High order numerical simulation of the Cox-Ingersoll-Ross process

James Foster

Supervised by Terry Lyons and Harald Oberhauser

University of Oxford

2019

Introduction

The Cox-Ingersoll-Ross (CIR) model is defined by the following SDE:

$$dy_t = a(b - y_t)\,dt + \sigma\sqrt{y_t}\,dW_t, \tag{1}$$
$$y_0 = \xi,$$

with the following parameters

- Mean reversion speed: $a > 0$
- Mean reversion level: $b > 0$
- Volatility: $\sigma > 0$

This diffusion is commonly used as a one-factor short rate model in mathematical finance for modelling interest rates and volatilities [1].

Introduction

Due to its analytically tractable vector fields, the SDE that governs the CIR process can also be written in Stratonovich form as follows:

$$\begin{aligned} dy_t &= a\left(\tilde{b} - y_t\right) dt + \sigma \sqrt{y_t} \circ dW_t, \\ y_0 &= \xi, \end{aligned} \tag{2}$$

where $\tilde{b} := b - \frac{1}{4a}\sigma^2$.

The above diffusion poses two challenges in a numerical simulation.

- The solution y is non-negative almost surely.
- The square root vector field is not globally Lipschitz.

Introduction

Our strategy for simulating the CIR process can be summarized as

1. Choose an initial step size to propagate the numerical solution.
2. Replace the Brownian motion W by a piecewise linear path $\widehat{W}$.
3. Along each piece of $\widehat{W}$ we can approximate (1) using the ODE:

$$d\widehat{y}_t = a\big(\tilde{b} - \widehat{y}_t\big)\,dt + \sigma\sqrt{\widehat{y}_t}\,d\widehat{W}_t. \tag{3}$$

4. Estimate the $L^2(\mathbb{P})$ error exhibited by the above approximation. If this estimator is above a tolerance, the step size is reduced. When this happens, W will be sampled over two half intervals.
5. In each step, we then discretize (3) with a suitable ODE solver.

Introduction

Before describing these steps, we note the Taylor expansion for (2).

$$
\begin{aligned}
y_t = y_s &+ \sigma\sqrt{y_s}\,W_{s,t} + a\left(\tilde{b} - y_s\right)h + \frac{1}{4}\sigma^2 W_{s,t}^2 \qquad (4)\\
&- a\sigma\sqrt{y_s}\int_s^t\int_s^u \circ\, dW_v\,du + \frac{a\sigma}{2\sqrt{y_s}}\left(\tilde{b} - y_s\right)\int_s^t\int_s^u dv \circ dW_u\\
&- \frac{1}{2}a^2\left(\tilde{b} - y_s\right)h^2 - \frac{1}{2}a\sigma^2\int_s^t\int_s^u\int_s^v \circ\, dW_r \circ dW_v\,du\\
&- \left(\frac{1}{2}a\sigma^2 + \frac{a\sigma^2}{4y_s}\left(\tilde{b} - y_s\right)\right)\int_s^t\int_s^u\int_s^v \circ\, dW_r\,dv \circ dW_u\\
&+ \text{higher order terms},
\end{aligned}
$$

for sufficiently small $h := t - s$.

Introduction

Similarly, $\widehat{y}$ gives this Taylor expansion but with $\widehat{W}$ integrals instead.

Therefore, we wish to construct $\widehat{W}$ to have the following properties:

- $W_{s,t} = \widehat{W}_{s,t}$.
- $\displaystyle\int_s^t W_{s,u}\,du = \int_s^t \widehat{W}_{s,u}\,du$.
- $\displaystyle\int_s^t W_{s,u}^2\,du = \int_s^t \widehat{W}_{s,u}^2\,du + R_{s,t}$,

where the error term satisfies both $\mathbb{E}\left[R_{s,t}\right] = 0$ and $\mathbb{E}\left[R_{s,t}^2\right] \sim O\left(h^4\right)$.

Of course, we should attempt to minimize $R_{s,t}$ as much as possible.

Introduction

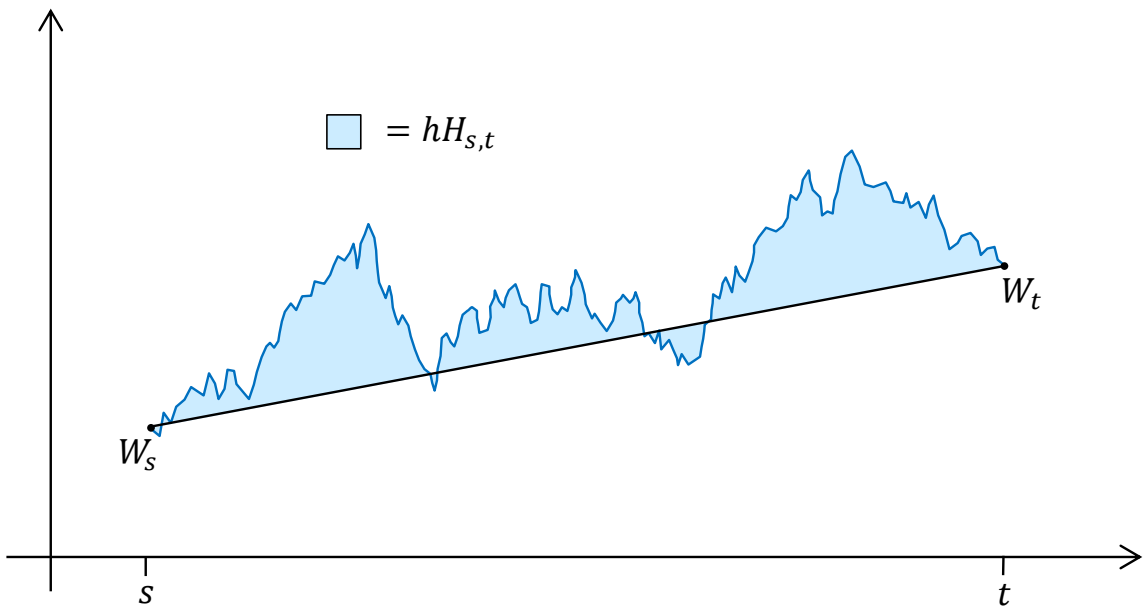
Thus to construct the path $\widehat{W}$, we will generate the stochastic area:

Definition

The rescaled **space-time Lévy area** of Brownian motion on $[s,t]$ is

$$H_{s,t} := \frac{1}{h}\int_s^t W_{s,u} - \frac{u-s}{h} W_{s,t}\, du,$$

where $h = t-s$. Then $W_{s,t}$ and $H_{s,t} \sim N\left(0, \frac{1}{12}h\right)$ are independent [2].



Introduction

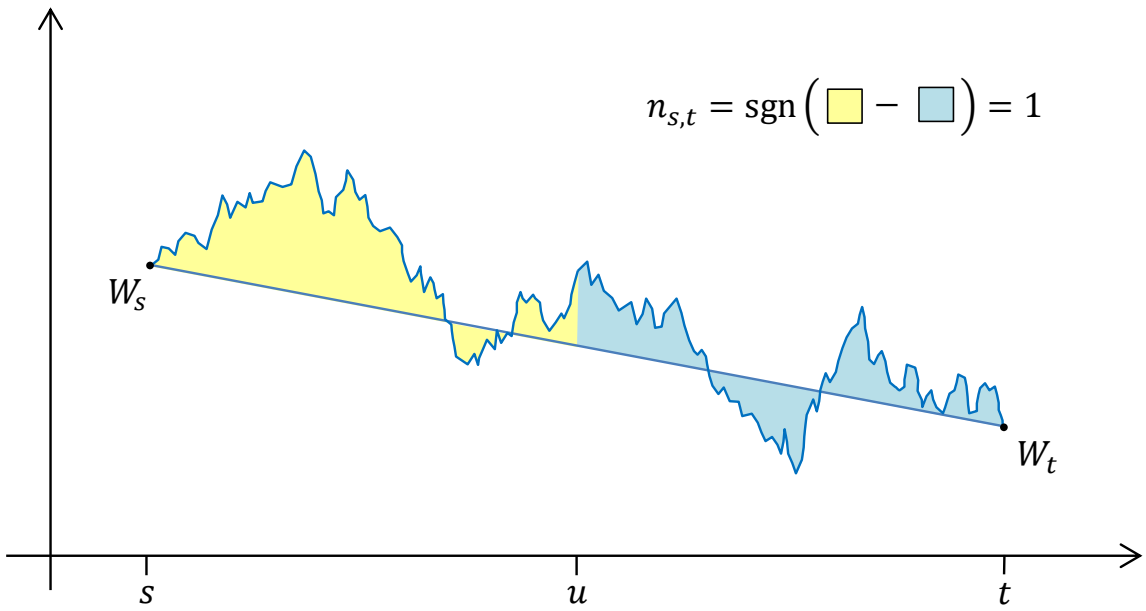
Since $\widehat{W}$ has three constraints, it will contain more than two pieces. Therefore, we shall generate further information about the path W.

Definition

The **space-time orientation** of a Brownian path on $[s,t]$ is given by

$$n_{s,t} := \operatorname{sgn}\left(H_{s,u} - H_{u,t}\right),$$

where $u := \frac{1}{2}(s+t)$.



Introduction

Although $n_{s,t}$ clearly follows the Rademacher distribution, it is not clear whether it can be generated in practice along with $(W_{s,t}, H_{s,t})$.

The result below shows $n_{s,t}$ has the required independence property.

Theorem

The variable $N_{s,t} := H_{s,u} - H_{u,t}$ is independent of $(W_{s,u}, W_{u,t}, H_{s,t})$.

Sketch Proof.

Arguing that $(H_{s,u}, H_{u,t})$ and $(W_{s,u}, W_{u,t})$ are independent is simple.
Using symmetry, one can prove that $N_{s,t}$ and $H_{s,t}$ are uncorrelated.
As $N_{s,t}$ and $H_{s,t}$ are jointly normal, they must be independent. □

Introduction

Due to the iterated integrals in (4), we will use the random variable:

Definition

The **space-space-time Lévy area** of Brownian motion over $[s,t]$ is

$$L_{s,t} := \frac{1}{6}\left(\int_s^t\int_s^u\int_s^v \circ dW_r \circ dW_v\, du - 2\int_s^t\int_s^u\int_s^v \circ dW_r\, dv \circ dW_u + \int_s^t\int_s^u\int_s^v dr \circ dW_v \circ dW_u\right).$$

Using several applications of integration by parts, we can then show

Theorem (Relationships between third order iterated integrals)

$$\int_s^t\int_s^u\int_s^v \circ dW_r \circ dW_v\, du = \frac{1}{6}hW_{s,t}^2 + \frac{1}{2}hW_{s,t}H_{s,t} + L_{s,t},$$

$$\int_s^t\int_s^u\int_s^v \circ dW_r\, dv \circ dW_u = \frac{1}{6}hW_{s,t}^2 - 2L_{s,t}.$$

Introduction

The key result which enables us to develop the numerical method is

Theorem (Conditional moments of space-space-time Lévy area)

$$\mathbb{E}\left[L_{s,t} \,|\, W_{s,t}, H_{s,t}, n_{s,t}\right] = \frac{1}{30}h^2 + \frac{3}{5}hH_{s,t}^2 - \frac{1}{8\sqrt{6\pi}}n_{s,t}\,h^{\frac{3}{2}}W_{s,t}, \quad (5)$$

$$\begin{aligned}\mathrm{Var}\left(L_{s,t} \,|\, W_{s,t}, H_{s,t}, n_{s,t}\right) = {} & \frac{11}{25200}h^4 + \left(\frac{1}{720} - \frac{1}{384\pi}\right)h^3 W_{s,t}^2 \quad (6)\\ & + \frac{1}{700}h^3 H_{s,t}^2 - \frac{1}{320\sqrt{6\pi}}n_{s,t}\,h^{\frac{7}{2}}W_{s,t}.\end{aligned}$$

Idea of Proof

We write $W = \widetilde{W} + Z$ where $\widetilde{W}_r = \mathbb{E}\left[W_r \,|\, W_{s,t}, H_{s,t}, N_{s,t}\right]$ for $r \in [s,t]$ and Z is a Gaussian process that is independent of $\left(W_{s,t}, H_{s,t}, N_{s,t}\right)$. This can be shown using the polynomial representation of W in [2]. The result then follows by direct (but also lengthy) calculations. □

Piecewise linear discretization of Brownian motion

Let $\widehat{W}$ denote the piecewise linear path on $[s,t]$ with three pieces that connect the points (s, W_s), $(s+ah, b)$, $(t-ah, c)$ and (t, W_t).

We wish to identify values of (a,b,c) so that $\widehat{W}$ has the integrals:

- $\int_s^t \widehat{W}_{s,u}\, du = \int_s^t W_{s,u}\, du.$
- $\int_s^t \widehat{W}_{s,u}^2\, du = \mathbb{E}\left[\int_s^t W_{s,u}^2\, du \,\middle|\, W_{s,t}, H_{s,t}, n_{s,t}\right].$

In addition, we will choose a so that $b-c$ is independent of $H_{s,t}$.

Note that the first integral equality is equivalent to the condition:

$$b+c = W_{s,t} + \frac{2}{1-a} H_{s,t}. \tag{7}$$

Piecewise linear discretization of Brownian motion

When $a = \frac{8-\sqrt{10}}{18}$, the second integral equality becomes equivalent to

$$b - c = aW_{s,t} \pm \sqrt{(1-a)^2 W_{s,t}^2 - \frac{3}{\sqrt{6\pi}} n_{s,t} \sqrt{h} W_{s,t} + \frac{4}{5} h}. \qquad (8)$$

Moreover, one can show that $b-c$ must depend on $H_{s,t}$ if $a \neq \frac{8-\sqrt{10}}{18}$.

Although the square root term in (8) is always positive, we have not determined whether it should be added to or subtracted from $aW_{s,t}$.

Thus, this sign will be chosen to additionally minimize the quantity:

$$\left| \int_s^t (u-s)\, \widehat{W}_{s,u}\, du - \mathbb{E}\left[\int_s^t (u-s)\, W_{s,u}\, du \,\middle|\, W_{s,t}, H_{s,t}, n_{s,t} \right] \right|$$

Piecewise linear discretization of Brownian motion

The above error can be explicitly computed for both choices of $b-c$.

Theorem

$$\mathbb{E}\left[\int_s^t (u-s)\, W_{s,u}\, du \,\middle|\, W_{s,t}, H_{s,t}, n_{s,t}\right] = \frac{1}{3}h^2 W_{s,t} + \frac{1}{2}h^2 H_{s,t} - \frac{1}{\sqrt{8\pi}} n_{s,t} h^{\frac{5}{2}}.$$

Proof.

The result follows using the same decomposition ($W = \widetilde{W} + Z$). □

Remark

The original motivation for including $n_{s,t}$ in the construction of $\widehat{W}$ was that it helped one decide which value of $b-c$ should be used.

In addition, $n_{s,t}$ is fast to generate and can be incorporated into a variable step size method as $N_{s,t} | n_{s,t}$ is half-normal (modulo sign).

Piecewise linear discretization of Brownian motion

Definition

Thus, we define the path $\widehat{W}$ using the parameters (a, b, c) satisfying

$$a = \frac{8 - \sqrt{10}}{18},$$

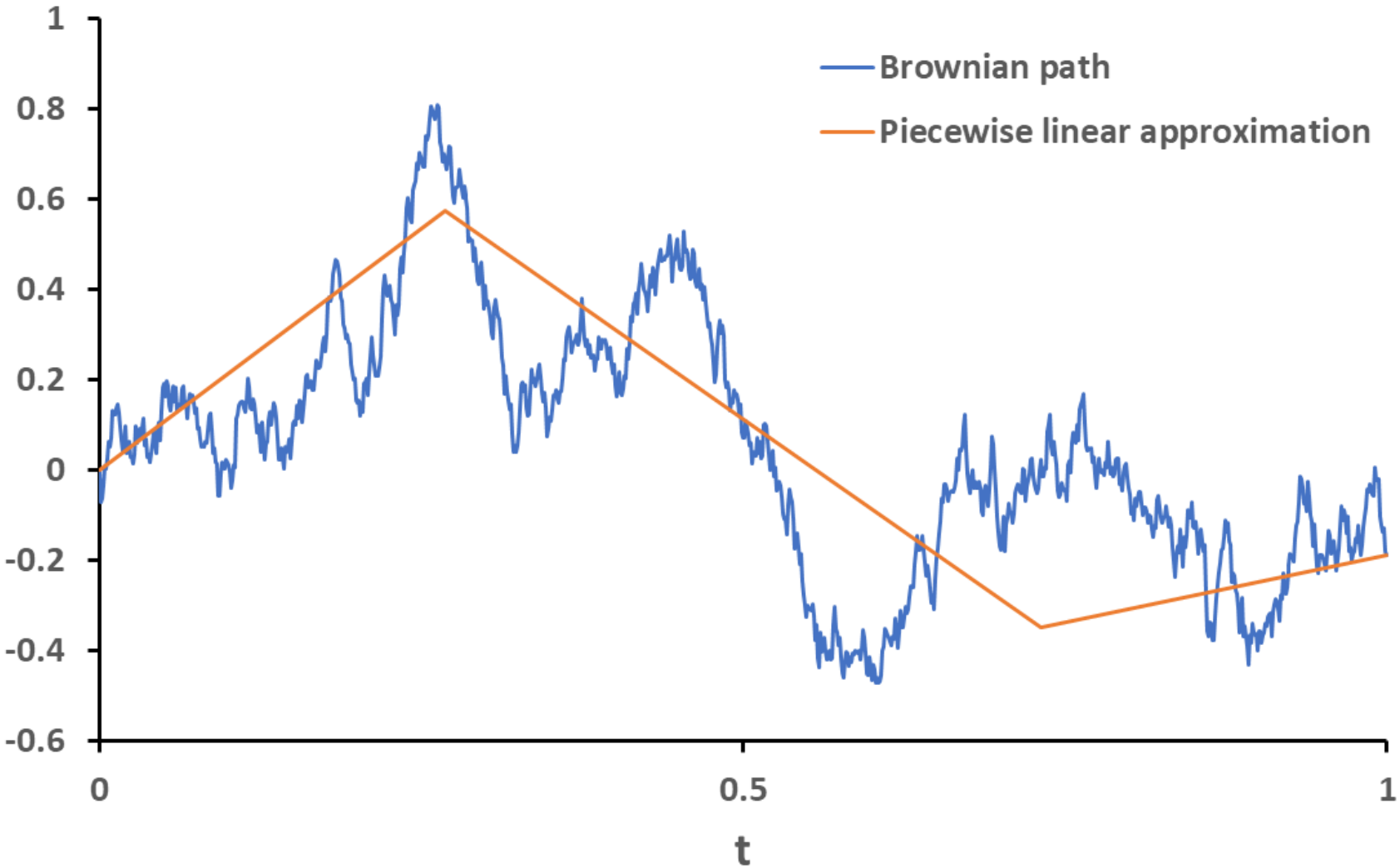
$$b + c = W_{s,t} + \frac{2}{1-a} H_{s,t},$$

$$b - c = aW_{s,t} + \varepsilon \sqrt{(1-a)^2 W_{s,t}^2 - \frac{3}{\sqrt{6\pi}} n_{s,t} \sqrt{h} W_{s,t} + \frac{4}{5} h},$$

where

$$\varepsilon = \begin{cases} 1 & \text{if } \frac{3}{2\sqrt{6\pi}} n_{s,t} \sqrt{h} \geq (1-a)^2 W_{s,t} \\ -1 & \text{if } \frac{3}{2\sqrt{6\pi}} n_{s,t} \sqrt{h} < (1-a)^2 W_{s,t} \end{cases}.$$

Piecewise linear discretization of Brownian motion



Discretization of the ODE approximation

Along each piece of $\widehat{W}$, we shall discretize (3) using a single step of

Definition (Third order A-stable implicit Runge-Kutta method)

This method is presented in [3] and defined by the Butcher tableau:

$$
\begin{array}{c|ccc}
0 & 0 & & \\
\frac{3+\sqrt{3}}{3} & \frac{3+\sqrt{3}}{6} & \frac{3+\sqrt{3}}{6} & \\
1 & \frac{3+\sqrt{3}}{12} & \frac{1-\sqrt{3}}{4} & \frac{3+\sqrt{3}}{6} \\
\hline
 & \frac{3+\sqrt{3}}{12} & \frac{1-\sqrt{3}}{4} & \frac{3+\sqrt{3}}{6}
\end{array}
$$

We will now see that this method is well suited for the CIR process.

Discretization of the ODE approximation

When applied to an ODE $y' = F(y)$ on the interval $[0,1]$, the above Runge-Kutta numerical solution Y is given by the implicit equations

$$\widetilde{Y} = Y_0 + \frac{3+\sqrt{3}}{6} F\big(Y_0\big) + \frac{3+\sqrt{3}}{6} F\big(\widetilde{Y}\big), \tag{9}$$

$$Y_1 = Y_0 + \frac{3+\sqrt{3}}{12} F\big(Y_0\big) + \frac{1-\sqrt{3}}{4} F\big(\widetilde{Y}\big) + \frac{3+\sqrt{3}}{6} F\big(Y_1\big). \tag{10}$$

When simulating the CIR diffusion on $[0,T]$, we will choose F to be the vector field governing the square root process $\widehat{z} := \sqrt{\widehat{y}}$ given by

$$F(z) := \frac{1}{2} a \left(\frac{\tilde{b}}{z} - z \right) \Delta t + \frac{1}{2} \sigma \Delta \widehat{W} \tag{11}$$

Discretization of the ODE approximation

Definition (Numerical scheme for a CIR process driven by $\widehat{W}$)

By applying this Runge-Kutta method to the ODE governed by F, we can approximate (3) along each piece of $\widehat{W}$ using the formulae:

$$\left(1+ca\Delta t\right)\widetilde{Y}-B_1\sqrt{\widetilde{Y}}-ca\tilde{b}\Delta t=0, \tag{12}$$

$$\left(1+ca\Delta t\right)Y_{t+\Delta t}-B_2\sqrt{Y_{t+\Delta t}}-ca\tilde{b}\Delta t=0, \tag{13}$$

where Y_t denotes the numerical solution of (3) computed at time t with the same initial value $Y_0 := y_0$ and the coefficients are given by

$$c := \frac{3+\sqrt{3}}{12},$$

$$B_1 := \sqrt{Y_t}+2cF\left(\sqrt{Y_t}\right)+c\sigma\Delta\widehat{W},$$

$$B_2 := \sqrt{Y_t}+cF\left(\sqrt{Y_t}\right)+\frac{1-\sqrt{3}}{4}F\left(\sqrt{\widetilde{Y}}\right)+c\sigma\Delta\widehat{W}.$$

Discretization of the ODE approximation

Therefore, each step of the proposed discretization of (2) requires three steps of the above numerical method (for each piece of $\widehat{W}$).

Remark

Since there is no $W_{s,t}^4$ term in the stochastic Taylor expansion of the CIR process, this Runge-Kutta method has a sufficiently high order.

In addition, Y is guaranteed to be positive provided that $a, \tilde{b}, y_0 > 0$.

The Runge-Kutta method is also A-stable and thus may be better suited for (3), which is a stiff equation, than explicit ODE solvers.

Estimating local $L^2(\mathbb{P})$ errors

Recall that the Taylor expansion of the SDE (2) can be expressed as

$$
\begin{aligned}
y_t = y_s &+ \sigma\sqrt{y_s}\,W_{s,t} + a\left(\tilde{b} - y_s\right)h + \frac{1}{4}\sigma^2 W_{s,t}^2 \qquad (14)\\
&- a\sigma\sqrt{y_s}\left(\frac{1}{2}hW_{s,t} + hH_{s,t}\right) + \frac{a\sigma}{2\sqrt{y_s}}\left(\tilde{b} - y_s\right)\left(\frac{1}{2}hW_{s,t} - hH_{s,t}\right)\\
&- \frac{1}{2}a^2\left(\tilde{b} - y_s\right)h^2 - \frac{1}{2}a\sigma^2\left(\frac{1}{6}hW_{s,t}^2 + \frac{1}{2}hW_{s,t}H_{s,t} + L_{s,t}\right)\\
&- \left(\frac{1}{2}a\sigma^2 + \frac{a\sigma^2}{4y_s}\left(\tilde{b} - y_s\right)\right)\left(\frac{1}{6}hW_{s,t}^2 - 2L_{s,t}\right)\\
&+ \text{higher order terms.}
\end{aligned}
$$

Similarly, $\widehat{y}$ gives this Taylor expansion but with $\widehat{W}$ integrals instead.

Estimating local $L^2(\mathbb{P})$ errors

Therefore the largest source of numerical error that is introduced by $\widehat{y}$ is due to the space-space-time Lévy area of $\widehat{W}$ not matching $L_{s,t}$.

Thus, the leading term in the variance expansion of the local error is

$$\left(\frac{a\tilde{b}\sigma^2}{2y_s}\right)^2 \mathrm{Var}\left(L_{s,t} \,|\, W_{s,t}, H_{s,t}, n_{s,t}\right). \tag{15}$$

This conditional variance can then be computed explicitly using (6).

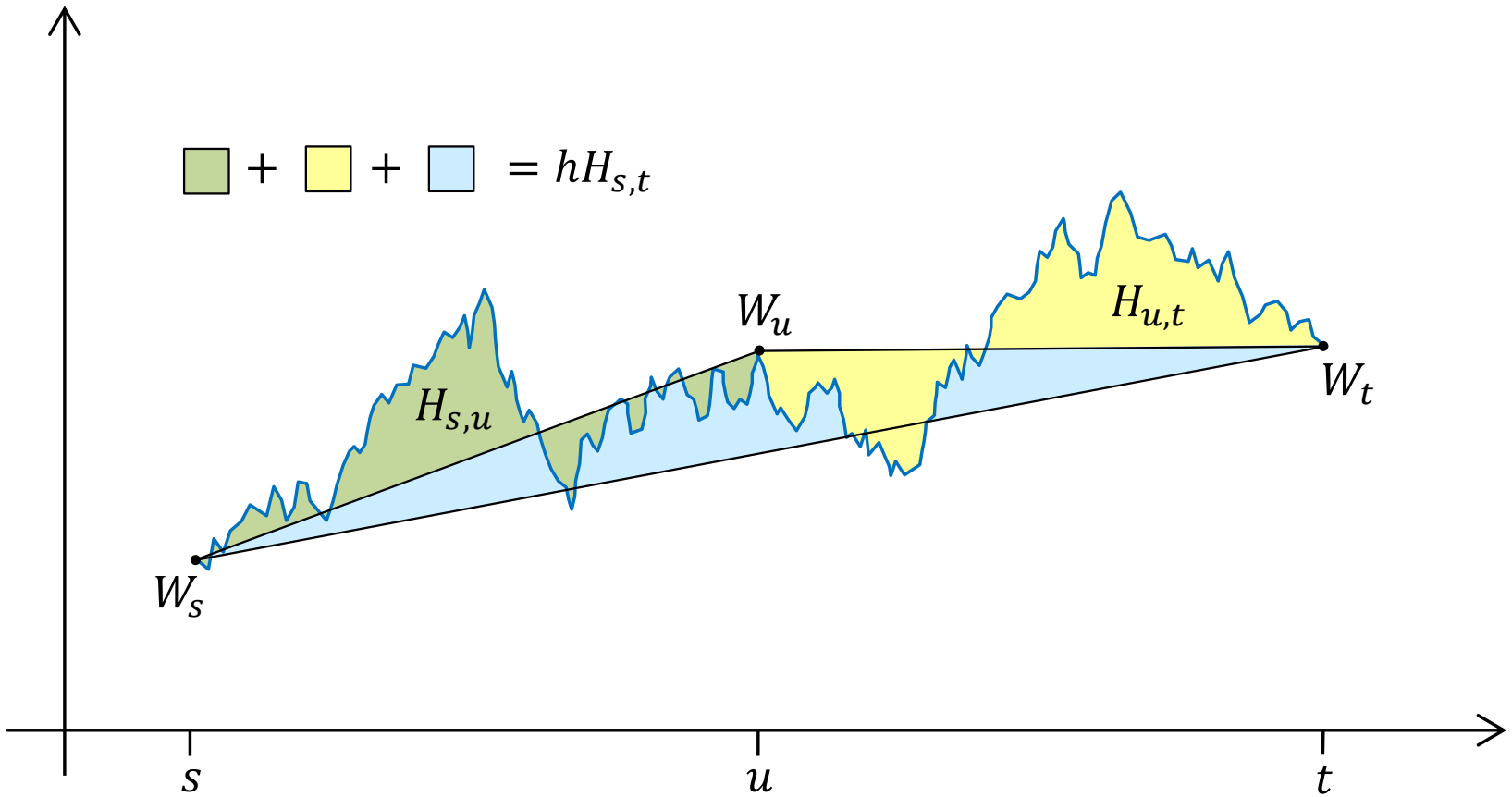
When simulating the CIR process, we aim to ensure that the above quantity is bounded above by Ch for some prespecified constant C.

In order to control this error estimator, we need a way to reduce h.

Refinement of increments and space-time areas

Question

What is the joint distribution of $\left(W_{s,u}, W_{u,t}, H_{s,u}, H_{u,t}\right)$ conditional on $\left(W_{s,t}, H_{s,t}\right)$ when $u = \frac{1}{2}(s+t)$?



Refinement of increments and space-time areas

Theorem (Dyadic refinement procedure for the pair $(W_{s,t}, H_{s,t})$)

$$W_{s,u} = \frac{1}{2}W_{s,t} + \frac{3}{2}H_{s,t} + Z_{s,u},$$

$$W_{u,t} = \frac{1}{2}W_{s,t} + \frac{3}{2}H_{s,t} - Z_{s,u},$$

$$H_{s,u} = \frac{1}{4}H_{s,t} - \frac{1}{2}Z_{s,u} + \frac{1}{2}N_{s,t},$$

$$H_{u,t} = \frac{1}{4}H_{s,t} - \frac{1}{2}Z_{s,u} - \frac{1}{2}N_{s,t},$$

where $W_{s,t}, H_{s,t}, Z_{s,u}, N_{s,t}$ *are independent random variables with*

$$Z_{s,u} \sim N\left(0, \frac{1}{16}h\right) \quad \textit{and} \quad N_{s,t} \sim N\left(0, \frac{1}{12}h\right).$$

Proof.

A refinement procedure for any $u \in (s,t)$ was established in [4]. □

Refinement of increments and space-time areas

Suppose that we have already obtained a sample of $\left(W_{s,t}, H_{s,t}, n_{s,t}\right)$.

The below procedure then generates these tuples on $[s,u]$ and $[u,t]$.

1. Since $N_{s,t}$ is independent of $\left(W_{s,t}, H_{s,t}\right)$ and $N_{s,t} = n_{s,t} \cdot |N_{s,t}|$, it can be sampled using $n_{s,t}$ and a half-normal random variable.

2. In addition, because $Z_{s,u}$ is independent of $\left(W_{s,t}, H_{s,t}, N_{s,t}\right)$ we can produce $\left(W_{s,u}, W_{u,t}, H_{s,u}, H_{u,t}\right)$ by a refinement procedure.

3. It is straightforward to generate both $n_{s,u}$ and $n_{u,t}$ as they are independent of the above and follow Rademacher distributions.

Variable step size control for the CIR process

Variable step size methodologies have seen widespread use within the numerical analysis of ODEs but appear less popular for SDEs.

That said, there are still numerous methods given in the literature.

Gaines and Lyons [5] proposed a step size control that aims for the steps to have uniform contribution to the variance of the global error.

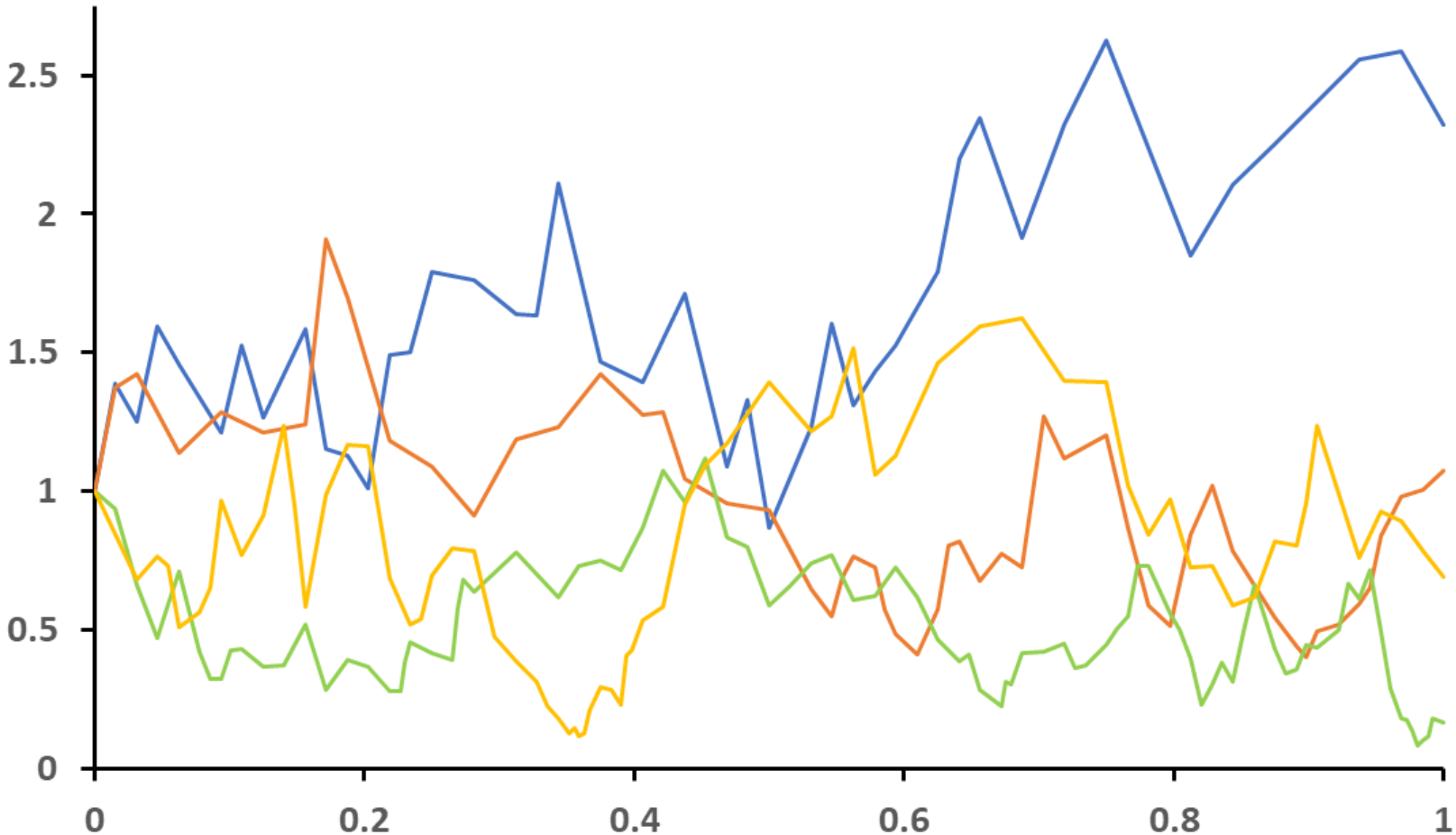
We shall emulate this approach and control each step size h so that

$$\left(\frac{a\tilde{b}\sigma^2}{2y_s}\right)^2 \mathrm{Var}\left(L_{s,t} \,|\, W_{s,t}, H_{s,t}, n_{s,t}\right) \leq Ch, \tag{16}$$

for all discretization intervals $[s,t]$ and some prespecified constant C.

Variable step size control for the CIR process

For simplicity, we will use the time horizon T as the initial step size.



CIR sample paths with $a = 1$, $b = 1$, $\sigma^2 = 3$ and the same control C.

Numerical results

We shall compare the proposed variable step size discretization with the best performing method given in [6] using the below parameters:

Low volatility case: $a = 1,\ b = 1,\ \sigma^2 = 1,\ y_0 = 1,\ T = 1.$

High volatility case: $a = 1,\ b = 1,\ \sigma^2 = 3,\ y_0 = 1,\ T = 1.$

The method that we are comparing against is given by the formula:

$$Y_{k+1} := \left(\frac{\frac{1}{2}\sigma W_{t_k,t_{k+1}} + \sqrt{Y_k} + \sqrt{\left(\frac{1}{2}\sigma W_{t_k,t_{k+1}} + \sqrt{Y_k}\right)^2 + 2a\tilde{b}h\left(1+\frac{1}{2}ah\right)}}{2\left(1+\frac{1}{2}ah\right)} \right)^2,$$

$$Y_0 := y_0,$$

for $k \in [0\,..\,N-1]$ where $t_k := kh$ and $h := \frac{T}{N}$.

Numerical results

We examine the strong and weak convergence using the estimators:

$$S_N := \sqrt{\mathbb{E}\left[\left(Y_N - Y_T^{\text{fine}}\right)^2\right]},$$

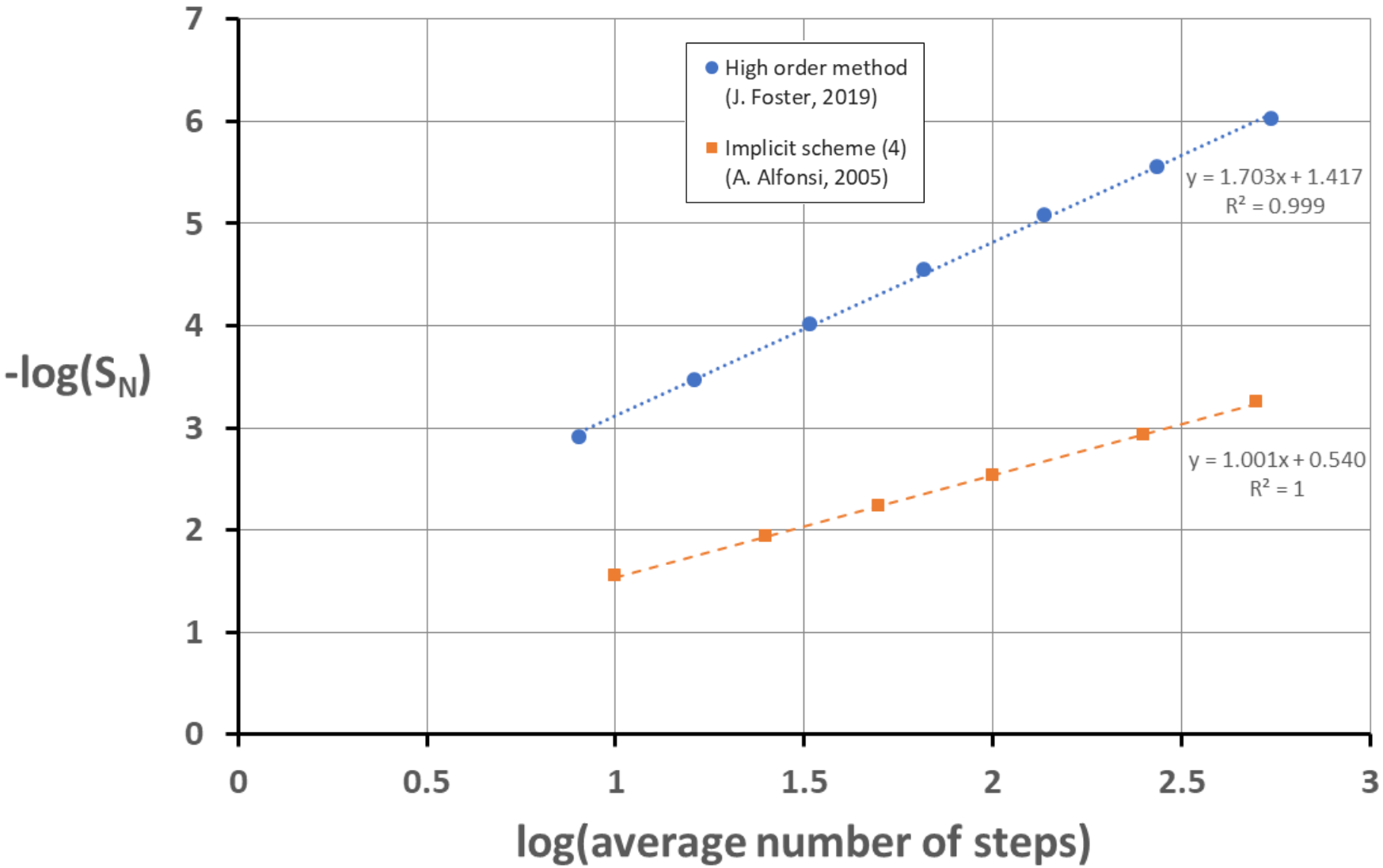
$$E_N := \left|\mathbb{E}\left[\left(Y_N - b\right)^+\right] - \mathbb{E}\left[\left(Y_T^{\text{fine}} - b\right)^+\right]\right|,$$

where the expectations are approximated by Monte-Carlo simulation and Y_T^{fine} denotes the numerical solution of (1) obtained at time T using this numerical scheme with a "fine" step size of $\min\left(\frac{h}{10}, \frac{T}{5000}\right)$.

We will compute both Y_N and Y_T^{fine} using the same Brownian paths.

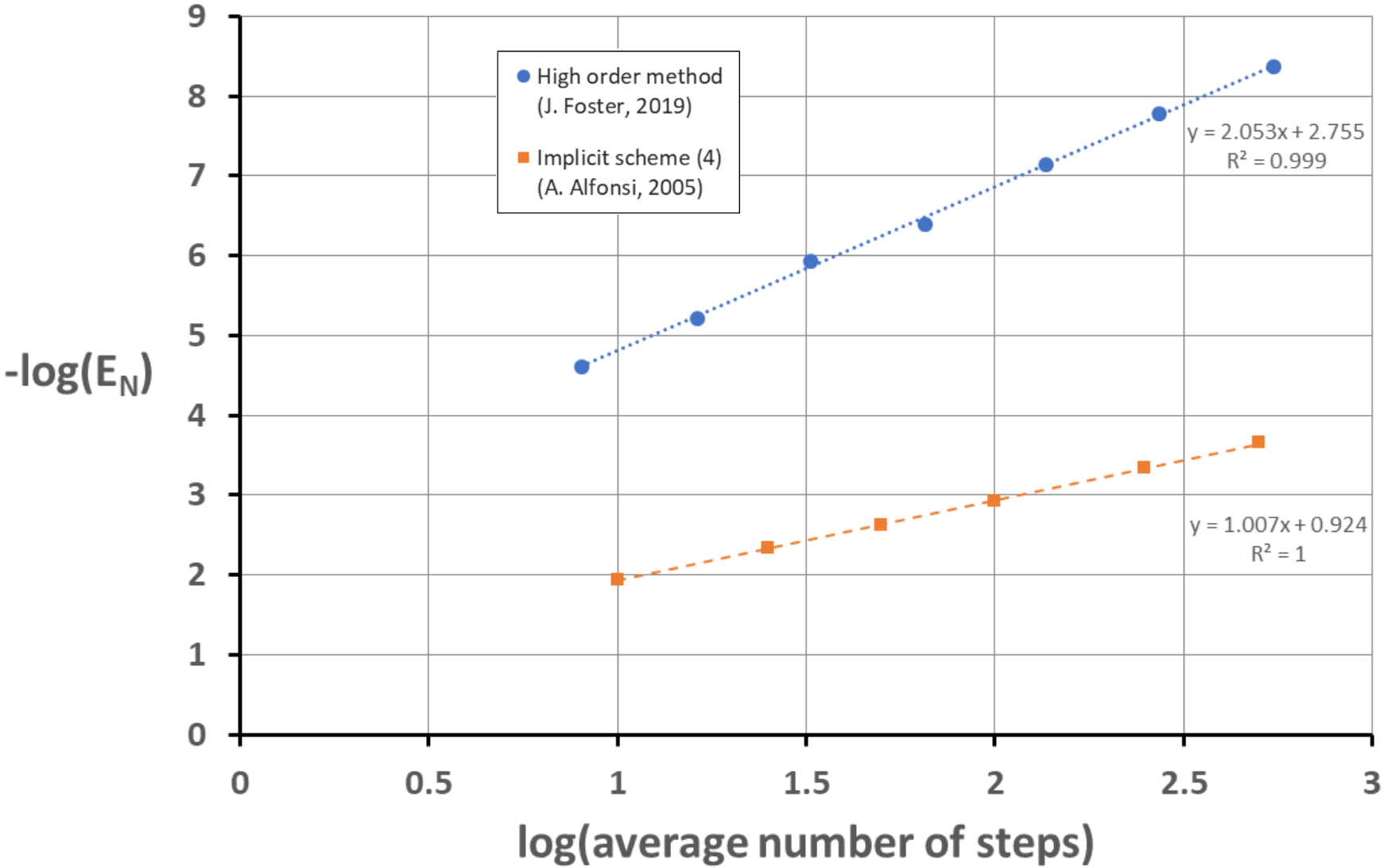
Similarly, we can estimate the above errors for the proposed method. The key difference is that now Y_T^{fine} is computed with variable steps that are bounded by $\min\left(\frac{h}{64}, \frac{T}{4096}\right)$, where h is the "crude" step size.

Numerical results



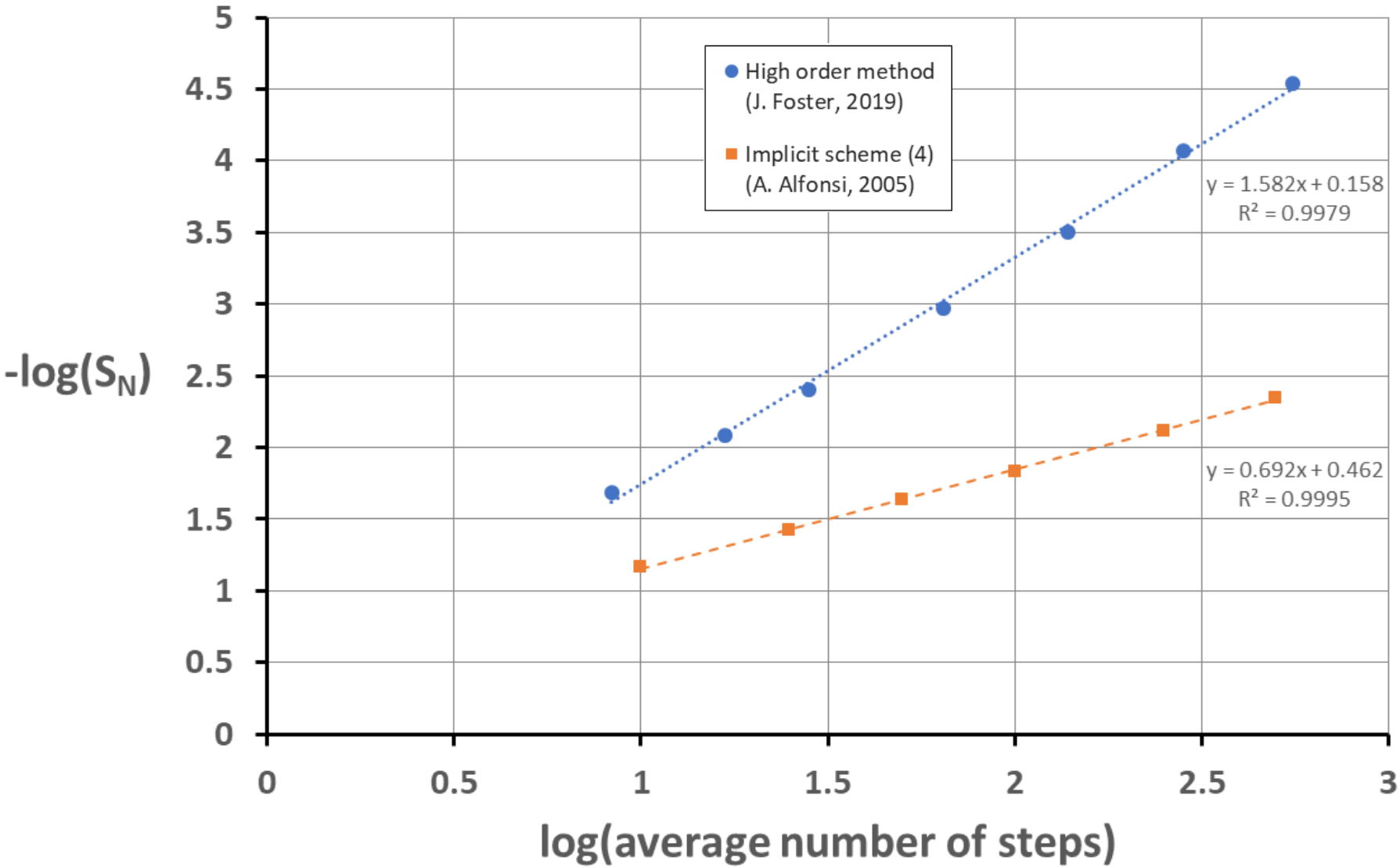
S_N computed using 100,000 sample paths in the low volatility case.

Numerical results



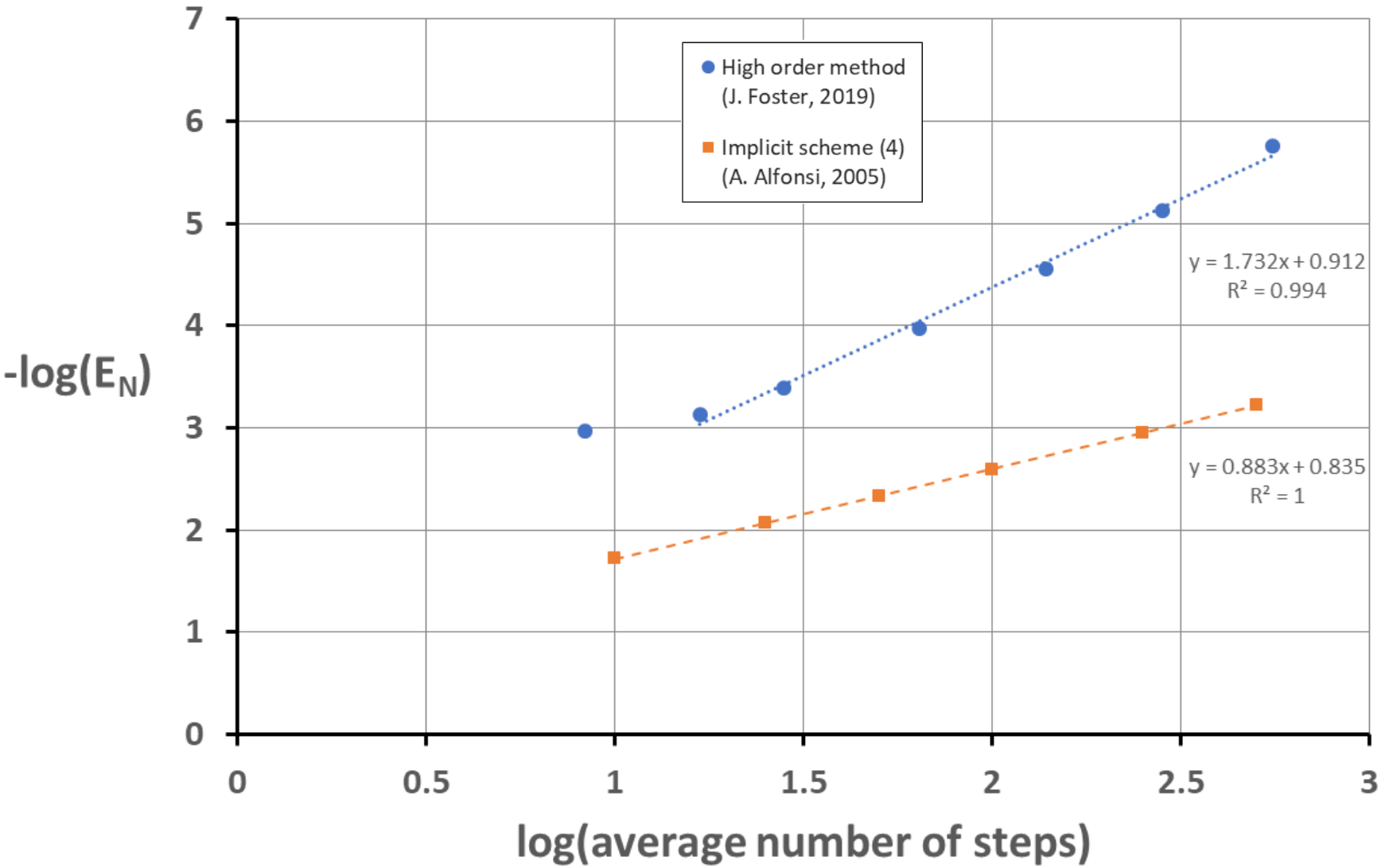
E_N computed using 100,000 sample paths in the low volatility case.

Numerical results



S_N computed using 100,000 sample paths in the high volatility case.

Numerical results



E_N computed using 100,000 sample paths in the high volatility case.

Conclusion and future work

By incorporating piecewise linear discretizations of Brownian motion into a variable step size framework, we have developed a high order (based on empirical evidence) numerical method for the CIR model.

Moreover, this research immediately leads to several open questions:

- Can one prove convergence for the ODE approximation (3)?
- Are there better piecewise linear paths for simulating SDEs?
- What improvements could be made to the step size control?
- How should this method be extended for the Heston model?

Thank you
for your attention!

References I

[1] D. Brigo and F. Mercurio, *Interest Rate Models: Theory and Practice*, Springer, 2001.

[2] J. Foster, T. Lyons and H. Oberhauser, *An optimal polynomial approximation of Brownian motion*, arXiv:1904.06998, 2019.

[3] C. A. Kennedy and M. H. Carpenter, *Diagonally Implicit Runge-Kutta Methods For Ordinary Differential Equations: A Review*, NASA Scientific and Technical Information, Hampton, 2016.

[4] P. M. Burrage and K. Burrage, *A variable stepsize implementation for stochastic differential equations*, SIAM Journal on Scientific Computing, Volume 24, 2002.

References II

[5] J. G. Gaines and T. J. Lyons, *Variable step size control in the numerical solution of stochastic differential equations*, SIAM Journal on Applied Mathematics, 1997.

[6] A. Alfonsi, *On the discretization schemes for the CIR (and Bessel squared) processes*, Monte Carlo Methods and Applications, Volume 11, 2005.