An optimal polynomial approximation of Brownian motion

James Foster

Supervised by Terry Lyons and Harald Oberhauser

University of Oxford

2019

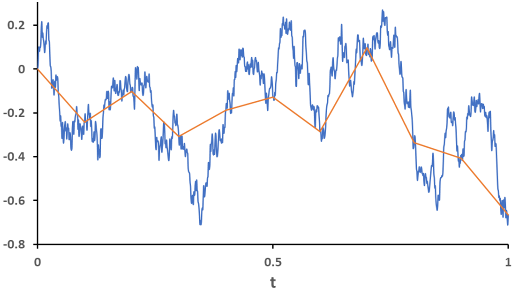
Introduction

Consider a standard real-valued Brownian motion W on an interval.

Theorem (Conditional expectation of Brownian motion)

For $0 \leq s < t$, *we have*

$$\mathbb{E}\left[W_u \,|\, W_s, W_t\right] = W_s + \frac{u-s}{t-s} \cdot W_{s,t}, \quad \forall u \in [s,t].$$



Question

Are there better discrete approximations of W than piecewise linear?

Introduction

The next simplest approximant would be the piecewise polynomial.

This was explored to an extent by Grebenkov, Belyaev and Jones in their 2015 paper "A multiscale guide to Brownian motion" (see [1]).

However they only prove the following theorem from [2] when $n \leq 2$.

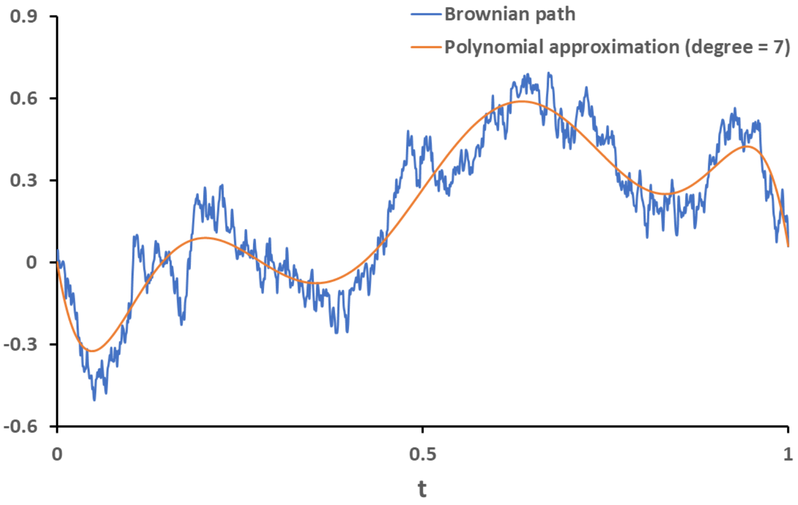
Theorem (Brownian motion as a polynomial with added noise)

Consider a standard Brownian motion W over the unit interval $[0,1]$. Let W^n be the unique n-th degree polynomial with a root at 0 and

$$\int_0^1 u^k \, dW_u^n = \int_0^1 u^k \, dW_u, \quad \text{for } k = 0, 1, \cdots, n-1.$$

Then $W = W^n + Z^n$, where Z^n is a centred Gaussian process that is independent of W^n.

Introduction



Introduction

In fact, we will be proving a stronger (and more useful) result which is a "polynomial" Karhunen-Loéve theorem for the Brownian bridge.

To achieve this, let B denote the standard Brownian bridge on $[0,1]$ and consider the Borel measure μ given by

$$\mu(a,b) := \int_a^b \frac{1}{x(1-x)}\,dx, \quad \text{for all open intervals } (a,b) \subset [0,1].$$

It's also worth mentioning that B is a square μ-integrable process as

$$\mathbb{E}\left[\int_0^1 (B_s)^2\,d\mu(s)\right] = \int_0^1 \mathbb{E}\left[(B_s)^2\right]\,d\mu(s) = \int_0^1 s(1-s)\cdot\frac{1}{s(1-s)}\,ds = 1.$$

Main result

Theorem (A Karhunen-Loéve theorem for the Brownian bridge)

There exists a family of polynomials $\{e_k\}_{k\geq 1}$ *with* $\deg(e_k) = k+1$ *and*

$$\int_0^1 e_i\, e_j\, d\mu = \delta_{ij},$$

such that

$$B = \sum_{k=1}^{\infty} I_k e_k,$$

where $\{I_k\}$ *denotes the collection of independent centered Gaussian random variables with*

$$I_k := \int_0^1 B_t \cdot \frac{e_k(t)}{t(1-t)}\, dt,$$

and

$$\mathrm{Var}(I_k) = \frac{1}{k(k+1)}.$$

Proof of main result

As with the standard argument, we define an integral operator from the Brownian bridge's covariance function $K_B(s,t) := \min(s,t) - st$.

$$T_K : L^2\left([0,1],\mu\right) \to L^2\left([0,1],\mu\right),$$
$$(T_K f)(t) := \int_0^1 K_B(s,t) f(s)\, d\mu(s).$$

Since T_K is continuous, we can apply Mercer's theorem for kernels.

This tells us that there is an orthonormal set $\{e_k\}_{k\geq 1}$ of $L^2([0,1],\mu)$ consisting of eigenfunctions for T_K and the associated sequence of eigenvalues $\{\lambda_k\}_{k\geq 1}$ is non-negative [3]. Moreover, any eigenfunction with non-zero eigenvalue is continuous and K_B can be expressed as

$$K_B(s,t) = \sum_{k=1}^{\infty} \lambda_k e_k(s) e_k(t). \tag{1}$$

Proof of main result

The main part is showing each e_k is a polynomial with degree $k+1$.

$$\begin{aligned}
T_K e_k = \lambda e_k &\implies \int_0^1 \frac{\min(s,t)-st}{s(1-s)}\, e_k(s)\, ds = \lambda_k e_k(t) \\
&\implies \int_0^t \frac{1-t}{1-s}\, e_k(s)\, ds + \int_t^1 \frac{t}{s}\, e_k(s)\, ds = \lambda_k e_k(t) \\
&\implies \int_0^t \frac{-1}{1-s}\, e_k(s)\, ds + \int_t^1 \frac{1}{s}\, e_k(s)\, ds = \lambda_k e_k'(t) \\
&\implies -\frac{1}{1-t}\, e_k(t) - \frac{1}{t}\, e_k(t) = \lambda_k e_k''(t).
\end{aligned}$$

So the eigenfunction e_k satisfies the following differential equation:

$$\lambda_k e_k''(t) = -\frac{1}{t(1-t)}\, e_k(t). \tag{2}$$

Proof of main result

For $x \in [0,1]$, we define the function

$$y_k(x) := e_k'\left(\frac{1}{2}(1+x)\right).$$

It can be shown from (2) that y_k satisfies the differential equation:

$$(1-x^2)\, y_k''(x) - 2x y_k'(x) + \frac{1}{\lambda_k} y_k(x) = 0.$$

Remarkably, this is the Legendre differential equation. It now follows from classical theory that $\frac{1}{\lambda_k} = k(k+1)$ and y_k is proportional to the k-th Legendre polynomial.

So e_k is a (normalised) shifted Jacobi polynomial with degree $k+1$.

Proof of main result

The result then follows from (1) and the orthogonality of $\{e_k\}$. $\square$

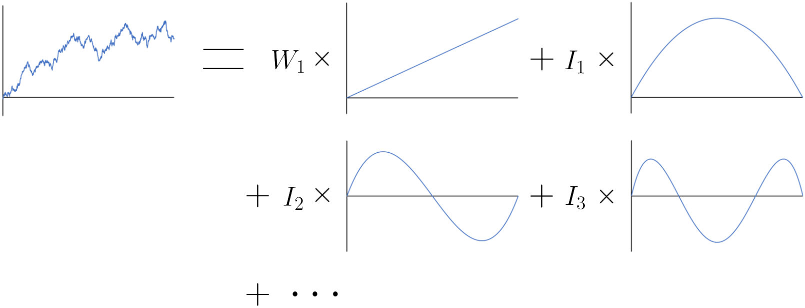
Similar to the standard Brownian bridge Karhunen-Loéve theorem, we can show $\{e_k\}$ is an optimal orthonormal basis of $L^2([0,1],\mu)$ for approximating B by truncated series expansions with respect to the following weighted $L^2(\mathbb{P})$ norm:

$$\|X\|_{L^2_\mu(\mathbb{P})} := \sqrt{\mathbb{E}\left[\int_0^1 (X_s)^2 \, d\mu(s)\right]},$$

where X is a square μ-integrable process.

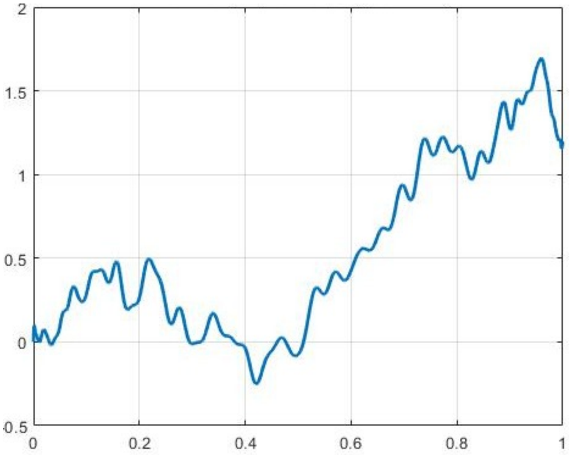
A decomposition of Brownian motion

Brownian motion is expressible as a sum of orthogonal polynomials with independent weights that capture specific features of the path.



Each weight is a sum of iterated time integrals of Brownian motion.

A Brownian polynomial (degree = 100)



These polynomials are straightforward to implement using Chebfun!

www.chebfun.org/examples/stats/RandomPolynomials.html

Brownian polynomials (degree = 2)

For developing numerical methods, we will use the below definitions:

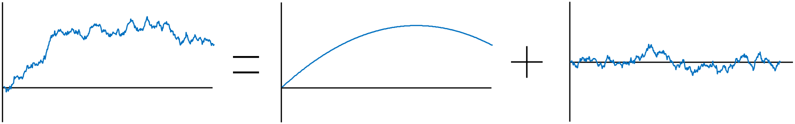
Definitions

The **standard Brownian parabola** $\widehat{W}$ denote the unique quadratic polynomial on $[0,1]$ with a root at 0 and that satisfies the following:

$$\widehat{W}_1 = W_1, \qquad \int_0^1 \widehat{W}_u\, du = \int_0^1 W_u\, du.$$

The **standard Brownian arch** is the Gaussian process $Z := W - \widehat{W}$. By the main theorem, Z is centered and has the covariance function

$$K_Z(s,t) = \min(s,t) - st - 3st(1-s)(1-t), \quad \text{for} \quad s,t, \in [0,1].$$



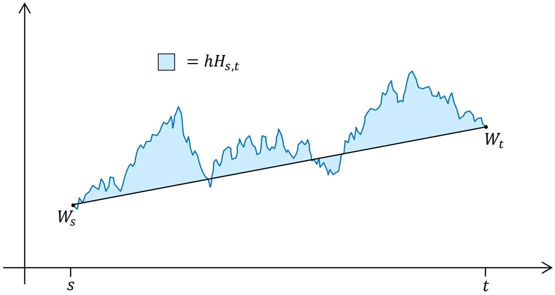
Brownian polynomials (degree = 2)

Definitions (continued)

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 $h = t - s$. As $e_1(t) = \sqrt{6}\, t(1-t)$, we see $H_{0,1}$ is equal to $\frac{\sqrt{6}}{6} I_1$ in the main result. So $H_{s,t} \sim N\left(0, \frac{1}{12}h\right)$ and is independent of $W_{s,t}$.



Applications to SDEs

Consider the (Stratonovich) stochastic differential equation given by

$$\begin{aligned} dy_t &= f_0(y_t)\,dt + f_1(y_t)\circ dW_t, \\ y_0 &= \xi, \end{aligned} \tag{3}$$

where the $f_i : \mathbb{R}^d \to \mathbb{R}^d$ denote smooth bounded vector fields on $\mathbb{R}^d$.

In order to simulate the above SDE on $[0, T]$, one typically samples the Brownian path over a uniform partition $\triangle_N = \{t_0 < t_1 < \cdots < t_N\}$.

Question

What is the best pathwise approximation of (3) that is measurable with respect to a discretization of the driving Brownian motion W?

Possible Answer

$$y_t^* := \mathbb{E}\Big[\, y_t \,\Big|\, y_0, W_{t_k, t_{k+1}}, H_{t_k, t_{k+1}} \text{ for } k \in [0 \,..\, N-1] \,\Big].$$

Applications to SDEs

Question

Can we derive high order numerical methods for approximating y^*?

In order to answer this, we consider the stochastic Taylor expansion:

$$
\begin{aligned}
y_t = y_s + f_0(y_s)\,h + f_1(y_s)\,W_{s,t} + (\cdots)\,W_{s,t}^2 + (\cdots)\,W_{s,t}^3 \\
+ (\cdots)\int_s^t\int_s^u \circ\, dW_v\, du + (\cdots)\int_s^t\int_s^u dv \circ dW_u + (\cdots)\,h^2 + (\cdots)\,W_{s,t}^4 \\
+ (\cdots)\int_s^t\int_s^u\int_s^v \circ\, dW_r \circ dW_v\, du + (\cdots)\int_s^t\int_s^u\int_s^v \circ\, dW_r\, dv \circ dW_u \\
+ (\cdots)\int_s^t\int_s^u\int_s^v dr \circ dW_v \circ dW_u + O\big(h^{\frac{5}{2}}\big),
\end{aligned}
\tag{4}
$$

where $(\cdots)$ denote terms involving f_0, f_1 as well as their derivatives.

Applications to SDEs

Thus approximating y^* is likely to require the following expectations

$$\mathbb{E}\left[\int_s^t\int_s^u\int_s^v \circ\, dW_r \circ dW_v\, du \,\middle|\, W_{s,t}, H_{s,t}\right],$$
$$\mathbb{E}\left[\int_s^t\int_s^u\int_s^v \circ\, dW_r\, dv \circ dW_u \,\middle|\, W_{s,t}, H_{s,t}\right],$$
$$\mathbb{E}\left[\int_s^t\int_s^u\int_s^v dr \circ dW_v \circ dW_u \,\middle|\, W_{s,t}, H_{s,t}\right].$$

Deriving explicit formulae for the above could lead to improvements for high order numerical methods (such as those proposed in [4, 5]).

By expressing the Brownian motion W as a (random) parabola plus independent noise, it is possible to obtain these integral estimators!

Applications to SDEs

Theorem (Conditional expectation of a Brownian time integral)

$$\mathbb{E}\left[\int_s^t W_{s,u}^2\,du\,\middle|\,W_{s,t}, H_{s,t}\right] = \frac{1}{3}hW_{s,t}^2 + hW_{s,t}H_{s,t} + \frac{6}{5}hH_{s,t}^2 + \frac{1}{15}h^2.$$

Proof.

By the natural Brownian scaling, it is enough to prove this on $[0,1]$.

$$\begin{aligned}
\mathbb{E}\left[\int_0^1 W_u^2\,du\,\middle|\,W_1, H_1\right] &= \mathbb{E}\left[\int_0^1 \left(\widehat{W}_u + Z_u\right)^2 du\,\middle|\,W_{s,t}, H_{s,t}\right]\\
&= \int_0^1 \widehat{W}_u^2\,du + 2\int_0^1 \widehat{W}_u\,\mathbb{E}\left[Z_u\right] du + \int_0^1 \mathbb{E}\left[Z_u^2\right] du\\
&= \int_0^1 \left(uW_1 + 6u(1-u)H_1\right)^2 du + 2\int_0^1 \widehat{W}_u \cdot 0\,du\\
&\quad + \int_0^1 u - u^2 - 3u^2(1-u)^2\,du.
\end{aligned}$$

The result now follows by evaluating the above integrals. □

Applications to SDEs

Definition

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

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

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

Theorem (Relationships between third order iterated integrals)

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

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

$$\int_s^t\int_s^u\int_s^v dr \circ dW_v \circ dW_u = \frac{1}{6}hW_{s,t}^2 - \frac{1}{2}hW_{s,t}H_{s,t} + L_{s,t}\,.$$

Applications to SDEs

The previous two theorems then directly give our second key result.

Theorem (Conditional expectation of a Brownian Lévy area)

$$\mathbb{E}\left[\,L_{s,t}\,|\,W_{s,t}, H_{s,t}\,\right] = \frac{3}{5}\,hH_{s,t}^2 + \frac{1}{30}\,h^2.$$

Remark

The first term corresponds to the parabola approximation whilst the second term captures a constant bias induced by the Brownian arch.

Whilst $L_{s,t}$ can be negative, this conditional expectation is positive. In the numerical example, the proposed method will exploit this fact.

In [4, 5], the area $L_{s,t}$ was approximated using $\mathbb{E}\left[\,L_{s,t}\,|\,W_{s,t}\,\right] = \frac{1}{12}\,h^2$.

Applications to SDEs

We will incorporate this estimator into the log-ODE method [4, 5].

Definition (A high order log-ODE method)

We can define a numerical solution $\{Y_k\}_{0\leq k\leq N}$ of (3) by fixing $Y_0 = \xi$ and for all $k \in [0\,..\,N-1]$, setting Y_{k+1} to be the solution at $u = 1$ of

$$\begin{aligned}\frac{dz}{du} &= f_0(z)\,h + f_1(z)\,W_{t_k,t_{k+1}} + \big[f_1, f_0\big](z)\cdot hH_{t_k,t_{k+1}} \\ &\qquad\qquad + \big[f_1,\big[f_1, f_0\big]\big](z)\cdot\mathbb{E}\big[L_{t_k,t_{k+1}}\,|\,W_{t_k,t_{k+1}}, H_{t_k,t_{k+1}}\big], \\ z_0 &= Y_k,\end{aligned} \tag{5}$$

where $h := \frac{T}{N}$, $t_k = kh$ and $[\cdot,\cdot]$ denotes the vector field Lie bracket.

Applications to SDEs

By considering the Taylor expansion of the ODE (5), one can prove:

Theorem

This new version of the log-ODE method convergences in a strong sense with order 1.5 and a weak sense with order 2. In other words,

1. *There exists a constant C, that is independent of N, such that*

$$\sqrt{\mathbb{E}\left[\left(Y_N^{log} - y_T\right)^2\right]} \leq C h^{\frac{3}{2}},$$

   for all sufficiently small step sizes $h = \frac{T}{N}$.

2. *For any polynomial p there exists a constant $C_p > 0$, such that*

$$\left|\mathbb{E}\left[p\left(Y_N^{log}\right)\right] - \mathbb{E}\left[p\left(y_T\right)\right]\right| \leq C_p h^2,$$

   for all sufficiently small step sizes $h = \frac{T}{N}$.

Numerical example: IGBM

We will demonstrate the effectiveness of the log-ODE method by discretizing Inhomogeneous Geometric Brownian Motion (IGBM).

$$dy_t = a\big(b - y_t\big)\,dt + \sigma y_t\,dW_t, \tag{6}$$

where $a, b \geq 0$ are mean reversion parameters and σ is the volatility.

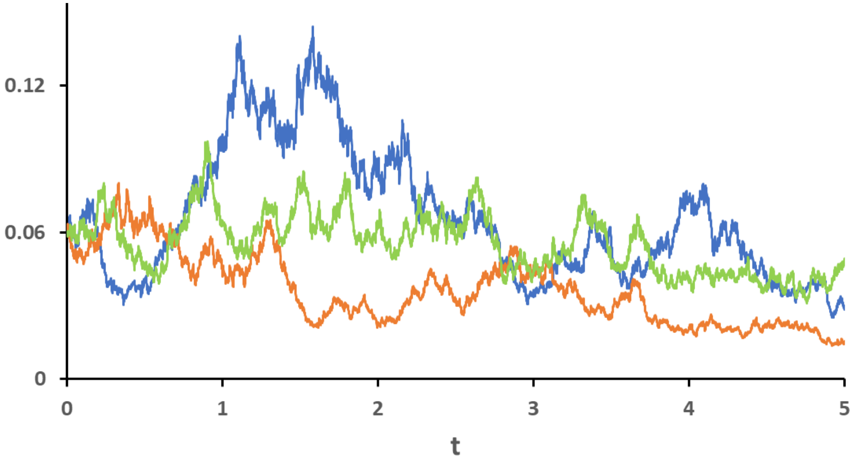
IGBM is an example of a short rate model and has seen attention recently in the literature as an alternative to popular models [6, 7].

Due to smooth vector fields, we can write (6) in Stratonovich form:

$$dy_t = \tilde{a}\big(\tilde{b} - y_t\big)\,dt + \sigma y_t \circ dW_t, \tag{7}$$

where $\tilde{a} := a + \frac{1}{2}\sigma^2$ and $\tilde{b} := \frac{2ab}{2a+\sigma^2}$ denote the "adjusted" parameters.

Numerical example: IGBM



Log-ODE sample paths of IGBM with $a = 0.1$, $b = 0.04$ and $\sigma = 0.6$.

Numerical example: IGBM

1. Log-ODE method

$$Y_{k+1}^{\mathsf{log}} := Y_k^{\mathsf{log}} e^{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}}$$
$$+ abh\left(1-\sigma H_{t_k,t_{k+1}} + \sigma^2\left(\frac{3}{5}H_{t_k,t_{k+1}}^2 + \frac{1}{30}h\right)\right)\frac{e^{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}}-1}{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}}.$$

2. Parabola-ODE method with 3-point Gauss-Legendre quadrature

$$Y_{k+1}^{\mathsf{para}} := e^{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}}\left(Y_k^{\mathsf{para}} + ab\int_{t_k}^{t_{k+1}} e^{\tilde{a}(s-t_k)-\sigma\widehat{W}_{t_k,s}}\,ds\right).$$

3. Piecewise linear method

$$Y_{k+1}^{\mathsf{lin}} := Y_k^{\mathsf{lin}} e^{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}} + abh\,\frac{e^{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}}-1}{-\tilde{a}h+\sigma W_{t_k,t_{k+1}}}.$$

4. Milstein method
5. Euler-Maruyama method

} with positive part taken if necessary.

Numerical example: IGBM

We examine the strong and weak convergence using the estimators:

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

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

where the expectations are approximated by Monte-Carlo simulation and Y_T^{fine} denotes the numerical solution of (7) obtained at time T using the log-ODE method with a "fine" step size of $\min\left(\frac{h}{10}, \frac{T}{1000}\right)$.

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

The experiment shall use the same parameter values as [6], namely $a = 0.1$, $b = 0.04$, $\sigma = 0.6$ and $y_0 = 0.06$. The end time will be $T = 5$.

Numerical example: IGBM

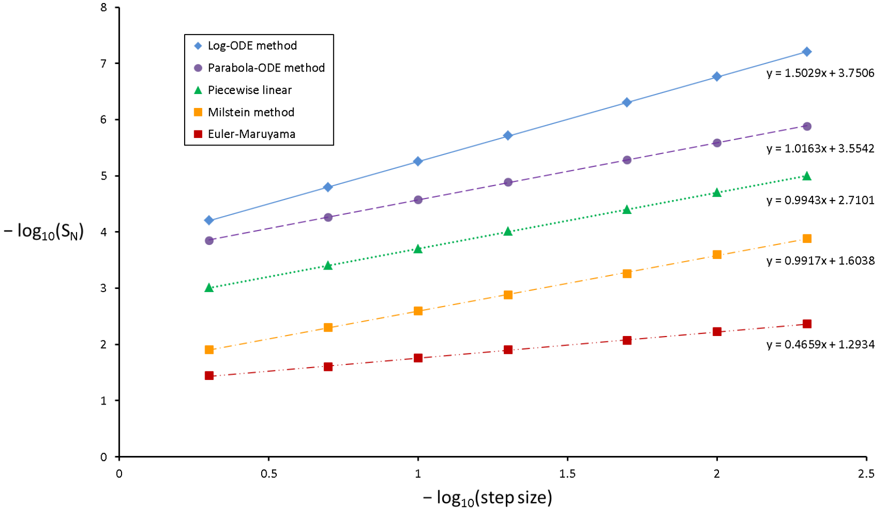


Figure: S_N computed with 100,000 sample paths using a step size $h = \frac{T}{N}$.

Numerical example: IGBM

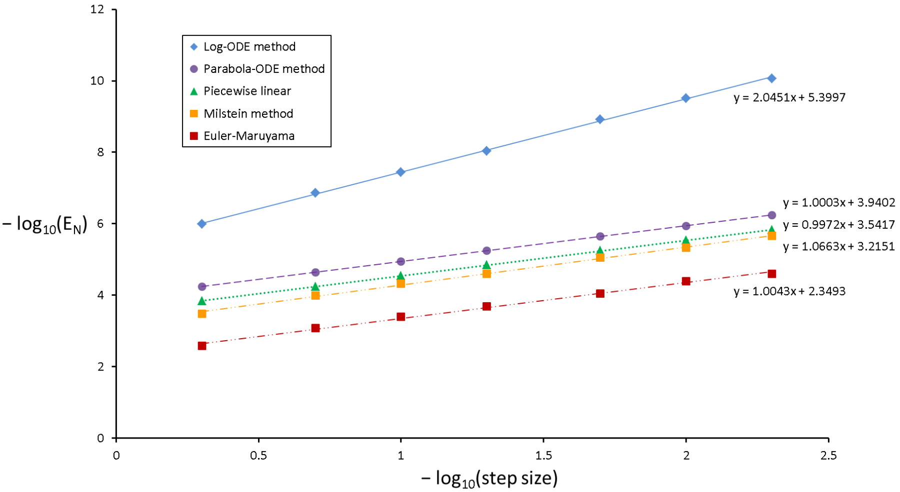


Figure: E_N computed with 500,000 sample paths using a step size $h = \frac{T}{N}$.

Numerical example: IGBM

Table: Estimated times for computing 100,000 sample paths that achieve a specified accuracy using a single-threaded C++ program on a desktop.

	Log-ODE	Parabola	Linear	Milstein	Euler
Estimated time to achieve an accuracy of $S_N = 10^{-4}$	0.179 (s)	0.405 (s)	1.47 (s)	15.4 (s)	0.437 (days)
Estimated time to achieve an accuracy of $S_N = 10^{-5}$	0.827 (s)	3.90 (s)	14.8 (s)	157 (s)	61.2 (days)

The above times were estimated from the graph and following table:

Table: Simulation times to compute 100,000 sample paths, with 100 steps for each path, by a single-threaded C++ program on a desktop computer.

	Log-ODE	Parabola	Linear	Milstein	Euler
Computation time (s)	2.44	2.95	1.48	1.18	1.17

Conclusion and future work

We have shown that Brownian motion can be expressed as a random polynomial (defined using certain integrals) plus independent noise.

By developing a state-of-the-art discretization of the IGBM process, we have demonstrated this result has applications in SDE numerics.

Moreover, this research immediately leads to several open questions:

- Can one extend the main theorem to other Gaussian processes?
- What are the most efficient Runge-Kutta methods for general SDEs which correctly use the new estimator for triple integrals?
- Do the polynomials give optimal approximations of Lévy area?

Thank you
for your attention!

(and see the paper on ArXiv or ResearchGate)

References I

[1] D. S. Grebenkov, D. Belyaev and P. W. Jones, *A multiscale guide to Brownian motion*, Journal of Physics A, Volume 49, 2015.

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

[3] C. Carmeli, A. De Vito, and E. Toigo, *Vector valued reproducing kernel Hilbert spaces of integrable functions and Mercer theorem*, Analysis and Applications, Volume 4, 2006.

[4] F. Castell and J. G. Gaines, *The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations*, Annales de l'Institut Henri Poincaré, Volume 32, 1996.

References II

[5] S. J. A. Malham and A. Wiese, *Stochastic Lie group integrators*, Siam Journal on Scientific Computing, Volume 30, 2008.

[6] L. Capriotti, Y. Jiang and G. Shaimerdenova, *Approximation Methods for Inhomogeneous Geometric Brownian Motion*, International Journal of Theoretical and Applied Finance, Volume 22, 2019.

[7] B. Zhao, C. Yan and S. Hodges, *Three One-Factor Processes for Option Pricing with a Mean-Reverting Underlying: The Case of VIX*, Financial Review, Volume 54, 2019.