

# Using the Chapman-Kolmogorov equation of random walks to identify drift and diffusion of the Fokker-Planck equation

**M. Annunziato**

*Dipartimento di Fisica "E. R. Caianiello"  
Università degli Studi di Salerno  
Via G. Paolo II 132  
Fisciano, 84084, Italy*

*mannunzi@unisa.it*

**A. Borzi**

*Institut für Mathematik  
Universität Würzburg  
Emil-Fischer-Strasse 30  
Würzburg, 97074, Germany*

*alfio.borzi@uni-wuerzburg.de*

---

## Abstract

A novel approach for the reconstruction of the drift and diffusion coefficients of the Fokker-Planck equation is presented. This approach is based on the Chapman-Kolmogorov equation of the inhomogeneous random walk related to the Fokker-Planck equation. Two numerical algorithms are formulated for the reconstruction problem. Results of numerical experiments demonstrate the ability of the proposed methods to solve this inverse problem also in the case of discontinuous coefficients.

**Keywords:** Reconstruction of diffusion and drift coefficients, random walk, Fokker-Planck equation, calibration.

**MSC:** 49J20, 49K20, 49N45, 65N21.

---

## 1. INTRODUCTION

The parabolic Fokker-Planck (FP) equation models the evolution of the probability density function (PDF) of Itô stochastic differential equations (SDEs) driven by drift and diffusion Wiener processes, and this fact can be easily verified by comparison of the PDFs obtained by direct simulation of the parabolic FP problem and of the Monte Carlo (MC) simulation of the related SDE. Actually, this MC approach can be used to numerically solve the FP problem.

Much more challenging is the inverse problem of obtaining the drift and diffusion coefficients based on the (partial) knowledge of the PDFs. For this purpose, some methods have been proposed that we can classify as follows.

Some methods belong to the class of PDE calibration problems. These methods provide satisfactory reconstruction of one of the two coefficients subject to appropriate regularity conditions; see, e.g.,

(Albeverio, Blanchard, Kusuoka, & Streit, 1989; Annunziato & Gottschalk, 2018; Banks, Tran, & Woodward, 1993; Dunker & Hohage, 2014; Egger & Engl, 2005; Jäger & Kostina, 2005; Lund, Hubbard, & Halter, 2014).

On the other hand, there are calibration methods based on discretely sampled observations of realizations of continuous-time drift-diffusion processes. These methods use kernel-based or histogram-based regression; see, e.g., (Comte, 2020; Comte & Genon-Catalot, 2025; Florens-Zmirou, 1989, 1993; Jamba, Subray, & Rodrigues, 2024; Jiang & Knight, 1997; Lamouroux & Lehnertz, 2009; Nicolau, 2003; Renò, 2008). Recently, machine learning techniques for reconstruction of FP drift and diffusion coefficients have been proposed; see, e.g. (Chen, Yang, Duan, & Karniadakis, 2021; Han et al., 2025), which could be extended based on our approach.

Our assessment of the existing literature is that the first class of PDE-based methods pursues a macroscopic approach and aims at identifying the coefficients from observations of the PDF. However, these methods seem unable to identify both coefficients and have difficulties in recovering less regular ones, like piecewise continuous functions. The other class of methods looks directly to path realizations of the stochastic process and therefore has a microscopic character. For this reason, they can be applied to recover the drift and diffusion coefficients only in a narrow window, and appear less accurate in the reconstruction of less regular coefficients, which seems to be also a limitation in the machine-learning approach.

We propose and validate a methodology that lies between the two classes mentioned above, in the sense that: 1) we assume that the stochastic process is approximated by an inhomogeneous random walk that is characterized by jump probabilities at each grid point where the random walk is defined; 2) we observe multiple realizations of the random walk and construct the corresponding numerical PDF; 3) we use this PDF as data for the Chapman-Kolmogorov equation of the random walk, and perform the reconstruction of the jump probabilities by inverting this equation. We infer the drift and diffusion coefficients of the FP equation from these jump probabilities.

Notice that in (Lund et al., 2014) a similar approach based on the inversion of the FP equation jointly to polynomial approximation is used. The authors discuss the volatility of the estimates of the coefficients for critical values the PDF and its derivative. However, the results seems limited to smooth drift and diffusion coefficients.

We remark that our strategy represents a consistent approximation of the continuous problem of inverting the Fokker-Planck equation. In fact, in the appropriate limit of a grid whose mesh size tends to zero, we have that 1) the random walk will convergence to the continuous stochastic process; 2) the numerical PDF converges to the continuous one; 3) the Chapman-Kolmogorov equation becomes the Fokker-Planck equation. The advantage of our novel heuristic approach is that working with a coarse grid drastically reduces the ill-posedness of the continuous inverse problem by acting as a regularization framework.

We implement our strategy in a reconstruction algorithm that requires to solve a linear algebraic system in a least squares sense. We successfully validate this algorithm with challenging test cases including the case where the drift and the diffusion are discontinuous.

In the next section, we illustrate our inhomogeneous random walk, and define the jump probabilities. In Section 3., we discuss our approach that determines the random walk parameters by inverting the Chapman-Kolmogorov equation based on a finite set of data. Our algorithm is presented in this section. Section 4. is devoted to the derivation of the Fokker-Planck equation as a continuous limit of the random walk. In Section 5., we present results of numerical experiments that successfully validate our methodology. A section of conclusion completes this work.

## 2. A RANDOM WALK WITH DRIFT

The discrete random walk is the basic stochastic process whose elementary set of events  $\Omega$  has two states, and two probabilities  $p \in [0, 1]$  and  $q \in [0, 1]$ , such that  $p + q = 1$ . Let  $W$  be the random variable

related to these events, which takes the values  $\Delta x$  and  $-\Delta x$  with probability  $p$  and  $q$  respectively.

On the discrete set of points  $\mathbb{M}_{\Delta x} \equiv \{x_m = m\Delta x, m \in \mathbb{Z}\}$ , we define the random walk (RW) on  $\mathbb{M}_{\Delta x}$  as the process resulting by summing up a discrete sequence of independent outcomes  $W_i$ ,  $i = 1, \dots, n$ . Therefore, if the process is initially placed at  $x^{(0)} \in \mathbb{M}_{\Delta x}$ , with probability 1, and at each step  $i$  the position is displaced according to the outcomes of independent random jumps  $W_i$ , then the value of the RW at the step  $n$  is given by the following sum:

$$X_n = x^{(0)} + \sum_{i=1}^n W_i. \quad (1)$$

In the case of a two-dimensional RW, we consider  $\mathbb{M}_{\Delta x, \Delta y}^2 \equiv \{(x_{m_1} = m_1 \Delta x, y_{m_2} = m_2 \Delta y), \text{ for } (m_1, m_2) \in \mathbb{Z}^2\}$  with mesh sizes  $\Delta x$  and  $\Delta y$ , and define the RW  $\mathbf{X}_n = (X_n, Y_n)$  as composition of two independent RW.

At each time step the process initially placed in  $(x_{m_1}, y_{m_2})$  moves towards one of the 4 nearest neighbour points of the grid according to the following random vector valued variable  $\sigma(\mathbf{x})$  defined as follows:

$$\sigma(\mathbf{x}) = \begin{cases} (+1, 0) & \text{with prob. } p_1(\mathbf{x}) \in [0, 1] \\ (0, -1) & \text{with prob. } p_2(\mathbf{x}) \in [0, 1] \\ (-1, 0) & \text{with prob. } p_3(\mathbf{x}) \in [0, 1] \\ (0, +1) & \text{with prob. } p_0(\mathbf{x}) \in [0, 1] \end{cases} \quad (2)$$

We have the normalizing condition:

$$p_0(\mathbf{x}) + p_1(\mathbf{x}) + p_2(\mathbf{x}) + p_3(\mathbf{x}) = 1. \quad (3)$$

The dependence of the jump probability on the grid point, defines an inhomogeneous RW. For a RW starting from  $\mathbf{X}_0 = \mathbf{x}^{(0)}$ , the position at the time  $n$  is

$$\mathbf{X}_n = \mathbf{x}^{(0)} + \Delta \mathbf{x} \circ \sum_{j=1}^n \sigma(\mathbf{X}_{j-1}), \quad (4)$$

where  $\circ$  is the Hadamard product. For convenience, in the following, in place of  $\sigma(\mathbf{x})$  we shall use the discrete random variable  $z(\mathbf{x}) \in \{0, 1, 2, 3\}$  to refer to the 4 possible outcomes with the corresponding probabilities  $p_z(\mathbf{x})$ . For example, let  $z = 2$  then according to (2) the outcome of  $\sigma$  is  $\sigma(z) = (0, -1)$ . Notice that since  $p_z(\mathbf{x})$ , i.e.  $\sigma(\mathbf{x})$ , is a function of the grid, then the RW is inhomogeneous.

In general, the RW is dispersive with finite velocity of propagation, that is, if we set the initial value of the RW concentrated at the point  $\mathbf{x}^{(0)}$ , then at the time step  $n$  the distribution of the process will be contained in the discrete domain  $[x^{(0)} - n\Delta x, x^{(0)} + n\Delta x] \subset \mathbb{M}_{\Delta x} \times [y^{(0)} - n\Delta y, y^{(0)} + n\Delta y] \subset \mathbb{M}_{\Delta y}$ , whereas outside this domain the distribution will be surely vanishing, which defines a natural boundary condition.

The further step of our analysis is writing the probability measure of the RW. For this purpose, let  $k_1$  be the number of positive jumps along the  $x$  direction,  $n_1$  be the total number of jumps along the  $x$  direction,  $k_0$  be the number of positive jumps along the  $y$  direction, and  $n_2$  be the total number of jumps along the  $y$  direction. For this RW, we can write the conditional probability to find the process at time  $n \geq 1$  to the point  $(x_m, y_l) = (x_{2k_1 - n_1} + x^{(0)}, y_{2k_0 - n_2} + y^{(0)})$  as follows:

$$\mathbb{P}\{\mathbf{X}_n = ((2k_1 - n_1)\Delta x, (2k_0 - n_2)\Delta y) + \mathbf{x}^{(0)} | \mathbf{X}_0 = \mathbf{x}^{(0)}\} = \sum_{\mathbf{z} = \mathbf{C}_4(n, n_1, n_2, k_1, k_0)} \prod_{i=0}^{n-1} p_{z_i}(\mathbf{x}^{(i)}) \Big|_{\mathbf{x}^{(i)} = \mathbf{x}^{(0)} + \Delta \mathbf{x} \circ \sum_{j=1}^i \sigma(\mathbf{z}_j)}, \quad (5)$$

where  $\mathbf{z} = (z_0, z_1, \dots, z_n)$  is a vector whose elements span over the set of all  $C_4(\dots)$  combinations of the digits 0, 1, 2, 3 of length  $n$ , with  $k_0$  the count of 0's,  $k_1$  the count of 1's,  $n_1$  is the count of the odd numbers 1 and 3,  $n_2$  the count of the even numbers 0 and 2, such that  $n = n_1 + n_2$ , and  $k_1 \leq n_1$ ,  $k_0 \leq n_2$ . As an example, for  $\mathbf{z} = 02213001$ , it is  $n = 8$ ,  $n_1 = 3$ ,  $n_2 = 5$ ,  $k_0 = 3$ ,  $k_1 = 2$ .

Notice that, from Eq. (5) for a given time step  $n$  and positive jumps  $k_0, k_1$  in the two dimensions, the displacement from the initial position is not uniquely determined, since there is one degree of freedom in the choice of  $n_1$  and  $n_2$  set by  $n = n_1 + n_2$ .

### 3. THE COMPUTATION OF THE RW PARAMETERS

By many repetitions of our RW process of time length  $n$ , starting from the same position, we can collect the probability distribution of the RW at the final time  $n$ . On the other hand, let the probability distribution at the final time be given, then our purpose is to calculate the values of the jump probabilities at each grid point.

The problem is to establish (2) for each point  $\mathbf{x}$  of the domain, when the 2D probability distribution  $f_{m,l}^n$ , of the particle to be located at  $(x_m, y_l) \in \mathbb{M}_{\Delta x, \Delta y}^2$ , is given for some values of the time step  $n$ . In this section, for the unknown probabilities we will use the notation  $p_{m,l}^{(s)} = p_s(x_m, y_l)$ .

Our approach is based on the 2-dimensional Chapman-Kolmogorov equation given by

$$f_{m,l}^{n+1} = f_{m-1,l}^n p_{m-1,l}^{(1)} + f_{m,l+1}^n p_{m,l+1}^{(2)} + f_{m+1,l}^n p_{m+1,l}^{(3)} + f_{m,l-1}^n p_{m,l-1}^{(0)}. \quad (6)$$

This equation governs the evolution of the probability such that the RW reaches the point  $(m, l)$  at time  $n + 1$  by jumping from the 4 nearest neighbours, according to the table (2). The total probability distribution is subject to conservativeness condition

$$\sum_{m,l} f_{m,l}^n = 1, \quad \text{for all } n.$$

Notice that the four unknowns are not constrained by the normalizing condition, since they are related to different locations of the grid, indeed the above condition applies to the  $p_{m,l}^{(s)}$  related to the same grid point.

One could propose to solve Eq. (6) for all  $(m, l)$  of the grid domain for fixed time steps  $n$  and  $n + 1$ , but this is not possible because the number of unknowns  $p_{m,l}^{(s)}$  is greater than the number of equations (6) for each grid point. In fact, suppose that at a certain time step the domain of the positive probabilities  $f_{m,l}^{n+1}$  is a square of side size  $M$  points, then there are  $M^2$  equations (6) and  $M^2$  normalizing equations (3) that is  $2M^2$ , whereas  $4M^2$  unknowns.

The number of the unknowns can be reduced if we consider the vanishing values of  $f_{m,l}^n$  outside the domain, that is,  $4M$  the number of the nearest neighbours of the perimeter of the domain. Hence, there are  $2M^2$  equations and  $4M^2 - 4M$  unknowns, which gives an underdetermined system of equations. Indeed, it can be shown that with an additional constraint to  $p_{m,l}^{(s)}$ , i.e. with only 2 degrees of freedom for the jump probabilities, it is possible to obtain a determined system of equations. Notice that this is also possible in the 1D case.

Therefore, in order to solve the inverse problem for the 2D-RW, we must include in the system so many time points to get enough equations to cover the number of unknowns. However, if the system of equations results rank-deficient a least squares method can be used to find a solution, eventually by including the normalizing constraints (3).

We can summarize the discussion above by noting that from Eq. (6) we get the following algebraic

problem for the unknown probabilities:

$$\begin{pmatrix} f_{m-1,l}^{n_1} & f_{m,l+1}^{n_1} & f_{m+1,l}^{n_1} & f_{m,l-1}^{n_1} \\ f_{m-1,l}^{n_2} & f_{m,l+1}^{n_2} & f_{m+1,l}^{n_2} & f_{m,l-1}^{n_2} \\ f_{m-1,l}^{n_3} & f_{m,l+1}^{n_3} & f_{m+1,l}^{n_3} & f_{m,l-1}^{n_3} \\ f_{m-1,l}^{n_4} & f_{m,l+1}^{n_4} & f_{m+1,l}^{n_4} & f_{m,l-1}^{n_4} \end{pmatrix} \begin{pmatrix} p_{m-1,l}^{(1)} \\ p_{m,l+1}^{(2)} \\ p_{m+1,l}^{(3)} \\ p_{m,l-1}^{(0)} \end{pmatrix} = \begin{pmatrix} f_{m,l}^{n_1+1} \\ f_{m,l}^{n_2+1} \\ f_{m,l}^{n_3+1} \\ f_{m,l}^{n_4+1} \end{pmatrix}, \quad (7)$$

We extend this equation to include more time points  $n_1, \dots, n_K$  with  $K \geq 4$ , so that it reads as follows:

$$\begin{pmatrix} f_{m-1,l}^{n_1} & f_{m,l+1}^{n_1} & f_{m+1,l}^{n_1} & f_{m,l-1}^{n_1} \\ f_{m-1,l}^{n_2} & f_{m,l+1}^{n_2} & f_{m+1,l}^{n_2} & f_{m,l-1}^{n_2} \\ f_{m-1,l}^{n_3} & f_{m,l+1}^{n_3} & f_{m+1,l}^{n_3} & f_{m,l-1}^{n_3} \\ \dots & \dots & \dots & \dots \\ f_{m-1,l}^{n_K} & f_{m,l+1}^{n_K} & f_{m+1,l}^{n_K} & f_{m,l-1}^{n_K} \end{pmatrix} \begin{pmatrix} p_{m-1,l}^{(1)} \\ p_{m,l+1}^{(2)} \\ p_{m+1,l}^{(3)} \\ p_{m,l-1}^{(0)} \end{pmatrix} = \begin{pmatrix} f_{m,l}^{n_1+1} \\ f_{m,l}^{n_2+1} \\ f_{m,l}^{n_3+1} \\ \dots \\ f_{m,l}^{n_K+1} \end{pmatrix}. \quad (8)$$

This overdetermined system of equations can be solved as a least-squares problem represented as  $\min_p \|Ap - b\|_2^2$ . That is, a system of  $K$  equations for 4 unknowns, solved by using the Moore-Penrose pseudoinverse of the coefficient matrix. However, this method does not guarantee that the solution satisfies the normalizing condition, which must be enforced after this calculation.

With this preparation, we can formulate the algorithm for the solution of our inverse problem as follows:

#### Algorithm 1

1. *Input data: probability distribution  $f_{m,l}^n$  of the RW, for  $n = n_1, n_2, \dots, n_K$  time steps. Set the matrix  $p_{m,l}^{(s)}$  to save the results. Set  $\text{count\_}p_{m,l}^{(s)} = 0$  matrix to flag the calculated values of  $p_{m,l}^{(s)}$ .*
2. *Start the iteration over all the values of  $(m, l)$  of the grid domain given by the size of  $f_{m,l}^n$ .*
3. *Build the matrices of Eq. (8) for a given  $(m, l)$  and for all  $n = n_1, n_2, \dots, n_K$ .*
4. *If the system matrix has not rank 4, then jump the next step.*
5. *Solve the system (8) with the Moore-Penrose pseudoinverse. Cumulates the calculated values of  $p_{m,l}^{(s)}$  and increments the counter  $\text{count\_}p_{m,l}^{(s)} + 1$ . During the iteration over all points  $(m, l)$  of the domain, the  $p_{m,l}^{(s)}$  could be calculated more than once repeatedly. The counters  $\text{count\_}p_{m,l}^{(s)}$  keep track of the calculated values of the corresponding  $p_{m,l}^{(s)}$ .*
6. *End of iteration 2.*
7. *Averages the calculated  $p$ :  $p_{m,l}^{(s)} / \text{count\_}p_{m,l}^{(s)}$ , in order to take in account for the eventually repetitions of the calculated values of  $p_{m,l}^{(s)}$ .*

The solution of the system at the step 5 can be executed by using available routines, e.g., `lsqminnorm`, `lsqnonneg` of Matlab®, or using the `lsqlin` function, setting the boundaries  $lb = 0$  and  $lu = 1$  for the unknown and the method `interior-point`.

The check on the rank at the step 4 is important because the system of equations can be rank deficient. This occurrence depends on vanishing values of  $f$  at the boundary of the domain depending on the choice of the time points  $n_1, \dots, n_K$ , and for resulting symmetries such that some equations of the system result linearly dependent. However, notice that this algorithm does not automatically guarantee the normalizing condition (3).

In order to guarantee fulfilment of the normalizing condition (3), we consider the system of Chapman-Kolmogorov equations for the 4 nearest neighbours of the point of calculation and require (3) for that point. The resulting system in the form  $A\tilde{p} = b$  is given by

$$A^{(n)} = \begin{pmatrix} f_{m,l}^n & f_{m+1,l+1}^n & f_{m+2,l}^n & f_{m+1,l-1}^n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{m-1,l-1}^n & f_{m,l}^n & f_{m+1,l-1}^n & f_{m,l-2}^n \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ f_{m-2,l}^n & f_{m-1,l+1}^n & f_{m,l}^n & f_{m-1,l-1}^n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{m-1,l+1}^n & f_{m,l+2}^n & f_{m+1,l+1}^n & f_{m,l}^n \end{pmatrix} \quad (9)$$

$$b^{(n)} = \begin{pmatrix} f_{m+1,l}^{n+1} \\ f_{m,l-1}^{n+1} \\ f_{m-1,l}^{n+1} \\ f_{m,l+1}^{n+1} \end{pmatrix} \quad (10)$$

$$\tilde{p} = \begin{pmatrix} p_{m,l}^{(1)} & p_{m+1,l+1}^{(2)} & p_{m+2,l}^{(3)} & p_{m+1,l-1}^{(0)} & p_{m-1,l-1}^{(1)} & p_{m,l}^{(2)} & p_{m+1,l-1}^{(3)} & p_{m,l-2}^{(0)} \\ p_{m-2,l}^{(1)} & p_{m-1,l+1}^{(2)} & p_{m,l}^{(3)} & p_{m-1,l-1}^{(0)} & p_{m-1,l+1}^{(1)} & p_{m,l+2}^{(2)} & p_{m+1,l+1}^{(3)} & p_{m,l}^{(0)} \end{pmatrix}^T \quad (11)$$

The matrices  $A^{(n)}$  and  $b^{(n)}$  are stack according to at least 4 values of  $n$ , so that we get a system of  $4K$  equations for 16 unknowns. The normalizing equation (3) reads as follows:

$$\sum_{s=0}^3 p_{m,l}^{(s)} = 1. \quad (12)$$

With these additions, the new algorithm including the normalizing conditions is given by

## Algorithm 2

1. Input data: probability distribution  $f_{m,l}^n$  of the RW, for  $n = n_1, \dots, n_K$  time steps with  $K \geq 4$ . Set the matrix  $p_{m,l}^{(s)}$  to save the results.
2. Iterate over all the values of  $(m, l)$  of the domain given by the size of  $f_{m,l}^n$ .
3. Build the matrices of Eq. (9) and and the vector (10).
4. If the resulting system matrix (9) has not rank 16, then jump the next two steps.
5. Include the normalizing condition (12) in the system, by appending the corresponding coefficients to (9) and (10), then solve it with the Moore-Penrose pseudoinverse.
6. From the calculated  $\tilde{p}$  of Eq. (11) save the values at the positions 1, 6, 11, 16 of that vector correspondingly to the matrix positions  $p_{m,l}^{(1)}, p_{m,l}^{(2)}, p_{m,l}^{(3)}, p_{m,l}^{(4)}$ .
7. End of iteration 2.

We remark that the outputs of our algorithms are post-processed in order to erase points that are subject to numerical errors due to a badly conditioned matrices. For this purpose we use the Matlab® (The MathWorks, 2021) function `filloutliers` with the options `'linear'`, `'movmedian'`, 15, `'ThresholdFactor'`, 0.7.

#### 4. FROM RANDOM WALKS TO FOKKER-PLANCK EQUATIONS

It is well known that the FP equations can be obtained by a suitable limiting process of RW models (Breitenbach, Annunziato, & Borzi, 2018; Cox & Miller, 1992). In this section, we refer to this fact in order to formulate the inverse problem for the continuous FP model.

Let  $f(x, y, t) \in \mathcal{M} \subseteq \mathbb{R}^2 \times \mathcal{T} \subseteq \mathbb{R}$  and  $p_s(x, y) \in \mathcal{M}$  be continuous and differentiable functions, by writing Eq. (6) in terms of these functions, we get

$$\begin{aligned} f(x_k, y_j, t_n + \Delta t) = & f(x_k + \Delta x, y_j, t_n) p_3(x_k + \Delta x, y_j) \\ & + f(x_k - \Delta x, y_j, t_n) p_1(x_k - \Delta x, y_j) \\ & + f(x_k, y_j - \Delta x, t_n) p_0(x_k, y_j - \Delta x) \\ & + f(x_k, y_j + \Delta x, t_n) p_2(x_k, y_j + \Delta x) \end{aligned} \quad (13)$$

where we have defined the uniformly spaced time grid  $\mathbb{T}_{\Delta t}$  of step size  $\Delta t$ , such that  $t_n = t_0 + n\Delta t$  for positive integer  $n$ , and used the same step size  $\Delta x$  for both space dimensions, i.e.  $(x_n, y_l) \in \mathbb{M}_{\Delta x, \Delta x}$ .

We define the dynamic parameters of the following limit equations:

$$\begin{aligned} a(x, y) &= p_1(x, y) + p_3(x, y) \\ \mu_1(x, y) &= p_1(x, y) - p_3(x, y) \\ \mu_2(x, y) &= p_0(x, y) - p_2(x, y). \end{aligned} \quad (14)$$

Therefore, we also have

$$\begin{aligned} p_1(x, y) &= \frac{1}{2}(a(x, y) + \mu_1(x, y)) \\ p_3(x, y) &= \frac{1}{2}(a(x, y) - \mu_1(x, y)) \\ p_2(x, y) &= \frac{1}{2}(-a(x, y) - \mu_2(x, y) + 1) \\ p_0(x, y) &= \frac{1}{2}(-a(x, y) + \mu_2(x, y) + 1) \end{aligned} \quad (15)$$

Further, due to the normalizing conditions, we have the following conditions

$$\begin{aligned} 0 &\leq a + \mu_1 \leq 2 \\ 0 &\leq a - \mu_1 \leq 2 \\ 0 &\leq 1 - a - \mu_2 \leq 2 \\ 0 &\leq 1 - a + \mu_2 \leq 2. \end{aligned} \quad (16)$$

By applying Taylor's expansion in (13) up to second order, we get

$$\begin{aligned} 0 = & \Delta t \partial_t f(x, y, t) + \Delta x [\partial_x (\mu_1(x, y) f(x, y, t)) + \partial_y (\mu_2(x, y) f(x, y, t))] \\ & - \frac{1}{2} \Delta t^2 \partial_t^2 f(x, y, t) - \frac{1}{2} \Delta x^2 [\partial_x^2 (a(x, y) f(x, y, t)) + \partial_y^2 ((1 - a(x, y)) f(x, y, t))] \\ & + o((\Delta x + \Delta t)^2), \end{aligned} \quad (17)$$

where  $o((\Delta x + \Delta t)^2)$  means an infinitesimal order superior with respect to  $(\Delta x + \Delta t)^2$ .

The classical approach to build the Fokker-Planck equation is the choice of the scaling in a such way that  $\Delta x^2 \propto \Delta t$ . Hence, we set the quadratic scaling  $\Delta x^2 = D\Delta t$  jointly to the dependence on the grid size  $\Delta x$  in the drift function, that is, the substitution  $\tilde{\mu}_i = \mu_i D / \Delta x$ , as well as  $\mu_i(x) \rightarrow \tilde{\mu}_i(x) \Delta x / D$ , where  $D$  is a positive diffusion constant. From Eq. (17) in the limit of vanishing  $\Delta t$  and neglecting the terms  $o((\Delta x + \Delta t)^2)$ , we obtain the FP equation

$$\begin{aligned} \partial_t f(x, y, t) - \frac{D}{2} \partial_x^2 (a(x, y) f(x, y, t)) - \frac{D}{2} \partial_y^2 ((1 - a(x, y)) f(x, y, t)) \\ + \partial_x (\tilde{\mu}_1(x, y) f(x, y, t)) + \partial_y (\tilde{\mu}_2(x, y) f(x, y, t)) = 0. \end{aligned} \quad (18)$$

This equation is completed with the initial

$$f(x, y, 0) = f_0(x, y) \quad (19)$$

and normalization

$$\int_{\mathcal{M}} f(x, y, t) dx dy = 1 \quad (20)$$

conditions. Notice, that if the sum in Eq. (3) is less than 1, i.e. there is a positive probability that the random walk does not jump, then the r.h.s. of Eq. (13) has the further addend  $f(x_k, y_j, t_n)(1 - \sum_{s=0}^3 p_s(x_k, y_j))$ . As a consequence, after the vanishing step grid limit, the FP equation (18) has two independent diffusion terms  $a_1(x, t)$  and  $a_2(x, t)$ . This means that the inverse problem could be solved, by using the random walk approximation, also for this more general case, albeit the numerical solution is much more sensitive to approximations due to the presence of one more unknown.

## 5. THE NUMERICAL SOLUTION OF THE INVERSE PROBLEM

Our goal is the estimation of the coefficients of the FP equation from observation of the PDF as follows: let the PDF  $f(x, y, t)$  be given for  $(x, y, t) \in \mathcal{M} \times \mathcal{T}$ , then find  $(\mu_1(x, y), \mu_2(x, y))$  and  $a(x, y)$  for the FP equation (18).

In practice, we use numerical PDFs in the form of histograms related to a binning process as usual in Monte Carlo experiments. Further, we work with synthetic data, which allows to test the accuracy and robustness of our reconstruction strategy. For this purpose, we choose the functions object of the reconstruction problem  $(\mu_1(x, y), \mu_2(x, y))$  and  $a(x, y)$ , then calculate numerically the PDF from the solution of the FP equation. Such a PDF is coarse grained in the form of a histogram in order to simulate the aforementioned case of measurements. The values  $f_{m,l}^n$ , needed as input data for our algorithms, are obtained by interpolation of the histogram values on the random-walk grid points. With this data, the algorithms calculate  $\mu_{m,l}^i \approx \mu_i(x_m, y_l)$  for  $i = 1, 2$ , and  $a_{m,l} \approx a(x_m, y_l)$ , on  $\mathbb{M}_{\Delta x, \Delta y} \times \mathbb{T}_{\Delta t}$ . We remark that this procedure to construct the values  $f_{m,l}^n$  is motivated by our purpose to establish convergence rates in the reconstruction process.

Next, we discuss results of numerical experiments obtained with Algorithm 2, since it shows overall better performance as discussed at the end of this section, where we compare both algorithms.

In the following test, we consider a problem with two discontinuities along  $x = 0$  and  $2x = y$ , the latter is a discontinuity that is not aligned with the RW grid. We set  $p_1(x, y) = 0.25 + 5 \cdot 10^{-3}|x| + 2 \cdot 10^{-2}\Theta(2x - y)$ ,  $p_2(x, y) = 0.2 + 2 \cdot 10^{-5}xy$ ,  $p_3(x, y) = 0.25 - 2 \cdot 10^{-2}\sin(y/5) - 5 \cdot 10^{-2}\Theta(x)$  and  $D = 1$  that corresponds to

$$\begin{aligned} \mu_1(x, y) &= 10^{-3}(5|x| + 20\Theta(2x - y) + 50\Theta(x) + 20\sin(y/5)) \\ \mu_2(x, y) &= 0.3 - (5 \cdot 10^{-3}|x| + 2 \cdot 10^{-2}\Theta(2x - y) - 5 \cdot 10^{-2}\Theta(x) \\ &\quad - 2 \cdot 10^{-2}\sin(y/5)) - 2 \cdot 10^{-5}xy \\ a(x, y) &= 0.5 + 5 \cdot 10^{-3}|x| + 2 \cdot 10^{-2}\Theta(2x - y) - 5 \cdot 10^{-2}\Theta(x) \\ &\quad - 2 \cdot 10^{-2}\sin(y/5) \end{aligned} \quad (21)$$

for Eq. (18), that is solved with a grid of  $800 \times 800 \times 400$  points for the domain  $(-40, 40)$  and time  $T = 20$ .

The initial PDF is  $f_0(x, y) = \frac{1}{6\pi L_0} e^{-(x^2+y^2)/(6L_0)}$ , with  $L_0 = 20$ .

The PDF is computed by solving the FP equation with the Chang-Cooper method (Annunziato & Borzi, 2013). Thereafter, a coarse grain binning of the PDF with a size of  $400 \times 400$  bins is performed, where a linear interpolation method is used.

With the aim to evaluate the convergence order of the reconstruction method, we repeat the procedure with three RW grid sizes, namely of 52, 100 and 200 points for the space domain. The time steps of discretization results from the scaling  $\Delta T = \Delta x^2/D$ . In the case of 200 points, with  $\Delta x = 0.4$ , having set

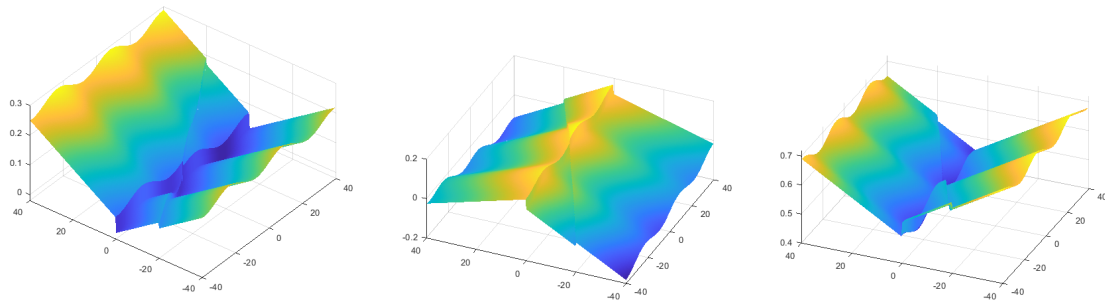


Figure 1: Shapes of the parameters of the Fokker-Planck equation. From left to right  $\mu_1(x, y)$ ,  $\mu_2(x, y)$ ,  $a(x, y)$ .

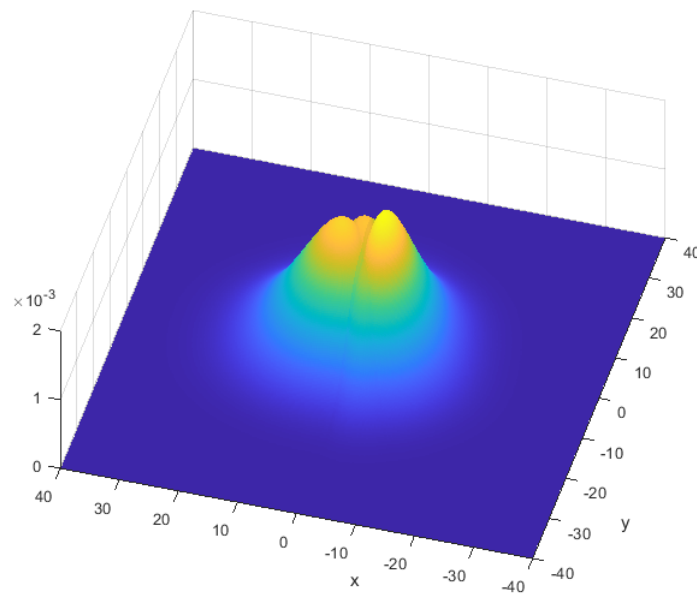


Figure 2: The binned PDF at time  $T = 20$ .

$D = 1$ , the time step of the RW results  $\Delta T = 0.16$ , that corresponds to 124 time intervals. Hence, upon the space-time grid of size  $200 \times 200 \times 124$  intervals, the PDF suitable for the input of the reconstruction algorithms is evaluated by using the linear interpolation of the binned PDF.

In Fig. 1, we show the shapes of  $\mu_1(x, y)$ ,  $\mu_2(x, y)$  and  $a(x, y)$  of Eq. (21). Notice that this functions have an entire line of discontinuity, which makes the application of other methods theoretically (at least) impossible.

In Fig. 2, we show the shape of the calculated and binned PDF at the final time  $T = 20$ . Notice the discontinuities on the surface due to the discontinuities in the drifts and diffusion coefficients.

In Fig. 3, we show the results for the reconstruction of  $\mu_1(x, 16.6581)$ . The calculated values are represented with circles, the post-processed with dots. This is valuable especially around the discontinuities of the function, where the calculation is affected from bigger numerical errors. Nevertheless, we can see the identification is in good agreement with the function used to generate the data.

In order to further illustrate the results of the reconstruction, we also plot the absolute error as the

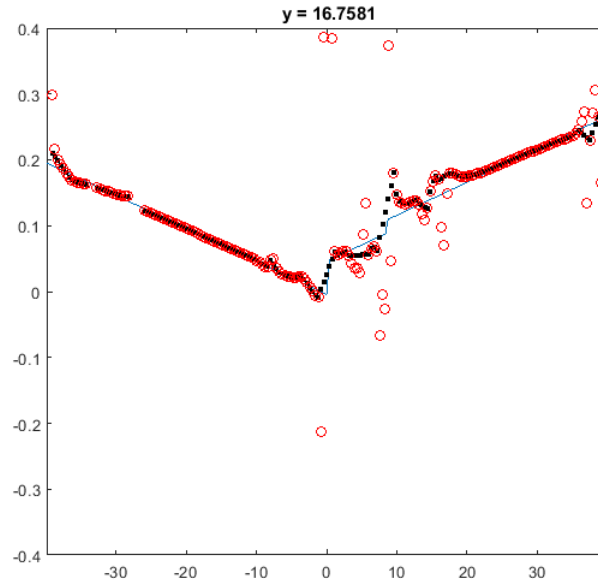


Figure 3: Plot of  $\mu_1(x, 16.7581)$  calculated with Alg. 2 for the first test. Circles are the calculated value, black dots represent the reconstruction after post-processing.

modulus of the difference between the given functions (21) and the calculated/post-processed. The same difference is evaluated with the  $L^2$ -norm for the convergence order estimation. In Fig. 4 the absolute error is depicted in colored regions of a level set representation.

Next, we perform a second test with a setting with the discontinuity along the line  $x = 0$  only. We have  $p_1(x, y) = 0.25 + 5 \cdot 10^{-3}|x|$ ,  $p_2(x, y) = 0.2 + 2 \cdot 10^{-5}xy$ ,  $p_3(x, y) = 0.25 - 2 \cdot 10^{-2} \sin(y/5) - 5 \cdot 10^{-2} \Theta(x)$  and  $D = 1$  that corresponds to

$$\begin{aligned}\mu_1(x, y) &= 10^{-3}(5|x| + 50 \Theta(x) + 20 \sin(y/5)) \\ \mu_2(x, y) &= 0.3 - (5 \cdot 10^{-3}|x| - 5 \cdot 10^{-2} \Theta(x) - 2 \cdot 10^{-2} \sin(y/5)) - 2 \cdot 10^{-5}xy \\ a(x, y) &= 0.5 + 5 \cdot 10^{-3}|x| - 5 \cdot 10^{-2} \Theta(x) - 2 \cdot 10^{-2} \sin(y/5)\end{aligned}\quad (22)$$

for Eq. (18). The other parameters of the numerical setting are the same of the former test. The resulting reconstruction is shown in Fig. 5. In Fig. 6 we see the absolute errors for the reconstructed  $\mu_1, \mu_2$  and  $a$  with the Alg. 2.

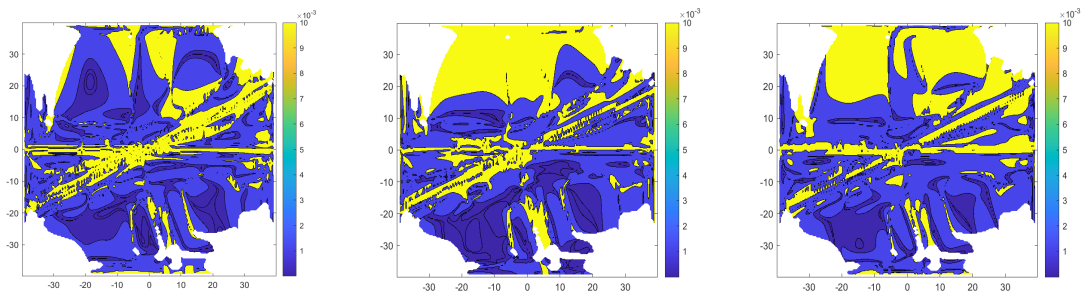


Figure 4: Plots of the absolute error between the given and reconstructed coefficients (21) as level set. From left to right those corresponding to  $\mu_1, \mu_2, a$ . The darker regions are values of errors in the interval  $(0, 10^{-4})$  and  $(10^{-4}, 10^{-3})$ , bright regions are  $(10^{-3}, 10^{-2})$  and greater than  $10^{-2}$ . In the white regions the calculation has not been completed due to rank deficiency of the matrix system.

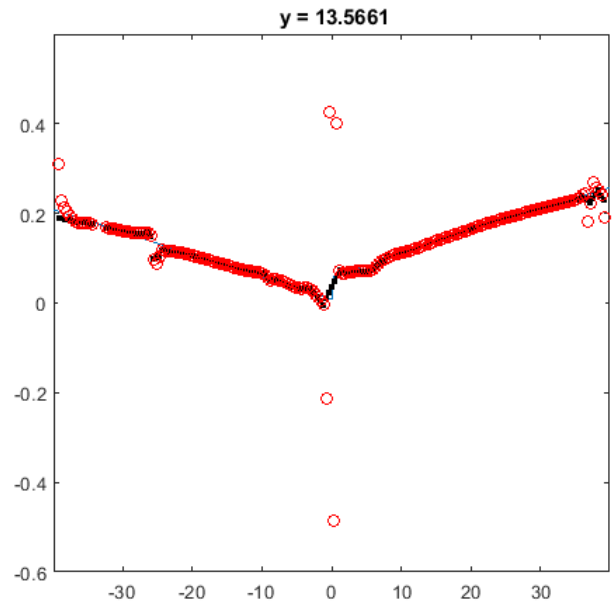


Figure 5: Plot of  $\mu_1(x, 13.5661)$  calculated with Alg. 2 for the second test. Circles are the calculated value, black dots represent the reconstruction after post-processing.

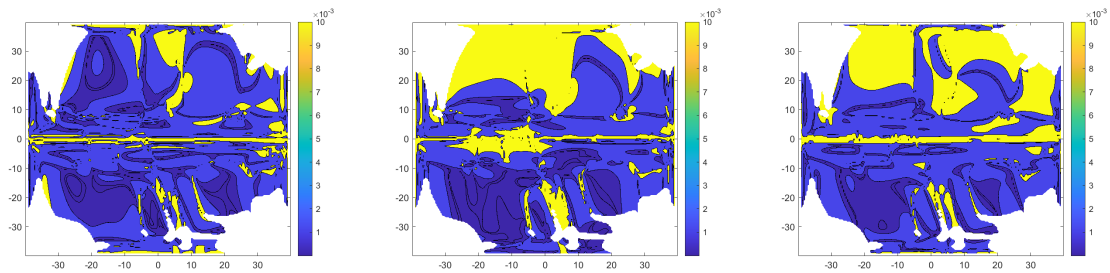


Figure 6: Plots of the absolute error between the given and reconstructed coefficients (22) as level set. From left to right those corresponding to  $\mu_1, \mu_2, a$ . The darker regions are values of errors in the interval  $(0, 10^{-4})$  and  $(10^{-4}, 10^{-3})$ , bright regions are  $(10^{-3}, 10^{-2})$  and greater than  $10^{-2}$ . In the white regions the calculation has not been completed due to rank deficiency of the matrix system.

Table 1: Convergence results for the setting (21). Left Alg. 1, right Alg. 2.

| $N_x$             | 52   | 100  | 200  | $N_x$             | 100  | 200  |
|-------------------|------|------|------|-------------------|------|------|
| $\epsilon(\mu_1)$ | 2.65 | 1.32 | 0.70 | $\epsilon(\mu_1)$ | 1.24 | 0.59 |
| $\epsilon(\mu_2)$ | 4.83 | 1.86 | 0.96 | $\epsilon(\mu_2)$ | 1.28 | 0.56 |
| $\epsilon(a)$     | 4.80 | 1.40 | 0.59 | $\epsilon(a)$     | 1.28 | 0.50 |

Table 2: Convergence results for the setting (22). Left Alg. 1, right Alg. 2.

| $N_x$             | 52   | 100  | 200  | $N_x$             | 100  | 200  |
|-------------------|------|------|------|-------------------|------|------|
| $\epsilon(\mu_1)$ | 2.09 | 0.70 | 0.33 | $\epsilon(\mu_1)$ | 0.69 | 0.31 |
| $\epsilon(\mu_2)$ | 3.06 | 1.55 | 1.02 | $\epsilon(\mu_2)$ | 1.22 | 0.92 |
| $\epsilon(a)$     | 4.16 | 1.30 | 0.49 | $\epsilon(a)$     | 1.19 | 0.43 |

Finally, we report in the tables the true  $L^2$ -errors  $\epsilon$  versus the grid size  $N_x$  of the random walk approximation. In Tab. 1 we report the errors for the setting of Eq. (21) for both the algorithms. Notice that we calculated the norms on the sub-domain  $(-20, 20) \times (-20, 20)$ , where the errors are less subject to fluctuations due to the vanishing values of the PDF on the boundary of its domain. We see that the error scales at least linearly with the mesh size, moreover the results of Alg. 2 are more accurate than the Alg. 1. Similar results are reported in Tab. 2 for the setting (22).

## 6. CONCLUSION

A novel numerical technique to reconstruct the drift and diffusion coefficients of the Fokker-Planck equation was presented. This technique was based on the Chapman-Kolmogorov equation of the inhomogeneous random walk related to the Fokker-Planck equation. Results of numerical experiments demonstrated the ability of the proposed method to solve this reconstruction problem also in the case of discontinuous coefficients. The proposed approach allows to calibrate Fokker-Planck equations with discontinuous drift and diffusion coefficients that appear in the modeling populations with sudden environmental changes, finance models with friction-like terms, and queueing theory.

## 7. REFERENCES

- Albeverio, S., Blanchard, P., Kusuoka, S., & Streit, L. (1989). An inverse problem for stochastic differential equations. *Journal of Statistical Physics*, 57(1), 347–356. doi: 10.1007/BF01023648
- Annunziato, M., & Borzi, A. (2013). A Fokker–Planck control framework for multidimensional stochastic processes. *Journal of Computational and Applied Mathematics*, 237(1), 487–507. doi: 10.1016/j.cam.2012.06.019
- Annunziato, M., & Gottschalk, H. (2018). Calibration of Lévy processes using optimal control of Kolmogorov equations with periodic boundary conditions. *Mathematical Modelling and Analysis*, 23(3), 390–413. doi: 10.3846/mma.2018.024
- Banks, H. T., Tran, H. T., & Woodward, D. E. (1993). Estimation of variable coefficients in the Fokker-Planck equations using moving node finite elements. *SIAM Journal on Numerical Analysis*, 30(6), 1574–1602.
- Breitenbach, T., Annunziato, M., & Borzi, A. (2018). On the optimal control of a random walk with jumps and barriers. *Methodology and Computing in Applied Probability*, 20(1), 435–462. doi: 10.1007/s11009-017-9565-4
- Chen, X., Yang, L., Duan, J., & Karniadakis, G. E. (2021). Solving inverse stochastic problems from discrete particle observations using the Fokker–Planck equation and physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(3), B811–B830. doi: 10.1137/20M1360153
- Comte, F. (2020). From regression function to diffusion drift estimation in nonparametric setting. *ESAIM: ProcS*, 68, 20–34.
- Comte, F., & Genon-Catalot, V. (2025). New results for drift estimation in inhomogeneous stochastic differential equations. *Journal of Multivariate Analysis*, 208, 105415. doi: 10.1016/j.jmva.2025.105415
- Cox, D., & Miller, H. (1992). *The Theory of Stochastic Processes*. London: Chapman & Hall.
- Dunker, F., & Hohage, T. (2014, Aug). On parameter identification in stochastic differential equations by penalized maximum likelihood. *Inverse Problems*, 30(9), 095001.
- Egger, H., & Engl, H. (2005, Apr). Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems*, 21(3), 1027.

Florens-Zmirou, D. (1989). Approximate discrete-time schemes for statistics of diffusion processes. *Statistics*, 20(4), 547–557.

Florens-Zmirou, D. (1993). On estimating the diffusion coefficient from discrete observations. *Journal of Applied Probability*, 30(4), 790–804.

Han, S., et al. (2025). System identification of Fokker–Planck equations via transport maps. *arXiv preprint*. Retrieved from <https://arxiv.org/abs/2507.15091> doi: 10.48550/arXiv.2507.15091

Jäger, S., & Kostina, E. (2005, Dec). Parameter estimation for forward Kolmogorov equation with application to nonlinear exchange rate dynamics. *Proceedings in Applied Mathematics and Mechanics*, 5, 745–746.

Jamba, N. T., Subray, N., & Rodrigues, H. M. (2024). Estimation for stochastic differential equation mixed models using approximation methods. *AIMS Mathematics*, 9(4), 7866–7894. doi: 10.3934/math.2024383

Jiang, G., & Knight, J. (1997). A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model. *Econometric Theory*, 13(5), 615–645.

Lamouroux, D., & Lehnertz, K. (2009). Kernel-based regression of drift and diffusion coefficients of stochastic processes. *Physics Letters A*, 373(39), 3507–3512.

Lund, S. P., Hubbard, J. B., & Halter, M. (2014, 11). Nonparametric estimates of drift and diffusion profiles via Fokker–Planck algebra. *The Journal of Physical Chemistry B*, 118(44), 12743–12749. doi: 10.1021/jp5084357

Nicolau, J. (2003). Bias reduction in nonparametric diffusion coefficient estimation. *Econometric Theory*, 19(5), 754–777.

Renò, R. (2008). Nonparametric estimation of the diffusion coefficient of stochastic volatility models. *Econometric Theory*, 24(5), 1174–1206.

The MathWorks. (2021). MATLAB (Version R2021a) [Computer software]: Natick, Massachusetts, United States: The MathWorks Inc.