > (1 \vee \frac{3}{4}p)\sigma^2$	1
Projected Euler (2016)	L_1	$\kappa\theta > \frac{5}{2}\sigma^2$ $\kappa\theta > \frac{3}{2}\sigma^2$ $\kappa\theta > \sigma^2$	1 $\frac{1}{2}$ $\frac{1}{6} \vee (\frac{1}{2} - \frac{\sigma^2}{2\kappa\theta+\sigma^2})$
Fully Truncated (2020)	L_p $p \in [2, \frac{2\kappa\theta}{\sigma^2} - 1)$	$\kappa\theta > \frac{3}{2}\sigma^2$	1/2
Symmetrized Milstein (2018)	L_p $p \geq 1$	$\kappa\theta > \frac{3}{2}\sigma^2(2[p \vee 2] + 1)$	1

References

A. Alfonsi,
Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process.
Statistics and Probability Letters, 83, pp. 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.
SIAM J. Financial Math. 7(1), pp. 993–1021, 2016

A. Cozma and C. Reisinger,
Strong order $1/2$ convergence of full truncation Euler approximations to the Cox-Ingersoll-Ross process.
IMA J. Numer. Anal. 40(1), pp. 358–376, 2020.

S. Dereich, A. Neuenkirch, and L. Szpruch,
An Euler-Type Method for The Strong Approximation of The Cox-Ingersoll-Ross Process,
Prof. R. Soc. Lond. Ser. A. Math. Phys. Eng. Sci. 468(2140), pp. 1105-1115, 2011.

R. Lord, R. Koekkoek, and D. van Dijk,
A comparison of biased simulation schemes for stochastic volatility models,
Quant. Finance, 10(2), pp. 177–194, 2010.

Construction of a splitting method

The transformed SDE

$$dY(t) = \left(\frac{\alpha}{Y(t)} - \beta Y(t)\right) dt + \gamma dW(t), \quad t \in [t_n, t_{n+1}].$$

has variation of constants form

$$Y(t) = e^{-\beta(t-t_n)} Y(t_n) + \int_{t_n}^{t} e^{-\beta(t-s)} \frac{\alpha}{Y(s)} ds + \gamma \int_{t_n}^{t} e^{-\beta(t-s)} dW(s),$$

The integral equation

$$z(t) = z(t_n) + \int_{t_n}^{t} \frac{\alpha}{z(s)} ds, \quad t \in [t_n, t_{n+1}],$$

has solution

$$z(t) = \sqrt{z(t_n)^2 + 2\alpha(t - t_n)}, \quad t \in [t_n, t_{n+1}].$$

Construction of a splitting method

The transformed SDE

$$dY(t) = \left(\frac{\alpha}{Y(t)} - \beta Y(t) \right) dt + \gamma dW(t), \quad t \in [t_n, t_{n+1}].$$

has variation of constants form

$$Y(t) = e^{-\beta(t-t_n)} Y(t_n) + \int_{t_n}^{t} e^{-\beta(t-s)} \frac{\alpha}{Y(s)} ds + \gamma \int_{t_n}^{t} e^{-\beta(t-s)} dW(s),$$

The integral equation

$$z(t) = z(t_n) + \int_{t_n}^{t} \frac{\alpha}{z(s)} ds, \quad t \in [t_n, t_{n+1}],$$

has solution

$$z(t) = \sqrt{z(t_n)^2 + 2\alpha(t - t_n)}, \quad t \in [t_n, t_{n+1}].$$

Our method is equivalent to the Lie-Trotter composition of the exact flows of the following subequations

$$dY^{[1]}(t) = \alpha \left(Y^{[1]}(t) \right)^{-1} dt, \quad Y^{[1]}(0) = Y_0^{[1]}$$
$$dY^{[2]}(t) = \gamma dW(t), \quad Y^{[2]}(0) = Y_0^{[2]};$$
$$dY^{[3]}(t) = -\beta Y^{[3]}(t) dt, \quad Y^{[3]}(0) = Y_0^{[3]}.$$

Construction of the splitting method

We propose the splitting approximation for Y:

$$Y_{n+1} = e^{-\beta \Delta t_{n+1}} \left(\sqrt{(Y_n)^2 + 2\alpha \Delta t_{n+1}} + \gamma \Delta W_{n+1} \right)$$

The CIR approximation preserves a.s. non-negativity:

$$X_{n+1} = (Y_{n+1})^2.$$

Can this method work outside the Feller region?

- The SDE for Y fails at $Y = 0$: soft zero.
- If $\alpha < 0$ (i.e. $4\kappa\theta < \sigma^2$) then must adapt the stepsize:

$$\Delta t_n < \frac{Y_n^2}{2|\alpha|}.$$

Construction of the splitting method

We propose the splitting approximation for Y:

$$Y_{n+1} = e^{-\beta \Delta t_{n+1}} \left(\sqrt{(Y_n)^2 + 2\alpha \Delta t_{n+1}} + \gamma \Delta W_{n+1} \right)$$

The CIR approximation preserves a.s. non-negativity:

$$X_{n+1} = (Y_{n+1})^2.$$

Can this method work outside the Feller region?

- The SDE for Y fails at $Y = 0$: soft zero.
- If $\alpha < 0$ (i.e. $4\kappa\theta < \sigma^2$) then must adapt the stepsize:

$$\Delta t_n < \frac{Y_n^2}{2|\alpha|}.$$

Adaptivity in the mesh

- Consider a mesh $\{0 =: t_0, t_1, \ldots, t_N := T\} \subset [0, T]$.
- Set $\Delta t_{n+1} := t_{n+1} - t_n$.
- $(\mathcal{F}_t)_{t \geq 0}$, the natural filtration of W, can be extended to any $\mathcal{F}_t$-stopping time τ by

$$\mathcal{F}_\tau := \{B \in \mathcal{F} \,:\, B \cap \{\tau \leq t\} \in \mathcal{F}_t\}.$$

- If each t_n is an $\mathcal{F}_t$-stopping time, this allows us to condition on $\mathcal{F}_{t_n}$ at any point on the random time-set $\{t_n\}_{n \in \mathbb{N}}$.

We need

- Δt_{n+1} to be $\mathcal{F}_{t_n}$-measurable and $N < \infty$ a.s.
- Maximum stepsize $\Delta t_{\max}$. No minimum stepsize.

Adaptivity in the mesh

- Consider a mesh $\{0 =: t_0, t_1, \dots, t_N := T\} \subset [0, T]$.
- Set $\Delta t_{n+1} := t_{n+1} - t_n$.
- $(\mathcal{F}_t)_{t\geq 0}$, the natural filtration of W, can be extended to any $\mathcal{F}_t$-stopping time τ by

$$\mathcal{F}_\tau := \{B \in \mathcal{F} \, : \, B \cap \{\tau \leq t\} \in \mathcal{F}_t\}.$$

- If each t_n is an $\mathcal{F}_t$-stopping time, this allows us to condition on $\mathcal{F}_{t_n}$ at any point on the random time-set $\{t_n\}_{n\in\mathbb{N}}$.

We need

- Δt_{n+1} to be $\mathcal{F}_{t_n}$-measurable and $N < \infty$ a.s.
- Maximum stepsize $\Delta t_{\max}$. No minimum stepsize.

Strong convergence

Suppose $\kappa\theta > \sigma^2$ so that Feller's condition ($2\kappa\theta > \sigma^2$) holds.

On adaptive mesh we can prove successively:

1. Uniform moment bound for scheme interpolant;
2. Strong bound in L_1 for error of cns extension of scheme;
3. Strong convergence in L_2 for linear interpolant of error.

Definition (Error interpolant)

$E_n := X(t_n) - X_n$, and define $(\mathcal{E}^2(t))_{t\in[0,T]}$ pathwise as

$$\mathcal{E}^2(s) := \frac{t_{n+1} - s}{\Delta t_{n+1}} E_n^2 + \frac{s - t_n}{\Delta t_{n+1}} E_{n+1}^2, \quad s \in [t_n, t_{n+1}], \quad a.s.$$

Strong convergence

Suppose $\kappa\theta > \sigma^2$ so that Feller's condition ($2\kappa\theta > \sigma^2$) holds.

On adaptive mesh we can prove successively:

1. Uniform moment bound for scheme interpolant;
2. Strong bound in L_1 for error of cns extension of scheme;
3. Strong convergence in L_2 for linear interpolant of error.

Definition (Error interpolant)

$E_n := X(t_n) - X_n$, and define $(\mathcal{E}^2(t))_{t\in[0,T]}$ pathwise as

$$\mathcal{E}^2(s) := \frac{t_{n+1} - s}{\Delta t_{n+1}} E_n^2 + \frac{s - t_n}{\Delta t_{n+1}} E_{n+1}^2, \quad s \in [t_n, t_{n+1}], \quad a.s.$$

Main result on strong convergence

Theorem

Suppose $\kappa\theta > \sigma^2$ and $\{t_0, t_1, \dots, t_N\}$ are selected so that

$$\max_n \Delta t_n \leq \Delta t_{max} < \min\left\{1, \frac{1}{\kappa}, \frac{1}{4\kappa|1-\kappa| + \theta\kappa^2}\right\}.$$

Then there exists a constant $C < \infty$ such that

$$\max_{t\in[0,T]} \mathbb{E}[\mathcal{E}^2(t)] \leq C\Delta t_{max}^{1/2}.$$

- Order $1/4$ strong L_2-convergence over a uniform mesh.
- Can reduce C if we choose (e.g.)

$$\Delta t_{n+1} = \frac{\Delta t_{\max}}{1 + 3\exp(-150 X_n)}.$$

Main result on strong convergence

Theorem

Suppose $\kappa\theta > \sigma^2$ *and* $\{t_0, t_1, \ldots, t_N\}$ *are selected so that*

$$\max_n \Delta t_n \leq \Delta t_{max} < \min\left\{1, \frac{1}{\kappa}, \frac{1}{4\kappa|1-\kappa| + \theta\kappa^2}\right\}.$$

Then there exists a constant $C < \infty$ *such that*

$$\max_{t\in[0,T]} \mathbb{E}[\mathcal{E}^2(t)] \leq C\Delta t_{max}^{1/2}.$$

- Order $1/4$ strong L_2-convergence over a uniform mesh.
- Can reduce C if we choose (e.g.)

$$\Delta t_{n+1} = \frac{\Delta t_{\max}}{1 + 3\exp(-150X_n)}.$$

Extension of scheme outside Feller's region

1. When $X_n < X_{\text{zero}}$, use splitting scheme with $\Delta t_{n+1} = \min\left\{0.95\frac{X_n}{2|\alpha|}, \Delta t_{\max}\right\}$.
2. When $X_n < X_{\text{zero}}$ switch off noise: $u'(t) = \kappa(\theta - u(t))$ has solution
$$u(t) = e^{-\kappa(t-t_n)}u(t_n) + \theta\left(1 - e^{-\kappa(t-t_n)}\right), \quad t \geq t_n.$$
3. Set $X_{\text{zero}} := \frac{1}{2}u(t_n + \Delta t_{\max})|_{u(t_n)=0}$, so
$$X_{\text{zero}} := \theta(1 - e^{-\kappa\Delta t_{\max}})/2.$$
As $\Delta t_{\max} \to 0$, $X_{\text{zero}} \to 0$.
4. When $X_n < X_{\text{zero}}$ we use the timestep
$$\Delta t_{n+1} = -\frac{1}{\kappa}\log\left(\frac{X_{\text{zero}} - \theta}{X_n - \theta}\right)$$
so that $X_{n+1} = X_{\text{zero}}$.
5. Then turn noise back on. Local error is preserved.

Extension of scheme outside Feller's region

1. When $X_n < X_{zero}$, use splitting scheme with $\Delta t_{n+1} = \min\left\{0.95\frac{X_n}{2|\alpha|}, \Delta t_{max}\right\}$.
2. When $X_n < X_{zero}$ switch off noise: $u'(t) = \kappa(\theta - u(t))$ has solution
$$u(t) = e^{-\kappa(t-t_n)}u(t_n) + \theta\left(1 - e^{-\kappa(t-t_n)}\right), \quad t \geq t_n.$$
3. Set $X_{zero} := \frac{1}{2}u(t_n + \Delta t_{max})|_{u(t_n)=0}$, so
$$X_{zero} := \theta(1 - e^{-\kappa\Delta t_{max}})/2.$$
As $\Delta t_{max} \to 0$, $X_{zero} \to 0$.
4. When $X_n < X_{zero}$ we use the timestep
$$\Delta t_{n+1} = -\frac{1}{\kappa}\log\left(\frac{X_{zero} - \theta}{X_n - \theta}\right)$$
so that $X_{n+1} = X_{zero}$.
5. Then turn noise back on. Local error is preserved.

Extension of scheme outside Feller's region

1. When $X_n < X_{\text{zero}}$, use splitting scheme with $\Delta t_{n+1} = \min\left\{0.95\frac{X_n}{2|\alpha|}, \Delta t_{\max}\right\}$.
2. When $X_n < X_{\text{zero}}$ switch off noise: $u'(t) = \kappa(\theta - u(t))$ has solution
$$u(t) = e^{-\kappa(t-t_n)}u(t_n) + \theta\left(1 - e^{-\kappa(t-t_n)}\right), \quad t \geq t_n.$$
3. Set $X_{\text{zero}} := \frac{1}{2}u(t_n + \Delta t_{\max})|_{u(t_n)=0}$, so
$$X_{\text{zero}} := \theta(1 - e^{-\kappa\Delta t_{\max}})/2.$$
As $\Delta t_{\max} \to 0$, $X_{\text{zero}} \to 0$.
4. When $X_n < X_{\text{zero}}$ we use the timestep
$$\Delta t_{n+1} = -\frac{1}{\kappa}\log\left(\frac{X_{\text{zero}} - \theta}{X_n - \theta}\right)$$
so that $X_{n+1} = X_{\text{zero}}$.
5. Then turn noise back on. Local error is preserved.

A numerical comparison of methods

We compare the numerical performance of the adaptive splitting method to

- Truncated Milstein
- Fully truncated
- Drift-implicit square-root Euler
- Projected square-root Euler

outside of the limits of our strong convergence theorem.

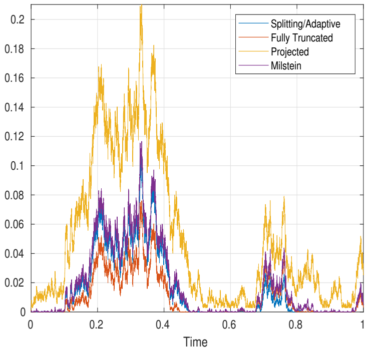
Truncated Milstein known to converge strongly across all parameter values: will be used as the reference solution.

Key σ values when $\kappa = 2$, $\theta = 0.02$

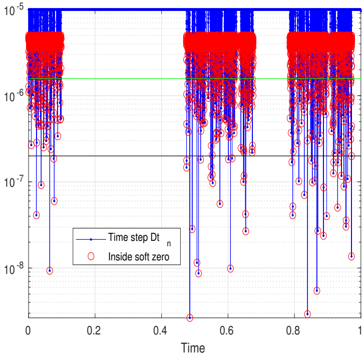
Description	σ	Value
Projected Euler limit rate $1/2$ Limit of theory for Truncated Euler (rate $1/2$)	$\sigma = \sqrt{2\kappa\theta/3}$	≈ 0.1633
Limit of theory for: - Splitting (rate $1/4$, $p = 1, 2$) - Drift Implicit (rate 1, $p < 2$) - Projected Euler (rate $1/6$, $p = 1$)	$\sigma = \sqrt{\kappa\theta}$	0.2
Feller boundary	$\sigma = \sqrt{2\kappa\theta}$	≈ 0.2828
$\alpha \leq 0$. Adaptivity/Soft-Zero required for splitting method	$\sigma \geq 2\sqrt{\kappa\theta}$	0.4

Sample paths and adaptive steps: fine mesh

$$\kappa = 2,\ \theta = 0.02,\ \sigma = 0.8,\ \Delta t_{\max} = 10^{-5}$$



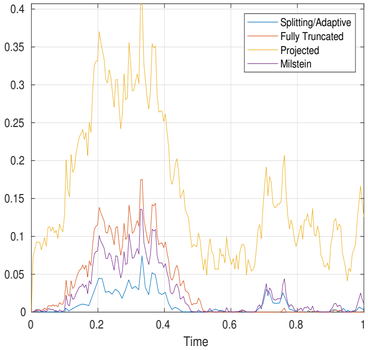
(a) Sample paths



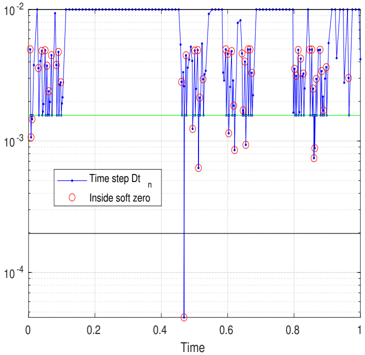
(b) Adaptive timesteps

Sample paths and adaptive steps: coarse mesh

$\kappa = 2,\ \theta = 0.02,\ \sigma = 0.8,\ \Delta t_{\max} = 10^{-2}$



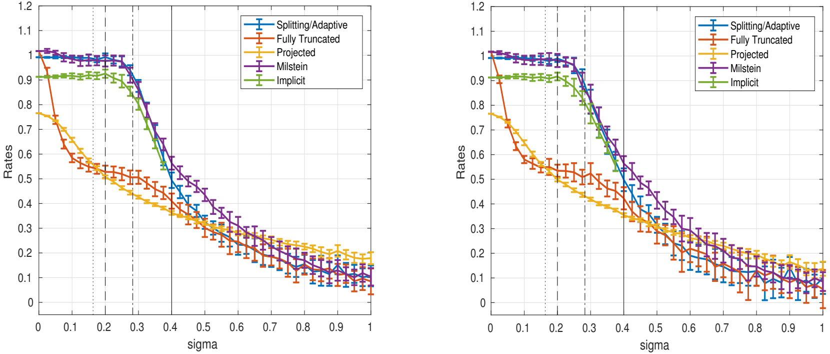
(a) Sample paths



(b) Adaptive timesteps

Decay in rates of convergence as σ increases

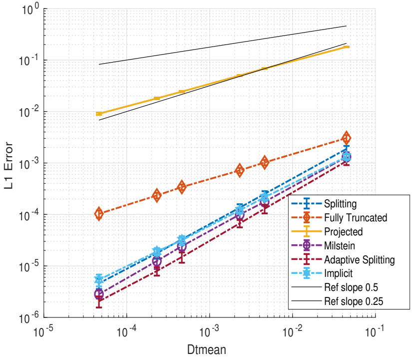
$\kappa = 2$, $\theta = 0.02$, $\sigma \in [0, 1]$. 20 groups of $M = 50$ samples.
Feller ends at $\sigma \approx 0.28$. $\alpha < 0$ when $\sigma > 0.4$.



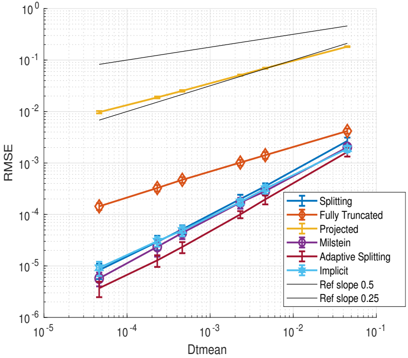
(a) L_1 error (b) L_2 error

Convergence plots for L_1 and L_2 strong error

$\kappa = 2$, $\theta = 0.02$, $\sigma = 0.3$, $M = 1000$.



(a) L_1 error



(b) L_2 error

CK and G. J. Lord, *An adaptive splitting method for the Cox-Ingersoll-Ross process*, Applied Numerical Mathematics **186** (2023), pp. 252–273.

- We propose a numerical method for CIR based upon
  - a square-root transformation;
  - variation of constants solution of the linear drift part;
  - exact solution of the nonlinear drift part.
- Strong $L_{1,2}$-convergence can be shown for small noise;
- Error constants can be reduced by use of an adaptive mesh.
- Large noise case: extend by use of a (different) adaptive mesh and the introduction of a “soft zero” region to capture deterministic dynamics close to zero.
- Numerical results competitive in large noise case, theoretical results needed.