Exact enumeration approach to first-passage time distribution of non-Markov random walks

Shant Baghram,¹ Farnik Nikakhtar,¹,² M. Reza Rahimi Tabar,¹,³,* S. Rahvar,¹,⁴ Ravi K. Sheth,²
Klaus Lehnertz,⁵,⁶,⁷ and Muhammad Sahimi⁸,⁹

¹Department of Physics, Sharif University of Technology, Tehran 11155-9161, Iran
²Center for Particle Cosmology, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, Pennsylvania 19104, USA
³Institute of Physics, Carl von Ossietzky University of Oldenburg, Carl von Ossietzky Straße 9-11, 26111 Oldenburg, Germany
⁴Department of Physics, College of Science, Sultan Qaboos University, P.O. Box 36, P.C. 123, Muscat, Sultanate of Oman
⁵Department of Epileptology, University of Bonn, Sigmund Freud Straße 25, 53105 Bonn, Germany
⁶Helmholtz Institute for Radiation and Nuclear Physics, University of Bonn, Nussallee 14-16, 53115 Bonn, Germany
⁷Interdisciplinary Center for Complex Systems, University of Bonn, Brühler Straße 7, 53175 Bonn, Germany
⁸Mok Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California 90089-1211, USA

(Received 13 August 2018; published 3 June 2019)

We propose an analytical approach to study non-Markov random walks by employing an exact enumeration method. Using the method, we derive an exact expansion for the first-passage time (FPT) distribution of any continuous differentiable non-Markov random walk with Gaussian or non-Gaussian multivariate distribution. As an example, we study the FPT distribution of the fractional Brownian motion with a Hurst exponent $H \in (1/2, 1)$ that describes numerous non-Markov stochastic phenomena in physics, biology, and geology and for which the limit $H = 1/2$ represents a Markov process.

DOI: 10.1103/PhysRevE.99.062101

I. INTRODUCTION

The concept of first passage refers to the crossing of a prespecified location or some sort of a threshold in a stochastic trajectory [1]. The distribution of the first-passage times (FPTs), which represents the probability of crossing the trajectory at a specific time or location [2,3] and depends on the nature of the stochastic process, plays a fundamental role in the theory of stochastic processes as well as in their applications. The FPT distribution makes it possible to investigate quantitatively the uncertainty in the properties of a stochastic system within a finite time. Two important applications are the extinction time of a disease in the models of epidemic processes [1], chemical [4], biology [5], and the analysis of 10⁶ trajectories.

Due to their very large number of applications, the FPT properties have been studied extensively and are well understood when the stochastic phenomena represent a Markov process. As a general rule, however, the dynamics of a given stochastic process in complex media is the result of its interactions with the environment around it, which may contain trapping sites, obstacles, moving parts, active pumps, etc. [28], and cannot be described as a Markov process. Indeed, although the evolution of the set of all microscopic degrees of freedom of a system is Markovian, the dynamics restricted only to the random walker is not [3,29,30]. Experimental realizations of non-Markov dynamics include diffusion of tracers in crowded narrow channels [31] and in complex fluids, such as nematics [32] and viscoelastic solutions [33,34], as well as the dark matter halo mass function [35]. Even in simple fluids, hydrodynamic memory influences various phenomena and, thus, non-Markov dynamics has been reported recently [36].

Using the inclusion-exclusion principle and an exact enumeration method, we derive in this paper the FPT distribution of a non-Markov random walk by assuming that the trajectory of the walk is differentiable at every point. As an example, we derive the FPT distributions of fractional Brownian motion (FBM) with a given Hurst exponent $H \in (0.5, 1)$. The analytical results are confirmed by extensive numerical simulation and the analysis of 10⁶ trajectories.

The rest of this paper is organized as follows. In the next section, we describe the exact enumeration approach to derive the FPT probability density of a non-Markovian random walk. We then drive, in Sec. III, an analytical expression for the FPT distribution of FBM. The results of numerical simulations

¹tabar@uni-oldenburg.de
¹moe@usc.edu
are presented in Sec. IV, whereas the paper is summarized in Sec. V. In the Appendix, we provide the details of the derivation of our results.

II. EXACT ENUMERATION METHOD FOR THE FPT PROBABILITY DENSITY

We define a general dynamical equation for a random walk $x(t)$, driven by a correlated nonstationary noise (velocity) $v(t)$, 

$$\frac{dx(t)}{dt} = v(t), \quad C(t, t') = \langle v(t)v(t') \rangle.$$  

(1)

$x(t)$ is assumed to be continuous and its derivative (velocity) $v(t)$ to be well defined at any time. The noise $v(t)$ has a zero mean and an arbitrary $n$-point joint distribution $p(v_n, t_n; \ldots; v_1, t_1; v_0, t_0)$. The correlation function $C(t, t')$ depends on both $t$ and $t'$. Because $x(t)$ is a stochastic process, each of its realizations reaches a given barrier $x = x_c$ for the first time at a different time $t$, giving rise to a FPT probability density $f(t)$. Consider the trajectories with the initial conditions $x(0) = x_0$ and $x(t_0) = v(t_0) = v_0$; crossing the barrier $x_c$ in the time interval $t$ and $t + dt$ with $v(t) > 0$. The crossing is equivalent to the conditions that $x(t) < x_c$ and $x_c < x(t + dt)$ [16,38]. If $x_c$ is constant, $x(t)$ will lie in the interval $x_c - v dt < x(t) < x_c$. Then, the probability that $x(t)$ satisfies the passage condition $x_c - v dt < x(t) < x_c$ is

$$P(x, v, t | x_0, v_0, t_0)dx = vP(x, v, t | x_0, v_0, t_0)dv,$$

where we kept the terms up to the order of $dt$. Since $v(t) > 0$ at $x_c$, we should integrate over all positive velocities. Therefore, the probability of crossing the barrier $x_c$ per unit time is given by [16]

$$n_1(x_c, t | x_0, v_0, t_0) = \int_0^\infty \int_0^\infty vP(x_c, v, t | x_0, v_0, t_0)dv.$$  

(2)

Equation (2) represents the rate of up-crossing rather than a density function and, thus, it is not normalized. We generalize Eq. (2) to the joint probability of multiple up-crossings, i.e., $x(t)$ crossing the barrier in each of the intervals $(t_1, t_1 + dt), \ldots, (t_p, t_p + dt)$ by integrating over all the crossing points $t_1, t_2, \ldots, t_p$,

$$n_p(x_c, t_p; \ldots; x_c, t_1 | x_0, v_0, t_0) = \int_0^\infty \int_0^\infty \int_0^\infty \ldots \int_0^\infty v_1v_p \ldots v_1$$

$$\times P(x_c, v_p, t_p; \ldots; x_c, v_1, t_1 | x_0, v_0, t_0).$$

(3)

Using Bayes’ theorem, one may substitute the conditional probability density in Eq. (3) with the joint probability density.

In Fig. 1, typical trajectories as well as the FPT distribution of the FBM for $x_c = 1$ with $x_0 = 0$ are presented. The trajectories are constructed using the Cholesky decomposition (see below).

A trajectory can cross $x_c$ several times (see the lower panel of Fig. 1). We relate the FPT distribution to the statistical properties of the up-crossings, which are considered as point processes with rates $n_p$, where $p$ refers to the number of up-crossing. To this end, we look for the fraction of all the trajectories that up-cross $x_c$ for the first time at time $t$ with the initial conditions $(x_0, v_0)$ at time $t_0$ and enumerate them in terms of $n_p$. To simplify the notation, we drop $x_c$ and the initial conditions.

The rate $n_1(t)$ is overcounted through the trajectories that had an up-crossing at shorter times $t_1 < t$. Therefore, we subtract their fraction from the first term. This stems from the fact that $n_1(t)$ is a local function in $t$, but there is no guarantee that a trajectory has not up-crossed before $t$. The overcounting implies that the main problem is a combinatorial counting. Thus, as an enumeration technique we use the inclusion-exclusion principle, one of the most useful principles of counting in combinatorics and probability. According to De Morgan’s laws, in the general and complementary form, the principle of inclusion-exclusion for finite sets $A_1, A_2, \ldots, A_n$ is expressed by

$$\bigcup_{i=1}^n A_i = U - \bigcup_{i=1}^n A_i \quad \sum_{i=1}^n n_i$$

$$= |U| - \sum_{i=1}^n |A_i| + \sum_{1 \leq i < j \leq n} |A_i \cap A_j| - \cdots$$

$$+ \sum_{1 \leq i < j < \cdots < n} (-1)^{n-i}[A_i \cap A_j \cap \cdots \cap A_n],$$

(4)

where $U$ is a finite universal set containing all the $A_i$ and $A_i$’s are the complement of $A_i$ in $U$. That the trajectories cross $x_c$ for the first time at time $t$ implies that they should not have been crossed at $x_c$ at shorter times. We consider $n_1(t)$ as the universal set and define the next subset by $A_1 = n_2(t, t_1)$, denoting the fraction of trajectories for which the up-crossing at time $t$ is not for the first time, and that they had a previous up-crossing at a shorter time $t_1 < t$. Then, the FPT distribution is given by $|\bigcup_{i=1}^n A_i|$ because only the trajectories that have a first up-crossing at time $t$ and do not belong to the
Table I. The exact enumeration method. The nth column corresponds to the nth term of the sum in Eq. (5).

| $f(t)$ | $= |U|$ | $- \sum_{i=1}^{n} |A_i|$ | $+ \sum_{1 \leq i \leq j \leq n} |A_i \cap A_j|$ | $- \sum_{1 \leq i \leq j \leq k \leq n} |A_i \cap A_j \cap A_k|$ |
|--------|--------|-----------------|------------------|------------------|
| $n_1(t)$ | $= n_1(t)$ | $- \int_{0}^{t} n_2(t_i, t) dt_i$ | $+ \frac{1}{2!} \int_{0}^{t} \int_{0}^{t} n_3(t_i, t_j, t) dt_i dt_j$ | $- \frac{1}{3!} \int_{0}^{t} \int_{0}^{t} \int_{0}^{t} n_4(t_i, t_j, t_k, t) dt_i dt_j dt_k$ |

where $n_{p+1}(t_1, t_2, \ldots, t_l)$ are given by the conditional probabilities (3). The factor $1/p!$ accounts for the number of permutations of the variables $t_1, \ldots, t_l$ with the signs explained in Table I. To calculate $n_p(t_1, \ldots, t_l)$, we consider the trajectories in the absence of $x_c$ and let them return after an up-crossing and, then, up-cross the barrier $p$ times. The correct counting of such multiple crossings yields the distribution $f(t)$ of the FPT. Equation (5) provides us with the FPT distribution [16].

$$f(t) = \frac{1}{p!} \int_{0}^{t} \cdots \int_{0}^{t} n_{p+1}(t_1, t_2, \ldots, t_l) dt_1 \cdots dt_l,$$

(5)

where $n_{p+1}(t_1, t_2, \ldots, t_l)$ are given by the conditional probabilities (3). The factor $1/p!$ accounts for the number of permutations of the variables $t_1, \ldots, t_l$ with the signs explained in Table I. To calculate $n_p(t_1, \ldots, t_l)$, we consider the trajectories in the absence of $x_c$ and let them return after an up-crossing and, then, up-cross the barrier $p$ times. The correct counting of such multiple crossings yields the distribution $f(t)$ of the FPT. Equation (5) provides us with the FPT distribution [16].

$$f(t) = n_1(t) - \int_{0}^{t} n_2(t, t_1) dt_1 + \frac{1}{2!} \int_{0}^{t} \int_{0}^{t} n_3(t, t_2, t_1) dt_1 dt_2 - \cdots$$

$$= \sum_{p=0}^{\infty} \frac{1}{p!} \int_{0}^{t} \cdots \int_{0}^{t} n_{p+1}(t_1, t_2, \ldots, t_l) dt_1 \cdots dt_l,$$

(5)

where $n_{p+1}(t_1, t_2, \ldots, t_l)$ are given by the conditional probabilities (3). The factor $1/p!$ accounts for the number of permutations of the variables $t_1, \ldots, t_l$ with the signs explained in Table I. To calculate $n_p(t_1, \ldots, t_l)$, we consider the trajectories in the absence of $x_c$ and let them return after an up-crossing and, then, up-cross the barrier $p$ times. The correct counting of such multiple crossings yields the distribution $f(t)$ of the FPT. Equation (5) provides us with the FPT distribution [16].

$$f(t) \approx n_1(t) \exp \left[ - \int_{0}^{t} n_1(t') dt' \right].$$

(7)

In the Hertz approximation [24,39,41], $f(t)$ factorizes to $n_1(t_1, \ldots, t_l)$.

The Hertz approximation is based on assuming that all the up-crossings are independent of each other and that the correlations between them are negligible. This leads to the following

$$f(t) = \psi(t) e^{-\psi(t)}.$$  

(6)

where $n_{p+1}(t_1, t_2, \ldots, t_l)$ are given by the conditional probabilities (3). The factor $1/p!$ accounts for the number of permutations of the variables $t_1, \ldots, t_l$ with the signs explained in Table I. To calculate $n_p(t_1, \ldots, t_l)$, we consider the trajectories in the absence of $x_c$ and let them return after an up-crossing and, then, up-cross the barrier $p$ times. The correct counting of such multiple crossings yields the distribution $f(t)$ of the FPT. Equation (5) provides us with the FPT distribution [16].

$$f(t) \approx n_1(t) \exp \left[ - \int_{0}^{t} n_1(t') dt' \right].$$

(7)

In the Hertz approximation [24,39,41], $f(t)$ factorizes to $n_1(t_1, \ldots, t_l)$.

In the Hertz approximation, we calculate exactly the first and the second terms of the expansion and approximate all the higher-order terms by the first two [35] with the corresponding FPT distribution being in the form of Eq. (6) with [39]

$$\psi_{St}(t) = - \int_{0}^{t} n_1(t') \ln \left[ 1 - \int_{0}^{t} R(t, t') n_1(t') dt' \right] dt,$$

(8)

where $R(t_1, t_2) = 1 - n_2(t_1, t_2)/[n_1(t_1)n_1(t_2)]$. For simplicity and in order to derive an expression for $f(t)$, we assume in the following that the velocity distribution is Gaussian.

III. ANALYTICAL DERIVATION OF THE FPT DISTRIBUTION OF THE FBM

We now derive the FPT distribution of the FBM with a Hurst exponent $H \in (0.5, 1)$, which is defined in terms of its nonstationary correlation function [42],

$$\langle x_H(t_1)x_H(t_2) \rangle = \frac{1}{2} \left[ |t_1|^{2H} + |t_2|^{2H} - |t_2 - t_1|^{2H} \right],$$

(9)

which is positive semidefinite (see the Appendix) with its first derivative (velocity) being the fractional Gaussian noise (FGN) $v_H(t)$ so that $\dot{x_H}(t) = v_H(t)$. Using physical arguments [43,44] as well as rigorous analysis [45], it was shown that
the scaling behavior of the FPT distribution of a FBM has the following long-time behavior:

\[ f(t) \sim t^{-H/2}. \] (10)

Given that the FBM and FGN have Gaussian distributions for \( x \) and \( v \), respectively, we determine \( n_1(t) \) and \( n_2(t_1, t_2) \) and, therefore, \( R(t_1, t_2) \) and the FPT distribution in the Hertz and Stratonovich approximations. It is straightforward to show that \( n_1(t) \) is given by the following expression:

\[
n_1(t) = p(x_c) \int_0^\infty dv \langle v|x_c \rangle dv, \tag{11}
\]

where \( p(v|x_c) \) is a Gaussian distribution with mean \( \langle v|x_c \rangle = x_c (v/\sqrt{\pi}) \) and variance \( \sigma^2(v|x_c) = H/\Gamma \) and \( \sigma^2 = H/\Gamma^2 \) in terms of the Hurst exponent \( H \). We find that the explicit expression for \( n_1(t) \) is given by

\[
n_1(t) = \frac{\Gamma^2}{2\pi H t^{2H/3}} \exp\left( -\frac{\gamma^2}{2} \right) \left[ \frac{H^2 t^{2H/3}}{\Gamma^2} \exp\left( -\frac{-\gamma^2}{2} \right) + \frac{H^2 x_c}{2\Gamma} t^{-H/2} \right], \tag{12}
\]

where \( \gamma = x_c / \Gamma t \). Using Eq. (7), we obtain the FPT distribution in the Hertz approximation, which, in general, is accurate for estimating the first peak of the FPT distribution, but it overestimates its tail. Similarly, we find that

\[
n_2(t_1, t_2) = \int_0^\infty dv \int_0^\infty dv' \langle v|x_c \rangle \langle v'|x_c \rangle dv \langle v|x_c \rangle dv' p(v|x_c, v, v') \]

\[ = p(x_c) \int_0^\infty dv \langle v|x_c \rangle p(x'|x_c, v) \]

\[ \times \int_0^\infty dv' \langle v'|x_c \rangle, \tag{12}
\]

where all the distributions in Eq. (12) are Gaussian. For example, \( p(x'|x_c, v) \) has the mean (see the Appendix for the variance),

\[ \langle x'|x_c, v \rangle = x_c \langle x' \rangle / \Gamma t^H + \left( v - \langle v|x_c \rangle \right) \left( \langle x' \rangle - \langle x' \rangle / \sigma^2 \right) s_{v|x_c}. \]

The correlation functions \( \langle x'x \rangle \) and \( \langle x'v \rangle \) are given by Eq. (A9) in the Appendix and \( s_{v|x_c} = H^2 t^{2H/3} / \Gamma^2 \). Having \( n_1(t) \) and \( n_2(t_1, t_2) \) enables one to determine \( R(t_1, t_2) \) and \( f(t) \) in the Hertz and Stratonovich approximations.

### IV. NUMERICAL RESULTS

Figure 1 presents the trajectories of a FBM process using the Cholesky decomposition [46,47] (see the Appendix) and their FPT distribution for \( H = 0.8, x_c = 1, \) and \( x_0 = 0 \). In Fig. 2, the FPT distributions of the FBM trajectories are plotted. The FPT is obtained from the Cholesky method. We also show in these plots the FPT distributions in the Hertz approximation, which deviate from the FPT directly computed using trajectories. For comparison, the theoretically predicted tails of the distributions, i.e., \( f(t) \sim t^{-H/2} \), are also plotted. Figure 2 indicates that the theoretical tails of the FPT in the long-time limit coincide with the FPT distributions computed using the trajectories. As already mentioned above, the Hertz approximation predicts correctly the location of the peak of the FPT distribution but underestimates the tails.

To derive the FTP distribution in the Stratonovich approximation with \( H = 0.8 \), one must calculate \( \psi_{\text{Str}}(t) \) via Eq. (9) and then use Eq. (6). To avoid any error from the numerical differentiation of \( \psi_{\text{Str}}(t) \), we determine the integrated FTP distribution via the term \( \exp[-\psi_{\text{Str}}(t)] \). In Fig. 3, the cumulative FTP distribution is presented for \( H = 0.6 \) and \( H = 0.8 \), indicating that the Hertz approximation deviates clearly from the results computed via the Cholesky decomposition. As shown in Fig. 2, the tail of \( f(t) \) in the Hertz approximation does not coincide completely with those obtained by the Cholesky decomposition. Higher-order approximations, such as the Stratonovich approximation are, therefore, needed, implying that \( n_2(t_1, t_2) \) should not be factorized as \( n_1(t_1)n_2(t_2) \). As shown in Fig. 3, the Stratonovich approximation provides better estimates for the FPT distributions.

One may define various measures to study the interdependence of the up-crossing events. The simplest measure is the Fano factor. Consider a time window \( T \) and count the mean number (and its variance) of up-crossing events for trajectories in the window. The Fano factor \( F(T) \) is defined as the variance of the number of up-crossing events in \( T \), divided by its mean number, and is written in terms of \( n_2(t_1, t_2) \) and \( n_1(t) \) [48]. More specifically, the Fano factor is given by

\[ F(T) = \frac{\text{Var}[N(T)]}{\text{E}[N(T)]}, \]

where \( N(T) \) is the number of up-crossing events in the time window \( T \).
because the integrals were computed by a Monte Carlo method. The Kolmogorov-Smirnov statistics for FPT distributions, derived from the Cholesky method, when compared with the Hertz approximation and the Stratonovich approximation, yields the values of 0.433 (p value = \(1.03 \times 10^{-6}\)), 0.221 (p value = 0.655), 0.294 (p value = \(1.24 \times 10^{-6}\)), and 0.256 (p value = 0.447) for \(H = 0.6\) and \(H = 0.8\), respectively.

\[
\mathcal{F} = \langle \Delta N^2 \rangle / \langle N \rangle \quad \text{with} \quad \langle \Delta N^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2, \quad \text{where} \quad \langle N \rangle = \int_0^T n_1(t) \, dt \quad \text{and} \quad \langle N^2 \rangle = \langle N \rangle + \int_0^T \int_0^T n_2(t_2, t_1) \, dt_2 \, dt_1 \quad [48].
\]

For independent point processes, i.e., \(n_2(t_2, t_1) = n_1(t_2) n_1(t_1)\), one has \(\mathcal{F} = 1\). Therefore, for a Poisson process \(\mathcal{F}(T) = 1\). By definition, \(\mathcal{F}(T) > 1\) and \(\mathcal{F}(T) < 1\) refer, respectively, to over- and underdispersion [49]. We plot in Fig. 4 the Fano factor versus the size of the time window \(T\), which indicates that, in the long-time limit, the up-crossing point processes are strongly overdispersed. This means that over such timescales \(n_2(t_i, t_j)\) should not be factorized as \(n_1(t_i)n_1(t_2)\) and that the Hertz approximation is not appropriate for estimating the tails of the FPT distribution. A very crucial point to point out is that, in the time span in which \(\mathcal{F} \sim 1\), the Hertz approximation is accurate and that it is very close to the Cholesky-derived FPT distribution. On the other, if the Fano factor deviates from unity, it is certain that the Hertz approximation is not suitable for describing the FPT, although \(\mathcal{F}\) cannot by itself quantify the accuracy of the Stratonovich approximation.

V. SUMMARY

The FPT distribution of the Markov random walks has been derived in the past. Here, we present an exact analytical expression for the FPT distribution of general non-Markov random walks. In principle, the FPT distribution of non-Markov processes may be obtained from the solution of the associated Fokker-Planck equation with absorbing boundaries in higher dimensions, resulting from the Markovian embedding of a non-Markov process [50]. Even the calculation of the mean FPT for a non-Markov process is, however, a rather difficult task since the corresponding boundary problem cannot be treated in a straightforward manner [51-55]. We presented a general method for deriving such analytical expressions. We then employed it to derive one (as a series) for the FPT distribution. This was performed by using an exact enumeration method based on combinatorics and the inclusion-exclusion principle, which can be generalized to include the FTP distribution of non-Markov random walks in higher dimensions. As an example, analytical results were presented for the FBM with the Hurst exponent \(H \in (0.5, 1)\), which is a non-Markov process with infinite-range memory and has wide applications in many disciplines [28]. The numerical results were also compared with two well-known approximations, namely, the Hertz and Stratonovich approximations, which revealed their shortcomings.

ACKNOWLEDGMENTS

S.B. acknowledges the partial support of Sharif University of Technology, Grant No. G960202 for this paper. F.N. acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1845298.

FIG. 4. The Fano factor for up-crossings of the trajectories of the FBM with a Hurst exponent \(H\) as a function of the window size \(T\). In the long-time limit, the up-crossing point processes are slightly overdispersed, whereas over short timescales, the Fano factor is equal to unity (red solid line), a hallmark of the Poisson process, and the variance of the up-crossing over such short timescales is equal to the mean.

FIG. 3. The cumulative FPT distributions in the Hertz (red dot-dashed lines) and Stratonovich (red dashed line with error regions) approximations for the FBM with \(H = 0.6\) and \(H = 0.8\). The black curve was computed by the Cholesky method. The errors are shown respectively.
APPENDIX

We provide the details of the main results presented in the main text of the paper.

1. Variance of the velocity of fractional Brownian motion

The time derivative (increments) of the FBM is the FGN and has the following correlation function:

\[
C_H(\tau, \delta) = \frac{\sigma^2 \delta^{2H-2}}{2} \left[ \left| \frac{\tau}{\delta} \right| + 1 \right]^{2H} + \left| \frac{\tau}{\delta} - 1 \right|^{2H}
\]

\[
- 2 \left[ \frac{\tau}{\delta} \right]^{2H},
\]

(A1)

where \(0 < H < 1\) and \(\tau = t_2 - t_1\). Here, \(\delta > 0\) is used for smoothing the FBM to make it numerically differentiable [42]. We note that, in the limit \(\tau \to 0\), the \(\delta\) dependence of \(\gamma^2 = (x/v)^2/(\langle x^2 \rangle)\) drops out. In the literature [42], there is no unique expression for \(\langle x^2 \rangle\). Here, by generating the FBM trajectories and numerically differentiating them for \(H \in (0.5, 1)\), the best fit is found to be \(\langle x^2 \rangle = c_0 + c_1 H^m\), where \(c_0 = -2.47 \pm 0.01, c_1 = 2.88 \pm 0.05,\) and \(m = -4.72 \pm 0.02\).

2. Fractional Gaussian noise

The stochastic representation of the FBM is given by

\[
B_H(t) = B_H(0) + \frac{1}{\Gamma(H + 1/2)} \int_{-\infty}^{t} \left[ (t - s)^{H-1/2} - (-s)^{H-1/2} \right] dW(s)
\]

\[
+ \int_{0}^{t} (t - s)^{H-3/2} dW(s) \right],
\]

(A2)

where \(dW(s)\) is a Wiener process that is written in terms of the Gaussian white noise \(\xi(s)\) as \(dW(s) = \xi(s) ds\). The FGN is then defined by \(G_H(t) = dB_H(t)/dt\). Taking the time derivative of Eq. (A2) yields

\[
G_H(t) = \frac{1}{\Gamma(H + 1/2)} \left[ \int_{-\infty}^{0} \left( H - 1/2 \right) (t - s)^{H-3/2} dW(s)
\right]

\[
+ \left[ (t - s)^{H-1/2} \xi(s) \frac{d}{dt} \right]_{s=0} \right] \frac{1}{\Gamma(H + 1/2)}
\]

\[
× \left[ \int_{0}^{t} \left( H - 1/2 \right) (t - s)^{H-3/2} dW(s) \right].
\]

(A3)

The second term on the right side of Eq. (A3) is not finite for \(H \in (0, 0.5)\). Therefore, the FBM has no well-defined velocity for the Hurst exponent in the range \((0, 0.5)\).

3. Proof for the variance of the FBM being positive semidefinite

A symmetric \(n \times n\) real matrix \(C\) is the covariance of some random (Gaussian) vector if and only if it is positive semidefinite, which means that

\[
z^T C z = \sum_{i=1}^{n} \sum_{j=1}^{n} z_i z_j C_{i,j} \geq 0 \quad \forall z_1, \ldots, z_n \in \mathbb{R},
\]

(A4)

where \(z\) is the aforementioned random vector. The FBM has a vanishing mean \([x(0) = 0]\), whereas its covariance is given by Eq. (10) of the main text for \((t_1, t_2) \geq 0\) and \(H \in (0, 1)\). We show that

\[
C(t_1, t_2) = \frac{1}{2} \left[ (|t_1|^{2H} + |t_2|^{2H} - |t_2|^{2H} - |x|^{2H}) \right]
\]

(A5)

is a covariance function. Consider the function,

\[
\Phi(t_2, r) = \left( t_2 - r \right)^{a_\alpha-1/2} - \left( -r \right)^{a_\alpha-1/2}
\]

(A6)

defined for all \(t_2 \geq 0\) and \(r \in \mathbb{R}\), where \(a_\alpha = \max(0, H)\) for all \(H \in \mathbb{R}\). Since \(H < 1\), we determine \(\int_{-\infty}^{\infty} \Phi(t_2, r)^2 dr < \infty\) and

\[
\int_{-\infty}^{\infty} \Phi(t_2, r) \Phi(t_1, r) dr = \kappa C(t_1, t_2) \quad \forall (t_1, t_2) \geq 0,
\]

(A7)

where \(\kappa\) is a positive and finite constant that depends only on \(H\). Therefore, we find

\[
\sum_{i=1}^{n} \sum_{j=1}^{n} z_i z_j C_{i,j} = \frac{1}{\kappa} \int_{-\infty}^{\infty} \left[ \sum_{i=1}^{n} z_i \Phi(t_1, r) \right]^2 dr \geq 0.
\]

4. Analytical expressions for \(n_1(t)\) and \(n_2(t, t')\)

with a Gaussian velocity

Due to the linearity of the system, all the joint probability densities are Gaussian and have the form

\[
P_{\alpha}(Q) = \frac{1}{(2\pi)^n/2 \det C_n} \exp \left( -\frac{Q C_n^{-1} Q}{2} \right).
\]

(A8)

Here, \(Q = [q_1(t_1), \ldots, q_n(t_n)]\) is a \(n\)-dimensional vector whose \(i\)th component is the coordinate \(x(t_i)\) or the velocity \(v(t_i)\) at time \(t_i\), and \(C_n\) is the symmetric \(n \times n\) correlation matrix whose entries are the correlation functions between the corresponding components of the vector \(Q; C_{ij} = C_{ji} = \langle q_i(t_1) q_j(t_1) \rangle\). Then, \(n_1(t)\) is obtained in closed analytical form

\[
n_1(t) = \frac{\Gamma^2}{2n H^2 H-1/2} \exp \left( -\frac{y^2}{2} \right) \left\{ \frac{H^2}{\Gamma^2} \exp \left( -\frac{y^2}{2} \gamma^2 \right) \right\}.
\]

(A9)

where \(y = x_c/t^H\) and \(\Gamma^2 = y^2/(1 - y^2)\). For the joint densities of multiple up-crossings \(n_{p,p_1, \ldots, p_{t_1}}\) no closed expression can be obtained. We evaluate the integral over \(v_1\) in Eq. (3) analytically and then perform numerical integration of the resulting expression over \(v_2, \ldots, v_p\) to determine \(n_{p,p_1, \ldots, p_{t_1}}\). The integrals over time in the expressions for \(f(t)\) are also evaluated numerically. For \(n_2(t, t')\), we compute the mean and variance of the conditional distributions,

\[
n_2(t, t') = \int_{0}^{\infty} v \, dv \, \int_{0}^{\infty} d v' \, p(x_c, x_c', v, v')
\]

\[
= p(x_c) \int_{0}^{\infty} d v \, v p(v | x_c) p(x'_c | x_c, v)
\]

\[
\times \int_{0}^{\infty} d v' \, v' p(v' | x'_c, x_c, v).
\]

(A10)
Assuming that $t' > t$, the correlations are given by
\[
\langle x'(t) \rangle = \frac{1}{2} \left[ t' + (t' - t) \right]^{2H},
\]
(A11)
\[
\langle x'(t) \rangle = H(t' - t)^{2H - 1},
\]
(A12)
\[
\langle v'(t) \rangle = H(t' - t)^{2H - 1} + H(t' - t)^{2H - 2},
\]
(A13)
\[
\langle v'(t) \rangle = H(2H - 1)(t' - t)^{2H - 2},
\]
(A14)
where, for example, $\langle x'(t) \rangle = \langle x(t')x(t) \rangle$. We also know that
\[
\sigma_x^2 = \langle x^2(t) \rangle = t^{2H}, \quad \langle xv(t) \rangle = H^{2H - 1}.
\]
(15)
Note that all the conditional distributions are Gaussians and, therefore, they are specified by their mean and variance. For example, for $p(x'|x, v)$, we have
\[
H_{x'|x, v} = \langle x'|x, v \rangle = \langle x'|x \rangle + \frac{\langle (x' - \langle x' \rangle)(v - \langle v \rangle) \rangle}{\sigma_{v|x}}(v - \langle v \rangle)
\]
\[
= t^{2H} + t^{2H} - (t' - t)^{2H} + \frac{\Gamma^2}{2t^{2H}} H^{2H - 1} \left[ H(t' - t)^{2H - 1} - \frac{H}{2t^H} + t^{2H} - (t' - t)^{2H} \right] \left( v - \frac{H}{t} x_r \right).
\]
(A16)
For $p(v'|x', x, v)$, we should calculate the mean and variance of $p(v'|x, v)$, which are given by
\[
\langle v'(t) \rangle = H(t' - t)^{2H - 1} + H(t' - t)^{2H - 2} - \frac{H}{t} \left[ H^{2H - 1} - H(t' - t)^{2H - 1} \right] \left( v - \frac{H}{t} x_r \right),
\]
(A18)
\[
\sigma_{v|x}^2 = \frac{\sigma_v^2}{\sigma_x^2} \left[ 1 - \frac{\langle xv' \rangle^2}{\sigma_x^2 \sigma_v^2} \right] - \frac{1}{\sigma_{v|x}} \left( \langle vv' \rangle - \frac{\langle xv' \rangle \langle xv \rangle}{\sigma_x^2} \right)^2 \]
\[
= \frac{\sigma_v^2}{\sigma_x^2} \left[ 1 - \frac{\langle xv' \rangle^2}{\sigma_x^2 \sigma_v^2} \right] - \frac{1}{\sigma_{v|x}} \left( \langle vv' \rangle - \frac{\langle xv' \rangle \langle xv \rangle}{\sigma_x^2} \right)^2 \]
(A19)
Now, for $p(v'|x', x, v)$, we obtain
\[
\mu_{v'|x', x, v} = \langle v'|x', x, v \rangle
\]
\[
= \langle v|x, v \rangle + \frac{1}{\sigma_{v|x}^2} \left[ \langle v'x \rangle - \frac{1}{1 - y^2} \left( \frac{\langle v'x \rangle \langle xx' \rangle}{\sigma_x^2} + \frac{\langle xx' \rangle \langle xv' \rangle}{\sigma_x^2} \right) - \frac{\langle v'x \rangle \langle xx' \rangle}{\sigma_x^2} \sigma_v^2 \right]
\]
\times (x_r - \langle x' | x, v \rangle),
\]
(A20)
\[
\sigma_{v|x,v}^2 = \frac{1}{\sigma_{v|x,v}^2} \left[ \langle v'x \rangle - \frac{1}{1 - y^2} \left( \frac{\langle v'x \rangle \langle xx' \rangle}{\sigma_x^2} + \frac{\langle xx' \rangle \langle xv' \rangle}{\sigma_x^2} \right) - \frac{\langle v'x \rangle \langle xx' \rangle}{\sigma_x^2} \sigma_v^2 \right] \sigma_{v|x,v}^2
\]
(A21)

5. The Cholesky decomposition

To compute the non-Markovian first up-crossing distribution for the FBM, we must generate trajectories with the correct ensemble properties. Here, we describe an algorithm to generate such trajectories. Equation (9) of the main text defined $C_{ij} \equiv C(t_i, t_j) = \langle x(t_i)x(t_j) \rangle$, the correlation between the $x(t)$ between times $t_i$ and $t_j$. The matrix $C$ is real, symmetric, and positive definite and, therefore, it has a unique decomposition $C = LL^T$ in which $L$ is a lower triangular matrix, which is known as Cholesky’s decomposition. We use $L$ to generate the ensemble of the trajectories as follows.

Consider, first, a vector $\xi$, which is Gaussian white noise with zero mean and unit variance (i.e., $\langle \xi_n \xi_m \rangle = \delta_{mn}$). If we generate the desired trajectories as
\[
x(t_i) = x_i = \sum_j L_{ij} \xi_j,
\]
(A22)
then $x_i$ will have the correlations of a random walk given by
\[
\langle x_i x_j \rangle = \sum_{m,n} L_{mn} L_{jm} \langle \xi_n \xi_m \rangle = LL^T = C.
\]
(A23)
Since $L$ is triangular, only a sum over $j \leq i$ is needed in the matrix calculations and, thus, the method is fast. Next, we provide a proof of the Cholesky decomposition and present it in terms of the correlation matrix. In a more general context, there is a sufficient condition for a square matrix to have a LU decomposition $C = LU$, where $L$ and $U$ are, respectively, the lower and upper triangular matrices of $C$. If all the $n$'s leading principal minors of the $n \times n$ matrix $C$ are nonsingular, then $C$ has an LU decomposition. Let us recall that the $k$th leading principal minor of $C$ is given by

$$C_k = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{12} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1k} & c_{2k} & \cdots & c_{kk} \end{pmatrix},$$

(A24)

where we have assumed that $C_1, C_2, \ldots, C_n$ are nonsingular. Using induction, it is not difficult to show that there is a LU decomposition for the correlation matrix. Using the symmetry of $C$, we write

$$LU = C = C^T = U^TL^T,$$

(A25)

which implies that

$$U(L^T)^{-1} = L^{-1}U^T.$$  

(A26)

The left side of the equation is upper triangular, whereas the right side is a lower triangular matrix. Consequently, there is a diagonal matrix $D$ such that $D = U(L^T)^{-1}$. Then, $U = DL^T$, which for the correlation matrix implies that $C = LDL^T$, where $D$ is a positive-definite matrix with its elements also being positive. Accordingly, we write $C$ as $C = LL^T$, where $L = LD^{1/2}$, which is the Cholesky decomposition.

It is clear that the matrix $L$ is a lower triangular matrix as well and can be used to transform independent normal variables into dependent multinormal variables, which is the main idea of the method we propose to construct the exact trajectories. The matrix $L$ is calculated by [40,41]

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ c_{12} \sqrt{1-c_{12}^2} & 0 & 0 & \cdots & 0 \\ c_{13} \frac{c_{23} - c_{12}c_{33}}{\sqrt{1-c_{12}^2}} & \sqrt{1-c_{33}R_1^{-1}c_{33}^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{1n} \frac{c_{2n} - c_{12}c_{3n}}{\sqrt{1-c_{12}^2}} & \frac{c_{3n} - c_{13}c_{3n}}{\sqrt{1-c_{33}^2}c_{3n}} & \cdots & \sqrt{1-c_{nn}R_n^{-1}c_{nn}^2} \end{pmatrix},$$

(A27)

where $R_n = c_{ij}^n, i,j=1$ is a positive-definite correlation matrix, $R^{-1}$ is its inverse, and $c_{ij} = (c_{1j}, c_{2j}, \ldots, c_{ij-1})$ for $j \geq i$ so that $c_i \equiv c_i^{ii}$. We note that for a semipositive definite matrix, we should remove the first row and first column of the matrix in order to have a positive-definite matrix and then apply the Cholesky decomposition.

Algorithmically, our Cholesky decomposition algorithm constructs $L$ as follows:

- **input $n, C_{ij}$**
- for $k = 1, 2, \ldots, n$ do
  - $L_{kk} \leftarrow (C_{kk} - \sum_{j=1}^{k-1} L_{kj}^2)^{1/2}$
  - for $i = k+1, k+2, \ldots, n$ do
    - $L_{ik} \leftarrow (C_{ik} - \sum_{j=1}^{k-1} L_{ij}L_{jk}) / L_{kk}$
  - end
- end
- output $L_{ij}$

All the trajectories for FBM in this paper were constructed using this algorithm.

---


