Revisiting Feller Diffusion: Derivation and Simulation

Ranjiva Munasinghe[*1]**, Leslie Kanthan**[2] **and Pathum Kossinna**[1]

[1]*MIND Analytics & Management*
10/1 De Fonseka Place, Colombo 5, Sri Lanka.
[2]*University College London*
Gower Street, London, WC1E 6BT, UK.

Review Paper

Abstract

We propose a simpler derivation of the probability density function of Feller Diffusion by using the Fourier Transform on the associated Fokker-Planck equation and then solving the resulting equation via the Method of Characteristics. We use the derived probability density to formulate an exact simulation algorithm whereby a sample path increment is drawn directly from the density. We then proceed to use the simulation to verify key statistical properties of the process such as the moments and the martingale property. The simulation is also used to confirm properties related to hitting time probabilities. We also mention potential applications of the simulation in the setting of quantitative finance.

Keywords: Feller Diffusion; Bessel Squared Process; Simulation; Absorbing Random Walks; Survival Probability

2010 Mathematics Subject Classification: 65C30; 60H35; 68U20; 65C05

1 Introduction

Feller Diffusion, which is also known as the Bessel Squared Process of dimension zero (FellerW, 1951; Feller, 1951, 1954; Gan & Waxman, 2015), is in the class of Ìto stochastic differential equations (SDEs) known as square root diffusions (Glasserman, 2004). In general, square root diffusions have a drift term in addition to a multiplicative square root noise term. Feller Diffusion is a special case in that it is driftless. Feller's motivation in this area was genetics (Feller, 1951) with natural extensions to biological systems and population dynamics. Rigourous studies of these processes can be found in references (Revuz & Yor, 1999; Göing-Jaeschke & Yor, 2003). Square root diffusions also have application in physics - in particular in the area of branching processes and percolation theory (Cardy, 1996) and phase transitions with absorbing states (Hinrichsen, 2000). In addition to the science applications listed above, there are applications of square root diffusions in quantitative finance - e.g.

Corresponding author: E-mail: ranjiva@mindlanka.org

the CEV Model for asset pricing (Cox, 1996; Hsu, Lin & Lee, 2008; Lindsay & Brecher, 2012) and the well-known CIR Model for interest rate modeling (Cox, Ingersoll & Ross, 1985; Wilmott, 2009; Hull, 2012).

The Îto SDE describing Feller Diffusion is given by:

$$\dot{x}(t) \quad = \quad \sigma\sqrt{x} \cdot \xi(t) \tag{1.1}$$

$$x(t_o) \quad = \quad x_o > 0. \tag{1.2}$$

Here $x(t)$ denotes the position of a random walker following the trajectory given by equation (1.1) with initial condition x_o at (initial) time t_o, and $\xi(t)$ is a Gaussian white noise independent of $x(t)$, with amplitude parameter σ, and the following correlation functions describing its statistics:

$$\langle \xi(t) \rangle_\xi \quad = \quad 0 \tag{1.3}$$

$$\langle \xi(t)\xi(t') \rangle_\xi \quad = \quad \delta(t - t'). \tag{1.4}$$

The averaging $\langle \cdots \rangle_\xi$ denotes an average with respect to the realizations of the stochastic process $\xi(t)$. Due to presence of the square root term the dynamics are such that the random walker's postion is always positive and if it reaches the origin it will stay there (i.e. the origin acts as an absorbing barrier).

The Fokker-Planck equation corresponding to the SDE (1.1) and initial condition (1.2) is given by

$$\frac{\partial}{\partial t}P(x,t) \quad = \quad \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}[xP(x,t)] \tag{1.5}$$

$$P(x,t_o) \quad = \quad \delta(x - x_o) \tag{1.6}$$

$$\int_0^\infty P(x,t)\,dx \quad = \quad 1. \tag{1.7}$$

Here $P(x,t) = P(x,t|x_o,t_o)$ describes the probability density of the position of the random walker at a particular time via the following relation:

$$P(x,t)\Delta x = \text{Probability walker} \in [x, x + \Delta x] \text{ at time } t$$

2 Derivation of the Probability Density

The first step in deriving the probability density is to use the Fourier transform on $P(x,t)$:

$$\hat{P}(k,t) = \int_{-\infty}^\infty P(x,t)e^{-ikx}dx = \int_0^\infty P(x,t)e^{-ikx}dx$$

The Fokker-Planck equation (1.5) now becomes

$$\frac{\partial}{\partial t}\hat{P}(k,t) \quad = \quad \frac{-i\sigma^2 k^2}{2}\frac{\partial}{\partial k}\hat{P}(k,t) \tag{2.1}$$

$$\hat{P}(k,t_o) \quad = \quad \exp\left[-ikx_o\right] \tag{2.2}$$

$$\hat{P}(k=0,t) \quad = \quad 1. \tag{2.3}$$

This partial differential equation (PDE) can be solved by the method of characteristics (please refer Appendix B) to yield

$$\hat{P}(k,t) = \exp\left[\frac{-ikx_o}{1 + \frac{ik\sigma^2}{2}(t - t_o)}\right] \tag{2.4}$$

Note that the initial condition given by equation (2.2) and normalization given by equation (2.3) are easily verified. The characteristic function $\varphi(k,t)$ can be easily derived by the transformation $k \to -k$ in the Fourier transform, i.e. using the relation:

$$\varphi(k,t) = \hat{P}(-k,t)$$

This gives the characteristic function for the probability density governing the SDE (1.1) with initial condition (1.2) as

$$\varphi(k, t) = \exp\left[\frac{ikx_o}{1 - \frac{ik\sigma^2}{2}(t - t_o)}\right] \tag{2.5}$$

This is the characteristic function of a non-central chi-squared distribution with zero degrees of freedom (please refer to Appendix D) where:

1. The random variable has been scaled such that $x \to \frac{4}{\sigma^2(t-t_o)}x$

2. The non-centrality parameter is $\lambda = \frac{4x_o}{\sigma^2(t-t_o)}$

3. $P(x, t | x_o, t_o) = g\left(\frac{4x}{\sigma^2(t-t_o)}; \frac{4x_o}{\sigma^2(t-t_o)}\right)$, where $g(x; \lambda)$ is the density function of a non-central chi-squared distribution with zero degrees of freedom with non-centrality parameter λ.

The non-centrality parameter (Siegel, 1979) enables us to estimate the probability that the random variable takes the value 0 as $e^{-\lambda/2}$. In terms of the stochastic process (1.1) this is the *absorption probability* $F(t)$, which is the probability the random walker gets absorbed (Gan & Waxman, 2015) at the origin at time t given the starting point x_o:

$$F(t) = \exp\left[\frac{-2x_o}{\sigma^2(t - t_o)}\right] \tag{2.6}$$

Note that in the limit $t \to \infty$ this probability will tend to 1 which means that all random walkers following process (1.1) will eventually be absorbed. The *survival function*, $S(t)$, is the probability the random walker is alive at time t and has the following relation $S(t) = 1 - F(t)$.

3 Sample Path Simulation

A non-central chi-squared distribution can be expressed as a Poisson compound mixture of central chi-square distributions (refer Appendix D) with even degrees of freedom (Siegel, 1979). Using this fact enables us to formulate a simple algorithm for the simulation of a random walk (Glasserman, 2004) corresponding to the SDE (1.1). We make use of items 1 (scaling) and 2 (non-centrality parameter) from the previous section along with the Markov property[†] (Oksendal, 2000) of the random walkers to describe the algorithm for generating a sample path:

1. For $i = 1, 2, \ldots, n$:

$$\begin{aligned} c &\leftarrow \frac{\sigma^2}{4\Delta t_i} \quad \text{where } \Delta t_i = t_i - t_{i-1} \\ \lambda &\leftarrow x_{i-1}/c \end{aligned}$$

2. Generate $N \sim \text{Poisson}(\lambda/2)$
3. If $N = 0$ set $X = 0$ otherwise generate $X \sim \chi^2_{2N}$
4. Set $x_i = cX$

The algorithm illustrated above was implemented in Rcpp (an implementation of C++ in R) and R was used in generating plots and data manipulation. The Rcpp implementation is given in Appendix E. A sample simulation of 100 such generated paths is shown in Figure 1.

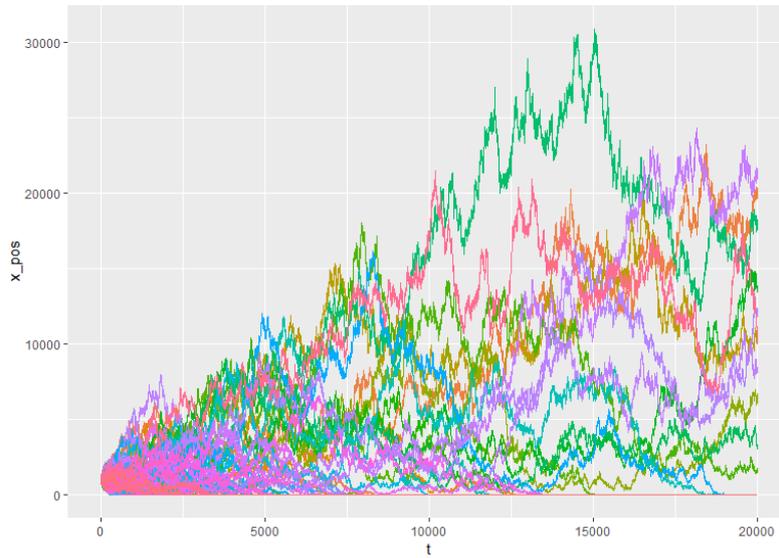[†] See also https://almostsure.wordpress.com/2010/07/28/bessel-processes/

Figure 1: Feller Sample Path Simulation with $t_0 = 0$, $t_n = 20000$, $\Delta t_i = 1$, $\sigma^2 = 1$ and $x_0 = 1000$ (100 paths). Note: Each path is represented by a different colour.

3.1 Examining Survival/Absorption Probabilities

Using our simulations we can explore several properties of Feller Diffusion paths (note: we will set $t_o = 0$ in both our simulation and computations henceforth). We begin by looking at the absorption probability (2.6) and estimate this via the evolution of the simulated surviving paths. Our simulations show that the more volatile the walker (i.e. larger values of σ), the lower the survival rate. They also show that for large times the survival drops. In both instances, the values are in close agreement with the theoretical values (Table 1 - using 10,000 simulations).

σ^2	Time (t)									
	100		1,000		5,000		10,000		20,000	
	Theo.	Sim.	Theo.	Sim.	Theo.	Sim.	Theo.	Sim.	Theo.	Sim.
0.1	0.00	0.00	0.00	0.00	1.83	1.70	13.53	13.47	36.79	36.11
1	0.00	0.00	13.53	13.31	67.03	67.08	81.87	82.10	90.48	90.68
10	13.53	13.02	81.87	81.84	96.08	96.12	98.02	98.03	99.00	99.03
100	81.87	82.10	98.02	97.97	99.60	99.61	99.80	99.76	99.90	99.88

Table 1: Percentage of paths absorbed at origin over time for different values of σ^2 with $t_0 = 0$, $\Delta t_i = 1$, $x_0 = 1000$. (10,000 paths)

We also observe that the further away from the origin the starting point (larger values of x_o) the better the chance of survival - as per the intuition from equation (2.6). These results are displayed in (Table 2 - using 10,000 simulations) and show close agreement with the theoretical values:

| x_0	Time (t)									
	100		1,000		5,000		10,000		20,000	
	Theo.	Sim.	Theo.	Sim.	Theo.	Sim.	Theo.	Sim.	Theo.	Sim.
10	81.87	82.10	98.02	97.97	99.60	99.61	99.80	99.76	99.90	99.87
100	13.53	13.02	81.87	81.85	96.08	96.11	98.02	98.04	99.00	99.03
1000	0.00	0.00	13.53	13.31	67.03	67.08	81.87	82.10	90.48	90.68
10000	0.00	0.00	0.00	0.00	1.83	1.79	13.53	13.51	36.79	36.07

Table 2: Percentage of paths absorbed at origin over time for different values of x_0 with $t_0 = 0$, $\sigma^2 = 1$, $\Delta t_i = 1$. (10,000 paths)
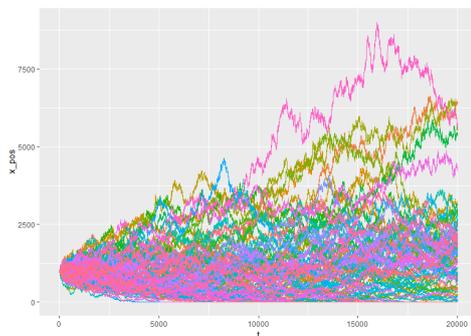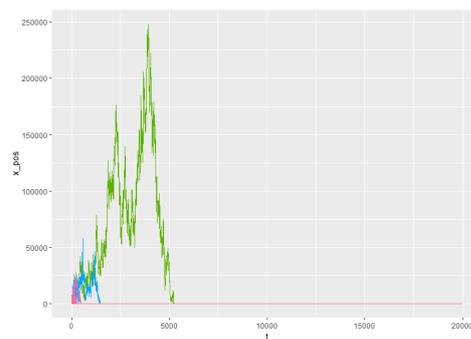

(a) Feller Sample Path Simulation with $\sigma^2 = 0.1$
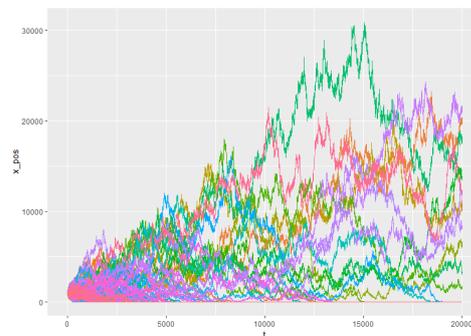
(b) Feller Sample Path Simulation with $\sigma^2 = 1$
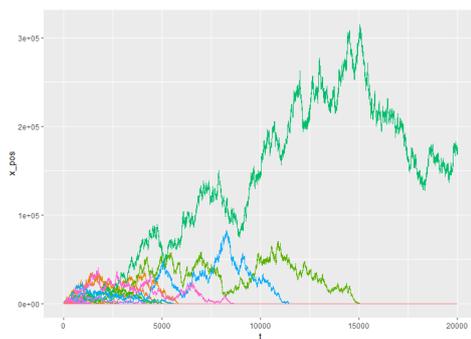
(c) Feller Sample Path Simulation with $\sigma^2 = 10$

(d) Feller Sample Path Simulation with $\sigma^2 = 100$

Figure 2: Feller Sample Path Simulations with $t_0 = 0$, $t_n = 20000$, $\Delta t_i = 1$, $x_0 = 1000$ and different values of σ^2 (100 paths). Note: Each path is represented by a different colour.

3.2 Martingale Property and Average Position

As equation (1.1) is a martingale (an SDE with no drift term), the expected position of a walker at time t to be x_o (calculation in the Appendix C). We can examine this via our simulation by evaluating the average position of several walkers at a particular time via:

$$\overline{x}(t) = \frac{1}{N} \sum_{m=1}^{N} x_m(t) \tag{3.1}$$

The fact that $\langle x(t) \rangle = x_o$ (given $x(0) = x_o$) might seem paradoxical given that the Feller process is one where eventually all walkers are absorbed at the origin. This can be understood through the simulation where the walkers that survive (even though this number will be small) will most likely be very far away from the origin and thus driving up the average given by equation (3.1). Results are displayed in Tables 3 and 4. Note the error bounds (confidence intervals) for the Monte Carlo estimates are in agreement with the theory (Glasserman, 2004) - these details can be found in Appendix C.

	Time (t)				
x_0	100	1000	5000	10000	20000
10	8.35	6.98	2.74	0.00	0.00
100	100.70	110.73	100.92	127.18	149.70
1000	1000.47	996.56	908.82	960.91	1011.71
10000	9956.68	9958.66	9752.99	9860.61	9775.20

Table 3: Average x value for varying x_0 values with $t_0 = 0$, $\sigma^2 = 10$, $\Delta t_i = 1$.
(10,000 paths)

	Time (t)				
x_0	100	1000	5000	10000	20000
10	10.07	11.07	10.29	12.68	14.92
100	100.05	99.72	90.96	96.19	99.01
1000	995.67	995.34	975.22	991.77	977.22
10000	9990.73	10000.84	10004.78	10049.72	9953.11

Table 4: Average x value for varying x_0 values with $t_0 = 0$, $\sigma^2 = 1$, $\Delta t_i = 1$.
(10,000 paths)

3.3 Hitting Time Distribution

From equation (2.6) we defined the absorption probability $F(t)$ and the corresponding survival probability $S(t)$, with the relation $S(t) = 1 - F(t)$. Given the nature of the Feller random walker, we can define these probabilities as:

$$S(t) = \Pr[x(t) > 0] \tag{3.2}$$
$$F(t) = \Pr[x(t) = 0] \tag{3.3}$$

We define a new random variable, the hitting time or first-passage time, denoted T^*:

$$T^* = \{\inf(t) : x(t) = 0\}$$

We can then define the hitting (first passage) time density by (Redner, 2001):

$$f(t) = -\frac{dS(t)}{dt}, \tag{3.4}$$

where the relation between $f(t)$ and T^* is given by

$$\text{Probability } T^* \in (t, t+dt) = f(t) \, dt$$

This in turn allows us to express the relations (3.2) and (3.3) in terns of relationships between T^* and $S(t)$, $F(t)$:

$$S(t) = \Pr[T^* > t] \tag{3.5}$$

$$F(t) = \Pr[T^* < t] = \int_0^t f(\tau) \, d\tau \tag{3.6}$$

Let us first turn our attention to a Brownian motion originating at x_o described by the SDE $\dot{x}(t) = \sigma\xi(t)$ with the origin set as an absorbing boundary. In this case, the hitting time density is given by (Redner, 2001):

$$f(t) = \frac{x_o}{\sqrt{2}\sigma t^{3/2}} \exp\left[\frac{-x_o^2}{2\sigma^2 t}\right], \tag{3.7}$$

whose integral over $t \in [0, \infty)$ is 1, indicating that absorption is certain. Paradoxically however the expected hitting time $\langle t \rangle = \langle T^* \rangle$ is infinite due to the large time power law behavior $f(t) \sim t^{-3/2}$ (Newman, 2005). Note the averaging here is with respect to the hitting time density $f(t)$. Setting $\frac{df}{dt} = 0$ for equation (3.7) yields the most likely (typical) hitting time

$$t^* = \frac{x_o^2}{3\sigma^2}$$

Now returning to our Feller walker we see an analogous situation to that of the absorbing Brownian walker. In this case the hitting time density is:

$$f(t) = \frac{2x_o}{\sigma^2 t^2} \exp\left[\frac{-2x_o}{\sigma^2 t}\right] \tag{3.8}$$

We saw earlier from equation (2.6) that absorption is certain, in terms of the density this is equivalent to the Brownian case with the integral of the hitting density over $t \in [0, \infty)$ equal to 1. As with the Brownian case, due to the large time power law (where now $f(t) \sim t^{-2}$) the expected hitting time is infinite:

$$\langle T^* \rangle = \langle t \rangle = \int_0^\infty t \, f(t) \, dt = \frac{2x_o}{\sigma^2} \int_0^\infty \frac{e^{-v}}{v} \, dv \tag{3.9}$$

Similarly setting $\frac{df}{dt} = 0$ for equation (3.8) yields the typical hitting time for the Feller walker as

$$t^* = \frac{x_o}{\sigma^2}$$

We use simulations to demonstrate the hitting time distributions for a Feller walker with $t_0 = 0$, $t_n = 20000$, $\Delta t_i = 1$, $x_0 = 100$ and $so^2 = 1, 10$ (Figure 3). A reference line (shown in red) has been marked on the plots to indicate the theoretical hitting time t^*. It can be seen that this appears at the peak of the histograms confirming this result. The theoretical distributions indicated in equation 3.8 are overlaid (in blue) as well to illustrate how closely the simulation agrees with the theoretical distribution.
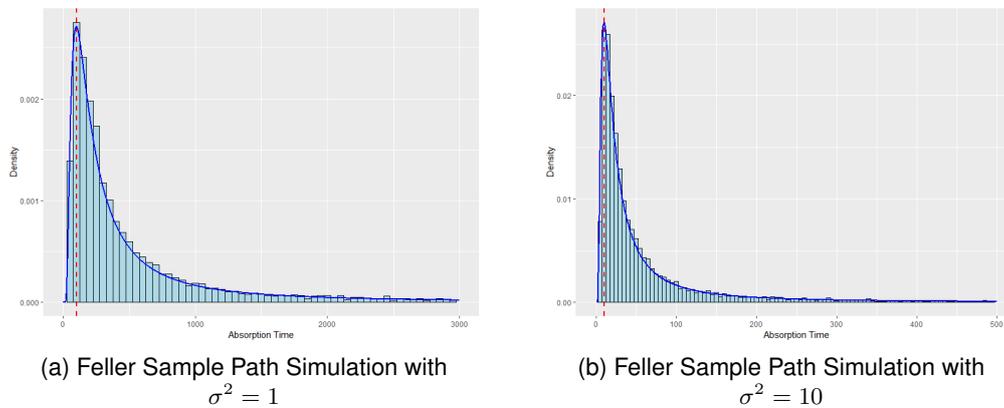
(a) Feller Sample Path Simulation with $\sigma^2 = 1$

(b) Feller Sample Path Simulation with $\sigma^2 = 10$

Figure 3: Hitting Time Distribution for Feller walkers with $t_0 = 0$, $t_n = 20000$, $\Delta t_i = 1$, $x_0 = 100$ and different values of σ^2.

4 Summary

In this article, we re-examine the derivation of the probability density of the Feller Diffusion equation. As opposed to the prior approaches (FellerW, 1951; Feller, 1951, 1954; Gan & Waxman, 2015), we use the Fourier transform on the Fokker-Planck equation corresponding to the Feller Diffusion and solve the resulting advection-type equation in Fourier space via the method of characteristics. The solution is easily mapped to the characteristic function of a scaled version of a non-central chi-squared distribution with zero degrees of freedom. As the theory of these distributions is well-known (Siegel, 1979), it is straight forward to interpret the results in terms of survival probabilities and further develop the algorithm for developing sample paths. We carry out such a simulation and i) find the simulated survival probability is very close to the theoretical value and ii) prove the martingale property by estimating the average position of the random walkers using Monte Carlo. We also find the hitting time distribution, including the typical hitting time shows good agreement between the theoretical results and simulation. For future work we plan to apply the Feller diffusion in the setting of quantitative finance where it has the potential to be used in the following modelling contexts: i) price dynamics of a risky asset or interest rate, ii) derivative pricing on risky underlying assets, iii) bond pricing with risky interest rates, iv) default probability estimation, v) survival analysis including hazard rate modelling and regression methods.

5 Acknowledgements

- Writing: RM/PK/LK

The authors declare that there is no conflict of interest regarding the publication of this article.

A   Fourier Transform of Fokker Planck Equation

In this section we look at deriving equation (2.1). We begin by re-writing equation (1.5) as

$$\frac{\partial}{\partial t}P(x,t) = \frac{1}{2}\sigma^2 x \frac{\partial^2}{\partial x^2}P(x,t) + \sigma^2 \frac{\partial}{\partial x}P(x,t) \tag{A.1}$$

Using the Fourier transform we obtain

$$\frac{\partial}{\partial t}\hat{P}(k,t) = \frac{1}{2}\sigma^2 \left[i\frac{\partial}{\partial k}(-k^2\hat{P}(k,t))\right] + ik\sigma^2 \frac{\partial}{\partial k}\hat{P}(k,t) \tag{A.2}$$

Note the use of following properties of Fourier transforms in deriving equation (A.2):

$$\frac{d^n G(x)}{dx^n} \to (ik)^n \hat{G}(k)$$

$$x \cdot G(x) \to i\frac{d}{dk}\hat{G}(k)$$

Expanding the right hand side of equation (A.2) and simplifying yields the result in equation (2.1):

$$\frac{\partial}{\partial t}\hat{P}(k,t) = \frac{-i\sigma^2 k^2}{2}\frac{\partial}{\partial k}\hat{P}(k,t)$$

B   Solution via Method of Characteristics

Let us solve the following PDE for $u(z,t)$ using the method of characteristics:

$$\frac{\partial u}{\partial t} + i\alpha z^2 \frac{\partial u}{\partial z} = 0 \tag{B.1}$$

$$u(z,t_o) = \psi(z) \tag{B.2}$$

The method of characteristics postulates a solution of the form $u(z(s),t(s))$ satisfying

$$0 = \frac{du}{ds} \tag{B.3}$$

$$= \frac{dt}{ds}\frac{\partial u}{\partial t} + \frac{dz}{ds}\frac{\partial u}{\partial z} \tag{B.4}$$

Now equation (B.3) implies that along the characteristic curve $(z(s),t(s))$, which is parametrized by s, the solution is constant, i.e.

$$u(s) = u(0) \Rightarrow u(z(s),t(s)) = u(z(0),t(0)) \tag{B.5}$$

To be explicit the $(z(s),t(s))$ and $(z(0),t(0))$ lie on the same characteristic curve. Matching the terms in equation (B.1) with equation (B.4) leads to the following system of ordinary differential equations (ODEs) for the characteristics:

$$\frac{dt}{ds} = 1 \quad t(s=0)=t_o \quad t(s)=t \tag{B.6}$$

$$\frac{dz}{ds} = -i\alpha z^2 \quad z(s=0)=z_o \quad z(s)=z \tag{B.7}$$

Equation (B.6) is solved by integrating as follows:

$$\int_{t_o}^{t} dt' = \int_{0}^{s} ds' \Rightarrow s = t - t_o \tag{B.8}$$

Equation (B.7) is solved by integrating as follows:

$$\int_{z_o}^{z} \frac{d\eta}{\eta^2} = -i\alpha \int_{0}^{s} ds' \Rightarrow z_o = \frac{z}{1 + i\alpha z s} = \frac{z}{1 + i\alpha z(t - t_o)} \tag{B.9}$$

Using equation (B.5) along with equations (B.8) and (B.9) enables us to obtain the solution

$$u(z, t) = u(z_o, t_o) = \psi(z_o) = \psi\left(\frac{z}{1 + i\alpha z(t - t_o)}\right) \tag{B.10}$$

We see that (B.10) satisfies the initial condition $u(z, t_o) = \psi(z)$. For the particular choice of initial condition (as in our problem):

$$\psi(z) = e^{-izx_o} \Rightarrow u(z, t) = \exp\left[\frac{-izx_o}{1 + i\alpha z(t - t_o)}\right] \tag{B.11}$$

We make the identifications to map back to our original problem:

$$\begin{aligned}
z &\rightarrow k \\
\alpha &\rightarrow \frac{\sigma^2}{2} \\
u(z, t) &\rightarrow \hat{P}(k, t) \\
\psi(z) &\rightarrow \exp[-ikx_o],
\end{aligned}$$

It is then straight forward to use (B.11)to arrive at the final answer given by equation (2.4):

$$\hat{P}(k, t) = \exp\left[\frac{-ikx_o}{1 + \frac{ik\sigma^2}{2}(t - t_o)}\right]$$

C Mean and Variance of Feller Diffusion Process

We may use the cumulant generating function $H(k, t)$ to derive the cumulants of the Feller Diffusion, where

$$H(k, t) = \ln \varphi(k, t) = \ln \hat{P}(-k, t) = \ln\langle e^{ikx} \rangle.$$

In particular the mean is given by

$$\begin{aligned}
\langle\, x(t) |\, x(0) = x_o \,\rangle &= -i\frac{\partial}{\partial k} H(k, t)\Big|_{k=0} \\
&= -i\frac{\partial}{\partial k}\left[\frac{-ikx_o}{1 + \frac{ik\sigma^2}{2}(t - t_o)}\right]\Big|_{k=0} \\
&= -i \cdot ix_o \cdot \left[1 + \frac{ik\sigma^2}{2}(t - t_o)\right]^{-2}\Big|_{k=0} \\
&= x_o \tag{C.1}
\end{aligned}$$

This matches our intuition as the random walk described by equation (1.1) is drift-less and thus a martingale. The variance of $x(t)$ is given by

$$
\begin{aligned}
\langle\, (x - \langle x \rangle)^2 \mid x(0) = x_o \,\rangle &= \left. (-i)^2 \frac{\partial^2}{\partial k^2} H(k, t) \right|_{k=0} \\
&= \left. -i \frac{\partial}{\partial k}\, x_o \cdot \left[1 + \frac{ik\sigma^2}{2}(t - t_o) \right]^{-2} \right|_{k=0} \\
&= \left. -ix_o \cdot i\sigma^2(t - t_o) \cdot \left[1 + \frac{ik\sigma^2}{2}(t - t_o) \right]^{-3} \right|_{k=0} \\
&= \sigma^2 x_o \cdot (t - t_o)
\end{aligned}
\tag{C.2}
$$

We use this value to obtain an error-bound on the Monte Carlo estimate of the average position, $\overline{x}(t)$ of a Feller random walker at time t given by equation(3.1). By Monte Carlo theory (Glasserman, 2004) $\overline{x}(t)$ is distributed (approximately) as a normal with a mean x_o (i.e. true value) and variance $\sigma^2 x_o t / N$. This enables us to obtain a confidence interval for x_o (Glasserman, 2004):

$$
\overline{x}(t) \pm \Phi^{-1}(1 - \alpha/2) \cdot \sqrt{\frac{\sigma^2 x_o t}{N}}
$$

Here Φ^{-1} is the inverse cumulative normal distribution and α is the confidence level. For the 95 percent confidence interval, we set $\alpha = 0.05$ and $\Phi^{-1}(0.975) = 1.96$.

To demonstrate this, we include here the results of the Monte Carlo simulations given in Section 3.2 with their 95% confidence intervals in Tables 5 and 6. It is noted that all the simulated results indicate that $\langle x(t) | x(0) = x_o \rangle = x_o$ within the error bounds.

x_0	Time (t)	$\overline{x}$	95% Confidence Interval	
			Lower Bound	Upper Bound
10	100	8.35	6.39	10.31
	1000	6.98	0.78	13.18
	5000	2.74	-11.12	16.60
	10000	0.00	-19.60	19.60
	20000	0.00	-27.72	27.72
100	100	100.70	94.50	106.90
	1000	110.73	91.13	130.33
	5000	100.92	57.09	144.75
	10000	127.18	65.20	189.16
	20000	149.70	62.05	237.35
1000	100	1000.47	980.87	1020.07
	1000	996.56	934.58	1058.54
	5000	908.82	770.23	1047.41
	10000	960.91	764.91	1156.91
	20000	1011.71	734.53	1288.89
10000	100	9956.68	9894.70	10018.66
	1000	9958.66	9762.66	10154.66
	5000	9752.99	9314.73	10191.25
	10000	9860.61	9240.81	10480.41
	20000	9775.20	8898.68	10651.72

Table 5: 95% Confidence Intervals of $\overline{x}$ values given in Table 3

x_0	Time (t)	$\overline{x}$	95% Confidence Interval	
			Lower Bound	Upper Bound
10	100	10.07	8.11	12.03
	1000	11.07	4.87	17.27
	5000	10.29	-3.57	24.15
	10000	12.68	-6.92	32.28
	20000	14.92	-12.80	42.64
100	100	100.05	93.85	106.25
	1000	99.72	80.12	119.32
	5000	90.96	47.13	134.79
	10000	96.19	34.21	158.17
	20000	99.01	11.36	186.66
1000	100	995.67	976.07	1015.27
	1000	995.34	933.36	1057.32
	5000	975.22	836.63	1113.81
	10000	991.77	795.77	1187.77
	20000	977.22	700.04	1254.40
10000	100	9990.73	9928.75	10052.71
	1000	10000.84	9804.84	10196.84
	5000	10004.78	9566.52	10443.04
	10000	10049.72	9429.92	10669.52
	20000	9953.11	9076.59	10829.63

Table 6: 95% Confidence Intervals of $\overline{x}$ values given in Table 4

D Zero Degrees of Freedom Non-central Chi-Squared Distribution

A non-central chi-squared distribution with zero degrees of freedom and non-centrality parameter λ (Siegel, 1979) is given by the density function:

$$g(x; \lambda) = e^{-\lambda/2} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j!} q(x; 2j) \tag{D.1}$$

Here $q(x; \nu)$ is the density function of a central chi-squared distribution with ν degrees of freedom.

$$q(x; \nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2} \tag{D.2}$$

Using the convention $q(x; 0) = \delta(x)$ we may re-write equation (D.1) as

$$g(x; \lambda) = e^{-\lambda/2} \left[\delta(x) + \sum_{j=1}^{\infty} \frac{(\lambda/2)^j}{j!} q(x; 2j) \right] \tag{D.3}$$

Equations (D.1) and (D.3) describe a compound Poisson mixture of central chi-square distributions with even degrees of freedom (Siegel, 1979). The corresponding characteristic function $\varphi_f(k; \lambda)$ can be derived from (D.3) very easily:

$$
\begin{aligned}
\varphi_g(k; \lambda) &= e^{-\lambda/2} \left[1 + \sum_{j=1}^{\infty} \frac{(\lambda/2)^j}{j!} \int_0^{\infty} q(x; 2j) e^{ikx} dx \right] \\
&= e^{-\lambda/2} \left[1 + \sum_{j=1}^{\infty} \frac{(\lambda/2)^j}{j!} (1 - 2ik)^{-j} \right] \\
&= \exp\left[\frac{ik\lambda}{1 - 2ik} \right]
\end{aligned}
\tag{D.4}
$$

We have used the fact that the characteristic function of $q(x; 2j)$ is $(1 - 2ik)^{-j}$ in the derivation. In addition, we note that we may write the density function in the alternative form (Siegel, 1979; Gan & Waxman, 2015)

$$g(x; \lambda) = e^{-\lambda/2}\delta(x) + \frac{1}{2}\sqrt{\frac{\lambda}{x}} e^{-(x+\lambda)/2} I_1\left(\sqrt{\lambda x}\right) \tag{D.5}$$

Here $I_1(z)$ denotes the modified Bessel function of order 1 which is given by

$$I_1(z) = \sum_{n=0}^{\infty} \frac{1}{n!(n+1)!} \left(\frac{z}{2}\right)^{2n+1} = \sum_{n=1}^{\infty} \frac{1}{n!(n-1)!} \left(\frac{z}{2}\right)^{2n-1} \tag{D.6}$$

Substituting equations (D.2) and (D.6) in equation (D.3) yields

$$
\begin{aligned}
g(x; \lambda) &= e^{-\lambda/2} \left[\delta(x) + \sum_{j=1}^{\infty} \frac{(\lambda/2)^j}{j!} \frac{x^{j-1}}{2^j \Gamma(j)} e^{-x/2} \right] \\
&= e^{-\lambda/2}\delta(x) + \frac{e^{-(x+\lambda)/2}}{x} \sum_{j=1}^{\infty} \frac{(\lambda x/4)^j}{j!(j-1)!} \\
&= e^{-\lambda/2}\delta(x) + \frac{e^{-(x+\lambda)/2}}{x} \frac{\sqrt{\lambda x}}{2} I_1\left(\sqrt{\lambda x}\right)
\end{aligned}
\tag{D.7}
$$

Re-arranging the last line leads to equation (D.5). The distribution has a straight forward interpretation where the the probability that random variable takes the value 0 is $e^{-\lambda/2}$ and the probability that the

random variable lies between $a(\neq 0)$ and b is given by the integral of the second term in (D.5) over the corresponding limits. If $a = 0$ we would include the δ function term in the integration. Finally, note for the Feller Diffusion process the density function (with t_o set to 0) is given by (Gan & Waxman, 2015)

$$P(x,t|x_o,0) = e^{-2x_o/\sigma^2 t} \cdot \delta(x) + \frac{2}{\sigma^2 t}\sqrt{\frac{x_o}{x}} \cdot e^{-2(x+x_o)/\sigma^2 t} \cdot I_1\left(\frac{4\sqrt{xx_o}}{\sigma^2 t}\right) \tag{D.8}$$

E Rcpp Implementation of Sample Path Simulation

The function CFellerPaths implemented below outputs a list containing time and the set of paths generated. The function allows to specify an initial seed, the number of paths required, a starting and ending time, Δt_i, the variance and the starting point.

```cpp
#include <Rcpp.h>
#include <math.h>
#include <random>

// [[Rcpp::plugins(cpp11)]]
using namespace Rcpp;

// [[Rcpp::export]]
List CFellerPaths(int seed, int num, int t0, int tn, double dt, double var,
double x0){
        int n = ceil(((double(tn - t0))/dt) + 1.0);

        NumericMatrix Paths(n, num);
        static std::mt19937 gen;

        NumericVector t(n);
        t[0] = t0;

        for(int j=0; j<num; j++){
                NumericVector x(n);
                x[0] = x0;

                gen.seed(seed);
                seed += 1;

                for(int i=1; i<n; i++){

                        double c = var/(4.0*dt);
                        double lambda = float(x[i-1])/c;

                        std::poisson_distribution<> pd((lambda/2.0));
                        double N = pd(gen);
                        if(N==0){
                                x[i] = 0.0;
                        }else{
                                std::chi_squared_distribution<> cd((2*N));
                                x[i] = c*cd(gen);
                        }
```

```
        }

            Paths(_,j) = x;
    }

    for(int i=1; i<n; i++){
            t[i] = t[i-1] + dt;
    }

    return(List::create(Named("t")=t,
                        Named("paths")=Paths));

}
```

References

Cardy J. Scaling and Renormalization in Statistical Physics. Cambridge: Cambridge University Press; 1996. (doi:10.1017/cbo9781316036440)

Cox JC. The Constant Elasticity of Variance Option Pricing Model. The Journal of Portfolio Management. 1996;22:15-17.

Cox JC, Ingersoll J, Ross S. A Theory of the Term Structure of Interest Rates. Econometrica. 1985;53(2):385-407. (doi:10.2307/1911242)

Docherty P, Dong Y, Song X, Tippett M. The Feller diffusion, filter rules and abnormal stock returns. European Journal of Finance. 2018;24(5):426-438. (https://doi.org/10.1080/1351847X.2017.1309328)

Feller W. Diffusion processes in one dimension. Transactions of the American Mathematical Society. 1954;77:1–31. (doi:10.1090/s0002-9947-1954-0063607-6)

Feller W. Diffusion processes in genetics. In: Neyman J, editor. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability 1950. Berkeley, Los Angeles (CA): University of California Press; 1951:227–246

Feller W. Two singular diffusion problems. Annals of Mathematics. 1951;54:173–182. (doi:10.2307/1969318)

Gan X, Waxman D. Singular Solution of the Feller Diffusion Equation via a Spectral Decomposition. Phys. Rev. E. 2015;91(1):012123. (doi:10.1103/physreve.91.012123)

Glasserman P. Monte Carlo Methods in Financial Engineering. New York: Springer; 2004. (doi:10.1007/978-0-387-21617-1)

Göing-Jaeschke A, Yor M. A survey and some generalizations of Bessel processes. Bernoulli. 2003;9(2):313-349. (doi:10.3150/bj/1068128980)

Hinrichsen H. Adv. Phys. 2000;49:815-918. (doi:10.1080/00018730050198152)

Hsu YL, Lin TI, Lee CF. Constant elasticity of variance (CEV) option pricing model: Integration and detailed derivation. Mathematics and Computers in Simulation. 2008;79(1):60-71. (doi.org/10.1016/j.matcom.2007.09.012)

Hull JC. Options, Futures, and Other Derivatives. 9th ed. Singapore: Pearson; 2012.

Lindsay AE, Brecher DR. Simulation of the CEV process and the local martingale property. Mathematics and Computers in Simulation. 2012;82(5):868-878. (doi.org/10.1016/j.matcom.2011.12.006)

Newman MEJ. Power Laws, Pareto Distributions And Zipf's Law. Contemporary Physics. 2005;46:323-351. (doi:10.1080/00107510500052444)

Oksendal B. Stochastic Differential Equations: An Introduction with Applications. 5th ed. New York: Springer-Verlag; 2000. (doi:10.1007/978-3-642-14394-6)

Redner S. A Guide to First-Passage Processes. Cambridge: Cambridge University Press; 2001. (doi:10.1017/cbo9780511606014)

Revuz D, Yor M. Continuous Martingales and Brownian Motion. 3rd ed. Berlin: Springer-Verlag; 1999. (doi:10.1007/978-3-662-06400-9)

Siegel AF. The Noncentral Chi-Squared Distribution with Zero Degrees of Freedom and Testing for Uniformity. Biometrika. 1979;66(2):381–386. (doi:10.1093/biomet/66.2.381)

Wilmott P. Frequently Asked Questions in Quantitative Finance. 2nd ed. Chichester: Wiley; 2009. (doi:978-0470748756)