Short-term Interest Rate Modeling

1 Introduction

Short-term interest rates are used to measure the yield of short-term borrowings taking place in the market. By analyzing the behavior of the interest rate curve, we can predict its further changes and get information about the behavior of financial markets. There are a lot of models that are used for short-term interest rates modeling. One of the most effective and popular, CIR process, had been introduced in the paper "A Theory of the Term Structure of Interest Rates" (1985) by John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross. The CIR process is defined by:

$$\begin{cases} dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}dW_t \\ X_0 = x_0 \end{cases} \tag{1}$$

where $x_0, a, b, \sigma \in \mathbb{R}$, $\sigma > 0$. We may interpret b as a mean value of the process, a as a speed of convergence to the mean value and σ as a volatility.

The aim of the project is to simulate this process, using different numerical schemes and compare them by speed of convergence and computational complexity. The simplest way to simulate a stochastic differential equation is to use Euler-Maruyama method, that comes from Euler method of simulation deterministic ordinary differential equation. We will introduce the scheme based on "An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations" (2001) by D. J. Higham and try to use it for CIR-process simulation. In order to explain how the scheme works, first of all let's write down the form of stochastic differential equation that we will simulate:

$$dX_t = f(X_t)dt + g(X_t)dW_t,\ X_0 = x_0,\ 0 \le t \le T, \tag{2}$$

where x_0 is a random variable and f and g are scalar functions.

First of all we should choose positive integer L and split interval $[0;T]$ into L equal parts, so each part Δt will be the size of $\frac{T}{L}$. Let $\tau_i = i\Delta t$, now for the discrete timesteps we can say, using (1), that:

$$X_{\tau_i} = X_{\tau_{i-1}} + \int_{\tau_{i-1}}^{\tau_j} f(X_s)ds + \int_{\tau_{i-1}}^{\tau_j} g(X_s)dW_s. \tag{3}$$

For $\tau_j - \tau_{j-1}$ small enough we may consider

$$\int_{\tau_{i-1}}^{\tau_j} g(X_s)dW_s \sim g(X_{\tau_{j-1}})(W_{\tau_j} - W_{\tau_{j-1}}), \tag{4}$$

so finally we get the sequence of random variables $X_1, X_2, ..., X_L$, where:

$$\hat{X}_i = \hat{X}_{i-1} + f(\hat{X}_{i-1})\frac{T}{L} + g(\hat{X}_{i-1})(W_i - W_{i-1}). \tag{5}$$

We will call this sequence a discrete EM approximation of initial process X_t.

After getting the scheme we can test it empirically: let's take Black-Scholes model:

$$\begin{cases} dX_t = \lambda X_t dt + \mu X_t dW_t \\ X_0 = x_0 \end{cases} \tag{6}$$

for which we know the exact solution:

$$X_t = x_0 e^{((\lambda - \frac{\mu^2}{2})t + \mu W_t)}$$

and simulate it with EM method.

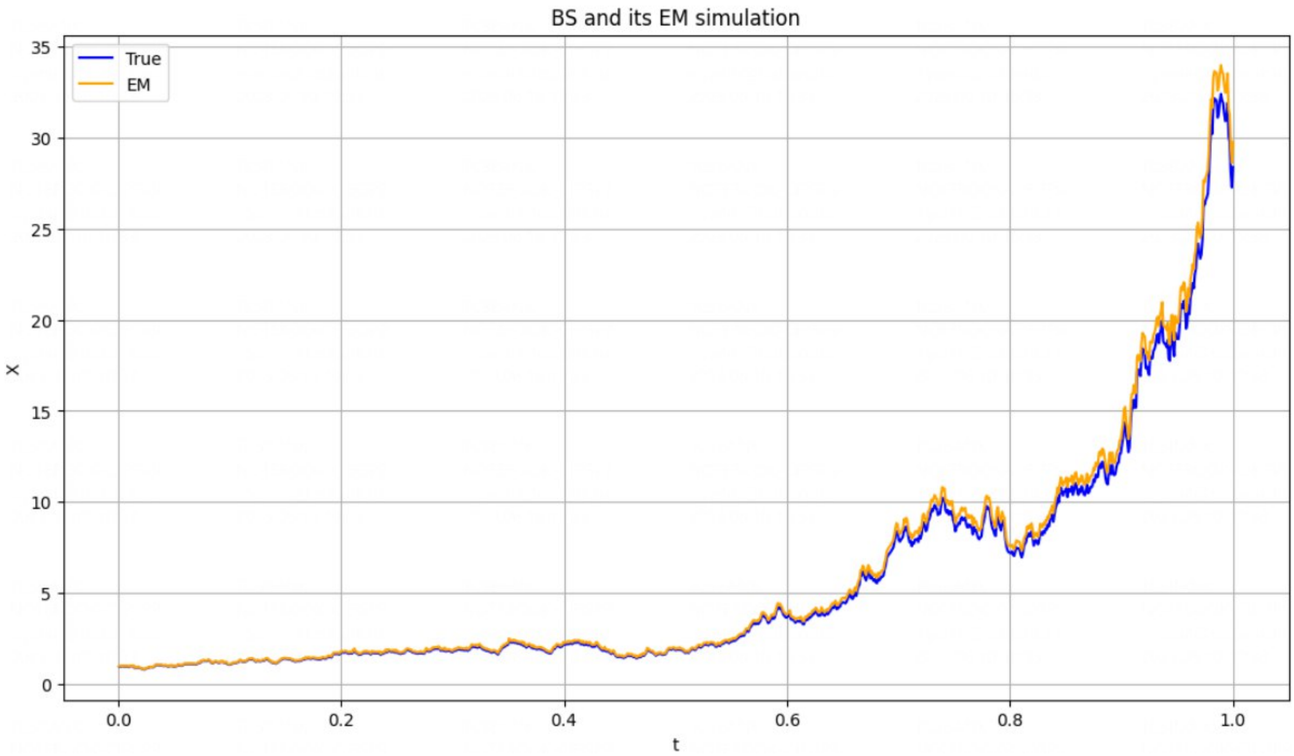

Figure 1: Black-Scholes model and EM simulation

Now we can visually check, that EM simulation coincides with the true solution of an SDE, but we have to specify strict criteria of similarity between two stochastic processes. Further in the work we will use strong convergence to check if the initial process and simulation converge to each other and if so, at what speed.

Definition 1.1 (Strong convergence)**.** Numerical scheme $\hat{X}_\tau$ strongly converge to initial process X_t if there exists constants C and γ such that

$$\sup_{n \in [0,L]} \mathbb{E}|X_t - \hat{X}_{n\Delta t}| \leq C(\Delta t)^\gamma$$

We will call constant γ a rate of convergence.

The intuition behind this definition is quite simple: we should calculate expected difference between initial process and its simulation at all points where our scheme exists and take the maximum. The expected value of $|X_t - \hat{X}_{n\Delta t}|$ will be the function of timestep Δt. We must try to bound this function with monomial of degree γ: the smaller γ the faster numerical scheme converge to the true solution. If we can not bound resulting function with monomial (for example $\mathbb{E}|X_t - \hat{X}_{b\Delta t}| \propto e^{\Delta t}$), we will say that our scheme diverges.

It can be shown that for some conditions for f and g in (2) the rate of convergence of Euler-Maruyama scheme is equal to $\frac{1}{2}$. But what if we try to use this scheme for CIR process? Let's put the coefficients into (2), so we get the scheme:

$$\hat{X}_{\tau_i} = \hat{X}_{\tau_{i-1}} + (a - k\hat{X}_{\tau_{i-1}})\frac{T}{L} + \sigma\sqrt{\hat{X}_{\tau_{i-1}}}(W_{\tau_i} - W_{\tau_{i-1}}) \tag{7}$$

EM scheme does not imply any restrictions on the solution, so in general it can be negative or even infinite depending on properties of simulated process. Since volatility function of the CIR process is a square root, its simulated solution must be nonnegative on the whole interval $[0, T]$. Hence we should extend the basic EM method to control the solution of the simulated process.

2 Truncated Euler-Maruyama method

First modification of basic Euler-Maruyama method we are about to use is presented in the paper "The truncated Euler–Maruyama method for stochastic differential equations with Hölder diffusion coefficients" (2020) by H. Yang et al. The idea of this method is to "truncate" initial drift and volatility functions f and g so we can control the moments of the process. First of all we should make an assumptions on coefficients and make sure that they are satisfied for a CIR process:

$$\begin{cases} dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}dW_t \\ X_0 = x_0 \end{cases} \tag{8}$$

Remark 2.1. Function $f(x) = a(b - x)$ satisfies the local Lipschitz condition: for any $R > 0$, there exists $K_R > 0$ such that

$$|f(x) - f(y)| \leq K_R|x - y|$$

for all $x, y \in \mathbb{R}$ with $max(|x|, |y|) \leq R$.

Proof. $|f(x) - f(y)| = |a(b - x) - a(b - y)| = |a| \cdot |x - y| \Rightarrow$ local Lipschitz condition is satisfied with $K_R = |a|$. □

Remark 2.2. Function $f(x) = a(b-x)$ satisfies the one-sided Lipschitz condition: there is a constant $L > 0$ such that

$$(x - y)(f(x) - f(y)) \leq L|x - y|^2$$

for all $x, y \in \mathbb{R}$.

Proof. $(x-y)(f(x)-f(y)) = (x-y)\cdot(a(b-x)-a(b-y)) = a\cdot|x-y|^2 \Rightarrow$ one-sided Lipschitz condition is satisfied with $L=|a|$. $\square$

Remark 2.3. Function $g(x)=\sigma\sqrt{x}$ satisfies the Hölder continuity condition: there are constants $L>0$ and $0\le\alpha<\frac{1}{2}$ such that

$$|g(x)-g(y)|\le L|x-y|^{\frac{1}{2}+\alpha}$$

for all $x,y\in\mathbb{R}_+$.

Proof.

$$|g(x)-g(y)|=\sigma|\sqrt{x}-\sqrt{y}|=\sigma\frac{|x-y|}{\sqrt{x}+\sqrt{y}}=\sigma|x-y|^{\frac{1}{2}}\frac{\sqrt{|x-y|}}{\sqrt{x}+\sqrt{y}}$$

$$(\sqrt{|x-y|})^2=|x-y|\le|x|+|y|=(\sqrt{|x|+|y|})^2\le(\sqrt{x}+\sqrt{y})^2\Rightarrow\sqrt{|x-y|}\le\sqrt{x}+\sqrt{y}$$

$$|g(x)-g(y)|=\sigma|x-y|^{\frac{1}{2}}\frac{\sqrt{|x-y|}}{\sqrt{x}+\sqrt{y}}\le\sigma|x-y|^{\frac{1}{2}}$$

Holder continuity condition is satisfied with $L=\sigma$ and $\alpha=0$ $\square$

Remark 2.4. For function $f(x)=a(b-x)$ there exists $\gamma>0$ and $H>0$ such that

$$|f(x)-f(y)|^2\le H(1+|x|^\gamma+|y|^\gamma)|x-y|^2$$

for all $x,y\in\mathbb{R}$.

Proof.

$$(f(x)-f(y))^2=a^2\cdot(x-y)^2\le H(1+|x|^\gamma+|y|^\gamma)|x-y|^2$$

$\forall\gamma>0$ and $H=a^2$. $\square$

In the paper of Yang et al it is proved that for any process $dx(t)=f(x(t))dt+g(x(t))dW(t)$ whose coefficients satisfies (2.1) and (2.2):

$$\mathbb{E}[\sup_{0\le t\le T}|x(t)|^p]<\infty \text{ for any } p>0.$$

Now we can start truncating initial functions to get the scheme. First we have to choose function $\mu:\mathbb{R}_+\to\mathbb{R}_+$ such that

$$\sup_{|x|<u}|f(x)|\le\mu(u), u\ge 1$$

Since our scheme is defined in $\mathbb{R}_+$, we may take $\mu(u)=|a|\cdot x+|ab|$. Now we can denote μ^{-1} as an inverse function of μ, it is strictly increasing, $\mu^{-1}(x)=\frac{1}{|a|}\cdot x-|b|$. Then we choose number $\Delta=\frac{N}{L}$. Here and further in work we will take $T=1$ and $L=2^{10}$, hence $\Delta t=\frac{T}{L}=\frac{1}{2^{10}}$. We also should define function $h:(0,\Delta]\to(0,\infty)$ such that

$$h(\Delta)\ge\mu(2),\ \lim_{\Delta\to0}h(\Delta)=\infty \text{ and } \Delta^{\frac{1}{4}}h(\Delta)\le 1,\ \forall\Delta\in(0,\Delta^*]$$

We will take $h(\Delta) = \frac{2|a|+|ab|}{\Delta}$, that satisfies all the condition above. For a given Δ we can define our truncated function f:

$$f_\Delta(x) = f(min(|x|, \mu^{-1}(h(\Delta))) \cdot sign(x))$$

Since function $g(x)$ satisfies linear growth condition ($g(x) = \sigma\sqrt{x} \leq \sigma x + 1, \forall x \in \mathbb{R}_+$) there is no need to truncate it.

Now we can define the Truncated scheme:

$$X_\Delta(t_{k+1}) = X_\Delta(t_k) + f_\Delta(X_\Delta(t_k))\Delta + g(X_\Delta(t_k))\Delta W_k \tag{9}$$

We can use it to simulate CIR process with parameters $a = 1$, $b = 2$, $\sigma = 1$ and $x_0 = 1$:

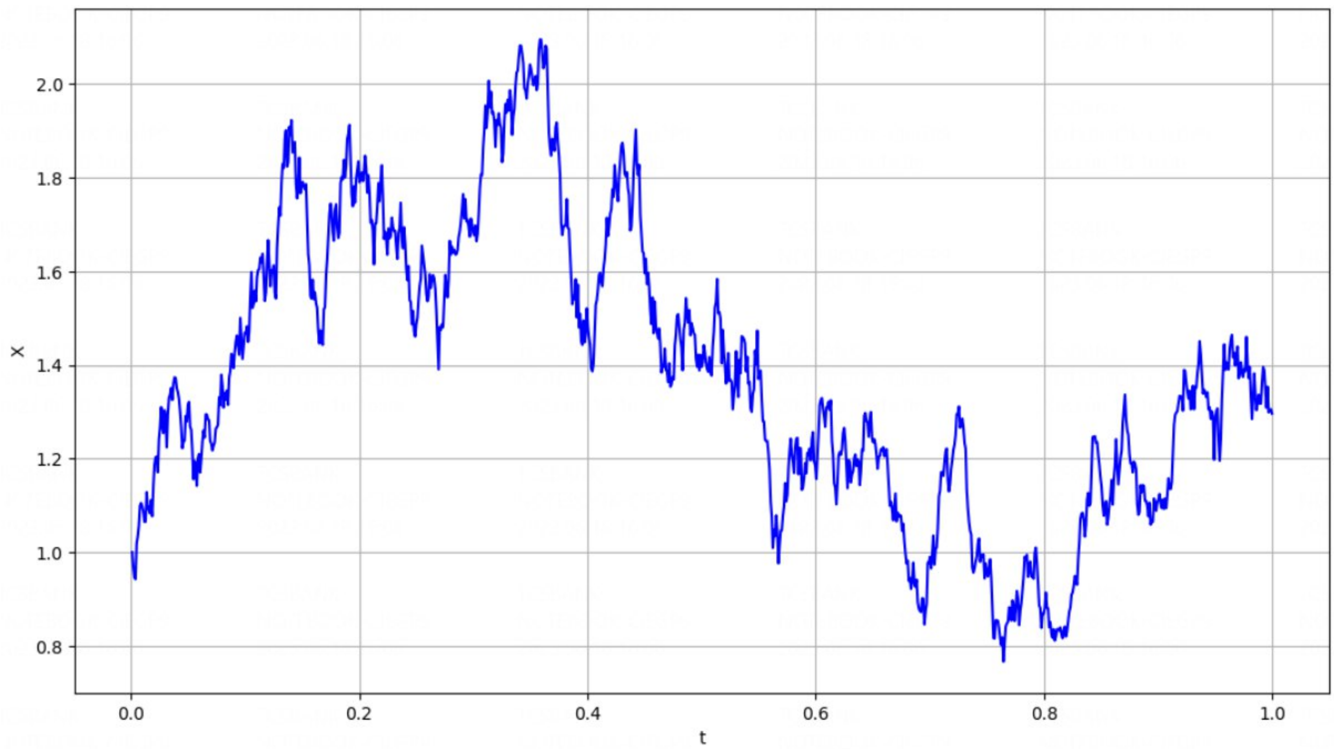

Figure 2: CIR process simulation with Truncated EM

We can visually check that the scheme converge to it's theoretical mean $b = 2$, but no too fast because $a = 1$. Also the trajectory is volatile enough because $\sigma = 1$. Hence we can empirically check that our scheme works, but we should prove it formally and find the rate of convergence. In Yang et al it is shown that for process with coefficients f and g, that satisfies Lemmas 2.1 - 2.4,

$$\mathbb{E}[\sup_{0\leq t\leq T} |X_t - \hat{X}_t|] \leq C(\frac{1}{ln\Delta^{-1}})^{\frac{1}{2}}$$

Note that generally speaking $\mathbb{E}[\sup ...] \neq \sup \mathbb{E}[...]$, that's why we will use these two types of strong convergence and compare them correctly.

3 Tamed Euler scheme

Let's move on to the next scheme, that is called Tamed Euler scheme. It was first proposed in the paper "Strong Convergence of an Explicit Numerical Method for SDES with Nonglobally LIPSCHITZ Continuous Coefficients" (2012) by M. Hutzenthaler, A. Jentzen and P. E. Kloeden. Before explaining idea of the scheme we should check if all initial conditions from the paper are satisfied: drift coefficient should be one-sided Lipschitz continuous and it's derivative should have at least polynomial growth. Diffusion coefficient should be globally Lipschitz continuous. The fact that $f(x) = a(b - x)$ is one-sided Lipschitz continuous is checked in the previous section. Since drift coefficient is linear function, it's derivative is a constant, so it doesn't grow polynomially.

Remark 3.1. Drift function $g(x) = \sigma\sqrt{x}$ is not globally Lipschitz continuous

Proof. Function f is called globally Lipschitz continuous if there exists constant C such that $|f(x) - f(y)| \leq C|x - y|$ for all $x, y \in \mathbb{R}$. Let's take $y = 0$ and check if there exists a constant such that $\sqrt{x} \leq C|x|$ for all $x \in \mathbb{R}_+$. On interval $(0, 1)$ this inequality is equivalent to $1 \leq C\sqrt{x}$. On this interval we can take any $C \in \mathbb{R}_+$ and see that there always exists x from $\mathbb{R}_+$ such that $1 > C\sqrt{x}$ because $\lim_{x\to 0} \sqrt{x} = 0$.

□

Two out of three conditions are not satisfied, so formally we can not use this scheme for CIR process modeling. But nevertheless we will describe the method in order to apply it to another SDE in the future and even try to use with CIR process to check what will go wrong.

The idea of method is as follows: we will bound our drift function f, like in the previous scheme, but we will do it differently:

$$f_\Delta(x) = \frac{\Delta \cdot f(x)}{1 + \Delta \cdot |f(x)|}$$

.

It is easy to see that $f_\Delta(x) < 1, \forall x \in \mathbb{R}$, so we can prevent our scheme from diverging. Now we can write down the discrete scheme:

$$\hat{X}_{n+1} = \hat{X}_n + \frac{\Delta \cdot f(\hat{X}_n)}{1 + \Delta \cdot |f(\hat{X}_n)|} + g(\hat{X}_n)\Delta W_n$$

Additionally, if we expand the tamed function $f(x)$ in $\frac{1}{N}$ for fixed $x \in \mathbb{R}$, we will get:

$$\hat{X}_{n+1} = \hat{X}_n + f(\hat{X}_n)\Delta + g(X_n)\Delta W - (\frac{T}{N})^2 \frac{f(\hat{X}_n) \cdot |f(\hat{X}_n)|}{1 + \Delta \cdot |f(\hat{X}_n)|} + O(\frac{1}{N^2}) \tag{10}$$

Note that (10) coincides with the Explicit Euler method up to terms of second order, but doesn't diverge, as we will show later.

Now we should interpolate our discrete scheme in order to compare it with an initial process. The simplest way to do it is to do it linearly:

$$\bar{X}_t = \hat{X}_n + \frac{(t - n\Delta) \cdot f(\hat{X}_n)}{1 + \Delta \cdot |\hat{X}_n|} + g(\hat{X}_n)\Delta W \tag{11}$$

$\forall t \in [n\Delta, (n+1)\Delta]$, $n \in [0, N]$.

Since the scheme is ready we try to use it for CIR process modeling with $a = 1$, $b = 2$ and $\sigma = 1$ (recall that formally we are not allowed to do it since initial conditions are not satisfied):

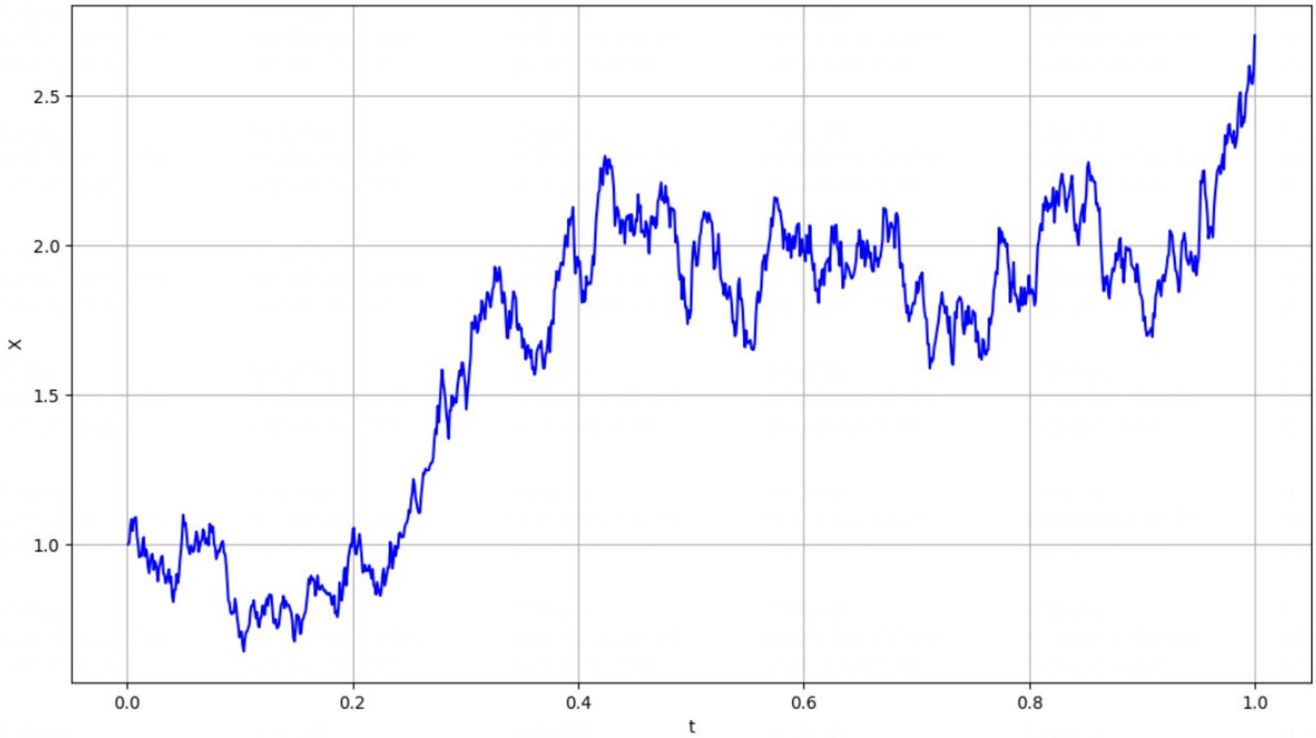

Figure 3: CIR process simulation with Tamed EM

As we can see, graph did not approach 0, so we may say that Tamed schemed worked out for CIR process with parameters written above. Our hypothesis is that Tamed scheme converges fast enough, so taking $L = 2^{10}$ allows to use coefficient that doesn't satisfy theoretical conditions, the scheme anyway will converge to initial process that is nonnegative by definition. We can visually test it by simulating same process with different number if timesteps L:

As we can see schemes with L equal to 2 to the power of 5, 6 and 10 are failed to converge. We may suppose that Tamed scheme will work with CIR process if x_0 and b are far enough from 0, and σ is not too large.

In order to measure the strong convergence we may refer to Hutzenthaler et al:

$$\mathbb{E}[\sup_{t \in [0,T]} |X_t - \bar{X}_t|] \leq C \cdot L^{-\frac{1}{2}}$$

We may point out that this convergence rate is equal to Euler scheme's and smaller, than Truncated scheme's.

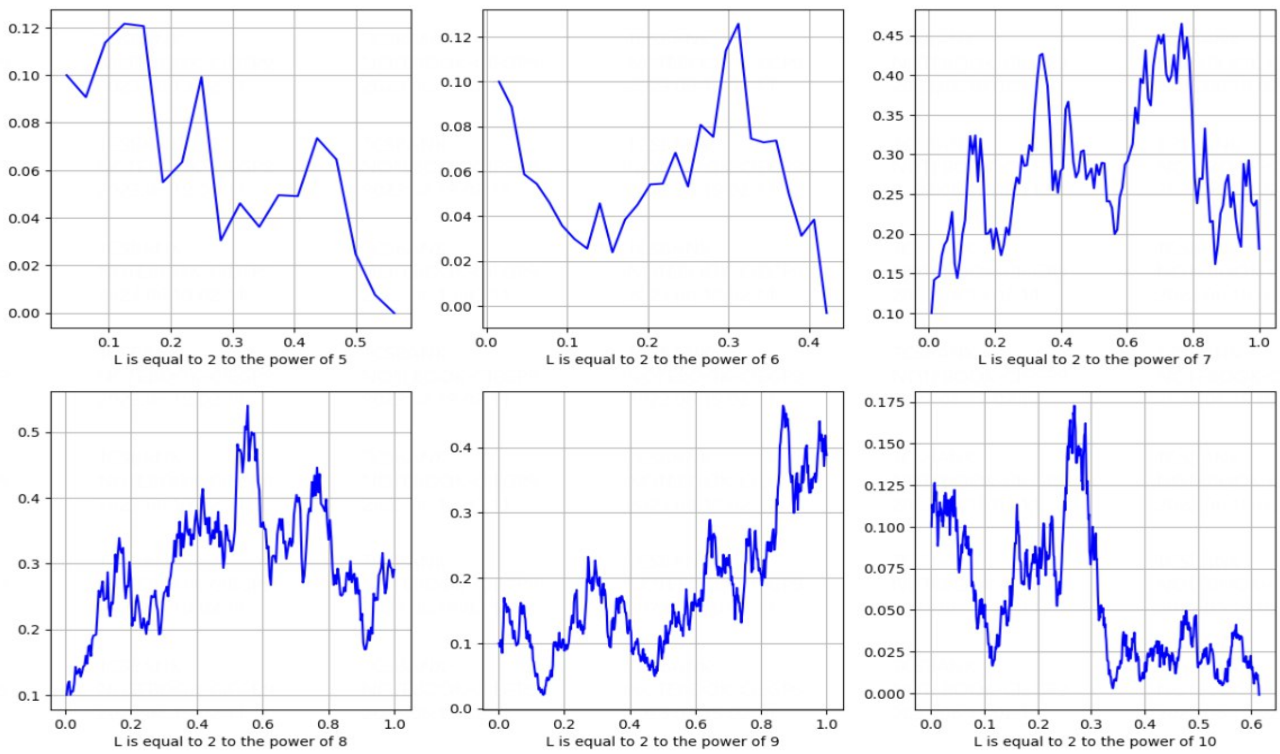

Figure 4: CIR process simulation with Truncated EM and different number of timesteps

4 Implicit scheme by Alfonsi

The last scheme we are going to use is "On the discretization schemes for the CIR (and Bessel squared) processes" (2005) by A. Alfonsi. The author developed his scheme only for CIR process, using:

$$\begin{cases} dX_t = (a - kX_t)dt + \sigma\sqrt{X_t}dW_t \\ x_0, \sigma, a \geq 0, k \in \mathbb{R} \\ 2a > \sigma^2. \end{cases} \tag{12}$$

Last condition is taken to make sure that the process stays positive on $[0, T]$. The idea of method is to try to expand the implicit scheme into quadratic equation and choose positive only solution:

$$X_t = x_0 + \int_0^t (a - kX_s)ds + \sigma \int_0^t \sqrt{X_s}dW_s$$

$$= x_0 + \lim_{n\to\infty}\left(\sum_{t_i<t}(a - kX_{t_{i+1}})\Delta + \sigma\sum_{t_i<t}\sqrt{X_{t_{i+1}}}(W_{t_{i+1}} - W_{t_i}) - \sigma\sum_{t_i<t}(\sqrt{X_{t_{i+1}}} - \sqrt{X_{t_i}})(W_{t_{i+1}} - W_{t_i})\right)$$

$$= x_0 + \lim_{n\to\infty}\left(\sum_{t_i<t}(a - \frac{\sigma^2}{2} - kX_{t_{i+1}})\Delta + \sigma\sum_{t_i<t}\sqrt{X_{t_{i+1}}}(W_{t_{i+1}} - W_{t_i})\right)$$

Based on this fact we can say write down a quadratic equation, which positive solution will be our discrete scheme:

$$\hat{X}_{t_{i+1}} = X_{t_i} + (a - \frac{\sigma^2}{2} - k\hat{X}_{t_{t+1}})\Delta + \sigma\sqrt{\hat{X}_{t_{i+1}}}(W_{t_{i+1}} - Wt_i) \tag{13}$$

Based on our previous calculations we can say that $\hat{X}_{t_i} \to X_{t_i}$ as number of timesteps L goes to infinity. So we may use it with $\Delta = \frac{T}{L}$ small enough. Hence we choose positive solution from equation (13):

$$\hat{X}_{t_{i+1}} = \Big(\frac{\sigma(W_{t_{i+1}} - W_{t_i}) + \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i} + (a - \frac{\sigma^2}{2})\Delta)(1 + k\Delta)}}{2(1 + k\Delta)}\Big)^2 \tag{14}$$

Formally we got the scheme, that is obviously nonnegative, but it's complicated to calculate. Later we will try to expand it, using Taylor series and some properties, that have been proven in the paper.

Remark 4.1. Suppose that $(\hat{X}_{t_i})$ is a nonnegative scheme, for which

$$\begin{cases} \hat{X}_{t_0} = x_0 \\ \hat{X}_{t+i+1} \leq (1 + \frac{b}{L})\hat{X}_{t_i} + \sigma\sqrt{\hat{X}_{t_i}}(W_{t_{i+1}} - W_{t_i}) + O(\frac{1}{L}) \end{cases} \tag{15}$$

is satisfied $\forall i \leq n - 1$. Then, scheme $(\hat{X}_{t_i})$ has uniformly bounded moments, so it means that $\hat{X}_{t_i} = O(1)$ $\forall t_i \in [0, T]$.

The proof of this lemma is complicated enough to use this fact only referring to original paper. We will use it to reduce some parts of the scheme to $O(1)$. Now we can define define hypothesis H_1 to explore strong convergence:

We will say that scheme $(\hat{X}_{t_i})$ satisfies H_1 if it is nonnegative scheme such that:

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_i} + (a - k\hat{X}_{t_i})\Delta + \sigma\sqrt{\hat{X}_{t_i}}(W_{t_{i+1}} - W_{t_i}) + m_{t_{i+1}} - m_{t_i} + O(\frac{1}{L^{\frac{3}{2}}})$$

where $m_{t_{i+1}} - m_{t_i}$ is a martingale of order 1 ($m_{t_{i+1}} - m_{t_i} = O(\frac{1}{L})$)

Now we can rewrite our scheme:

$$\hat{X}_{t_{i+1}} = \Big(\frac{\sigma(W_{t_{i+1}} - W_{t_i}) + \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i} + (a - \frac{\sigma^2}{2})\Delta)(1 + k\Delta)}}{2(1 + k\Delta)}\Big)^2$$

$$= \frac{1}{4(1 + k\Delta)^2}\Big(\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + \sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i} + (a - \frac{\sigma^2}{2})\Delta)(1 + k\Delta) +$$

$$+2\sigma(W_{t_{i+1}} - W_{t_i})\sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i} + (a - \frac{\sigma^2}{2})\Delta)(1 + k\Delta)}\Big)$$

We may observe, using (4.1), that

$$\sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}_{t_i} + (a - \frac{\sigma^2}{2})\Delta)(1 + k\Delta)} = O(\frac{1}{\sqrt{L}})$$

, so our scheme looks like

$$\hat{X}_{t_{i+1}} = \frac{1}{1 + k\Delta}\hat{X}_{t_i} + \frac{\sigma}{(1 + k\Delta)^{\frac{3}{2}}}\sqrt{\hat{X}_{t_i}}(W_{t_{i+1}} - W_{t_i}) + O(1/L)$$

Using this fact we can prove that $(\hat{X}_{t_i})$ satisfies H_1 as it is done in the paper of Alfonsi. using this, we may state, that for out implicit scheme $(\hat{X}_{t_i})$:

$$\sup_{0 \leq i \leq n} \mathbb{E}(|X_{t_i} - \hat{X}_{t_i}|) \leq \frac{C}{ln(L)}$$

$$\mathbb{E}[\sup_{0 \leq i \leq n} |X_{t_i} - \hat{X}_{t_i}|] \leq \frac{C}{\sqrt{ln(L)}}$$

Now we can test the scheme for CIR process with $a = 3$, $k = 2$ and $\sigma = 1$:

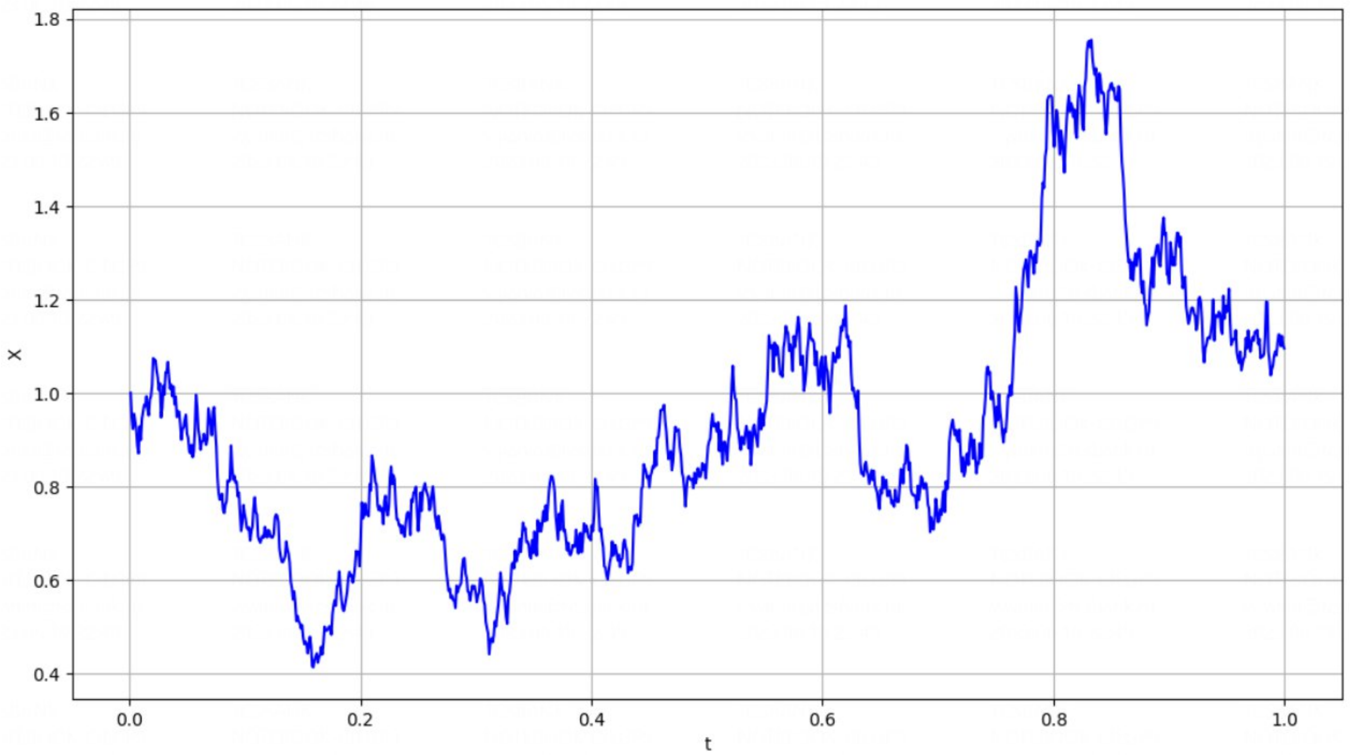

Figure 5: CIR process simulation with Implicit scheme by Alfonsi

5 Conclusion

The aim of the work was to explore different modifications of Euler-Maruyama method, that we can use to model CIR process. To sum up, we may write down a simple table:

Scheme	$\mathbb{E}[\sup\limits_{0\le t\le T} \lvert X_t - \hat{X}_t\rvert]$	$\sup\limits_{0\le t\le T} \mathbb{E}[\lvert X_t - \hat{X}_t\rvert]$	Works for CIR
EM	?	$\frac{C}{L^{\frac{1}{2}}}$	-
Tamed	$\frac{C}{\sqrt{L}}$	?	-
Truncated	$\frac{C}{\sqrt{lnL}}$	?	+
Alfonsi	$\frac{C}{\sqrt{lnL}}$	$\frac{C}{lnL}$	+

The main result we got: Truncated scheme and scheme by Alfonsi are faster than others and can be used to simulate CIR process. As can be seen from the graphs they look quite familiar with original process and converge fast enough, so in the same graph two schemes are close enough to each other. Tamed scheme converges slower, but it has different advantages: it's supposed to work with process with polynomial coefficients, whose coefficients grows supelinearly.

In the future I plan to continue working with this theme, adjusting the topic. First of all, we should use another process to simulate short-term interest rate and compare results with CIR. Second, we should add an economic application: get real data and, estimate parameters and choose metric to estimate an error between model's prediction and real data.

References

[1] Aur´elien Alfonsi (2005): On the discretization schemes for the CIR (and Bessel squared) processes

[2] Desmond J. Higham (2001): An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations

[3] Hao Yang, Fuke Wu, Peter E. Kloeden, Xuerong Mao (2020): The truncated Euler–Maruyama method for stochastic differential equations with Hölder diffusion coefficients

[4] John C. Cox, Jonathan E. Ingersoll, Stephen A. Ross (1985): A Theory of the Term Structure of Interest Rates

[5] Martin Hutzenthaler, Aenulf Jentzen, Peter E. Kloeden (2012): Strong convergence of an explicit numerical method for SDES with nonglobally Lipschitz continuous coefficients

[6] Sotirios Sabansis (2016): Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficient

[7] Xuerong Mao (2015): The truncated Euler–Maruyama method for stochastic differential equations