An Exploratory Survey of Stochastic Differential Equations from the Perspective of Numerical Simulation via the Euler-Maruyama Method with an application in Financial Time-Series

Zenas Huang and Timothy Russell

December 3, 2018

1 Introduction

A Stochastic Differential Equation (SDE) is an extension of classical differential equations to include the influence of random white noise perturbations. Simplistically, it is useful to consider them as a type of inhomogeneous differential equations in which the forcing term is now a random process in lieu of the usual kinds of functions in classical mechanics.

Why study stochastic differential equations? Because the behavior of many phenomena of interest in physics, engineering, climate science, finance, and many other fields are generally not nicely behaved; that is they tend to operate with some observable random effects or error that is not easily accounted for by contemporary deterministic models. Moreover, often times when it is not possible to directly model phenomena due to prohibitive informational costs that prevent knowledge of all the relevant processes underlying some observed behavior, it is useful to use stochastic differential equations as an approximation to those hidden processes; in fact this technique of modeling uncertain or latent processes as random variables is typical of Bayesian hierarchical modeling and epistemology. To summarize, we note that very often the data produced by phenomena "in the wild" is generally messy and that human existence is a condition often characterized by imperfect knowledge and finite computational resources for which stochastic differential equations represent one possible viable approximation to understanding phenomena of interest which cannot be accurately described under contemporary deterministic methods.

As SDEs comprise their own rich field, a general treatment of the subject would require extensive rigorous probability theory, measure theory, and martingale theory. However, since we intend our work to be an introductory exposition, we will bypass much of this theory and refer to it only as needed where the interested reader is invited to explore our references. Moreover, we will limit our discussion to focus only on a special particular case of SDE problem, namely

we are interested in the linear SDE of the form:

$$dX(t) = f(X(t))dt + g(X(t))dW(t), X(0) = X_0, 0 \leq t \leq T$$

The paper is organized as follows: first we begin section II by developing necessary definitions and concepts and discuss how Brownian Motion is regarded a solution to our linear scalar-coefficient SDE and its relation to the heat equation PDE; then in Section III we discuss the necessity of numerical solution methods for SDEs and review the Euler-Maruyama (EM) Method (the topic of our primary source paper) and discuss some relationships to the deterministic Euler Method; then in Section IV we deal with the notions of weak and strong convergence of the EM method; next in Section V we consider a brief alternative perspective of Brownian motion; and finally in Section VI we deal with a real-world financial application of SDEs by applying the EM-method to simulate historical price behavior of Amazon stock providing some context on its application in finance, discussion of results, and avenues for future exploration.

2 Main Definitions and Concepts

As discussed, it can be helpful to consider SDEs as a form of inhomegeneous DE with a random process as the forcing function. The process of interest for us is a standard Brownian motion process which we define, for some finite time interval $[0, T]$, to be a random variable W(t) that depends continuously on t and meets the following conditions:

1. W(0) = 0 (i.e. with probability 1 no noise at initial time)
2. For any two distinct times s and t such that $0 \leq s < t \leq T$, W(t) - W(s) is normally distributed with mean zero and variance equal to t - s. Accordingly, $W(t) - W(s) \sim \sqrt{t-s}N(0, 1)$
3. For $0 \leq s < t < u < v \leq T$, the increments given by W(t) - W(s) and W(v) -W(u) are independent events.

2.1 Physical origins of Brownian Motion

For our linear SDE,

$$dX(t) = f(X(t))dt + g(X(t))dW(t), X(0) = X_0, 0 \leq t \leq T$$

We are interested in the case where $f(X(t)) = \lambda X(t)$ and $g(X(t) = \mu X(t)$ with λ, μ real constants. This SDE has a known solution given by the Brownian process described by:

$$X(t) = X(0)exp((\lambda - \frac{\mu^2}{2})t + \mu W(t))$$

In order to provide some context for the Brownian process, it is possible to interpret Brownian motion as a solution to the heat equation. To see this, we

review an informal summary of Albert Einstein's 1905 derivation from *On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat*, whose argument proceeds as follows from [Kozdron, 3-4]:

1. Let K be some number of particles suspended in a liquid.
2. Within a short time interval T, the x-coordinate of any single particle will change by ε , where ε has a different value for every particle but ε obeys a certain probability distribution φ.
3. During the interval T, we will informally let dK = K$\varphi(\varepsilon)$dε express the infinitesimal of particles that experience some displacement between ε and $\varepsilon + \Delta\,\epsilon$ so that we have dK = K$\varphi(\varepsilon)$dε where φ (ε) differs from zero only where ε is sufficiently small and, by construction, satisfies φ (-ε) = φ (ε), and $\int_R \varphi(\varepsilon)d\epsilon = 1$ and $\int_R \varepsilon\varphi(\varepsilon)d\epsilon = 0$ by assumption.
4. Then in order to find the distribution of particles T moments in the future, we look to find the integral $f(x, t+T) = \int_R f(x+\varepsilon, t)\varphi(\varepsilon)d\varepsilon$ and taking α^2 to be our usual conductivity constant, we can obtain via Taylor approximations that $f(x, t+T) = f(x,t) + \frac{\partial f}{\partial t}T$ and $f(x+\varepsilon, t) = f(x,t) + \frac{\partial f}{\partial x}\varepsilon + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\varepsilon^2$.
5. Proceeding by substitution we obtain that $f(x, t+T) = \int_R (f(x,t) + \frac{\partial f}{\partial x}\varepsilon + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\varepsilon^2)\varphi(\varepsilon)d\epsilon = f(x,t) + \alpha^2\frac{\partial^2 f}{\partial x^2}T$ by properties of $\varphi(\varepsilon)$
6. Then by setting the two taylor approximations equal to each other we obtain $\frac{\partial f}{\partial t} = \alpha^2\frac{\partial^2 f}{\partial x^2}$ which is the well known form of the heat equation; where it can be seen by differentiation and substitution into the PDE that $f(x,t) = \frac{K}{2\alpha\sqrt{\pi t}}exp(-\frac{x^2}{4\alpha^2 t})$ is a solution to the heat equation and is of precisely the same form as our Brownian motion solution.

Thus, the fact that Brownian motion of particles is a solution to the heat equation is consistent with our intuition of heat as molecular motion and illustrates its ubiquity in natural phenomena.

A remark on φ. In most synopses of Einstein's proof, φ is understood as either the distribution of displacement values ε or else referred to as the "probability distribution" of ε and is said to satisfy φ (-ε) = φ (ε) by assumption. The justifications for this come from statistical mechanics and the laws of thermodynamics. In Einstein's original derivation, he considered the motion of small particles whose movements were mutually independent and whose motions in consecutive t-coordinates were also independent which suggests it is likely that Einstein's derivation was relying upon the ergodic hypothesis in thermodynamics which yields the long-run stationary distribution of system states (in this case the ε-position distribution of the particle) to be a uniform distribution.

2.2 The Ito Stochastic Integral and SDE Solution

In order to address the issue of how our standard BM can be regarded as a solution to our linear-scalar coefficient SDE, we require a new form of integral. The reason for this is that in trying to solve our SDE:

$$dX(t) = f(X(t))dt + g(X(t))dW(t), X(0) = X_0, 0 \leq t \leq T$$

We find that although the integral of the first term is amenable to classical integration methods such as Lebesgue integration, the integration of the second term presents a challenge as our Brownian process W is continuous over our domain but nowhere differentiable and also unbounded. As a result Brownian motion is too irregular to define partitions for Riemannian-Stieljes or Lebesgue integrals. These issues motivated the development of what are called stochastic integrals; of particular interest to us will be the Ito stochastic integral whose full definition is typically presented with three parts involving probability theory, martingales, and the development of a special class of functions to which we refer the reader to [Oksendal, p. 26] and [Steele,1-6]

But as we are primarily interested in numerical simulation, we will rely on a definition of the ito integral that is common for numerical methods which proceeds by analogy to the Riemann, as the limit of the following sum:

$\sum_{j=0}^{N-1} h(t_j)(W(t_{j+1}) - W(t_j))$

For some Lipschitz continuous function h(t) and for discrete times $t_j = j\delta t$ as we take $\delta t \to 0$.

A key distinction in this integration method is that the evaluation of W(t) in the sum now makes the result of our stochastic integral a random variable as distinct from classical integrals that generally resolve to a scalar value. Another frequently used type of stochastic integral is the Stratonovich integral which, for our purposes, mainly differs in that the Stratonovich evaluates W(t) at the midpoint of each sub-partition whereas the Ito integral evaluates the sums from the left endpoint. However, because of unbounded variation from W(t) at every t-coordinate, these different sums will not tend to converge as we take increasing refinements of the partition which illustrates another key difference from Riemann and Lebesgue integration methods where the choice of points in the partition could be arbitrary while the choice of points in our partition in the stochastic integral will now alter the sums through our W(t) term. In some ways, the development of stochastic calculus can be considered an extension of classical methods in order to deal with pathological functions of unbounded variance by using a notion of almost sure convergence to the distribution of a random variable in lieu of the traditional convergence of upper and lower sums to a single value.

In taking the limit of our sum, we first simplify as follows:

$$\sum_{j=0}^{N-1} W(t_j)(W(t_{j+1}) - W(t_j))$$
$$= \sum_{j=0}^{N-1} W(t_{j+1})W(t_j) - W(t_j)^2$$
$$= \frac{1}{2} \sum_{j=0}^{N-1} (2W(t_{j+1})W(t_j) - 2W(t_j)^2)$$
$$= \frac{1}{2} \sum_{j=0}^{N-1} (W(t_{j+1})^2 - W(t_j)^2 - (W(t_{j+1}) - W(t_j))^2)$$
$$= \frac{1}{2}(W(T)^2 - W(0)^2 - \sum_{j=0}^{N-1} (W(t_{j+1}) - W(t_j))^2)$$

Where it can be shown through the moments of the distribution of W(t) that the random variable resulting from the sum in our stochastic integral will have $\mathbb{E}[\sum_{j=0}^{N-1} (W(t_{j+1}) - W(t_j))^2] = T$, the end point time, and with variance of O(δt)

So for sufficiently small δt we can expect the sum to have an expected value close to the constant T which, for the Ito integral, gives us the following upon taking the expected value of each of the terms in our sum:

$$\int_0^T W(t)dW(t) = \frac{1}{2}W(T)^2 - \frac{1}{2}T$$

The Ito Integral now enables us to address the problematic second integral component of our SDE by allowing us to integrate the white noise and obtain a random variable perturbation that follows a distribution and whose realized values we add to the first term of our solution's path. In fact, In discussing solutions to our linear scalar autonomous SDE, there are generally two such notions: weak and strong solutions. Given a probability space (Ω,S,P) and a Brownian process W(t) for all t $\geq$ 0, then a continuous stochastic process X(t) with drift and diffusion coefficients f(X(s)) and g(X(s)) respectively, and initial value X_0 , if for all $t > 0$:
$X(t) = X_0 + \int_t^0 f(X(s))ds + \int_t^0 g(X(s))dW(s)$
where each of the integrals are assumed to exist is said to be a solution of the SDE above. Moreover, we say that given the process X(t), then it is a strong solution if X(t) is a process that is defined for a given probability space. While the concept of a weak solution is a process X(s) defined on any probability space such that X(t) still solves the SDE. To put it more succinctly, X(t) will be a strong solution when we are constructing it from a noise process, Weiner process W(t) in our case, that we know in advance whereas in the case of a weak solution, then provided only knowledge of the coefficient functions f(X(s)) and g(X(s)), if we can choose a path and noise pair, X(s) and W(t) respectively, defined over any probability space, such that the SDE is satisfied, then it is a weak solution.

One consequence of having a strong solution is that we also expect our solution to be "unique" in the sense that provided another such process Y(t) that

satisfies the SDE, then:

$$P(\sup_{s \leq t} |X_s - Y_s| > 0) = 0$$

Which means that we expect that for for all t in our domain, then the probability of the two processes evaluating to different points at that time is effectively zero. In other words, the two paths will be equal "almost surely" (i.e. with probability 1) across our time domain.

In our case, we expect to have a strong solution as our Brownian process is well understood. Moreover, we will appeal to a powerful result in stochastic calculus known as Ito's Existence and Uniqueness Theorem which guarantees that we can expect to have a unique strong solution provided that our f(X(t)) and g(X(t)) satisfy lipschitz continuity and are linear grown functions, which indeed obtains because for our case these functions are linear (i.e. scalar multiples, μ X(t) and σX(t)) and we are only considering our solution X(t) on the closed and bounded time domain interval [0,1]. The complete proof of the Ito Theorem involves a non-trivial amount of functional analysis and some background on adapted processes or non-anticipative functions from probability theory. We will not review the proof whose argument can be found in [Chun Yip, 108-112] and [Fernandez]. For our purposes, it is sufficient to invoke the result of Ito's uniqueness and existence theorem so that we may be assured that we have convergence in distribution (i.e. almost sure convergence, which also implies convergence in probability) of our stochastic integral term to a limiting random variable whose distribution is known to us, in order that our linear SDE is guaranteed to have a unique strong solution.

3 Numerical Algorithm: Euler-Maruyama Method

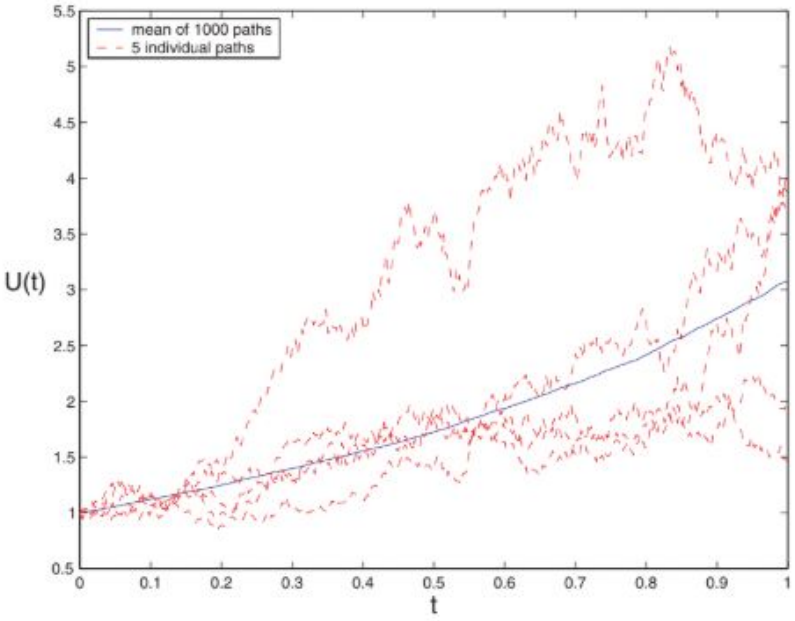
3.1 Why numerical methods?

Due to the irregularity of white noise in our SDE, it is often difficult to obtain closed form analytic solutions for SDEs, hence numerical solvers are often used. Even when dealing with a linear SDE where the noise process follows a well known as in the case of Brownian motion, our applications sometimes require that we make use of information about specific paths for which we must resort to numerical simulation to instantiate.

Furthermore, numerical methods are used because often times our models are often over simplified equations that explain things within an error tolerance that we cannot often know. But even with such simplified equations we quickly get to a point where we cannot calculate a analytic solution–i.e. n-body problem after $n \geq 3$. This is the case even with aggressive simplification. If we want our models to have any relevance to these phenomena then we need a means by which to compactly express the influence of these random noise processes in our models.

3.2 The Euler-Maruyama Method

The EM method is an extension of the euler method where we draw a normally distributed random number for a given partition of our interval forming a discretized brownian path; we then add in that number at each step of the EM method running the method once per path. Since we now have a distribution applied to our partition our result will have a distribution as well. Consequently we sample the space by performing the integration many times and taking the expected value as can be seen in [Higham, Fig 2]:



The function $u(W(t))$ averaged over 1000 discretized Brownian paths and along 5 individual paths, from `bpath3.m`.

Each run from t over [0,1] samples from a discrete brownian path which is nothing more than our specific random scalar component over one iteration, connected via lines for readability as in [Higham, Fig 1].

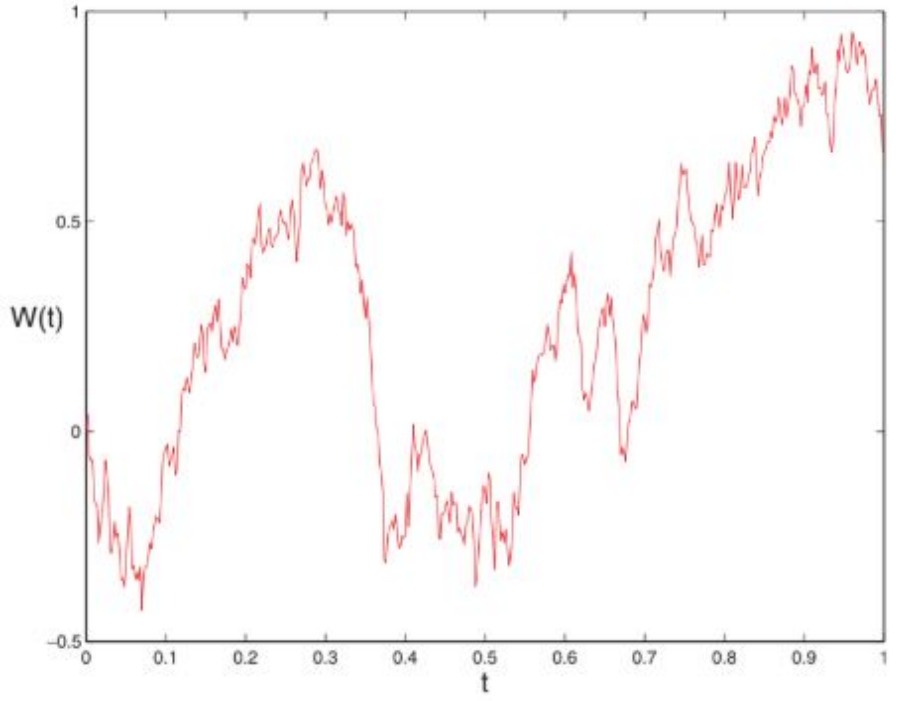


Fig. 1 *Discretized Brownian path from* `bpath1.m` *and* `bpath2.m`.

Continuing with the analogy to the Riemann/Lebesgue we can evaluate our sum similarly except that at each partition we now incorporate a noise element which makes our sums vary according to a normal distribution. So if we take the partition width to 0 we are taking into ever more noise at each step; noise which is only weakly bound by variance/confidence interval.

We now turn to examining the main lines of the EM code provided by Higham.

```
%EM  Euler-Maruyama method on linear SDE
%
% SDE is  dX = lambda*X dt + mu*X dW,   X(0) = Xzero,
%      where lambda = 2, mu = 1 and Xzero = 1.
%
% Discretized Brownian path over [0,1] has dt = 2^(-8).
% Euler-Maruyama uses timestep R*dt.

randn('state',100)
lambda = 2; mu = 1; Xzero = 1;    % problem parameters
T = 1; N = 2^8; dt = 1/N;
dW = sqrt(dt)*randn(1,N);         % Brownian increments
W = cumsum(dW);                   % discretized Brownian path

Xtrue = Xzero*exp((lambda-0.5*mu^2)*([dt:dt:T])+mu*W);
plot([0:dt:T],[Xzero,Xtrue],'m-'), hold on

R = 4; Dt = R*dt; L = N/R;        % L EM steps of size Dt = R*dt
Xem = zeros(1,L);                 % preallocate for efficiency
Xtemp = Xzero;
for j = 1:L
   Winc = sum(dW(R*(j-1)+1:R*j));
   Xtemp = Xtemp + Dt*lambda*Xtemp + mu*Xtemp*Winc;
   Xem(j) = Xtemp;
end

plot([0:Dt:T],[Xzero,Xem],'r--*'), hold off
xlabel('t','FontSize',12)
ylabel('X','FontSize',16,'Rotation',0,'HorizontalAlignment','right')

emerr = abs(Xem(end)-Xtrue(end))
```

Code from [Higham, Listing 5]

The specific SDE of interest is $dX = \lambda X dt + \mu X dW$ from the black scholes asset price model. The following setup within the code establishes the parameters of our discretized path for t $\in$ [0,1] we have $T = 1$, $N = 2^8$, step size $dt = 1/N$, and
$dW = \sqrt{dt} * randn(1, N)$ and $W = sum(dW)$
where L is the number of Euler steps of size $Dt = R * dt$ with R some arbitrary constant of increment usually chosen to be an integer.
Then the main lines of the algorithm are as follows:

$forj = 1 : L$
$Winc = sum(dW(R * (j - 1) + 1 : R * j))$;
$Xtemp = Xtemp + Dt\lambda Xtemp + \mu XtempWinc$
$Xem(j) = Xtemp$;
end

Where $Xtemp = Xtemp + Dt\lambda Xtemp$ is the Euler portion of the step as we discussed in class and μ Xtemp Winc is the portion of our calculation and one will note that Winc here is the $W(t_{j-1}) - W(t_j)$. So at each step it resembles the normal Euler method but now with noise times the old solution added in.

$Xzero = X(0) = x_0 = initialcondition$, here it is a real number: in our application example it is the price at the beginning of our time interval of roughly a year. In general Xzero is a measure of the initial state of the system we are provided, thus it is often a source of error as well.

Step size is the division of the interval. t $\in$ [0,1] step size R*dt with n being the number of partitions. Step size raises a lot of theoretical issues we do not have the time to fully explore; but in general we will want small step size for greater accuracy. However this will only hold if the noise is not too large relative to drift and diffusion coefficients discussed earlier. Making the step size smaller significantly increases computation time as well as compounding machine computation errors, making it impractical to run the error to 0. As can be seen in [Higham, Fig. 4], note that we are working with a scale of 10^-1 to 10^-4 level of errors. This is far from the 10^-16 or more error tolerance many applications may desire.

4 Strong and Weak Convergence

From generalizing the same process we use for analyzing the stability of Euler method from class, we can develop an understanding of what it means to say the EM method converges. As a quick sketch of the idea, we note we can discretize a space by n $\in$ Z and apply Taylor's Theorem. to approximate differential equations. Thus we apply an appropriate finite difference approximation to an equation of interest keeping in mind at least the order of the truncation or local error. Doing this we may write out a system of n equations and n unknowns developed from moving one iteration at a time and applying a consistent order to our partitioning. For convenience we place this in matrix form: $x_{n+1} = Ax_n$. This is nothing more than the classical Finite difference approach for approximating DEs.

We may ask what happens as we take t to infinity. Do we converge and if so to what? Through analogy with other iterative methods and by Theorem we have stability if and only if the spectral radius of A ¡ 1. Where spectral $radius = \rho = max|eigenvalue(A)|$. Stable here means that the error does not accumulate over the interval we are approximating on. We may run the partition to 0 and still not run the error to 0 though, because at each step we may think of the error as being acted upon by a linear operator. In class this operator was a derivative, in our matrix case it is an eigenvalue. The same idea holds, if the operator multiplying the error is too large then the error grows and our method does not converge to the exact solution–note with a distribution applied to each step we are only converging to a distribution. For understanding the eigenvalues of our resulting system of equations, we apply

Gerschgorin's Theorem to matrix A representing the system, and using the standard trigonometric sine sum formula, one may bound A's eigenvalues which then allows us to bound the stability. In the Euler case the order of $\frac{1}{2}$ means that if we want to decrease the error by a factor of 100 we need to make the step size 100^2 times smaller with computation scaling as much.

Note that the EM is not unconditionally convergent. Because of the brownian motion portion we apply extra conditions on the method compared to the deterministic case. Not only do we require the appropriate order or differentiability, but we also require the function in question not change too fast. (Lipschitz continuity) Strong convergence seen as the "mean of the error" serves as the stability error from the Euler method discussed before. But because one distribution leads to another, we also have the concept of weak converge, which simply put allows us to state that our distribution in the answer is bounded by an appropriately smooth class of, often, polynomial functions, such as the class of p functions as in [Higham, Eq 5.5]. It can often be easier to work with simpler functions describing a bound we are confident in than it is to work with specific, more accurate, but complicated functions–it can be useful to polynomially bound noise about a model as long as one remembers to account for the unbounded nature of the brownian motion.

5 Karhunen-Loève Expansion

In the theory of stochastic processes, there is an analogous extension of Fourier Series known as the Karhunen-Loeve Theorem by which it is possible to represent certain Stochastic Procsses as an infinite linear combination of orthogonal functions. The Karhunen-Loeve Theorem states that for X_t a zero-mean, square integrable stochastic process over a probability space that is indexed over a compact interval and which has a continuous, symmetric function K_x(s,t), also known as a mercer kernel, then the stochastic process can be represented by an infinite linear combination of orthogonal eigenfunctions e_k which are determined by mercer kernel, which happens to be the covariance matrix K_x(s,t) of the stochastic process. Then our stochastic process admits of the following representation:

$X_t = \sum_{k=1}^{\infty} Z_k e_k(t)$ with $Z_k = \int_a^b X_t e_k(t) dt$ a stochastic integral

where the random variable coefficients will each have mean zero, be mutually independent, and have a variance equal to it's associated eigenvalue.

We review the KL-expansion derivation for our standard Brownian motion/Weiner Process here for the time domain [0,1]. Typically, the eigenvalue problem in the first line of our proof can be found from Mercer's Theorem in functional analysis and the Homogeneous Fredholm Integral Equations of the Second Kind. However, for expository simplicity we may proceed by heuristic analogy to the eigenvalue problem in the finite dimensional case where we are taking the matrix

product $\mathbf{A}x = \lambda x$, where in setting up the KL-expansion, the $\mathbf{A}$ matrix is now a covariance matrix K_x and the vector x is represented by our vector of basis functions e(s). However, because the Wiener process is continuous everywhere we now have a continuous covariance function and so the matrix product now becomes an integration which gives us the eigenvalue problem

$\int_0^1 K_w(s,t)e(s)ds = \lambda e(s)$ where s, t are points from our time domain.

Given any two t-coordinates t and s in our compact time domain [0,1], it has also been shown as a standard result that the covariance function K_w for our Weiner process is K_w = cov(W_t,W_s) = min(s,t) . So the eigenvalue problem becomes:

$\int_0^1 min(s,t)e(s)ds = \lambda e(t)$

$\int_0^t se(s)ds + \int_t^1 e(s)ds = \lambda e(t)$

where taking t = 0, yields the boundary condition that e(0) = 0. Then by differentiating once with respect to t we obtain:

$\int_t^1 e(s)ds = \lambda e'(t)$

which, taking t = 1 implies the second boundary condition e'(1) = 0. Differentiating twice with respect to t, thus gives us that:

$\frac{\partial^2}{\partial t^2}(\int_0^t se(s)ds + \int_t^1 e(s)ds) = -e(t)$

Thus we have the ODE: $-e(t) = \lambda e''(t)$ which has the general solution given by:

$e(t) = Asin(\frac{t}{\sqrt{\lambda}}) + Bcos(\frac{t}{\sqrt{\lambda}})$

Then, by application of our boundary conditions, we have that $e(0) = 0$ which implies that B = 0. Then differentiating once, we apply e'(1) = 0 to find:

$e'(1) = Acos(\frac{1}{\sqrt{\lambda}}) = 0$

This implies that our eigenvalues are $\lambda_k = (\frac{1}{(k-\frac{1}{2})\pi})^2$ for all natural numbers and our corresponding eigenfunctions are $e_k(t) = Asin((k-\frac{1}{2})\pi t)$ where $A = \sqrt{2}$ is chosen as a normalizing constant, whose derivation can be seen in the full proof [Giambartolomei, 48-50]. Thus, KL-expansion over [0,1] gives us a sequence of random variable coefficients so that the Wiener Process is expressible as the linear combination:

$$W_t = \sum_{k=1}^{\infty} \sqrt{2} Z_k \frac{sin((k-\frac{1}{2})\pi t}{(k-\frac{1}{2})\pi}$$

6 Application

6.1 Application Context in Financial Econometrics

The use of SDEs in Quantitative Finance spans over a century beginning with Louis Bachelier's 1900 thesis *Theory of Speculation* in which he independently arrived at the mathematical model of Brownian motion five years before Einstein uncovered their significance in particle physics. From an Economics standpoint, Bachelier's primary contribution was that if asset prices displayed any discernible pattern aside from long-term growth trends associated with macroeconomic shocks, then trading activity would reduce any such potential gains to zero. In other words, the market fully prices in all available pertinent information so that prices always reflect fair-value meaning that one cannot ever "predict the market" or engage in "market-timing" to affect any substantive gains. Bachelier's thesis essentially demonstrated the arbitrary character of asset markets; this insight was later developed into the financial theory known as the Random Walk Hypothesis, a precursor to the prominent modern economic theory known as the Efficient Markets Hypothesis (EMH), developed and promoted by Economists Eugene Fama and Burton Malkiel, which states that assets always trade at a fair market value so that there do not exist any arbitrage opportunities to exploit for gain by engaging in trading or active management.

Nevertheless, the "unpredictability" of the stock market does not make our application nugatory. In fact, there continues to be lively debate among many academic economists, financial mathematicians, and industry practitioners disputing the extent to which the EMH is true. Irrespective of how much one "buys" the idea that the stock market is predictable/unpredictable, many financial management firms still utilize software packages based on SDEs as tools for risk management in which the sampled Brownian paths are used to assist in forecasting the potential for portfolio losses, absent any non-systematic, exogenous market shocks (i.e. whereas systematic or market risks is what the SDE is intended to model). We demonstrate this type of use in the following application.

6.2 Simulating Price Behavior for Amazon Shares

6.2.1 Methodology

For our application of the Euler-Maruyama Method, we follow the example provided in Listing 5 of Des. Higham's paper where the EM algorithm is applied, via monte carlo simulation, to generate sample path solutions to a linear SDE of the form:

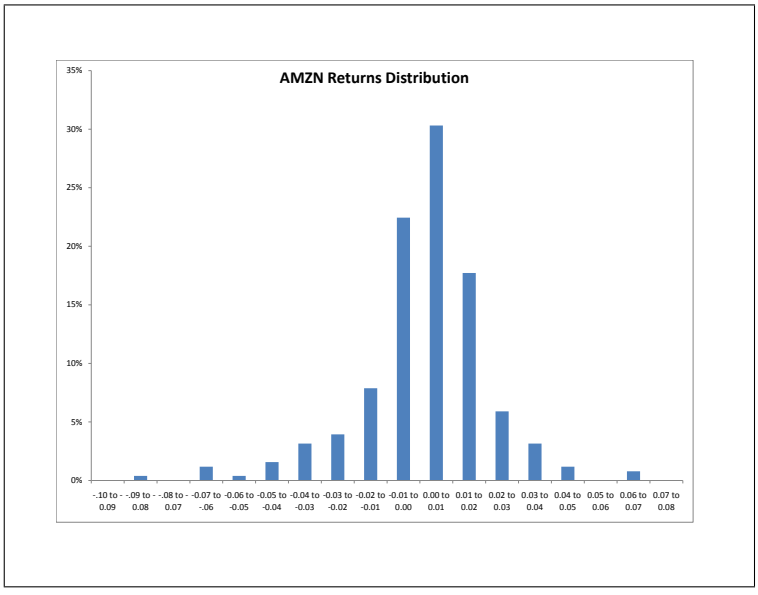
$$dX(t) = \lambda X(t)dt + \mu X(t)dW(t),\ X(0) = X_0$$

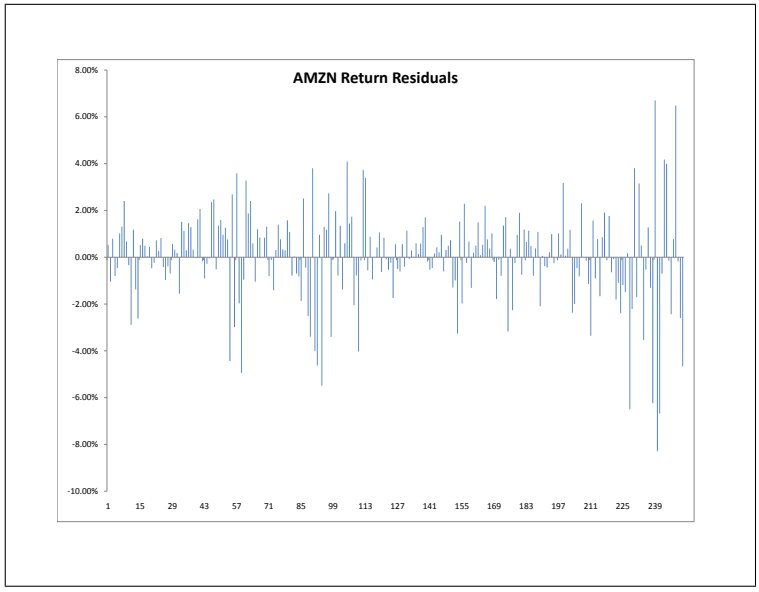
where λ and μ are some real constants so that we are guaranteed a strong unique solution by Ito's Existence and Uniqueness theorem. It is also well known that this SDE shows up in the derivation of the Black-Scholes PDE model in mathematical finance and that this SDE has an exact solution which is the geometric brownian motion process that is described by:

$$X(t) = X(0)exp((\lambda - \frac{\mu^2}{2})t + \mu W(t))$$

As this stochastic process is often used to model stock prices in the field of mathematical finance, we adapted the Euler-Maruyama code to simulate the the historical path of share prices for Amazon.com (AMZN) over the roling 252-day trading year period from 11-13-2017 to 11-12-2018 using data from nasdaq.com and rescaling the period to the [0,1] time domain. As the goal was to get the EM method to instantiate paths which were reasonable approximations of the the real data, the task was to obtain accurate estimates for the λ and μ constants in our model.

First, as a preliminary step, we check that our AMZN price data satisfies our normal distribution assumptions. To do so, we conducted an informal test by examining the distribution of the log returns for consecutive days in our data and checking that the residual daily returns do not show any discernible correlation.





The normality condition on our log returns data thus being satisfied, we proceed to estimate our model parameters. First, we determine parameters for drift and volatility on an intra-day basis by first determining the average and variance of our daily returns for the whole 252 day period which we found to be an average of $\overline{r} = 0.15\%$ and a variance $\sigma_r^2 = 0.04\%$. This enables us to obtain our intra-day parameters where the data gave us $\lambda_t = 0.001666$ and $\mu_t = 0.019604$ by applying the following formulas (whose form comes from the application of the Stochastic result known as Ito's Lemma, an analogue to Taylor expansions for functions of random variables, to our linear SDE and whose derivation is discussed in [Dmouj, 14-22]):

$\lambda_t = \overline{r} + \frac{\sigma_r^2}{2}$
$\mu_t = \sigma_r$

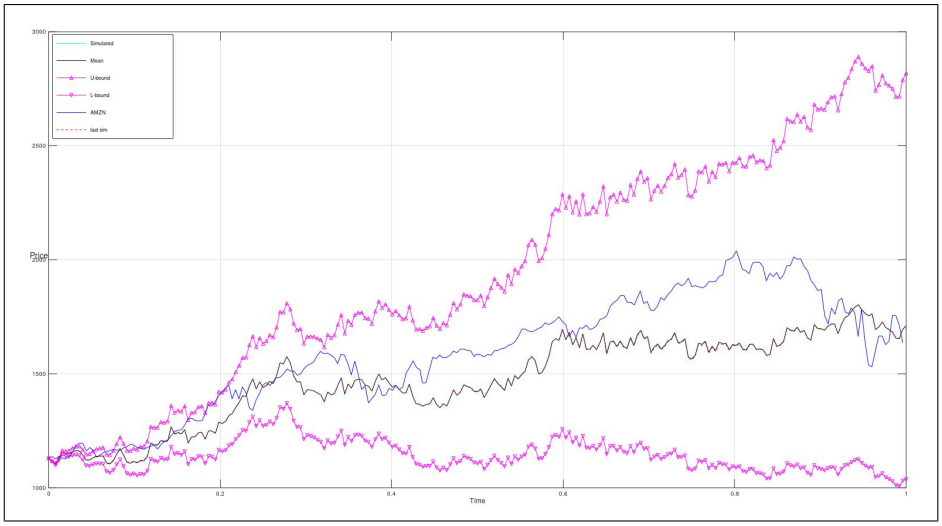
Then by simply scaling the intra-day drift and volatility by the length of our entire period, we obtain:

$\lambda_{252} = 252\lambda_t = 0.419713$ and $\mu_{252} = \sqrt{252}\sigma_r = 0.311199$

which are our drift and volatility parameters respectively for the entire period.

6.2.2 Results & Discussion

Using these parameter point estimates, we then simulated a single price path, as illustrated in the below graph and found its error with respect to the EM algorithms constructed exact solution and the error of our simulated path with respect to the Amazon price path which we sought to emulate.



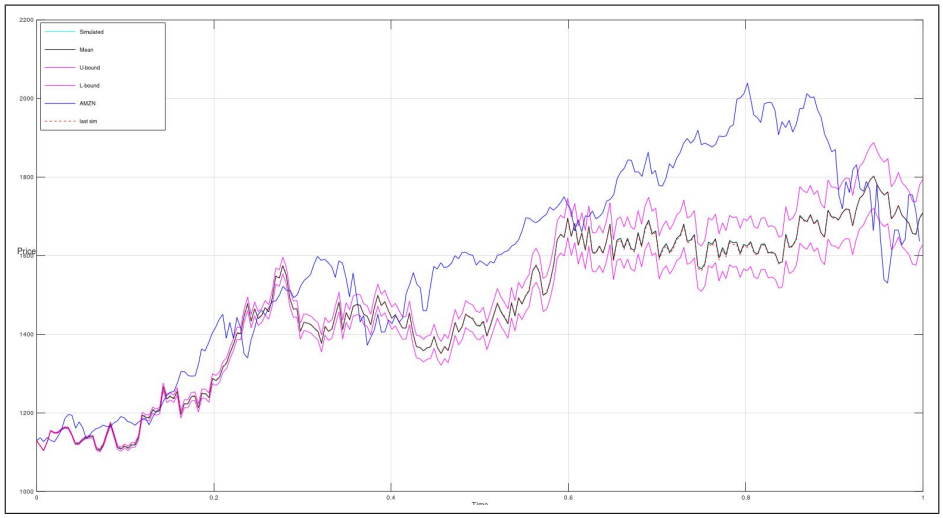
The blue line representing our historical AMZN price data and the black line representing the mean path (which overlaps substantially with and occludes our simulated path since M = 1).
In addition we observe an upper and lower bound end price point range from [$1037.0, $2818.8] where the AMZN endpoint price of $1636.8 is contained with this interval. We also find a difference error of $74.565 between the endpoint AMZN price and our final simulated price of $1711.4 which is a $\sim 4.4\%$ difference.

In addition to this, we include a confidence interval about our point-estimated brownian path which should contain 34% of paths as, M, the number of sampled paths $M \to \infty$. This confidence bound also gives us an endpoint price range of [$1037, $2818.8] containing the actual AMZN endprice of $1636.8 within it.

We report an endpoint error of 1.7399 for our simulated path relative to the constructed exact solution, which is somewhat larger than the error we would expect based on the strong convergence of EM and the results from Higham's own simulation results. We suspect that the reason for this is likely due to a number of factors such as the presence of sampling error in our point estimates, the larger step size we use due to the constraints of our data, and accessory sources of error from float-point arithmetic roundoff and potential random number bias. Nevertheless, in accordance with Higham's emperical results we expect a decay in the sampling error which we can confirm through our experiments showing that sampling M=2 paths we find the endpoint error between our simulated path and the exact solution declines to 0.54833 and further to 0.14436 with M=5. Thus, in the limit we would expect to be in conformity with the strong convergence of the Euler-Maruyama method.

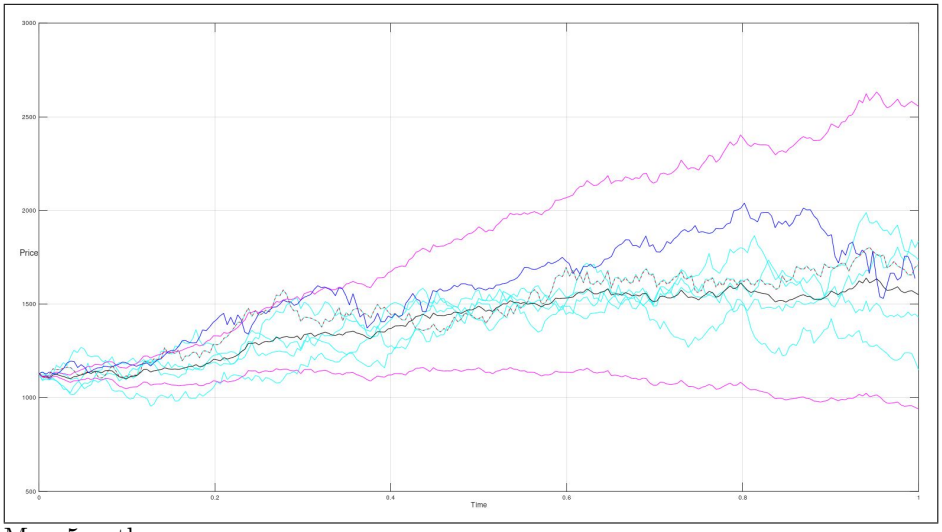
Although the confidence bound is useful in risk management for forecasting a range of potential outcomes, in practice a narrower precision bound may

sometimes be used in projecting into the short-term future possible price outcomes when the assumption is that intraday volatility is constant or at least not expected to change substantially from its historical value, this enables us to construct precision bounds by using the intraday volatility as standard error:

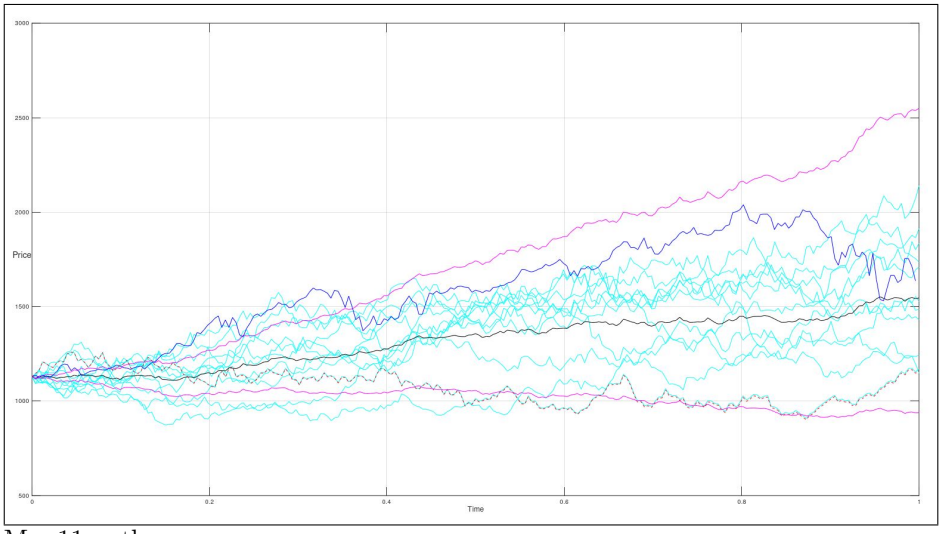


Where the upper and lower magenta bounds about our simulated mean path now give us a much narrower endpoint interval [$1627.9,$1795.6] that contains our AMZN endpoint price of $1636.8.

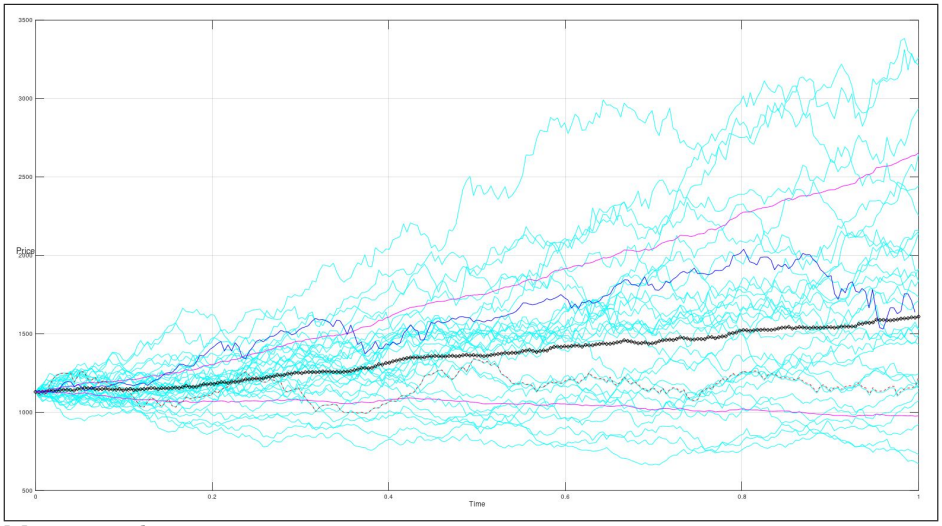
Lastly, since Brownian motion is a zero-mean stochastic process, we should expect that as the number of paths sampled $M \to \infty$, i.e. the noise from our process should thus tend to zero. We verify this empirically below for the cases of 5, 11, 30, and 50 sampled paths where the mean of our sampled paths begins to smooth out and resemble an exponential curve as greater number of paths are sampled thereby recovering an approximation to the deterministic solution to the first component of our SDE which is our familiar exponential ODE solution.



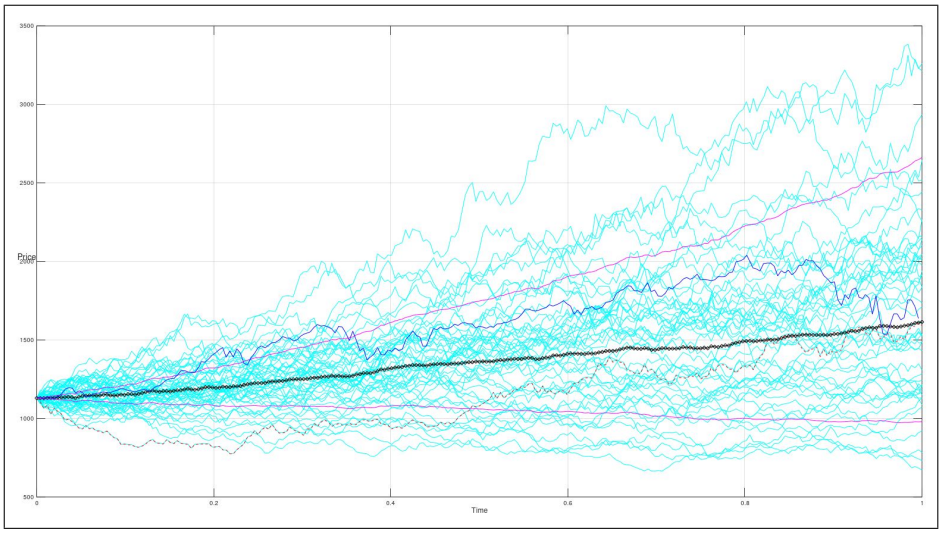
M = 5 paths



M= 11 paths



M = 30 paths



M = 50 paths

6.3 Conclusion & Future work

Our project has served as an introductory exploration into the extension of differential equations for the modeling of irregular processes through a numerical methods vantage point on the field of stochastic differential equations. In so doing we relied heavily on the exposition provided by Desmond Higham's numerical methods paper but have also found other relationships to the field of classical ordinary and partial differential equations through Einstein's molecular-kinetic theory of heat, the Random Fourier Series of the Karhunen-Loève Expansion,

the similarities of the Euler-Maruyama method with the standard Euler method, and in the empirical case of recovering the deterministic ODE exponential solution via repeated sampling of Brownian paths with mean-zero noise. Apart from those relationships, we have also investigated the use of the Euler-Maruyama method as a simulation and forecasting tool in the field of Financial econometrics and attempted to replicate this application by estimating and simulating a time-series dataset of Amazon.com shareprices. As always, there are numerous avenues for future work. Further investigation into the error analysis and convergence of the numerical method, particularly for the higher-order Milstein method is one such area. Another promising area for future work is the potential to adapt the parameter estimation methodology into an online machine learning program whereby parameter point estimates can be dynamically updated in real time as new pricing data is made available. Lastly, as we mentioned in our introduction we have for the most part omitted the theoretical perspective on stochastic calculus which represents another important avenue for future work, particularly as they are crucial to the theory of bayesian variational inference, stochastic optimization in markov chain monte carlo techniques, and for popular contemporary algorithms such as stochastic gradient descent.

7 References & Sources

1. Higham, Desmond J. *An Algorithmic Introduction to Numerical Solution of Stochastic Differential Equations*, https://epubs.siam.org/doi/pdf/10.1137/S0036144500378302

2. Kozdron, Michael. *Brownian Motion and the Heat Equation*, http://stat.math.uregina.ca/~kozdron/Research/UgradTalks/BM_and_Heat/heat_and_BM.pdf

3. Sarkka, Simo. *Applied Stochastic Differential Equations*, https://users.aalto.fi/~ssarkka/course_s2012/pdf/sde_course_booklet_2012.pdf

4. Oksendal, Bernt. *Stochastic Differential Equations: An Introduction with Applications*, http://th.if.uj.edu.pl/~gudowska/dydaktyka/Oksendal.pdf

5. Steele, J.M. *Ito Calculus*, http://www-stat.wharton.upenn.edu/~steele/Publications/PDF/EASItoCalculus.pdf

6. Giambartolomei, Giordano. *The Karhunen-Loéve Theorem*, https://amslaurea.unibo.it/10169/1/Giambartolomei_Giordano_Tesi.pdf

7. https://www.encyclopediaofmath.org/index.php/Stochastic_differential_equation

8. https://nptel.ac.in/courses/111103022/module6/lec4.pdf

9. Sauer, Timothy. *Numerical Solution of Stochastic Differential Equations in Finance*, http://math.gmu.edu/~tsauer/pre/sde.pdf

10. Wang, Limin. *Karhunen-Loeve Expansions and their Applications*, http://etheses.lse.ac.uk/2950/1/U615901.pdf

11. *The Karhunen-Loeve Transform*, http://fourier.eng.hmc.edu/e161/lectures/klt/node3.html

12. Driscoll, Toby. *Mercer's theorem and the Karhunen-Loeve Expansion*,

http://www.chebfun.org/examples/stats/MercerKarhunenLoeve.html
13. Jandhyala, V.K. and Jiang, P.L. *Eigenvalues of a Fredholm Integral operator and Applications to Problems of Statistical Inference*, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.43.6117&rep=rep1&type=pdf
14. Pavliotis, G.A. *Stochastic Processes and Applications*, http://wwwf.imperial.ac.uk/~pavl/PavliotisBook.pdf
15. Liden, Joel. *Stock Price Predictions using a Geometric Brownian Motion*, http://www.diva-portal.org/smash/get/diva2:1218088/FULLTEXT01.pdf
16. *Stochastic Processes with Focus in Petroleum Applications*, http://marcoagd.usuarios.rdc.puc-rio.br/stochast.html#forecasting
17. Dmouj, Abdelmoula. *Stock price modelling: Theory and Practice*, https://beta.vu.nl/nl/Images/werkstuk-dmouj_tcm235-91341.pdf pg.
18. Yeo, Dominic. *Eventually Almost Everywhere: A blog about probability and olympiads*,https://eventuallyalmosteverywhere.wordpress.com/tag/pathwise-uniqueness/
19. Lalley, Steven. *Stochastic Differential Equations*, http://www.stat.uchicago.edu/~lalley/Courses/385/Old/SDE.pdf
20. ReiB, Markus. *Stochastic Differential Equations Lecture Notes for Courses given at Humboldt University Berlin and University of Heidelberg*, https://www.math.uni-heidelberg.de/studinfo/reiss/sode-lecture.pdf
21. Whitt, Ward. *A Quick Introduction to Stochastic Calculus*, http://www.columbia.edu/~ww2040/4701Sum07/lec0813.pdf
22. Herzog, Florian. *Stochastic Differential Equations*, https://www.ethz.ch/content/dam/ethz/special-interest/mavt/dynamic-systems-n-control/idsc-dam/Lectures/Stochastic-Systems/SDE.pdf
23. Zimmer, Alec. *Stochastic Calculus*, http://www.math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Zimmer.pdf
24. Hormann, Siegfried. *Stochastic Calculus (Lecture #3)*, http://homepages.ulb.ac.be/~shormann/teaching/SC_3.pdf
25. *Brownian motion: Heuristic Motivation*, http://www.stat.wisc.edu/~yzwang/Ito.pdf
26. *Almost Sure: The Stochastic Integral*, https://almostsure.wordpress.com/2010/01/03/the-stochastic-integral/
27. Chun Yip, Yau. *Ito's Stochastic Calculus*,http://www.sta.cuhk.edu.hk/cyyau/course/2013f_rmsc4005/chapter6.pdf
28. Fernandez, Arturo. *Ito Calculus: An Abridged Overview*, https://www.stat.berkeley.edu/~arturof/Teaching/STAT150/Notes/III_Ito_Calculus.pdf
29. Time Series Data for our application was sourced from https://www.nasdaq.com/symbol/amzn/historical.