Sampling fractional Brownian motion in presence of absorption: a Markov Chain method

Alexander K. Hartmann*

Institute of Physics, University of Oldenburg, Oldenburg, Germany

Satya M. Majumdar and Alberto Rosso

Université Paris-Sud, CNRS, LPTMS, UMR 8626, Orsay F-91405, France.

(Dated: March 7, 2013)

We study fractional Brownian motion (fBm) characterized by the Hurst exponent H. Using a Monte Carlo sampling technique, we are able to numerically generate fBm processes with an absorbing boundary at the origin at discrete times for a large number of 10^7 time steps even for small values like $H = 1/4$. The results are compatible with previous analytical results that the distribution of (rescaled) endpoints y follow a power law $P_+(y) \sim y^\phi$ with $\phi = (1 - H)/H$, even for small values of H. Furthermore, for the case $H = 0.5$ we also study analytically the finite-length corrections to the first order, namely a plateau of $P_+(y)$ for $y \to 0$ which decreases with increasing process length. These corrections are compatible with the numerical results.

I. INTRODUCTION

The Brownian motion plays a key role in modern theoretical physics, as it explains many effects observed in physical systems. It is currently used in various fields of science to understand, for instance, the trend of financial markets, the dynamics of complex molecules within the cell or the animals food-searching strategies. In order to describe the fluctuations in these systems, it is often necessary to go beyond the Brownian motion and consider random walkers whose mean square displacement grows over time in a nonlinear way. The term used to refer to this situation is *anomalous diffusion*, in particular *sub diffusion* if the mean square displacement grows less than linearly, *super diffusion* if it is faster.

In practice, anomalous diffusion occurs whenever the process $x(\tau)$ is self affine (at least at large time) with a characteristic value of the so-called *Hurst exponent* $H \neq 1/2$, so that the displacement grows with time as τ^H. A remarkable example of process displaying anomalous diffusion is the fractional Brownian motion (fBm), originally introduced by Mandelbrot [1]. This process is self-affine Gaussian process with $0 < H < 1$. A Gaussian process is completely defined by its autocorrelation function, which for fBm writes as

$$\langle x(\tau)x(\tau') \rangle = \tau^{2H} + (\tau')^{2H} - |\tau - \tau'|^{2H} . \qquad (1)$$

The brackets $\langle \ldots \rangle$ refer to an ensemble average over the realizations of the Gaussian processes. The strength of the correlation is described by the Hurst exponent. Note that Eq. (1) implies $\langle [x(\tau_1) - x(\tau_2)]^2 \rangle = 2|\tau_1 - \tau_2|^{2H}$. This means, $H = 1/2$ corresponds to the Brownian motion (standard diffusion), $H > 1/2$ to super-diffusive paths, while $H < 1/2$ correspond to sub-diffusive paths.

Recently these random walks have found to be relevant for many physical applications. Among them we mention the fluctuations of a tagged monomer of an equilibrated Rouse chain [2, 3] or of a tagged particle in the one dimensional system [4, 5]. In both cases the motion of the tagged object can be modeled as a fractional Brownian motion with $H = 1/4$. Other physical processes such as the mechanical unzipping of DNA [6] or the translocation of biomolecules through nanopores [2, 7–9] can be well described by fBm diffusion.

Despite the large number of cases where fBm is observed, very little is known about the properties of this process when is confined in a domain of the space, which is the case for many of the above mentioned applications. In presence of boundaries the translational invariance is lost and analytical representations like the fractional Langevin equation are of little help. In these situations, numerical simulations remain the option to answer many concrete questions arising from biology and physics [10–12]. Here we present a new numerical method to study these processes in presence of boundaries. We will study in detail the case where there is an absorbing wall in $x = 0$: we thus consider only the paths that remain positive up to the final time τ. Recently, analytical predictions [7, 13] were obtained for the distribution $P_+(y)$ of rescaled motion endpoints $y \sim x/(\tau^H)$ at end time τ. A possible numerical strategy consists in the direct sampling of L-step fBm paths $x_0, x_1, \ldots, x_L$ at discrete times, starting at $x_0 = 0$. This strategy is demanding, in particular for $H < 1/2$, since in presence of an absorbing boundary the success probability of generating a non-absorbed trajectory is very small. Hence, such simulations were restricted to a small number L of discrete steps. Here, using a Markov Chain approach, we were able to generate long fBm processes up to $L = 10^7$ dis-

*Electronic address: a.hartmann@uni-oldenburg.de

crete steps for values such as $H = 1/4$, $1/2$ and $2/3$.

The outline of the remainer of the paper is as follows: Next, we present our numerical approach and then we present our numerical results: First - in order to verify that our approach is working- we consider the case of Brownian motion ($H = 1/2$), where analytical results for the finite-length corrections are available. Furthermore we study the cases $H = 1/4$ and $H = 2/3$ as examples for the two regions $H < 1/2$ and $H > 1/2$. In both cases the results are compatible with previous analyitcial predictions. Finally, we summarize our results.

II.   NUMERICAL METHODS

To generate fBm processes on a computer, we study discrete-time random walks with suitable correlations. It is useful to introduce the increments of the random walk, namely $\Delta x_l = x_{l+1} - x_l$. For Gaussian processes the increments are Gaussian variables defined by their auto-correlation function. Using Eq. (1) we can compute the autocorrelation function of the Gaussian increments:

$$
\begin{aligned}
C_{l+m,l} &\equiv \langle \Delta x_{l+m} \Delta x_l \rangle \qquad\qquad (2) \\
&= |m+1|^{2H} - 2|m|^{2H} + |m-1|^{2H} \equiv C(m),
\end{aligned}
$$

We note that this function is independent of the initial time l. Matrices having this property are called *Toeplitz matrices*. Moreover thus the increments are identical Gaussian numbers with variance $\sigma^2 = 2$ displaying power law correlations. Taking the limit $m \to \infty$ it is easy to extract the power law decay of these correlations. For super diffusive fBm ($H > 1/2$), $C(m)$ is positively correlated with a decay as $m^{-2(1-H)}$. Positive correlation means that there is a high probability to observe a long sequence of increments of same sign. For sub diffusive fBm ($H < 1/2$), $C(m)$ is negatively correlated and decay as $-m^{-2(1-H)}$. Negative correlation means that there is a high probability to observe a long sequence of increments of oscillating sign.

The direct generation of L steps with increment correlation Eq.(2) is straightforward, in principle. The starting point is a vector $\xi = (\xi_0, \xi_1, \ldots, \xi_{L-1})$ of L independent and identically distributed (iid) Gaussian (mean zero, variance one) numbers $\sim G(0, 1)$. For the uncorrelated case ($H = 1/2$) one could directly use the random numbers, multiplied by $\sqrt{2}$ to obtain the right $C_{l,l}$, as increments of the fBm processes, i.e., $x^{\mathrm{uncorr}}(L) = \sum_{l=0}^{L-1} \sqrt{2}\xi_l$.

For the case $H \neq 1/2$, since C as a correlation matrix is positive semi-definite, there exist a matrix A such that $C = A^2$. Thus, one could use $\Delta x = A\xi$ to obtain a random vector with the desired property Eq. (2). Nevertheless, this is too time consuming, since it requires diagonalizing a $L \times L$ C matrix once ($\sim L^3$ operations) and, for each process, the multiplication with the $L \times L$ matrix, A ($\sim L^2$ operations) [26]. This is not feasible, in practice, given the sizes $L = 10^7$ we study here.

Instead, we generate random increments which are approximately correlated according to Eq.(2) by generating a periodic increment sequence of period L', with $L' \geq 2L$. The correlation of this periodic sequence are encoded in a covariance matrix $\mathcal{C}_{l,l+m} = \mathcal{C}(m)$ of size $L' \times L'$ built using the original covariance C and defined as:

$$
\begin{aligned}
\mathcal{C}(m) &= C(m) \quad \text{for } m = 0, \ldots, L'/2 - 1 \qquad (3) \\
\mathcal{C}(m) &= C(L' - m) \quad \text{for } m = L'/2, L' - 1\,.
\end{aligned}
$$

Toeplitz matrices displaying this periodicity are called *circulant* matrices. For the actual analysis of the numerical simulations we consider only the first L steps $\Delta x_1, \Delta x_2, \ldots, \Delta x_L$. If L is large, the correlation between the first and the last increment is small and the periodicity has no large influence. The advantage of this approach is that the periodicity of the matrix $\mathcal{C}$ allows the application of Fast Fourier Transformation (FFT) to generate the fBms [14]. The FFT is performed in $\sim L' \log(L')$ operations, technically we use the GNU Scientific Library (GSL) [15]. Let be $\hat{c}_k$ the FFT of $\mathcal{C}(m)$ from Eq. (3), i.e., $\hat{c}_k = \sum_{m=0}^{L'-1} \mathcal{C}(m) e^{-2\pi i \frac{k}{L'} m}$. Since $\mathcal{C}(m)$ is symmetric and positive, the coefficients $\hat{c}_k$ are real positive numbers. The generation of the correlated random numbers works as follows.

(i) the starting point are L' independent and identically distributed (iid) Gaussian numbers (with zero mean and variance one).

$$
\xi = (\xi_0, \xi_1, \ldots, \ldots, \xi_{L-1}).
$$

(ii) we define

$$
\hat{\delta}_k = \sqrt{L'\hat{c}_k}\,\xi_k \quad (k = 0, \ldots, L' - 1) \qquad (4)
$$

which are real numbers as well and where the factor $\sqrt{L'}$ takes into account the correct normalization.

(iii) The vector of the increments is obtained after back transforming the vectors of $\hat{\delta}_k$:

$$
\delta_l = \frac{1}{L'} \sum_{k=0}^{L'-1} \hat{\delta}_k e^{2\pi i \frac{l}{L'} k}\,, \qquad (5)
$$

and taking real and imaginary part:

$$
\Delta x_l = \mathrm{Re}\,\{\delta_l\} + \mathrm{Im}\,\{\delta_l\} \qquad (6)
$$

with $l = 1, 2, \ldots, L'$.

It is easy to check that these three steps lead to the desired correlation. Using $\xi_k = \xi_k^*$ and $\langle \xi_k \xi_{k'} \rangle = \delta_{k,k'}$ one arrives at $\langle \delta x_l^{(*)} \delta x_j^{(*)} \rangle = C(\pm i \pm j)$ where the first and second signs are $+$ for the case of having δx_l and δx_j on the left, respectively, and $-$ for the conjugate complex, respectively Thus, using $\mathrm{Re}(z) = (z + z^*)/2$, $\mathrm{Im}(z) = (z - z^*)/2i$, and $C(m) = C(-m)$ one obtains finally, as desired

$$
\begin{aligned}
&\langle \Delta x_{l+m} \Delta x_l \rangle = \\
&\frac{1}{4} \langle (\delta_{l+m} + \delta_{l+m}^* - i\delta_{l+m} + i\delta_{l+m}^*)(\delta_l + \delta_l^* - i\delta_l + i\delta_l^*) \rangle \\
&\qquad\qquad\qquad\qquad\qquad\qquad\quad = C(m)\,.
\end{aligned}
$$

A numerical test of the method is shown in Fig. 1, which proves that indeed the generated ranomd numbers follow Eq. (2).
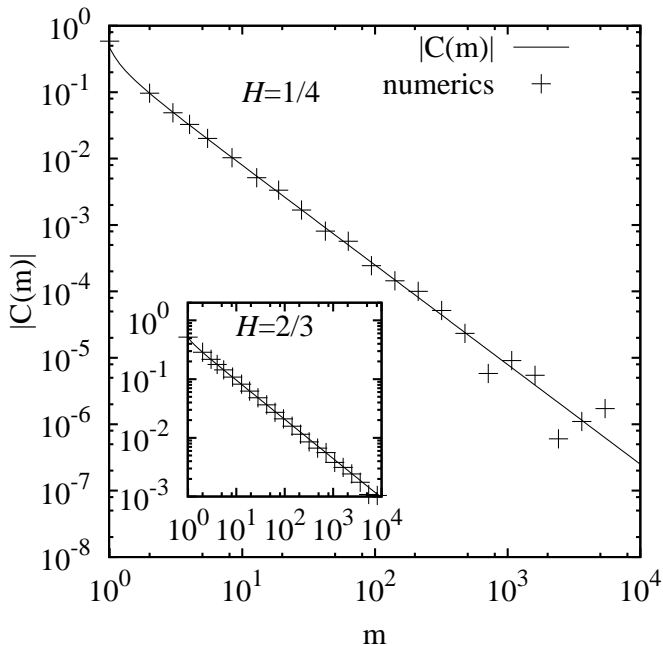

FIG. 1: Correlation between increments. Shown are the wanted function $C(m)$ and the numerical data for $L = 10^4$. The main plot is for $H = 1/4$, which the inset displays $H = 2/3$.

For a direct simulation of fBm processes, one generates a vector of L' real random numbers, constructs the vector of complex numbers, ξ, uses the transformation Eqs. (4),(6) to obtain the correlated increments Δx_l and finally

$$x \equiv x(L) = \sum_{l=1}^{L} \Delta x_l \qquad (7)$$

Nevertheless, since we use an absorbing boundary at $x = 0$, most of the time at least one of the intermediate steps will visit the negative half axis, i.e. $\sum_{l=1}^{\ell} \Delta x_l < 0$ (for some $\ell \leq L$), and the obtained value $x(L)$ does not contribute to the distribution $P_+(y)$. The probability of being not absorbed, i.e., the *persistence* (or survival probability), behaves like

$$S(x_0 = 0, L) \sim L^{-\theta} \qquad (8)$$

with θ being the persistence exponent known to be $\theta = 1 - H$ [3]. Hence, for the case $L = 10^7$ and $H = 1/4$, which we study here, we obtain $P_0(L) \approx 10^{-5}$. This means, a direct simulation is not feasible.

To circumvent this problem, we performed a Markov-chain Monte Carlo simulation with the configuration space being the set of all *feasible* random vectors ξ. Feasible means that the resulting fBm process (after FFT to generate the correlation of the increments Δx) is not absorbed. The simulation must be initialized with an allowed configuration, namely we start from a random vector $\xi^{(0)}$ and a corresponding correlated increment $\Delta x^{(0)}$ such that the resulting process is not absorbed. In practice, for $H \leq 1/2$, we facilitate the generation of a feasible initial configuration by sampling from a shifted Gaussian $G(\bar{\xi}, 1)$ ($\bar{\xi} > 0$) and repeat the search for an initial configurations until one feasible increment vector is found. This initial configuration is clearly biased, but does not have influence on the final result since only after some sufficient equilibration time we start to sample the observables.

Each Monte Carlo step $\xi^{(t)} \rightarrow \xi^{(t+1)}$ consists of changing a fraction p of randomly chosen entries of the configuration $\xi^{(t)}$, the new entries being iid G(0,1), resulting in a trial configuration ξ^{trial}. Then, again after using FFT to introduce the correlation, we obtain Δx^{trial}: if the resulting fBm $\{\sum_{l=1}^{\hat{L}} \tilde{\Delta} x_l^{\text{trial}}\}$ ($\hat{L} = 1, \ldots, L$) is absorbed, the trial configuration is rejected, i.e., $\xi^{(t+1)} = \xi^{(t)}$. If the resulting fBm is allowed, the trial configuration is feasible, hence it is accepted, i.e., $\xi^{(t+1)} = \xi^{\text{trial}}$. This approach satisfies detailed balance, hence converges to the correct distribution: The distribution of the configurations is given by a product of Gaussians over the space of feasible configurations, i.e.,

$$P(\xi) = \prod_{i=1}^{L'} \left(\frac{1}{\sqrt{2\pi}} \exp(-\xi_i^2/2) \right) I_\xi$$

where the indicator function I_ξ is 1 if ξ is feasible, i.e., the resulting fBm is not absorbed, and 0 else. Hence, if a certain fraction p of the entries of ξ is replaced to yield ξ', the resulting change of weight is given by

$$w(\xi \rightarrow \xi') = \prod_{j}' \left(\frac{1}{\sqrt{2\pi}} \exp(-\xi_j'^2/2) \right) I_{\xi'}$$

where the product runs over the changed entries. This change of weight is symmetric to the exchange $\xi \leftrightarrow \xi'$, hence detailed balance is fulfilled: $P(\xi)w(\xi \rightarrow \xi') = P(\xi')w(\xi' \rightarrow \xi)$.

The Markov chain in the configuration space is reflected by the sequence of the endpoints $x^{(0)}(L), x^{(1)}(L), \ldots$ of our Monte Carlo simulation. Here we studied the statistics of the rescaled variable $y = x(L)/(\sigma L^H)$ Since we are interested in the behavior of $P_+(y)$ near $y = 0$, we also used a bias $b(y) = y^{-a}$ ($a > 0$), by imposing an additional Metropolis criterion [16–18] and accepting a feasible configuration with the probability $p_{\text{accept}} = \min\{1, b(y^{\text{trial}})/b(y^{(t)})\}$. This drives the simulation into the range of interest. We adjusted the fraction p of changed entries such that the total acceptance probability of an MC step is near 0.5. Hence, for each value H of the Hurst exponent and each length L, we had to find a suitable value $p = p(H, L)$ empirically.

In Fig. 2 a sample trajectory in the space of (rescaled) endpoints is shown for the case $H = 1/2$ and $L = 10^7$.
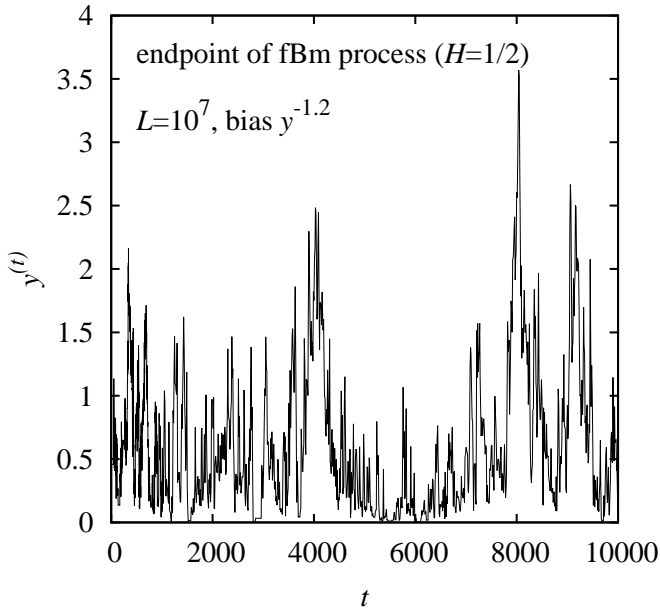
FIG. 2: Sample trajectory of a Monte Carlo simulation: End-point $y^{(t)}$ of a non-absorbed fBm ($H = 1/2$) as a function of the MC time t. A bias $\sim y^{-1.2}$ is used to increase the statistics near $y = 0$.

Via a bias $b(y) \sim y^{-1.2}$ the simulation is concentrated near $y \approx 0$.

Concerning equilibration of our Monte Carlo Simulation, we found that typically, for the longest fBm processes, after 1000 sweeps we do not find any sign of the initial configuration. After disregarding this initial bunch of Monte Carlo sweeps, we measured histograms [19] of the rescaled endpoints of the processes. In case a bias is applied, the histograms have to be multiplied by $b^{-1} = y^a$ and normalized to get the final distributions $P_+(y)$.

III. RESULTS

We have performed simulations to generate fBms for values of the Hurst exponent $H = 1/4$, $H = 1/2$ and $H = 2/3$, of lengths $L = 10^3$, 10^4, 10^5, 10^6 and 10^7, respectively. For the rescaling, we used $a = 2$ ($H = 1/4$), $a = 1.2$ ($H = 1/2$) and $a = 0.5$ ($H = 2/3$). For each case, we determined the parameter p, such that the acceptance probability is (very roughly) about 0.5. The values we used are shown in table I.

Note that for $H = 1/4$ and $H = 1/2$, we have restricted the simulations to fBm processes with $y > 0.0001$ to prevent the simulation being caught near $y = 0$ due to a very small acceptance ration via the rescaling factor in that region.

First, to verify our method, we studied the case of standard random walks, $H = 1/2$. For this case it is possible to know the first corrections to the continuum

L	$H = 1/4$	$H = 1/2$	$H = 2/3$
10^3	0.010	0.030	0.40
10^4	0.020	0.020	0.40
10^5	0.020	0.020	0.20
10^6	0.010	0.020	0.10
10^7	0.015	0.015	0.05

TABLE I: Value of the Monte Carlo parameter p for different lengths of the fBms and different values of the Hurst exponent H.
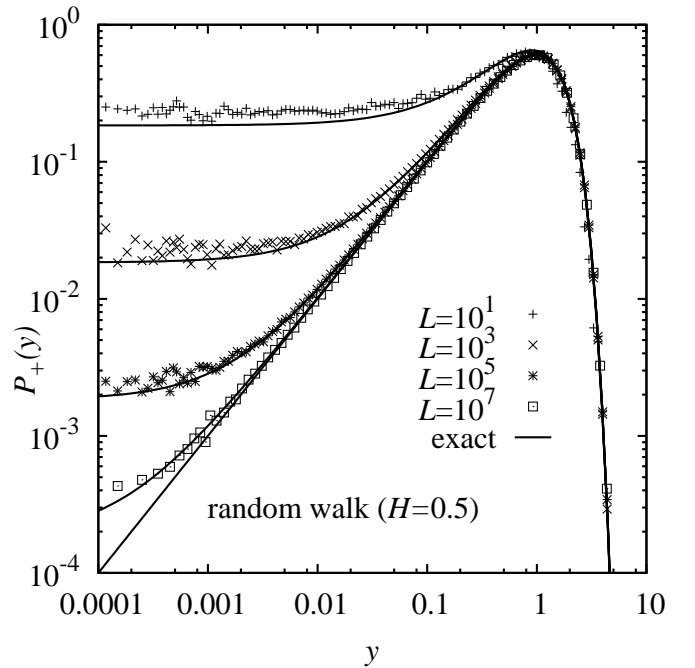

FIG. 3: Distribution $P_+(y)$ of endpoints for non-absorbed fBms (Hurst exponent $H = 1/2$)

limit behavior (see Appendix):

$$P_+(y, L) = f_0(y) - \frac{c}{\sqrt{L}} f_1(y) + \dots \quad (9)$$

where L is the number of increments, the constant c depends on the increments distribution of the random walk. For Gaussian numbers (zero mean, unit variance) we have $c = \zeta(1/2)/\sqrt{2\pi} \sim -0.582597\dots$ and the scaling functions are:

$$f_0(y) = y\,e^{-y^2/2} \quad (10)$$
$$f_1(y) = \left(1 - \frac{2y}{\pi}\right) e^{-y^2/2}\,.$$

The rescaled distributions $P_+(y)$ of the endpoints are shown, together with the predictions of Eq.(9) valid for large L, in Fig. 3. One is able to see strong finite-size effects for small values $y \to 0$, where a plateau is visible. For increasing length L, the plateau decreases as $c/\sqrt{L}$ and the data approaches better and better the continuum limit scaling function f_0.

We conclude that for a generic fBm for which first corrections to the continuum limit behavior is not known, the plateau should also vanish when the size of the system is large and for very long walks ($L = 10^7$) the continuum limit behavior is displayed over several order of magnitudes.
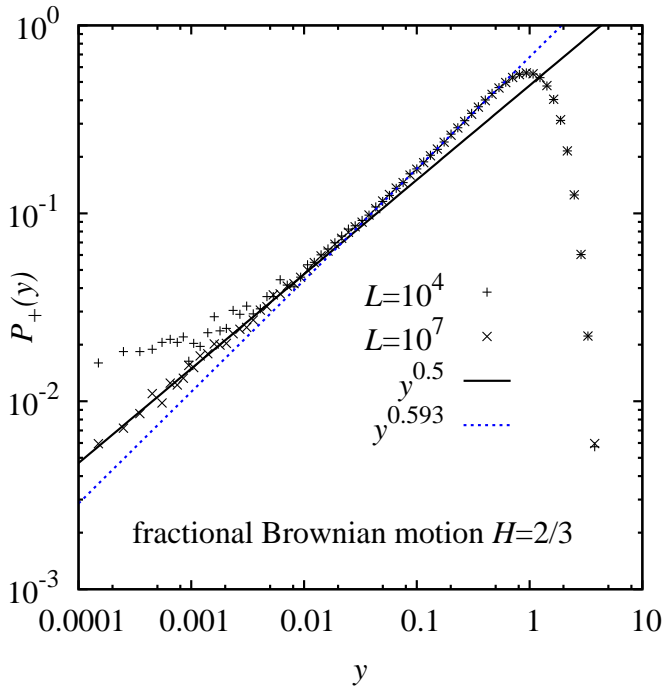

FIG. 4: (color online) Distribution $P_+(y)$ of endpoints for non-absorbed fBms (Hurst exponent $H = 2/3$)

Based on scaling arguments it has been conjectured [7] that, in the continuum limit $P_+(y)$ vanishes as y^ϕ with $\phi = (1 - H)/H = \theta/H$ for $y \to 0$, θ being the persistence defined via Eq. (8). This conjecture was confirmed by an epsilon expansion around the Brownian solution obtained thanks to a field theory calculation [13]. The numerical check of the conjecture for values of H far from $1/2$ remains very challenging. We first consider the discrete random walk with $H = 2/3$. In this case, since the persistence is decreasing not very fast ($\theta = 1/3$), numerical results were obtained for moderate lengths $L = 2 \times 10^4$ by direct simulations [13], which were compatible with the analytics. Here, we were able to study this case again. Our results, up to a length of $L = 10^7$, confirm the analytics with much better accuracy, see Fig. 4.

Finally, we turn to the most difficult case whith $H = 1/4$ where we expect that $P_+(y)$ vanishes as y^3 as $y \to 0$. Direct simulations on this process are not conclusive and a scaling behavior $\sim y^2$ is consistent with the data [20, 21]. Using our Markov chain approach we can see the finite size effects remain important even for long processes (see Fig. 5). Even if we are not able to extract the correct continuum limit behavior, we can show that the apparent scaling exponent grows with the size of the system. When $L = 10^7$ the best fit give $\sim y^{2.59}$, much
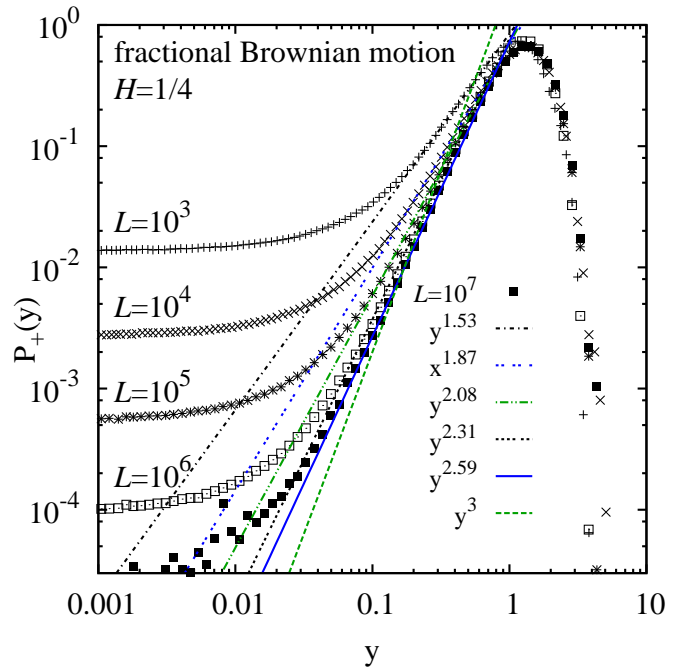

FIG. 5: (color online) Distribution $P_+(y)$ of endpoints for non-absorbed fBms (Hurst exponent $H = 1/4$)

more close to the expected y^3 behavior than previous simulations.

IV. SUMMARY

We have introduced a Markov-chain Monte Carlo approach to study numerically fractional Brownian motion in the presence of an absorbing boundary via generating finite-step random walks with correlated disorder. Our approach allowed us to study long walks up to $L = 10^7$ steps. For the test case $H = 1/2$ the result for the distribution $P_+(y)$ of the rescaled endpoints y of the walks agrees in the limit $L \to \infty$ with the exacty analytic result. We also derived analytical expression for the fite-length corrections, which turn also to be compatible with the numerical results, better with increasing step number L. Finally, we studied the cases $H = 1/4$ and $H = 2/3$ where we find for $y \to 0$ a power law behavior $P_+(y) \sim y^\phi$. For the case $H = 2/3$, the exponent is compatible with the analytical prediction $\phi = (1 - H)/H$. For the case $H = 1/4$ we also find a finite-size dependece of the effective exponents with a strong trend towards the predicted value $\phi = 3$. In particular, from our results it is clear that $\phi > 2$.

V. ACKNOWLEDGMENTS

APPENDIX A: DERIVATION OF EQ. 8 AND EQ. 9

We consider a random walk starting at the origin. Its position at discrete time steps evolves via

$$x_n = x_{n-1} + \eta_n \tag{A1}$$

starting from $x_0 = 0$. The random variables η_n's are independent and identically distributed noises, each drawn from a symmetric and continuous probability density function (pdf) $f(\eta)$. Let $p_L(x)$ denote the probability density that the particle arrives at x at step L while staying above 0 at all intermediate steps. An exact expression for $p_L(x)$, or rather for its generating function, is known explicitly for arbitrary jump density $f(\eta)$ and is given by [22]

$$\int_0^\infty dx\, e^{-\lambda x} \sum_{L=0}^\infty p_L(x)\, s^L = \phi(s, \lambda) \tag{A2}$$

with

$$\varphi(s, \lambda) = \exp\left(-\frac{\lambda}{\pi} \int_0^\infty \frac{\ln\left[1 - s\hat{f}(k)\right]}{k^2 + \lambda^2}\, dk \right) \tag{A3}$$

where $\hat{f}(k) = \int_{-\infty}^\infty e^{i k \eta} f(\eta)\, d\eta$ is the Fourier transform of the noise density. Our goal is to extract the leading (and subleading) scaling behavior of $p_L(x)$ for large L from Eq. (A2).

To make progress, it is useful to consider an alternative expression for $\phi(s, \lambda)$ derived in Ref. [23], valid for all $f(\eta)$'s with a finite variance $\sigma^2 = \int_\infty^\infty \eta^2 f(\eta)\, d\eta$,

$$\phi(s, \lambda) = \frac{1}{\left[\sqrt{1-s} + \sigma \lambda \sqrt{s/2}\right]} \times$$
$$\times \exp\left[-\frac{\lambda}{\pi} \int_0^\infty \frac{dk}{\lambda^2 + k^2} \ln\left(\frac{1 - s\hat{f}(k)}{1 - s + s\sigma^2 k^2/2} \right) \right]. \tag{A4}$$

We next consider the scaling limit when $x \to \infty$, $L \to \infty$, with the ratio $y = x/\sqrt{L}$ fixed. In the Laplace place, this

corresponds to taking the limit $\lambda \to 0$, $s \to 1$, keeping the ratio $\lambda/\sqrt{1-s}$ fixed. Taking this scaling limit in Eq. (A4), one gets

$$\phi(s, \lambda) \to \frac{1 - c\lambda}{\sqrt{1-s} + \sigma\lambda/\sqrt{2}} \tag{A5}$$

where c is a constant with the following expression [23, 24]

$$c = \frac{1}{\pi} \int_0^\infty \frac{dk}{k^2} \ln\left[\frac{1 - \hat{f}(k)}{\sigma^2 k^2/2} \right]. \tag{A6}$$

Substituting the scaling-limit expression of $\phi(s, \lambda)$ from Eq. (A5) on the right hand side of Eq. (A2) and inverting the Laplace transform with respect to λ gives,

$$\sum_{L=0}^\infty p_L(x)\, s^L \approx \frac{\sqrt{2}}{\sigma} \left[1 + \frac{\sqrt{2}}{\sigma} c\sqrt{1-s} \right] e^{-\sqrt{2(1-s)}\, x/\sigma}, \tag{A7}$$

valid in the scaling limit $s \to 1$, $x \to \infty$ but keeping the product $\sqrt{1-s}\, x$ fixed. Next, one can invert this generating function with respect to s, using Cauchy's inversion formula. Skipping details, we find that the two leading terms, in the scaling limit where $x \to \infty$, $L \to \infty$, but keeping $y = x/\sqrt{L}$ fixed, are given by

$$p_L(x) \approx \frac{1}{\sigma^2 \sqrt{\pi L}} \left[y\, e^{-y^2/2\sigma^2} - \frac{c}{\sqrt{L}}\, e^{-y^2/2\sigma^2} \right]. \tag{A8}$$

The conditional probability $P_L(x)$ (probability density to reach the position x given that it has survived up to L steps) is defined as

$$P_L(x) = \frac{p_L(x)}{\int_0^\infty p_L(x)\, dx}. \tag{A9}$$

Substituting the scaling behavior for $p_L(x)$ from Eq. (A8) in the above definition, we find that $P_L(x)$ has the following scaling behavior

$$P_L(x) \to \frac{1}{\sqrt{L}}\, P_+(y, L) \tag{A10}$$

with $y = x/\sqrt{L}$ and

$$P_+(y, L) = f_0(y) - \frac{c}{\sqrt{L}}\, f_1(y) + O(1/L) \tag{A11}$$

where

$$f_0(y) = \frac{y}{\sigma^2}\, e^{-y^2/2\sigma^2} \tag{A12}$$

$$f_1(y) = e^{-y^2/2\sigma^2} - \frac{2}{\pi\sigma}\, y\, e^{-y^2/2\sigma^2} \tag{A13}$$

and the constant c is given by Eq. (A6). For the special case of the Gaussian jump density, $f(\eta) = e^{-\eta^2/2}/\sqrt{2\pi}$

(with $\sigma^2 = 1$), one can evaluate the constant c in Eq. (A6) explicitly [23]

$$c = \frac{\zeta(1/2)}{\sqrt{2\pi}} = -0.582597\ldots \qquad (A14)$$

In this case, in particular, putting $y = 0$ we get

$$P_+(0, L) \approx -\frac{c}{\sqrt{L}} = \frac{0.582597\ldots}{\sqrt{L}}, \qquad (A15)$$

which is consistent with our simulations.

[1] B. B. Mandelbrot and J. W. van Ness, SIAM Review **10**, 422 (1968).

[2] Y. Kantor and M. Kardar, Phys. Rev. E **69**, 021806 (2004), URL http://link.aps.org/doi/10.1103/PhysRevE.69.021806.

[3] J. Krug, H. Kallabis, S. N. Majumdar, S. J. Cornell, A. J. Bray, and C. Sire, Phys. Rev. E **56**, 2702 (1997), URL http://link.aps.org/doi/10.1103/PhysRevE.56.2702.

[4] T. E. Harris, Journal of Applied Probability **2**, pp. 323 (1965), ISSN 00219002, URL http://www.jstor.org/stable/3212197.

[5] S. N. Majumdar and M. Barma, Phys. Rev. B **44**, 5306 (1991), URL http://link.aps.org/doi/10.1103/PhysRevB.44.5306.

[6] J.-C. Walter, A. Ferrantini, E. Carlon, and C. Vanderzande, Phys. Rev. E **85**, 031120 (2012), URL http://link.aps.org/doi/10.1103/PhysRevE.85.031120.

[7] A. Zoia, A. Rosso, and S. N. Majumdar, Phys. Rev. Lett. **102**, 120602 (2009), URL http://link.aps.org/doi/10.1103/PhysRevLett.102.120602.

[8] D. Panja and G. T. Barkema, The Journal of Chemical Physics **132**, 014902 (pages 10) (2010), URL http://link.aip.org/link/?JCP/132/014902/1.

[9] D. Panja, G. T. Barkema, and R. C. Ball, Journal of Physics: Condensed Matter **19**, 432202 (2007), URL http://stacks.iop.org/0953-8984/19/i=43/a=432202.

[10] R. García-García, A. Rosso, and G. Schehr, Phys. Rev. E **81**, 010102 (2010), URL http://link.aps.org/doi/10.1103/PhysRevE.81.010102.

[11] G. Oshanin, A. Rosso, and G. Schehr (2012).

[12] R. Cakir, P. Grigolini, and A. A. Krokhin, Phys. Rev. E **74**, 021108 (2006), URL http://link.aps.org/doi/10.1103/PhysRevE.74.021108.

[13] K. J. Wiese, S. N. Majumdar, and A. Rosso, Phys. Rev. E **83**, 061141 (2011), URL http://link.aps.org/doi/10.1103/PhysRevE.83.061141.

[14] L. E. Wittig and A. K. Sinha, J. Acoust. Soc. Amer. **58**, 630 (1975), URL http://link.aip.org/link/?JAS/58/630/1.

[15] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, M. Booth, and F. Rossi, *GNU Scientific Library Reference Manual* (Network Theory Ltd., Bristol, UK, 2006), URL http://www.gnu.org/software/gsl/.

[16] A. K. Hartmann, Phys. Rev. E **65**, 056102 (2002).

[17] A. Engel, R. Monasson, and A. K. Hartmann, J. Stat. Phys. **117**, 387 (2004).

[18] A. K. Hartmann, Eur. Phys. J. B **84**, 627 (2011).

[19] A. K. Hartmann, *Practical Guide to Computer Simulations* (World Scientific, Singapore, 2009).

[20] Y. Kantor and M. Kardar, Phys. Rev. E **76**, 061121 (2007), URL http://link.aps.org/doi/10.1103/PhysRevE.76.061121.

[21] A. Amitai, Y. Kantor, and M. Kardar, Phys. Rev. E **81**, 011107 (2010), URL http://link.aps.org/doi/10.1103/PhysRevE.81.011107.

[22] V. V. Ivanov, Astron. Astrophys. **286**, 328 (1994).

[23] A. Comtet and S. N. Majumdar, Journal of Statistical Mechanics: Theory and Experiment **2005**, P06013 (2005), URL http://stacks.iop.org/1742-5468/2005/i=06/a=P06013.

[24] S. N. Majumdar, A. Comtet, and R. M. Ziff, Journal of Statistical Physics **122**, 833 (2006), ISSN 0022-4715, URL http://dx.doi.org/10.1007/s10955-005-9002-x.

[25] T. Dieker, URL http://www2.isye.gatech.edu/~adieker3/fbm/index.html.

[26] Better results for direct sampling can be obtained by making use of the fact for fBm the matrix C is also a Toeplitz matrix. For Toeplitz matrices efficient numerical methods allow to avoid the full diagonalization of C. An example is given by the Levinson algorithm (for a practical implementation of Levinson's algorithm see [25].). However here we use a Markov Chain sampling (not compatible with the Levinson algorithm) which is more efficient in presence of absorption.