The Spectral Hermite Transform as a Means of Simulating Brownian Motion

Report for:
NERS 590
Stochastic Differential Equations

Submission Date
April 23rd, 2024

Author
M.A. Johnson
University of Michigan - Ann Arbor

UNIVERSITY OF MICHIGAN

The object of my graduate research has been the "<u>Spec</u>tral <u>Her</u><u>m</u>it<u>e</u> <u>T</u>ransform", or "Sechmet" (where the *ch* sounds like the German *doch*). It is defined as a linear combination over a grid of N regularly-spaced Hermite function nodes, where each node contains a sum of Hermite functions from order 0 up through order M. In formulas, this looks like

$$\begin{aligned}
\text{𓌂}_N^M\Big[f(\bar{x})\Big](x) &= \mu \sum_{n=0}^{N-1} \sum_{m=0}^{M} f_n^m \, \mathcal{H}_\sigma^m\Big(\frac{x - n\,\mu}{\sigma}\Big), \\
f_n^m &= \frac{1}{m!} \int_{-\infty}^{\infty} f(\bar{x}) \, \mathcal{H}_\sigma^m\Big(\frac{\bar{x} - n\,\mu}{\sigma}\Big) \, \mathrm{d}\bar{x}, \text{ and} \\
\mathcal{H}_C^m(x) &= \Big(-\frac{\mathrm{d}}{\mathrm{d}x}\Big)^m \Big[\frac{1}{C\sqrt{2\pi}} \, \mathrm{e}^{-\frac{x^2}{2}}\Big] = -\frac{\mathrm{d}}{\mathrm{d}x}\Big(\mathcal{H}_C^{m-1}(x)\Big).
\end{aligned} \tag{1}$$

Here, 𓌂 is the "sekhem" hieroglyph; the notation is intended to suggest that the Sechmet is a transformation of $f(\bar{x})$ [in the barred variable] into a new function in x [without the barred variable]; μ represents the grid spacing between nodes; σ gives the width of each Hermite function; and $\delta_x = \frac{\mu}{\sigma}$ (which is typically of order unity) determines the extent to which neighboring Hermite functions overlap. I have previously proven that

$$\text{𓌂}_N^M\Big[f(\bar{x})\Big](x) \approx K_M \int_{-\infty}^{\infty} f(\bar{x}) \, \mathcal{H}_{2\sigma}^0\Big(\frac{x - \bar{x}}{2\sigma}\Big) \, \mathrm{sinc}\Big(\frac{x - \bar{x}}{\sigma_M}\Big) \, \mathrm{d}\bar{x}, \tag{2}$$

for $0 \le x \le N\mu$, where $K_M = 4\sqrt{\frac{M+0.75}{2\pi}}$, $\sigma_M = \frac{\sigma}{\sqrt{M+0.75}}$, and $\mathrm{sinc}(x) \equiv \frac{\sin(x)}{x}$ is the unnormalized sinc function. Equation (2) shows that the Sechmet acts as a sinc-Gaussian filter of width $\sim\sigma$ applied to $f(x)$, i.e. it passes signals of frequency $\ll \frac{1}{\sigma}$, blocks signals of frequency $\gg \frac{1}{\sigma}$, and partially attenuates signals of frequency $\sim \frac{1}{\sigma}$. (Note that since δ_x is typically of order unity, we could replace occurrences of σ in the above with μ.)

I have used the Sechmet in my research as a way of decomposing an arbitrary function into a linear combination of Hermite functions, which then have a wealth of useful properties for constructing a numerical scheme. However, my numerical experiments have indicated that assigning Gaussian-distributed random numbers to the coefficients f_n^m provides an efficient algorithm for Brownian motion simulations! (This came as quite a surprise to me, because I didn't create the Sechmet with SDEs in mind.) The goal of this paper will be to demonstrate that if we define $f(x)$ via

$$\begin{aligned}
f'(x) &= \mu \sum_{n=0}^{N-1} \sum_{m=0}^{M} \frac{f_n^m}{\sqrt{m!\,\mu}} \, \mathcal{H}_\sigma^m\Big(\frac{x - n\,\mu}{\sigma}\Big), \\
f(x) &= \int_0^x f'(y) \, \mathrm{d}y, \text{ and} \\
f_n^m &\sim \mathcal{N}(0; 1),
\end{aligned} \tag{3}$$

then $f(x) \approx B_x$ down to a scale of roughly μ, as Equation (2) would predict. Inserting the Hermite function definition in Equation (1) into the middle line of Equation (3), and noting that $\mathcal{H}_1^{-1}\big(\frac{-n\mu}{\sigma}\big) - \mathcal{H}_1^{-1}\big(\frac{x-n\mu}{\sigma}\big)$ looks like a unit step function at $x = n\,\mu$ that has been Gaussian-smoothed to a scale of σ,

$$\begin{aligned}
f(x) &= \int_0^x \mu \sum_{n=0}^{N-1} \sum_{m=0}^{M} \frac{f_n^m}{\sqrt{m!\,\mu}} \, \mathcal{H}_\sigma^m\Big(\frac{y - n\,\mu}{\sigma}\Big) \, \mathrm{d}y \\
&= \sqrt{\mu} \sum_{n=0}^{N-1} \sum_{m=0}^{M} \frac{f_n^m}{\sqrt{m!}} \Big[\mathcal{H}_1^{m-1}\Big(\frac{-n\,\mu}{\sigma}\Big) - \mathcal{H}_1^{m-1}\Big(\frac{x - n\,\mu}{\sigma}\Big)\Big] \\
&= \sqrt{\mu} \sum_{n=0}^{N-1} f_n^0 \Big[\mathcal{H}_1^{-1}\Big(\frac{-n\,\mu}{\sigma}\Big) - \mathcal{H}_1^{-1}\Big(\frac{x - n\,\mu}{\sigma}\Big)\Big] + \sum_{m=1}^{M} \frac{f_n^m}{\sqrt{m!}} \Big[\mathcal{H}_1^{m-1}\Big(\frac{-n\,\mu}{\sigma}\Big) - \mathcal{H}_1^{m-1}\Big(\frac{x - n\,\mu}{\sigma}\Big)\Big]
\end{aligned}$$

$$
\begin{aligned}
&\approx \sqrt{\mu}\sum_{n=0}^{N-1} f_n^0\,\mathrm{u}(x-n\,\mu) + \sum_{m=1}^{M}\frac{f_n^m}{\sqrt{m!}}\left[\mathcal{H}_1^{m-1}\!\Big(\frac{-n\,\mu}{\sigma}\Big) - \mathcal{H}_1^{m-1}\!\Big(\frac{x-n\,\mu}{\sigma}\Big)\right] \\
&= \left(\sqrt{\mu}\sum_{n=0}^{\lfloor x/\mu\rfloor} f_n^0\right) + \left(\sqrt{\mu}\sum_{n=0}^{N-1}\sum_{m=1}^{M}\frac{f_n^m}{\sqrt{m!}}\left[\mathcal{H}_1^{m-1}\!\Big(\frac{-n\,\mu}{\sigma}\Big) - \mathcal{H}_1^{m-1}\!\Big(\frac{x-n\,\mu}{\sigma}\Big)\right]\right) \\
&\equiv f_{\text{LR}}(x) + f_{\text{SR}}(x).
\end{aligned} \tag{4}
$$

Equation (4) shows that there are two components to $f(x)$: the long-range component $f_{\text{LR}}(x)$ and the short-range component $f_{\text{SR}}(x)$. All of the terms in $f_{\text{SR}}(x)$ contain a super-exponential decay factor that only allows each term to contribute if x is within a few σ-lengths of that term's center at $x = n\,\mu$; thus, $f_{\text{SR}}(x)$ exists only to increase the short-range noise of $f(x)$. By contrast, $f_{\text{LR}}(x)$ is roughly a sum over all of the $m = 0$ coefficients from node $n = 0$ up through node $n = \lfloor x/\mu \rfloor$, so $f_{\text{LR}}(x)$ incorporates all of the long-range information in $f(x)$ on the interval $[0, x]$. It will thus suffice to show that $f_{\text{LR}}(x)$ approximately satisfies the requirements of Brownian motion given in [1] to prove that $f(x)$ approximates Brownian motion, since in the limit $N \to \infty$, the contributions from $f_{\text{SR}}(x)$ will vanish. These four requirements on $f(x)$ are:

- $f(0) = 0$. This is trivially true from the definition of $f(x)$.
- $f(x)$ is continuous. This is trivially true from the definition of $f(x)$.
- $f(x)$ has independent non-overlapping increments. For $0 \le y_2 < x_2 < y_1 < x_1 \le 1$, since $f_{\text{LR}}(x_1 - y_1)$ and $f_{\text{LR}}(x_2 - y_2)$ are approximately sums over distinct sets of independently-generated random variables, the increments $y_1 \le x \le x_1$ and $y_2 \le x \le x_2$ must be approximately independent. This approximation becomes exact as $N \to \infty$ with fixed δ_x, or equivalently, as $\mu \to 0$ and $\sigma \to 0$.
- $f(x) - f(y) \sim f(x-y) \sim \mathcal{N}(0; x-y)$. Let x and y satisfy $0 \le y < x \le 1$; then

$$
\begin{aligned}
f_{\text{LR}}(x) - f_{\text{LR}}(y) &\approx \left[\sqrt{\mu}\sum_{n=0}^{\lfloor x/\mu\rfloor} f_n^0\right] - \left[\sqrt{\mu}\sum_{n=0}^{\lfloor y/\mu\rfloor} f_n^0\right] \\
&= \sqrt{\mu}\sum_{n=\lfloor y/\mu\rfloor+1}^{\lfloor x/\mu\rfloor} f_n^0 \\
&\Big\downarrow \text{ sums over same-sized sequences of } f_n^0 \text{ are distributionally equivalent} \\
&\stackrel{\text{d}}{=} \sqrt{\mu}\sum_{n=0}^{\lfloor x/\mu\rfloor-\lfloor y/\mu\rfloor-1} f_n^0 \\
&\approx f_{\text{LR}}(x-y-\mu) \\
&\approx f_{\text{LR}}(x-y).
\end{aligned}
$$

Moreover, since $\frac{f_{\text{LR}}(x)}{\sqrt{\mu}}$ is approximately a sum over $\left\lfloor \frac{x}{\mu} \right\rfloor$ normally-distributed random variables with mean 0 and variance 1, the Central Limit Theorem says that when $x \gg \mu$, we have

$$
\begin{aligned}
\frac{f_{\text{LR}}(x)}{\sqrt{\mu}\cdot\sqrt{\left\lfloor \frac{x}{\mu}\right\rfloor}} &\approx \mathcal{N}(0;1) \\
f_{\text{LR}}(x) &\approx \sqrt{\mu\left\lfloor\frac{x}{\mu}\right\rfloor}\cdot\mathcal{N}(0;1) \\
&\approx \sqrt{x}\cdot\mathcal{N}(0;1) = \mathcal{N}(0;x).
\end{aligned}
$$

Thus, $f(x) - f(y) \approx f(x-y) \approx \mathcal{N}(0; x-y)$, and all of these approximations become exact as $N \to \infty$ with fixed δ_x, or equivalently, as $\mu \to 0$ and $\sigma \to 0$.

We thus see that the Sechmet representation of $f(x)$ has all the necessary properties to simulate Brownian motion down to a length scale of μ. I have observed in numerical experiments that increasing M gives better Brownian motion properties (in e.g. the integrals to be discussed shortly), but increasing M beyond about 4 adds computational cost without significantly improving these properties. Similarly, increasing δ_x above 1 seems to help with Brownian motion properties, but this effect caps out for values of δ_x around 2 to 3.

With these facts in mind, let us take $N = 5{,}000$, $M = 4$, and $\delta_x = 2.0$ for $f(x)$ on the unit interval and use MATLAB's randn() function for $f(x)$'s normally-distributed random coefficients. This produces the following approximation to Brownian motion:

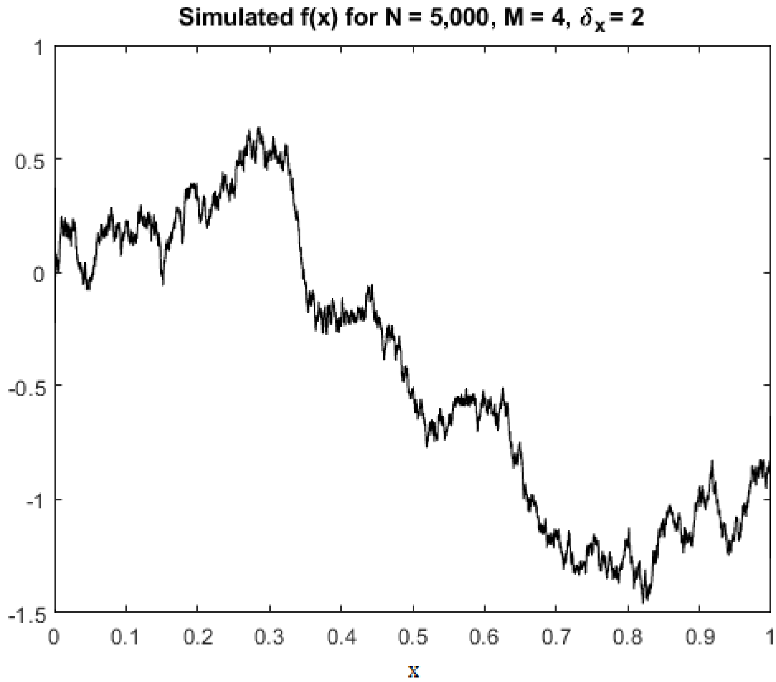

Figure 1: A Sechmet simulation of Brownian motion over the unit interval.

Figure 1 has the correct qualitative appearance; let's test whether $f(x)$ follows some simple, known integral properties via numerical quadrature. If $f(x) \approx B_x$ is true, then we should observe that $f(x)$ approximately obeys

$$\textstyle\int_I \equiv \int_0^x B_s \, \mathrm{d}B_s = \frac{1}{2}\left(B_x^2 - x\right) \quad \text{(Itô's Integral), and} \tag{5}$$

$$\textstyle\int_S \equiv \int_0^x B_s \circ \mathrm{d}B_s = \frac{1}{2} B_x^2 \quad \text{(Stratonovich's Integral),} \tag{6}$$

depending, respectively, on whether we use a left-hand Riemann rule or a centered Riemann rule for numerical integration. For the $f(x)$ in Figure 1, analytical calculation gives Equations (5) and (6) as $\int_I \approx \frac{1}{2}\left(f(1)^2 - 1\right) = -0.12170462$ and $\int_S \approx \frac{1}{2} f(1)^2 = 0.37829537$ when integrating up to $x = 1$. If we take the left-hand (i.e. Itô style) Riemann sum over Q sub-intervals of $f(x)$ on $0 \le x \le 1$ to be

$$\Sigma_I = \sum_{n=0}^{Q-1} f\left(x_n^L\right) \Big(f\left(x_n^R\right) - f\left(x_n^L\right)\Big), \tag{7}$$

and if we take the centered (i.e. Stratonovich style) Riemann sum to be

$$\Sigma_S = \sum_{n=0}^{Q-1} f\left(x_n^M\right) \Big(f\left(x_n^R\right) - f\left(x_n^L\right)\Big), \tag{8}$$

where in both Equations (7) and (8) we define

$$x_n^L = \frac{n}{Q}\,,$$
$$x_n^R = \frac{n+1}{Q}\,, \text{ and}$$
$$x_n^M = \frac{n+0.5}{Q}\,, \tag{9}$$

then we find the following:

Q		Σ_I	$\Sigma_I - \int_I$	$\Sigma_I - \int_S$		Σ_S	$\Sigma_S - \int_S$
50		-0.12336329	-0.00165867	0.50165867		0.36344031	-0.01485506
100		-0.12476829	-0.00306367	0.50306367		0.44360958	0.06531420
200		-0.11784864	0.00385597	0.49614402		0.30480419	-0.07349118
500		-0.07685401	0.04485060	0.45514939		0.35517576	-0.02311961
1,000		-0.09530174	0.02640288	0.47359711		0.38752541	0.00923003
2,000		-0.10844267	0.01326195	0.48673804		0.39848862	0.02019324
5,000		-0.04977184	0.07193277	0.42806722		0.38118503	0.00288966
10,000		0.07599573	0.19770035	0.30229964		0.37830450	0.00000912
20,000		0.20949159	0.33119621	0.16880378		0.37829447	-0.00000090
50,000		0.30861696	0.43032158	0.06967841		0.37829523	-0.00000014
100,000		0.34329727	0.46500189	0.03499810		0.37829534	-0.00000003

Table 1: Numerical results for the Itô style Riemann sums Σ_I and Stratonovich style Riemann sums Σ_S using Q integration intervals. These are compared with the differences between the Riemann sums and the analytically calculated Itô/Stratonovich integrals $\int_I = -0.12170462$ and $\int_S = 0.37829537$. These results are shown graphically in Figure 2.

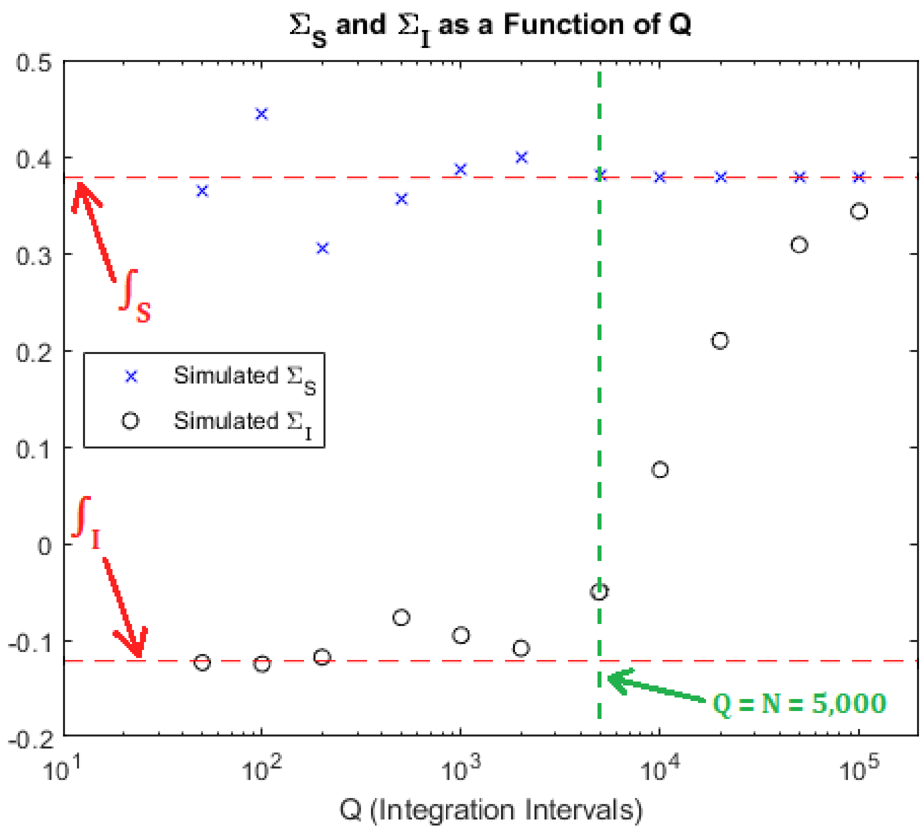

Figure 2: Plots of the Itô style Riemann sums Σ_I and Stratonovich style Riemann sums Σ_S using Q integration intervals. These correspond to the data in Table 1.

The data in Table 1 and Figure 2 show that there are three distinct regimes for Σ_I and two distinct regimes for Σ_S. For the Itô style Riemann sums: when $Q \ll N$, we have $\Sigma_I \approx \int_I$ to about 3 to 4 digit accuracy; for $Q \lesssim N$, we have $\Sigma_I \approx \int_I$ to about 1 to 2 digits of accuracy; and finally, when $Q \gtrsim N$, Σ_I slowly converges toward $\int_S$ as Q is increased. For the Stratonovich style Riemann sums: when $Q \lesssim N$, we have $\Sigma_S \approx \int_S$ to about 1 to 2 digit accuracy, while for $Q \gtrsim N$, Σ_S rapidly converges toward $\int_S$ as Q is increased.

These results make intuitive sense for the most part. For $Q \ll N$, there are many Sechmet nodes within one integration interval, and so the sampled distribution appears to have the jagged, nowhere-differentiable properties of Brownian motion within each interval to an acceptably high degree. This allows Σ_I and Σ_S to converge reasonably close to their theoretically predicted values (albeit with significant noise, since we're integrating over Brownian motion). However, for $Q \gtrsim N$, the interval of integration contains less than one Sechmet node, so the distribution inside the interval of integration loses the characteristic shape of Brownian motion, and it starts to look more and more like an infinitely-differentiable function as we zoom in. In particular, this causes $f\left(x_n^L\right) \to f\left(x_n^M\right)$ as $Q \to \infty$, which is why the Itô style and Stratonovich style Riemann sums converge in this limit. (They converge to $\int_S$ of course, since the Stratonovich integral recovers the classical result.) A left-hand Riemann sum converges much more slowly than a centered Riemann sum when integrating a twice-differentiable function, which is where the difference in convergence rates comes from in the limit $Q \to \infty$. The only tidbit in Table 1 and Figure 2 that doesn't make intuitive sense is why $\Sigma_I \approx \int_I$ to 3 to 4 digit accuracy when $50 \leq Q \leq 200$, while $\Sigma_S \approx \int_S$ to only 1 to 2 digit accuracy in these cases. Future work could investigate this phenomenon.

Regardless of the small curiosity above, the experimental results demonstrate that the Sechmet form $f(x)$ is a good approximation to Brownian motion at a length scale longer than the nodal spacing, μ, while $f(x)$ behaves like a classically-integrable function at a length scale smaller than μ. This is potentially quite advantageous: I have derived an extremely fast algorithm for performing exact, analytical integral calculations over functions in Sechmet form; this algorithm could be used to integrate exactly over (any function of!) $f(x)$, which would then give the Stratonovich integral over Brownian motion for any length scale longer than μ. If needed, the Stratonovich integral could then be easily transformed into the Itô integral. The Sechmet thus appears to offer an efficient way of generating Brownian motion and performing subsequent integral calculations for numerical research.

References

[1] T. Mikosch, *Elementary Stochastic Calculus*. Pg. 33. World Scientific (1998). ISBN 9810235437.