PHYSICAL REVIEW E **86**, 011119 (2012)

Record statistics for multiple random walks

Gregor Wergen*
Institut für Theoretische Physik, Universität zu Köln, 50937 Köln, Germany

Satya N. Majumdar† and Grégory Schehr‡
Laboratoire de Physique Théorique et Modèles Statistiques, UMR 8626, Université Paris Sud 11 and CNRS, Bâtiment 100, Orsay F-91405, France

(Received 25 April 2012; published 18 July 2012)

We study the statistics of the number of records $R_{n,N}$ for N identical and independent symmetric discrete-time random walks of n steps in one dimension, all starting at the origin at step 0. At each time step, each walker jumps by a random length drawn independently from a symmetric and continuous distribution. We consider two cases: (I) when the variance σ^2 of the jump distribution is finite and (II) when σ^2 is divergent as in the case of Lévy flights with index $0 < \mu < 2$. In both cases we find that the mean record number $\langle R_{n,N}\rangle$ grows universally as $\sim\alpha_N\sqrt{n}$ for large n, but with a very different behavior of the amplitude α_N for $N > 1$ in the two cases. We find that for large N, $\alpha_N \approx 2\sqrt{\ln N}$ independently of σ^2 in case I. In contrast, in case II, the amplitude approaches to an N-independent constant for large N, $\alpha_N \approx 4/\sqrt{\pi}$, independently of $0 < \mu < 2$. For finite σ^2 we argue—and this is confirmed by our numerical simulations—that the full distribution of $(R_{n,N}/\sqrt{n} - 2\sqrt{\ln N})\sqrt{\ln N}$ converges to a Gumbel law as $n \to \infty$ and $N \to \infty$. In case II, our numerical simulations indicate that the distribution of $R_{n,N}/\sqrt{n}$ converges, for $n \to \infty$ and $N \to \infty$, to a universal nontrivial distribution independently of μ. We discuss the applications of our results to the study of the record statistics of 366 daily stock prices from the Standard & Poor's 500 index.

DOI: 10.1103/PhysRevE.86.011119 PACS number(s): 05.40.−a, 02.50.Sk

I. INTRODUCTION

A record is an entry in a series of events that exceeds all previous entries. In recent years there has been a surge of interest in the statistics of record-breaking events, both from the theoretical point of view as well as in multiple applications. The occurrence of record-breaking events has been studied, for instance, in sports [1,2], in evolution models in biology [3,4], in the theory of spin glasses [5,6], and in models of growing networks [7]. Recently there has been some progress in understanding the phenomenon of global warming via studying the occurrence of record-breaking temperatures [8–11].

More precisely, let us consider a sequence or a discrete-time series of random variables $\{x(0),x(1),x(2),\ldots,x(n)\}$ with $n+1$ entries. This sequence may represent, for example, the daily maximum temperature in a city or the daily maximum price of a stock. A record is said to happen at step m if the mth member of the sequence is bigger than all previous members, that is, if $x(m) > x(i)$ for all $i = 0,1,2,\ldots,(m-1)$. Let R_n denote the number of records in this sequence of $n+1$ entries. Clearly, R_n is a random variable whose statistics depends on the joint distribution of $P(x(0),x(1),\ldots,x(n))$ of the members of the sequence. When the members of the sequence are independent and identically distributed (i.i.d.) random variables, each drawn from a distribution $p(x)$, that is, the joint distribution factorizes, $P(x(0),x(1),\ldots,x(n)) = \prod_{i=0}^{n} p(x(i))$, the record statistics is well understood from classical theories [12–14]. In particular, when $p(x)$ is a continuous distribution, it is known that the distribution of record number $P(R_n,n)$ is universal for all n, that is, independent of the parent distribution $p(x)$. The average number of records up to step n, $\langle R_n\rangle = \sum_{m=1}^{n+1} 1/m$ for all n and the universal distribution, for large n, converges to a Gaussian distribution with mean $\approx\ln(n)$ and variance $\approx\ln n$.

While the statistical properties of records for i.i.d. random variables (RVs) are thus well understood for many years, numerous questions remain open for more realistic systems with time-dependent or correlated RVs. In principle, there are many different ways to generalize the simple i.i.d. RV scenario described above. For instance, one can consider time series of RVs that are independent, but not identically distributed. One example for this case is the so called linear drift model with RV's from probability distributions with identical shape, but with a mean value that increases in time. This model was first proposed in the 1980s [15] and was recently thoroughly analyzed in Refs. [16–18]. In 2007 Krug also considered the case of uncorrelated RV's from distributions with increasing variance [4].

Another possible generalization is the one where RVs are correlated. Perhaps the simplest and the most natural model of correlated RVs is an n-step one dimensional discrete-time random walk with entries $\{x(0) = 0,x(1),x(2),\ldots,x(n)\}$ where the position $x(m)$ of the walker at discrete time m evolves via the Markov jump process,

$$x(m) = x(m-1) + \eta(m), \tag{1}$$

with $x(0) = 0$ and $\eta(m)$ represents the random jump at step m. The noise variables $\eta(m)$'s are assumed to be i.i.d. variables, each drawn from a symmetric distribution $f(\eta)$. For instance, it may include Lévy flights where $f(\eta) \sim |\eta|^{-1-\mu}$ for large η with the Lévy index $0 < \mu < 2$, which has a divergent second moment. Even though this model represents a very simple

*gw@thp.uni-Koeln.de
†satya.majumdar@u-psud.fr
‡gregory.schehr@u-psud.fr

Markov chain, statistical properties of certain observables associated with such a walk may be quite nontrivial to compute, depending on which observable one is studying [19–21]. For instance, in recent years there has been a lot of interest in the extremal properties of such random walks. These include the statistics of the maximal displacement of the walk up to n steps with several applications [21–25] and the order statistics, that is, the statistics of the ordered maxima [26,27], as well as the universal distribution of gaps between successive ordered maxima of a random walk [27].

The statistics of the number of record-breaking events in the discrete-time random walk process in Eq. (1) has also been studied in a number of recent works with several interesting results [28–32]. In 2008, Majumdar and Ziff computed exactly the full distribution $P(R_n,n)$ of the record number up to n steps and found that when the jump distribution $f(\eta)$ is continuous and symmetric, the record number distribution $P(R_n,n)$ is completely universal for all n, that is, independent of the details of the jump distribution [28]. In particular, for instance, the Lévy flight with index $0 < \mu < 2$ [thus with a divergent second moment of the jump distribution $f(\eta)$] has the same record number distribution as for a Gaussian walk [with a finite second moment of $f(\eta)$]. This is a rather amazing result and the deep reason for this universality is rooted [28] in the so called Sparre Andersen theorem [33]. In particular, for large n, $P(R_n,n) \sim n^{-1/2} G(R_n/\sqrt{n})$, where the scaling function $G(x) = e^{-x^2/4}/\sqrt{\pi}$ is universal [28]. The mean number of records $\langle R_n \rangle \approx \sqrt{4n/\pi}$ for large n [28]. In contrast, this universal result does not hold for symmetric but discontinuous $f(\eta)$. For example, if $f(\eta) = \frac{1}{2}\delta(\eta-1) + \frac{1}{2}\delta(\eta+1)$, then x_m represents the position of a random walker at step m on a one dimensional (1D) lattice with lattice spacing 1. In this case, the mean number of records still grows as $\sqrt{n}$ for large n but with a smaller prefactor, $\langle R_n \rangle \approx \sqrt{2n/\pi}$ [28].

These results were later generalized to several interesting cases, for instance, to the record statistics of 1D random walk in presence of an external drift [29,30] and 1D continuous-time random walk with a waiting-time distribution between successive jumps [31]. The record statistics of the distance traveled by a random walker in higher dimensions with and without drift has been studied numerically in the context of contamination spread in porous medium [32]. In [30], it was also found that the record statistics of stock markets is very similar to the ones of biased random walks.

While in Refs. [28–32] the record statistics of a single discrete-time random walker was studied, the purpose of this article is to generalize these results to the case where one has N independent 1D discrete-time random walks. In this N-walker process, a record happens at an instant when the maximum position of all the walkers at that instant exceeds all its previous values. We will see that despite the fact that the walkers are independent, the record statistics is rather rich, universal, and nontrivial even in this relatively simple model.

Let us first summarize our main results. We derive asymptotic results for the mean record number $\langle R_{n,N} \rangle$ up to time n and also discuss its full distribution. It turns out that for $N > 1$, while the full universality with respect to the jump distribution found for $N = 1$ case is no longer valid, there still remains a vestige of universality of a different sort. In our analysis, it is important to distinguish two cases: case (I), where the jump distribution $f(\eta)$ has a finite variance $\sigma^2 = \int_{-\infty}^{\infty} \eta^2 f(\eta)\, d\eta$; and case (II), where σ^2 is divergent as in the case of Lévy flights with Lévy index $0 < \mu < 2$. In both cases, we find that the mean record number $\langle R_{n,N} \rangle$ grows *universally* as $\sim \alpha_N \sqrt{n}$ for large n. However, the N dependence of the prefactor α_N, in particular for large N, turns out to be rather different in the two cases

$$\alpha_N \xrightarrow[N\to\infty]{} \begin{cases} 2\sqrt{\ln N} \text{ in case I} \quad (\text{independent of } \sigma^2), \\ \\ 4/\sqrt{\pi} \text{ in case II} \quad (\text{independent of } \mu). \end{cases} \tag{2}$$

In addition, we also study the distribution of the record number $R_{n,N}$. For finite σ^2 we argue and confirm numerically that the distribution of the RV $(R_{n,N}/\sqrt{n} - 2\sqrt{\ln N})\sqrt{\ln N}$ converges to the Gumbel law asymptotically for large n and N (see Sec. II for details). In contrast, in case II, we find numerically that the distribution of $R_{n,N}/\sqrt{n}$ converges, for large n and N, to a nontrivial distribution independent of the value of $0 < \mu < 2$ (see Sec. II for details). We were, however, unable to compute this asymptotic distribution analytically and it remains a challenging open problem. Finally, we discuss the applications of our results to the study of the record statistics of 366 daily stock prices from the Standard & Poor's 500 index [34]. We analyze the evolution of the record number in subsets of N stocks that were randomly chosen from this index and compare the results to our analytical findings. While the strong correlations between the individual stocks seem to play an important effect in the record statistics, the dependence of the record number on N still seems to be the same as in the case of N independent random walkers.

The rest of the paper is organized as follows. In Sec. II, we define the N-walker model precisely and summarize the main results obtained in the paper. In Sec. III, we present the analytical calculation of the mean number of records for multiple random walkers, in both cases where σ^2 is finite (case I) and σ^2 is infinite (case II). Section IV is devoted to an analytic study of the distribution of the record number in the case where σ^2 is finite. In Sec. V we present a thorough numerical study of the record statistics of multiple random walks, and in Sec. VI we discuss the application of our results to the record statistics of stock prices. Finally, we conclude in Sec. VII and present the technical details of some of the analytical computations concerning the computation of the mean number of records and the distribution of the record number for lattice random walks in the three Appendixes A, B, and C.

II. RECORD STATISTICS FOR MULTIPLE RANDOM WALKS: THE MODEL AND THE MAIN RESULTS

Here we consider the statistics of records of N independent random walkers all starting at the origin 0. The position $x_i(m)$ of the ith walker at discrete time step m evolves via the Markov evolution rule

$$x_i(m) = x_i(m-1) + \eta_i(m), \tag{3}$$

where $x_i(0) = 0$ for all $i = 1,2,\ldots,N$ and the noise $\eta_i(m)$'s are i.i.d. variables (independent from step to step and from

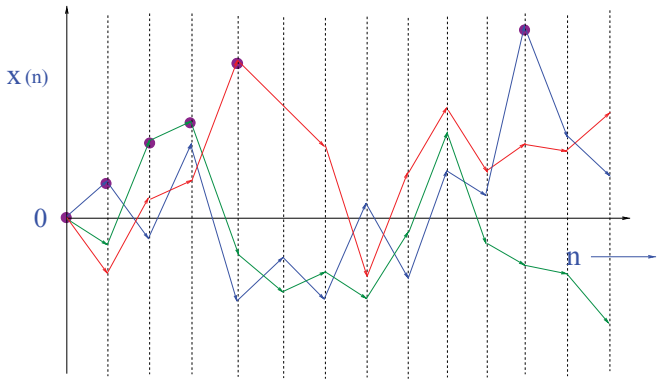

FIG. 1. (Color online) Schematic trajectories of $N = 3$ random walkers. Each walker starts at the origin and evolves via the Markov jump process in Eq. (3). A record happens at step m if the maximum position at step m exceeds its previous values, i.e., $x_{\max}(m) > x_{\max}(k)$ for all $k = 0,1,2,\ldots,(m-1)$. The record values are shown by solid circles.

walker to walker), each drawn from a symmetric distribution $f(\eta)$. We are interested in the record statistics of the composite process. More precisely, consider at each step m, the maximum position of all N random walkers

$$x_{\max}(m) = \max\left[x_1(m), x_2(m), \ldots, x_N(m)\right]. \tag{4}$$

A record is said to happen at step m if this maximum position at step m is bigger than all previous maximum positions, that is, if $x_{\max}(m) > x_{\max}(k)$ for all $k = 0,1,\ldots,(m-1)$ (see Fig. 1). In other words, we are interested in the record statistics of the stochastic discrete-time series $\{x_{\max}(m)\}$, with the convention that the initial position $x_{\max}(0) = 0$ is counted as a record. Note that even though the position of each walker evolves via the simple independent Markovian rule in Eq. (3), the evolution of the maximum process $\{x_{\max}(m)\}$ is highly non-Markovian and hence is nontrivial.

Let $R_{n,N}$ denote the number of records up to step n for this composite N-walker process. Clearly, $R_{n,N}$ is a RV and we are interested in its statistics. For a single walker $N = 1$, we have already mentioned that the probability distribution of the record number $R_{n,1}$ is completely universal, that is, independent of the jump distribution $f(\eta)$ as long as $f(\eta)$ is symmetric and continuous [28]. In particular, for example, the record number distribution is the same for simple Gaussian walkers as well for Lévy flights with index $0 < \mu < 2$. Here we are interested in the opposite limit when $N \to \infty$.

We find that while the complete universality of the record statistics is no longer true for $N > 1$, a different type of universal behavior emerges in the $N \to \infty$ limit. In this large N limit, there are two universal asymptotic behaviors of the record statistics depending on whether the second moment $\sigma^2 = \int_{-\infty}^{\infty} \eta^2 f(\eta)\, d\eta$ of the jump distribution is finite or divergent. For example, for Gaussian, exponential, uniform jump distributions σ^2 is finite. In contrast, for Lévy flights where $f(\eta) \sim |\eta|^{-\mu-1}$ for large η with the Lévy index $0 < \mu < 2$, the second moment σ^2 is divergent. In these two cases, we find the following behaviors for the record statistics.

Case I (σ^2 finite). In this case, we consider jump distributions $f(\eta)$ that are symmetric with a finite second moment $\sigma^2 = \int_{-\infty}^{\infty} \eta^2 f(\eta)\, d\eta$. In this case, the Fourier transform of the jump distribution $\hat{f}(k) = \int_{-\infty}^{\infty} f(\eta)\, e^{ik\eta}\, d\eta$ behaves, for small k, as

$$\hat{f}(k) \approx 1 - \frac{\sigma^2}{2} k^2 + \cdots. \tag{5}$$

Examples include the Gaussian jump distribution, $f(\eta) = \sqrt{a/\pi}\, e^{-a\eta^2}$, exponential jump distribution $f(\eta) = (b/2)\exp[-b|\eta|]$, uniform jump distribution over $[-l,l]$, etc. For such jump distributions, we find that for large number of walkers N, the mean number of records grows asymptotically for large n as

$$\langle R_{n,N}\rangle \xrightarrow[N\to\infty]{n\to\infty} 2\sqrt{\ln N}\,\sqrt{n}. \tag{6}$$

Note that this asymptotic behavior is universal in the sense that it does not depend explicitly on σ as long as σ is finite.

Moreover, we argue that for large N and large n, the scaled RV $R_{n,N}/\sqrt{n}$ converges, in distribution, to the Gumbel form, that is,

$$\text{Prob}\left[\frac{R_{n,N}}{\sqrt{n}} \leqslant x\right] \xrightarrow[N\to\infty]{n\to\infty} F_1[(x - 2\sqrt{\ln N})\sqrt{\ln N}], \quad \text{where} \quad F_1(z) = \exp[-\exp[-z]]. \tag{7}$$

Indeed, for large N and large n, the scaled variable $R_{n,N}/\sqrt{n}$ converges, in distribution, to the maximum of N independent RVs,

$$\frac{R_{n,N}}{\sqrt{n}} \xrightarrow[N\to\infty]{n\to\infty} M_N, \quad \text{where} \quad M_N = \max(y_1, y_2, \ldots, y_N), \tag{8}$$

where $y_i \geqslant 0$'s are i.i.d. non-negative RVs, each drawn from distribution $p(y) = \frac{1}{\sqrt{\pi}}\, e^{-y^2/4}$ for $y \geqslant 0$ and $p(y) = 0$ for $y < 0$.

Case II (σ^2 divergent). In this case we consider jump distributions $f(\eta)$ such that the second moment σ^2 is divergent. In this case, the Fourier transform $\hat{f}(k)$ of the noise distribution behaves, for all k, as

$$\hat{f}(k) = 1 - |a\,k|^{\mu} + \cdots, \tag{9}$$

where $0 < \mu < 2$. Examples include Lévy flights where $f(\eta) \sim |\eta|^{-\mu-1}$ with the Lévy index $0 < \mu < 2$. For the noise distribution in Eq. (9), we find, quite amazingly, that in the large N and large n limit, the record statistics is (i) completely universal, that is, independent of μ and a; and (ii) more surprisingly, and unlike in Case I, the record statistics also becomes independent of N as $N \to \infty$. For example, we prove that for large N, the mean number of records grows asymptotically with n as

$$\langle R_{n,N}\rangle \xrightarrow[N\to\infty]{n\to\infty} \frac{4}{\sqrt{\pi}}\sqrt{n}, \tag{10}$$

which is exactly twice that of one walker; that is, $\langle R_{n,N\to\infty}\rangle = 2\langle R_{n,1}\rangle$ for large n. Similarly, we find that the scaled variable $R_{n,N}/\sqrt{n}$, for large n and large N, converges to a universal distribution

$$\text{Prob}\left[\frac{R_{n,N}}{\sqrt{n}} \leqslant x\right] \xrightarrow[N\to\infty]{n\to\infty} F_2(x), \tag{11}$$

which is independent of the Lévy index μ as well as of the scale a in Eq. (9). While we have computed this universal

distribution $F_2(x)$ numerically rather accurately, we were not able to compute its analytical form.

III. MEAN NUMBER OF RECORDS FOR MULTIPLE WALKERS

Let $R_{n,N}$ be the number of records up to step n for N random walkers, that is, for the maximum process $x_{\max}(n)$. Let us write

$$R_{m,N} = R_{m-1,N} + \xi_{m,N}, \tag{12}$$

where $\xi_{m,N}$ is a binary RV taking values 0 or 1. The variable $\xi_{m,N} = 1$ if a record happens at step m and $\xi_{m,N} = 0$ otherwise. Clearly, the total number of records up to step n is

$$R_{n,N} = \sum_{m=1}^{n} \xi_{m,N}. \tag{13}$$

So, the mean number of records up to step n is

$$\langle R_{n,N} \rangle = \sum_{m=1}^{n} \langle \xi_{m,N} \rangle = \sum_{m=1}^{n} r_{m,N}, \tag{14}$$

where $r_{m,N} = \langle \xi_{m,N} \rangle$ is just the record rate, that is, the probability that a record happens at step m. To compute the mean number of records, we first evaluate the record rate $r_{m,N}$ and then sum over m.

To compute $r_{m,N}$ at step m, we need to sum the probabilities of all trajectories that lead to a record event at step m. Suppose that a record happens at step m with the record value x (see Fig. 2). This corresponds to the event that one of the N walkers (say the dashed trajectory in Fig. 2), starting at the origin at step 0, has reached the level x for the first time at step m, while the rest of the $N-1$ walkers, starting at the origin at step 0, have all stayed below the level x until the step m. Also, the walker that actually reaches x at step m can be any of the N walkers. Finally, this event can take place at any level $x > 0$ and one needs to integrate over the record value x. Using the independence of N walkers and taking into account the event detailed above, one can then write

$$r_{m,N} = N \int_0^\infty p_m(x)\, [q_m(x)]^{N-1}\, dx, \tag{15}$$

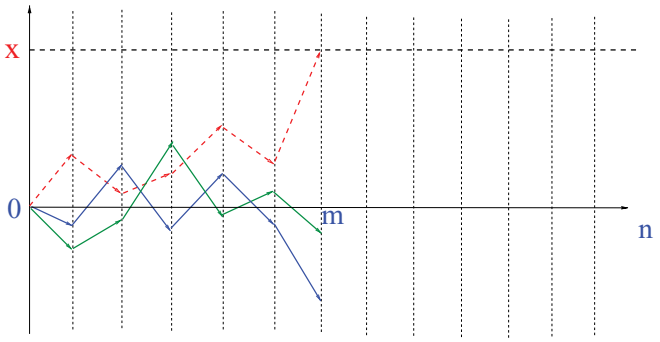

FIG. 2. (Color online) A record happens at step m with record value x for $N = 3$ walkers, all starting at the origin. This event corresponds to one walker (the dashed line) reaching the level x for the first time at step m while the other walkers stay below the level x up to step m.

where $q_m(x)$ denotes the probability that a single walker, starting at the origin, stays below the level x up to step m and $p_m(x)$ is the probability density that a single walker reaches the level x for the first time at step m, starting at the origin at step 0. Note that $q_m(x)$ can also be interpreted as the cumulative distribution of the maximum of a single walker's (say the ith walker) positions (starting at 0) up to step m

$$q_m(x) = \text{Prob}[\max_{0 \leqslant k \leqslant m} \{x_i(k)\} \leqslant x]. \tag{16}$$

The two quantities $p_m(x)$ and $q_m(x)$ can be reinterpreted in terms of slightly more familiar objects via the following observation. Note that by shifting the origin to the level x and using the time-reversal property of the trajectory of a single random walker, it is easy to see that $p_m(x)$ is just the probability density that a single walker, starting at the origin at step 0, reaches x at step m while staying positive at all intermediate steps. By a similar shift of the origin to level x and using the reflection symmetry of the trajectories around the origin, it is clear that $q_m(x)$ can be interpreted as the probability that a single walker, starting at an initial position $x > 0$ at step 0, stays positive (i.e., does not cross the origin) up to step m. This is then the familiar persistence or the survival probability of a single random walker [21]. In fact, both of these quantities $p_m(x)$ and $q_m(x)$ can be regarded as special cases of the more general restricted Green's function in the following sense. Consider a single random walker starting at position x at step 0 and evolving its position via successive uncorrelated jumps as in Eq. (1). Let $G_+(y,x,m)$ denote the probability density that the walker reaches $y > 0$ at step m, starting at $x > 0$ at step 0, while staying positive at all intermediate steps. The subscript $+$ denotes that it is indeed the restricted Green's function counting only the trajectories that reaches y at step m without crossing the origin in between. It is then clear from our discussion above that

$$p_m(x) = G_+(x,0,m), \tag{17}$$

$$q_m(x) = \int_0^\infty G_+(y,x,m)\, dy. \tag{18}$$

In the second line, the survival probability $q_m(x)$ is obtained from the restricted Green's function by integrating over all possible final positions of the walker. Note also, from Eqs. (17) and (18), that the survival probability starting exactly at the origin is

$$q_m(0) = \int_0^\infty p_m(x)\, dx. \tag{19}$$

Hence, if we know the restricted Green's function $G_+(y,x,m)$, we can, in principle, compute the two required quantities $p_m(x)$ and $q_m(x)$. Using the Markov evolution rule in Eq. (1), it is easy to see that the restricted Green's function $G_+(y,x,m)$ satisfies an integral equation in the semi-infinite domain [21],

$$G_+(y,x,m) = \int_0^\infty G_+(y',x,m-1)\, f(y-y')\, dy', \tag{20}$$

starting from the initial condition, $G_+(y,x,0) = \delta(y-x)$. Such integral equations over the semi-infinite domain are called Wiener-Hopf equations and are notoriously difficult

to solve for arbitrary kernel $f(z)$. Fortunately, for the case when $f(z)$ represents a continuous and symmetric probability density as in our case, one can obtain a closed form solution for the following generating function (rather its Laplace transform) [35]:

$$\int_0^\infty dy\, e^{-\lambda y} \int_0^\infty dx\, e^{-\lambda_0 x} \left[\sum_{m=0}^\infty G_+(y,x,m)\, s^m\right] = \tilde{G}(\lambda,\lambda_0,s) = \frac{\phi(s,\lambda)\,\phi(s,\lambda_0)}{\lambda+\lambda_0}, \tag{21}$$

where

$$\phi(s,\lambda) = \exp\left[-\frac{\lambda}{\pi}\int_0^\infty \frac{\ln[1-s\,\hat{f}(k)]}{\lambda^2+k^2}\,dk\right] \quad \text{and} \quad \hat{f}(k) = \int_{-\infty}^\infty f(x)\, e^{i\,k\,x}\,dx. \tag{22}$$

While the formula in Eq. (21) is explicit, it is rather cumbersome and one needs further work to extract the asymptotic behavior of $p_m(x)$ and $q_m(x)$ from this general expression. To make progress, one can first make a change of variable on the left hand side (lhs) $\lambda_0 x = z$ and then take the $\lambda_0 \to \infty$ limit. Using $\phi(s,\lambda_0 \to \infty) = 1$ and the definition $G_+(y,0,m) = p_m(y)$, and replacing y with x, we then obtain the following relation:

$$\sum_{m=0}^\infty s^m \int_0^\infty p_m(x)\, e^{-\lambda x}\,dx = \phi(s,\lambda), \tag{23}$$

where $\phi(s,\lambda)$ is given in Eq. (22). Similarly, by taking the $\lambda \to 0$ limit on the lhs of Eq. (21), using the asymptotic behavior in Eq. (A2) together with the definition $q_m(x) = \int_0^\infty G_+(y,x,m)\,dy$ and replacing λ_0 with λ, one obtains

$$\sum_{m=0}^\infty s^m \int_0^\infty q_m(x)\, e^{-\lambda x}\,dx = \frac{1}{\lambda\sqrt{1-s}}\,\phi(s,\lambda). \tag{24}$$

The formula in Eq. (24) is known in the literature as the celebrated Pollaczek-Spitzer formula [36,37] and has been used in a number of works to derive exact results on the maximum of a random jump process [23,38–40]. Interestingly, this formula has also been useful to compute the asymptotic behavior of the flux of particles to a spherical trap in 3D [24,41,42].

Let us also remark that by making a change of variable $\lambda x = y$ on the lhs of Eq. (24) and taking $\lambda \to \infty$, one obtains the rather amazing universal result for all m,

$$\sum_{m=0}^\infty q_m(0)\, s^m = \frac{1}{\sqrt{1-s}} \Longrightarrow q_m(0) = \binom{2m}{m}\frac{1}{2^{2m}}, \tag{25}$$

which is known as the Sparre Andersen theorem [33]. In particular, for large m, $q_m(0) \approx 1/\sqrt{\pi m}$. Note that for the case of a single walker $N = 1$, it follows from Eq. (15) that the record rate at step m is simply given by

$$r_{m,1} = \int_0^\infty p_m(x)\,dx = q_m(0) = \binom{2m}{m}\frac{1}{2^{2m}} \xrightarrow{m\to\infty} \frac{1}{\sqrt{\pi m}}, \tag{26}$$

where we have used Eq. (19) and the Sparre Andersen theorem (25). Thus, one obtains the rather surprising universal result for the $N = 1$ case: For all continuous and symmetric jump distributions, the mean number of records up to step n, $\langle R_{1,N}\rangle = \sum_{m=1}^n r_{m,1}$, is universal for all n and grows as $\sqrt{4n/\pi}$ for large n [28]. The universality in this case can thus be traced back to Sparre Andersen theorem.

In contrast, for $N > 1$, we need the full functions $p_m(x)$ and $q_m(x)$ to compute the record rate in Eq. (15). This is hard to compute explicitly for all m. However, one can make progress in computing the asymptotic behavior of the record rate $r_{m,N}$ for large m and large N, as we show below. It turns out that for large m, the integral in Eq. (15) is dominated by the asymptotic scaling behavior of the two functions $p_m(x)$ and $q_m(x)$ for large m and large x. To extract the scaling behavior of $p_m(x)$ and $q_m(x)$, our starting point would be the two equations (23) and (24). The next step is to use these asymptotic expressions in the main formula in Eq. (15) to determine the record rate $r_{m,N}$ at step m for large m and large N. The procedure to extract the asymptotics is somewhat subtle and algebraically cumbersome. To facilitate an easy reading of the paper, we relegate this algebraic procedure in the appendixes. Here we just use the main results from these appendixes and proceed to derive the results announced in Eqs. (6) and (10). The asymptotic behavior of $p_m(x)$ and $q_m(x)$ depend on whether $\sigma^2 = \int_{-\infty}^\infty \eta^2\, f(\eta)\,d\eta$ is finite or divergent. This gives rise to the two cases mentioned in Sec. II.

Case I (σ^2 finite). In this case, we show in Appendix A that in the scaling limit $x \to \infty$, $m \to \infty$ but keeping the ratio $x/\sqrt{m}$ fixed, $p_m(x)$ and $q_m(x)$ approach the following scaling behavior:

$$p_m(x) \to \frac{1}{\sqrt{2\sigma^2 m}}\, g_1\left(\frac{x}{\sqrt{2\,\sigma^2\, m}}\right), \tag{27}$$

where $g_1(z) = \frac{2}{\sqrt{\pi}}\, z\, e^{-z^2}$,

$$q_m(x) \to h_1\left(\frac{x}{\sqrt{2\,\sigma^2\, m}}\right), \tag{28}$$

where $h_1(z) = \mathrm{erf}(z)$,

where $\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-u^2}\,du$. Note that $dh_1(z)/dz = g_1(z)/z$.

Case II (σ^2 divergent). For the case when the Fourier transform of the jump distribution $\hat{f}(k)$ has the small k behavior as in Eq. (9), we show in Appendix A that in the scaling limit when $x \to \infty$, $m \to \infty$, but keeping the ratio $x/m^{1/\mu}$ fixed,

$$p_m(x) \to \frac{1}{m^{1/2+1/\mu}}\, g_2\left(\frac{x}{m^{1/\mu}}\right), \tag{29}$$

$$q_m(x) \to h_2\left(\frac{x}{m^{1/\mu}}\right). \tag{30}$$

While it is hard to obtain explicitly the full scaling functions $g_2(z)$ and $h_2(z)$ for all z, one can compute the large z asymptotic behavior and obtain

$$g_2(z) \underset{z\to\infty}{\sim} \frac{A_\mu}{z^{1+\mu}}, \tag{31}$$

$$h_2(z) \underset{z\to\infty}{\sim} 1 - \frac{B_\mu}{z^\mu}, \tag{32}$$

where the two amplitudes are

$$A_\mu = \frac{2\mu}{\sqrt{\pi}}\,\beta_\mu, \tag{33}$$

$$B_\mu = \beta_\mu, \tag{34}$$

with the constant β_μ having different expressions for $0 < \mu < 1$ and $1 \leqslant \mu < 2$,

$$\beta_\mu = \frac{a^\mu}{\pi\Gamma(1-\mu)}\int_0^\infty \frac{u^\mu}{1+u^2}\,du \quad \text{for} \quad 0 < \mu < 1, \tag{35}$$

$$\beta_\mu = \frac{2a^\mu}{\pi\Gamma(2-\mu)}\int_0^\infty \frac{u^\mu}{(1+u^2)^2}\,du \quad \text{for} \quad 1 \leqslant \mu < 2. \tag{36}$$

The expressions above [Eqs. (35) and (36)] can be written in a unified way for any $0 < \mu < 2$ as

$$\beta_\mu = \frac{a^\mu}{2\Gamma(1-\mu)\cos(\frac{\mu\pi}{2})} = \frac{a^\mu\Gamma(\mu)\sin(\frac{\mu\pi}{2})}{\pi}, \tag{37}$$

where, in the last equality, we have used $\Gamma(1-\mu)\Gamma(\mu) = \frac{\pi}{\sin\mu\pi}$. We recall that here we are considering discrete-time random walks (1). In the continuous-time random walk framework, with an exponential waiting time between jumps, the quantity $g_2(z)$ was studied in Ref. [43]. By performing an asymptotic analysis similar to the one presented in Appendix B, the authors showed that $g_2(z)$ behaves, for large z, like in Eq. (31) with the same exponent, albeit with a different amplitude. On the other hand, the exact asymptotic result (31), together with Eq. (37), can also be used to study the normalized probability distribution function (pdf) $\tilde{p}_m(x)$ of the position after m steps, with the condition that the walker stays positive at all intermediate steps, which was recently studied in Ref. [40]. It reads

$$\tilde{p}_m(x) = \frac{p_m(x)}{\int_0^\infty p_m(x)dx} \to \frac{1}{m^{1/\mu}}\tilde{g}_2\left(\frac{x}{m^{1\mu}}\right), \qquad \tilde{g}_2(z) = \sqrt{\pi}\,g_2(z), \tag{38}$$

where we have used the Sparre Andersen theorem $\int_0^\infty p_m(x)\,dx = q_m(0) \sim 1/\sqrt{\pi m}$ for large m. From Eq. (31), one obtains the large z behavior of $\tilde{g}_2(z)$ as

$$\tilde{g}_2(z) \underset{z\to\infty}{\sim} \frac{\tilde{A}_\mu}{z^{1+\mu}}, \quad \tilde{A}_\mu = \frac{2a^\mu\sin\left(\frac{\mu\pi}{2}\right)\Gamma(\mu+1)}{\pi}, \tag{39}$$

where we have used $\mu\Gamma(\mu) = \Gamma(\mu+1)$. On the other hand, if one considers the probability density function $P_m(x)$ of the position of a free Lévy random walk after m steps, with a jump distribution as in Eq. (9) after m steps, it assumes the scaling form, valid for large m, $P_m(x) \sim m^{-1/\mu}\,p(x/m^{1/\mu})$, where the asymptotic behavior is given by

$$p(z) \underset{z\to\infty}{\sim} \frac{C_\mu}{z^{1+\mu}}, \quad C_\mu = \frac{a^\mu\sin\left(\frac{\mu\pi}{2}\right)\Gamma(\mu+1)}{\pi}. \tag{40}$$

Therefore, the above result (39) establishes that $\tilde{A}_\mu = 2C_\mu$: This result was recently obtained analytically in perturbation theory for μ close to 2, $2-\mu \ll 1$, and conjectured to hold for any μ, on the basis of thorough numerical simulations [40]. Here this result is established exactly for any $\mu \in (0,2)$. While the large z behavior of $g_2(z)$ is the most relevant one for our study, we mention, for completeness, that its small z behavior was also studied in Refs. [43,44], yielding $g_2(z) \sim z^{\mu/2}$. Finally, we remark that the asymptotic behavior of $h_2(z)$ for large z has been computed in great detail recently in Ref. [39]; only the first two leading terms are presented in Eq. (32) here.

We are now ready to use these asymptotic behaviors of $p_m(x)$ and $q_m(x)$ in Eq. (15) to deduce the large m behavior of the record rate. Noting that for large m, the integral is dominated by the scaling regime, we substitute in Eq. (15) the scaling forms of $p_m(x)$ and $q_m(x)$ found in Eqs. (27), (28), (29), and (30). We then get, for large m,

$$r_{m,N} \approx \frac{N}{\sqrt{m}}\int_0^\infty g(z)\,[h(z)]^{N-1}\,dz, \tag{41}$$

where $g(z) = g_{1,2}(z)$ and $h(z) = h_{1,2}(z)$ depending on the two cases. So we notice that in all cases the record rate decreases as $m^{-1/2}$ for large m, albeit with different N-dependent prefactors in the two cases. Hence, the mean number of records $\langle R_{n,N}\rangle$ up to step n grows, for large n, as

$$\langle R_{n,N}\rangle \approx \alpha_N\sqrt{n}, \quad \text{where} \quad \alpha_N = 2N\int_0^\infty g(z)\,[h(z)]^{N-1}\,dz. \tag{42}$$

Next we estimate the constant α_N for large N. We first note that α_N in Eq. (42) can be expressed as

$$\alpha_N = 2\int_0^\infty \frac{g(z)}{h'(z)}\frac{d}{dz}\{[h(z)]^N\}\,dz, \tag{43}$$

where $h'(z) = dh/dz$. Noticing that $h(z)$ is an increasing function of z approaching 1 as $z \to \infty$, the dominant contribution to the integral for large N comes from the large z regime. Hence, we need to estimate how the ratio $g(z)/h'(z)$ behaves for large z. Let us consider the two cases separately.

Case I (σ^2 finite). In this case, we have explicit expressions for $g_1(z)$ and $h_1(z)$, respectively, in Eqs. (27) and (28). Hence, we get

$$\alpha_N = 2\int_0^\infty dz\,z\,\frac{d}{dz}[\mathrm{erf}(z)]^N \tag{44}$$

$$= \int_0^\infty dy\,y\,\frac{d}{dy}[\mathrm{erf}(y/2)]^N. \tag{45}$$

The rhs of Eq. (45) has a nice interpretation. Consider N i.i.d. positive RVs $\{y_1, y_2, \ldots, y_N\}$, each drawn from the distribution $p(y) = \frac{1}{\sqrt{\pi}}\,e^{-y^2/4}$ for $y \geqslant 0$ and $p(y) = 0$ for $y < 0$. Let M_N denote their maximum. Then the cumulative distribution function (cdf) of the maximum is given by

$$\mathrm{Prob}[M_N \leqslant y] = \left[\int_0^y p(y')\,dy'\right]^N = [\mathrm{erf}(y/2)]^N. \tag{46}$$

The probability density of the maximum is then given by $\frac{d}{dy}[\mathrm{erf}(y/2)]^N$. Hence, the rhs of Eq. (45) is just the average value $\langle M_N\rangle$ of the maximum. This gives us an identity for all N,

$$\alpha_N = \langle M_N\rangle. \tag{47}$$

From the standard extreme value analysis of i.i.d. variables [45], it is easy to show that to leading order for large N, $\langle M_N\rangle \approx 2\sqrt{\ln N}$, which then gives, via Eq. (42), the leading

asymptotic behavior of the mean record number,

$$\langle R_{n,N}\rangle \xrightarrow[N\to\infty]{n\to\infty} 2\sqrt{\ln N}\,\sqrt{n}. \tag{48}$$

Case II (σ^2 divergent). To evaluate α_N in Eq. (43), we note that when σ^2 is divergent, unlike in Case I, we do not have the full explicit form of the scaling functions $g_2(z)$ and $h_2(z)$. Hence, evaluation of α_N for all N is difficult. However, we can make progress for large N. As mentioned before, for large N, the dominant contribution to the integral in Eq. (43) comes from large z. For large z, using the asymptotic expressions in Eqs. (31) and (32), we get

$$\frac{g_2(z)}{h_2'(z)} \xrightarrow{z\to\infty} \frac{A_\mu}{\mu\, B_\mu} = \frac{2}{\sqrt{\pi}}, \tag{49}$$

where we have used Eqs. (33) and (34) for the expressions of A_μ and B_μ. We next substitute this asymptotic constant for the ratio $g_2(z)/h_2'(z)$ in the integral on the rhs of Eq. (43). The integral can then be performed trivially and we get, for large N,

$$\alpha_N \xrightarrow{N\to\infty} \frac{4}{\sqrt{\pi}}. \tag{50}$$

From Eq. (42) we then get for the mean record number

$$\langle R_{n,N}\rangle \xrightarrow[N\to\infty]{n\to\infty} \frac{4}{\sqrt{\pi}}\,\sqrt{n}. \tag{51}$$

In contrast to case I in Eq. (48), here the mean record number becomes independent of N for large N.

IV. THE DISTRIBUTION OF THE NUMBER OF RECORDS FOR FINITE σ^2

In the previous section, we performed a very precise study of the mean number of records $\langle R_{n,N}\rangle$ up to step n, in both cases where σ^2 is finite and divergent. In the present section, we investigate the full pdf of the record number $R_{n,N}$. However, we have been able to make analytical progress for the record number distribution only in case I, where σ^2 is finite, to which we restrict ourselves below.

The clue that leads to an analytical computation of the record number distribution in case I (finite σ^2) is actually already contained in the exact expression of the mean record number in Eqs. (42) and (45) which reads for large n, $\langle R_{n,N}\rangle \approx \alpha_N\sqrt{n}$. The quantity $\alpha_N\sqrt{n}$ has a precise physical meaning. To uncover this, let us first define a new RV,

$$Y_{n,N} = \max_{0\leqslant m\leqslant n} x_{\max}(m) = \max_{0\leqslant m\leqslant n}\ \max_{1\leqslant i\leqslant N}[x_i(m)], \tag{52}$$

which denotes the maximum of all the walkers *up to* step n. Let us first compute its average value. Let $m_i(n) = \max_{0\leqslant m\leqslant n}\{x_i(m)\}$ denote the maximum of the ith walker up to step n. Its cumulative distribution is given in Eq. (16), $\mathrm{Prob}\,[m_i(n)\leqslant x] = q_n(x)$. Taking average in Eq. (52) and exchanging the two maximums, it follows that

$$\langle Y_{n,N}\rangle = \langle \max_{1\leqslant i\leqslant N}[m_i(n)]\rangle. \tag{53}$$

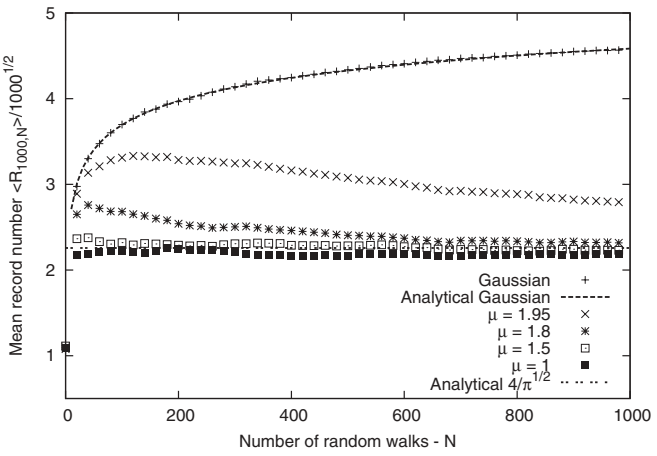

FIG. 3. Rescaled mean record number $\langle R_{n,N}\rangle/\sqrt{n}$ for a fixed $n = 1000$ plotted against the number of random walkers N. We performed simulations with jump distributions of the type $f(\eta) \sim |\eta|^{-\mu-1}$ and different $\mu = 1$, 1.5, 1.8, and 1.95 and for the Gaussian jump distribution with zero mean and $\sigma^2 = 1$. The Gaussian case is compared to our analytical finding for the finite σ^2 case [Eq. (48)], which is given by the dashed line. The horizontal dotted black line gives the analytical result for the infinite σ^2 regime [Eq. (51)]. With increasing N all $\langle R_{n,N}\rangle/\sqrt{n}$ with $\mu < 2$ approach this line of value $4/\sqrt{\pi}$.

Using the independence of the walkers and the fact that $m_i(n)$'s are identically distributed via the cdf $q_n(x)$, one gets

$$\langle Y_{n,N}\rangle = N\int_0^\infty x\, q_n'(x)\,[q_n(x)]^{N-1}\,dx. \tag{54}$$

Finally, substituting the asymptotic large n behavior of $q_n(x)$ from Eq. (28) on the rhs of Eq. (54) and comparing it to the expression of α_N in Eq. (45), we get

$$\langle Y_{n,N}\rangle \approx \frac{\sigma}{\sqrt{2}}\,\alpha_N\,\sqrt{n}. \tag{55}$$

Given the fact that $\langle R_{n,N}\rangle \approx \alpha_N\sqrt{n}$ for large n then leads to the relation

$$\lim_{n\to\infty}\frac{\langle R_{n,N}\rangle}{\sqrt{n}} = \lim_{n\to\infty}\frac{\sqrt{2}}{\sigma}\,\frac{\langle Y_{n,N}\rangle}{\sqrt{n}}. \tag{56}$$

Thus the mean record number is asymptotically identical to the mean global maximum of the N-walker process, up to the rescaling factor $\sigma/\sqrt{2}$. In this section, we would like to argue that actually the asymptotic proportionality between $R_{n,N}$ and $Y_{n,N}$ is true at the level of the full distribution, not just the average. In other words,

$$\lim_{n\to\infty}\frac{R_{n,N}}{\sqrt{n}} \equiv \lim_{n\to\infty}\frac{\sqrt{2}}{\sigma}\,\frac{Y_{n,N}}{\sqrt{n}}, \tag{57}$$

where $\equiv$ indicates that the RVs on both sides of Eq. (57) have the same probability distribution. While this identity can be rigorously proved for $N = 1$ as shown below, we have not been able to prove it rigorously for $N > 1$. However, our numerical simulations strongly favor this conjecture as demonstrated later in Sec. V A (see Fig. 4). If one accepts this relation, then one can compute the large n distribution of $R_{n,N}$ since it is easy to compute the distribution of $Y_{n,N}$ for large n.

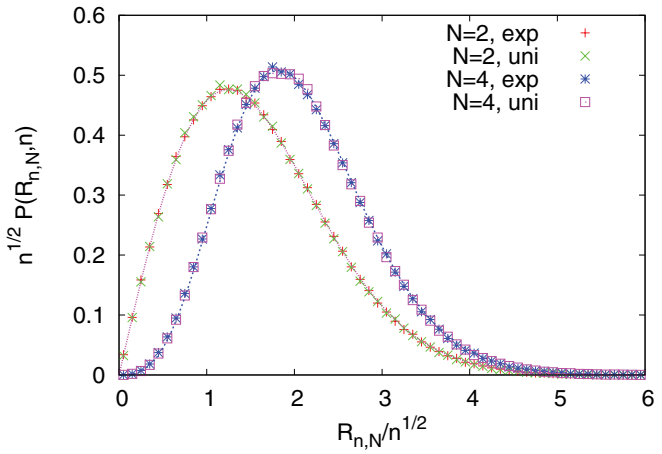

FIG. 4. (Color online) Scaled pdf $\sqrt{n}P(R_{n,N},n)$ as a function of $R_{n,N}/\sqrt{n}$ for $N = 2$ and $N = 4$ independent of random walks of length $n = 10^4$. For each value of N, we show the result for the case where the jumps are distributed exponentially ("exp") and uniformly between $-1/2$ and $1/2$ ("uni"). The dotted lines correspond to the result in Eq. (62) which we conjectured to be the exact one. There are no fitting parameters.

This is the strategy that we adopt in this section to compute the distribution of $R_{n,N}$ for large n.

Before computing the distribution of $Y_{n,N}$, let us just make one useful detour which also provides a hint to the guess in Eq. (57). It turns out that a proportionality relation similar to Eq. (57) can actually be rigorously proved for lattice random walks. Let us first demonstrate this below before we come back to the continuous jump density case.

Equivalence between $R_{n,N}$ and $Y_{n,N}$ for lattice random walks. Consider the simple lattice random walk model where a walker steps one step to the right or left with equal probability, starting at the origin. Consider N independent walkers, all starting at the origin at the same time. The position of the the ith walker again evolves by the Markov rule in Eq. (3), but the noise $\eta_i(m)$'s are now i.i.d. RVs with a distribution $f(\eta) = \frac{1}{2}\delta(\eta - 1) + \frac{1}{2}\delta(\eta + 1)$. Note that since $f(\eta)$ is not continuous, the results and theorems stated for continuous jump densities in the previous section are no longer valid here. Consider now the time evolution of the two random processes: the record process $R_{n,N}$ and the global maximum $Y_{n,N}$ defined in Eq. (52). At the next time step $(n + 1)$, if a new site on the positive axis is visited by any of the walkers for the first time, the process $Y_{n,N}$ increases by 1; otherwise its value remains unchanged. Whenever this event happens, that is, a new site on the positive side is visited for the first time, one also has a record event, that is, the process $R_{n,N}$ also increases by 1. Otherwise $R_{n,N}$ remains unchanged. Thus, the two random processes $Y_{n,N}$ and $R_{n,N}$ are completely locked with each other at all steps: Whenever one of them increases by unity at a given step the other does the same simultaneously and when one of them does not change, the other also remains unchanged. In other words, for every realization, we have, $Y_{n+1,N} - Y_{n,N} = R_{n+1,N} - R_{n,N}$. Now, initially all walkers start at the origin indicating $Y_{0,N} = 0$ while $R_{0,N} = 1$ since the initial point is counted as a record by convention. This allows us to write the following identity for all n and N

$$R_{n,N} = Y_{n,N} + 1 = \max_{0\leqslant m\leqslant n} \max_{1\leqslant i\leqslant N} [x_i(m)] + 1. \tag{58}$$

Thus, in this case one proves rigorously the identity $R_{n,N} \equiv Y_{n,N}$ for large n and the proportionality constant is just 1. Note that this cannot be considered as a special case of the proposed identity in Eq. (57) valid for continuous jump densities. The difference is in the proportionality constant. For example, using $\sigma = 1$ for lattice random walks, it is clear that there is an additional factor of $\sqrt{2}$ in Eq. (57), whereas the constant is actually 1. For this lattice walk model, one can compute the full distribution of $Y_{n,N}$ and hence that of $R_{n,N}$ exactly for all n as demonstrated in Appendix C.

Similarly, one can consider other lattice models, for example, with a jump density of the form $f(\eta) = [6\delta(\eta) + \delta(\eta + 2) + \delta(\eta - 2)]/8$, which is normalized and has $\sigma = 1$. In this case, following similar arguments as in the previous case, it is easy to see that the relationship between $R_{n,N}$ and $Y_{n,N}$ is now given by $R_{n,N} = Y_{n,N}/2 + 1$. Thus, for large n, $R_{n,N} \approx Y_{n,N}/2$. Thus, this case also cannot be considered as a special case of Eq. (57), as the proportionality constant does not match.

Back to the continuous jump densities. These two lattice examples discussed above are actually quite instructive. While we cannot use these results for the case of continuous jump densities, they serve as a useful pointer to the guess that for large n, $R_{n,N}$ simply becomes proportional to $Y_{n,N}$ even for continuous jump densities: $R_{n,N} \equiv cY_{n,N}$. The proportionality constant c can then be fixed from the exact relation between their averages in Eq. (56): $c = \sqrt{2}/\sigma$ and this then leads to the conjecture in Eq. (57).

Let us now see the consequence of this relation by first computing the distribution of $Y_{n,N}$ for large n. This can be done very simply using the independence of the walkers. According to Eq. (52),

$$Y_{n,N} = \max_{1\leqslant i\leqslant N} [m_i(n)], \tag{59}$$

where we recall that $m_i(n)$ is the maximum of the ith walker up to step n and is distributed via the cdf $\text{Prob}\,[m_i(n) \leqslant x] = q_n(x)$. Thus,

$$\text{Prob}[Y_{n,N} \leqslant y] = \prod_{i=1}^{N} \text{Prob}[m_i(n) \leqslant y] = [q_n(y)]^N. \tag{60}$$

Substituting the asymptotic large n scaling behavior of $q_n(y)$ from Eq. (28) and setting $y = (\sigma\sqrt{n/2})\,x$ gives

$$\text{Prob}\left[\frac{\sqrt{2}}{\sigma}\frac{Y_{n,N}}{\sqrt{n}} \leqslant x\right] \underset{n\to\infty}{\longrightarrow} \text{erf}\left(\frac{x}{2}\right). \tag{61}$$

Exploiting the relationship in Eq. (57) then provides an exact asymptotic result for the record number distribution

$$\text{Prob}\left(\frac{R_{n,N}}{\sqrt{n}} \leqslant x\right) \underset{n\to\infty}{\longrightarrow} \left[\text{erf}\left(\frac{x}{2}\right)\right]^N. \tag{62}$$

Let us point out that for $N = 1$, this result was already proved in Ref. [28]. But we believe, supported by our numerical simulations presented in Fig. 4 that the result in Eq. (62) is actually valid for all $N \geqslant 1$.

In Sec. V B we demonstrate that this conjecture is well supported by our numerical simulations. From Eq. (60), one can then use standard results for the extreme statistics of independent RVs [45] to obtain that for large N and large n, the scaled RV $R_{n,N}/\sqrt{n}$ (properly shifted and scaled) converges, in distribution, to the Gumbel distribution as announced in Eq. (7).

In the case of divergent σ^2, the two RVs $R_{n,N}$ and $Y_{n,N}$ do not seem to be related in any simple way. This is already evident from the result for the mean record number $\langle R_{n,N}\rangle$ in Eq. (51) for $0 \leqslant \mu < 2$. As $n \to \infty$, $\langle R_{n,N}\rangle \approx 4/\sqrt{\pi}\,\sqrt{n}$ where the prefactor does not depend on N for large N. In contrast, using standard extreme value statistics [45], it is easy to show that the mean value $\langle Y_{n,N}\rangle \sim (N\,n)^{1/\mu}$ for large n and N with $0 < \mu < 2$. This rather different asymptotic behavior of the mean thus already rules out any relationship between $R_{n,N}$ and $Y_{n,N}$ for case II. So, for this case, our result for the distribution of $R_{n,N}$ is only restricted to numerical simulations that are presented in the next section.

V. NUMERICAL SIMULATIONS

In this section we present the results of our numerical simulations of N independent random walks both with a finite σ^2 (case I) and with a divergent σ^2 (case II) and compare them with our analytical results. In Sec. V A we study the statistics of the record numbers (both its mean value and its full distribution). Since, at least in case I, the mean record number is strongly related to the expected maximum of the process we then analyze the evolution of the largest of the N random walkers. This is done in Sec. V B. We find that the statistics of the maximum significantly differs between the cases I and II. Finally, in Sec. V C we consider the correlations between individual record events in the two different cases and show that, at least asymptotically, these correlations are not different from the case of only one single random walker.

A. Statistics of the record numbers

Case I (σ^2 finite). In Fig. 3, we show our numerical results for $\langle R_{n,N}\rangle$ for σ^2 finite, which were obtained by a direct simulation of the jump process in Eq. (1) with $n = 10^4$ steps, with a Gaussian distribution of the jump variables $\eta_i(m)$'s (mean 0 and $\sigma^2 = 1$). These results have been obtained by averaging over 10^3 different realizations of the random walks. These data, on Fig. 3 are indexed by the label "Gaussian." Our numerical data show a very nice agreement with our analytical result obtained in Eq. (42), yielding $\langle R_{n,N}\rangle/\sqrt{n} = \alpha_N$. The large N behavior of α_N can be easily obtained by a saddle-point analysis, yielding

$$\frac{\langle R_{n,N}\rangle}{\sqrt{n}} = 2\sqrt{\ln N} - \frac{\ln(\ln N)}{2\sqrt{\ln N}} + O[(\ln N)^{-1/2}]\,. \tag{63}$$

It turns out that for $N \sim 1000$, the subleading corrections (63) are still sizable.

We have also computed numerically the distribution of the (scaled) record numbers $R_{n,N}/\sqrt{n}$ and compared it to our conjecture in Eq. (62). The results of this comparison, for different values of $N = 2,4$ are shown in Fig. 4, where the data were obtained by averaging over 5×10^4 realizations of independent random walks of $n = 10^4$ steps. The data, for two different continuous jump distributions (exponential and uniform) show a very nice agreement with our analytical prediction in Eq. (62). As mentioned above, one expects that this distribution will converge, for $N \to \infty$ to a Gumbel distribution (7), albeit with strong finite N effects.

Case II (σ^2 divergent). In Fig. 3, we show our numerical results for $\langle R_{n,N}\rangle$ for σ^2 divergent, obtained by a direct simulation of the random walk (1) where the distribution of $\eta_i(m)$'s has a power law tail $f(\eta) \sim |\eta|^{-1-\mu}$ with $\mu < 2$, and different values of μ. The data presented there were also obtained by averaging over 10^3 different realizations of random walks, with 10^4 steps. These data show that, in this case, $\langle R_{n,N}\rangle/\sqrt{n}$ approaches a constant value for fixed (and large) n and $N \to \infty$, which is fully consistent with the value of $4/\sqrt{\pi}$ obtained analytically in Eq. (51). The simulations also show how the speed of this convergence is modified when $\mu < 2$ is varied. While for small $\mu \ll 2$, $\langle R_{n,N}\rangle/\sqrt{n}$ approaches the universal limit value of $4/\sqrt{\pi}$ very quickly, we find a slower convergence for $2 - \mu \ll 1$.

The numerical computation of the distribution of the (scaled) record number $R_{n,N}/\sqrt{n}$ is of special interest because an analytical study of it, beyond the first moment, is still lacking. The two plots on Fig. 5 show our numerical results for this distribution, which were obtained by averaging over 10^4 independent random walks of length $n = 10^4$.

The left panel in Fig. 5 shows the rescaled distribution of $R_{n,N}/\sqrt{n}$ at a fixed time step $n = 10^3$ and $\mu = 1$. Apparently all curves for $N = 10$, 10^2, 10^3, and 10^4 collapse on one line. In the inset of the left panel we kept $n = 10^3$ and $N = 10^3$ fixed and varied μ: one can see that all the cdfs collapse. These results suggest that

$$\text{Prob}\left[\frac{R_{n,N}}{\sqrt{n}} \leqslant x\right] \underset{N\to\infty}{\overset{n\to\infty}{\longrightarrow}} F_2(x), \tag{64}$$

where the limiting distribution function $F_2(x)$ is independent of $\mu < 2$. We tried to guess the analytic form of $F_2(x)$ by comparing with several known continuous distributions that are defined for positive real numbers. We are certain that $F_2(x)$ is not a Gaussian distribution. By far the best results were obtained by fitting with a Weibull distribution

$$F_2(x) = 1 - \exp\left(-(\lambda x)^k\right), \tag{65}$$

with two free real parameters $\lambda > 0, k > 0$. Fitting with the least-squares method gives values of $\lambda \approx 0.8944 \pm 0.0003$ and $k = 2.558 \pm 0.003$. The right panel in Fig. 5 compares this fit with the distribution obtained from a simulation of $N = 10^4$ random walks of length $n = 10^4$. While the agreement, both for the cdf and the pdf is quite good, there are still some deviations between the two, particularly for small values close to zero. We are not sure if this difference is a finite N effect or if the real limit distribution of $R_{n,N}/\sqrt{n}$ for $N \to \infty$ effectively differs from a Weibull distribution.

B. Temporal evolution of two stochastic processes: The record number $R_{n,N}$ and the global maximum $Y_{n,N}$ up to step n

In the case of jump distributions with a finite second moment σ^2 (case I), we have shown that the mean $\langle Y_{n,N}\rangle$ of the maximum of all walkers up to step n and the mean

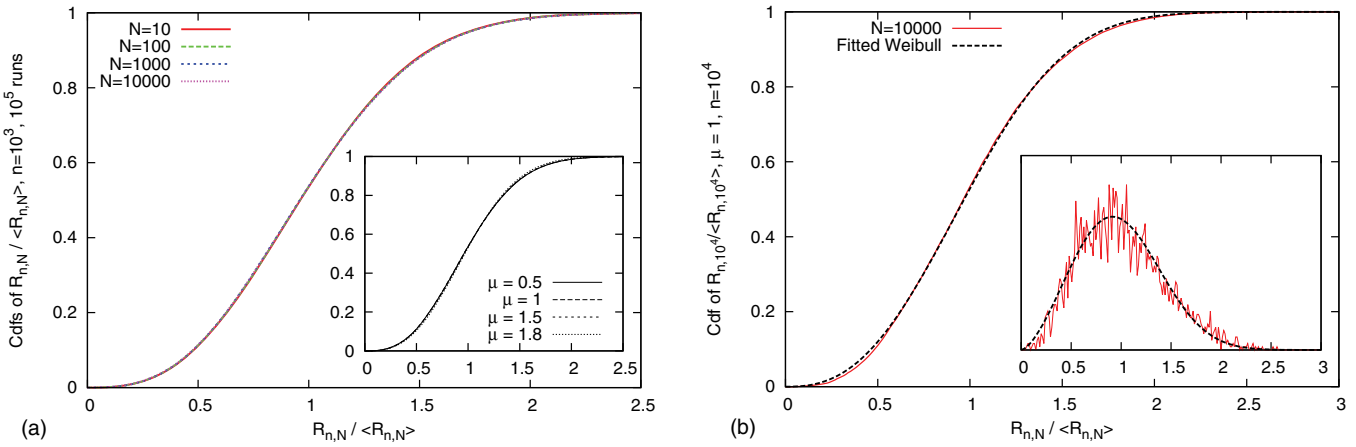

FIG. 5. (Color online) (a) Cumulative distribution function of the scaled variable $R_{n,N}/\sqrt{n}$ for the Lévy index $\mu = 1$, which approaches the universal distribution $F_2(x)$. The figure gives results for a fixed $n = 10^3$ and different N, for each N we performed 10^5 simulations. The inset gives simulations of the cdf for fixed $N = 10^3$ while Lévy index is varied. (b) The cdf for $\mu = 1$ and $N = 10^4$. We have fitted the data with a Weibull distribution as in Eq. (65), where the fitting parameters were $\lambda \approx 0.8944 \pm 0.0003$ and $k = 2.558 \pm 0.003$. The inset gives the corresponding curves for the pdfs.

record number $\langle R_{n,N} \rangle$ are proportional to each other in the limit $n \to \infty$; both grow with $\sqrt{n \ln N}$. In contrast, for Lévy walks with index $\mu < 2$ (case II) the relation between these two observables does not hold any more and the mean record number grows much slower than the maximum (see the discussion at the end of Sec. IV). To illustrate the similarities and differences in the growth rate of $R_{n,N}$ and $Y_{n,N}$ in the two cases (I and II), we compare their respective time evolution for four different samples in Fig. 6. On the left panel, we consider the Gaussian jump distribution with zero mean and $\sigma^2 = 1$ and we see that the process $R_{n,N}$ and $(\sqrt{2}/\sigma)\, Y_{n,N}$ become identical very quickly. Moreover, the two processes evolve almost in a deterministic fashion with growing n and hardly fluctuate from one sample to another. In contrast, on the right panel where we plot the two processes for $\mu = 1$, their behavior change drastically. First of all, the two processes $R_{n,N}$ and $Y_{n,N}$ do not seem to have relation to each other. While $R_{n,N}$ again evolves almost deterministically and in a self-averaging way, the trajectories of the process $Y_{n,N}$ differ strongly from one sample to another and $Y_{n,N}$ is clearly non-self-averaging. In particular, the process $Y_{n,N}$ can, like in a single Lévy flight, perform very large jumps exceeding its previous value by several orders of magnitude.

C. Correlations between record events

An important feature of the record statistics of a single random walk ($N = 1$) is its renewal property, which leads to the fact that each time after a record event, the record statistics is, in some sense, *reset*. After a record event the process evolves as a new process with the record value as its new origin.

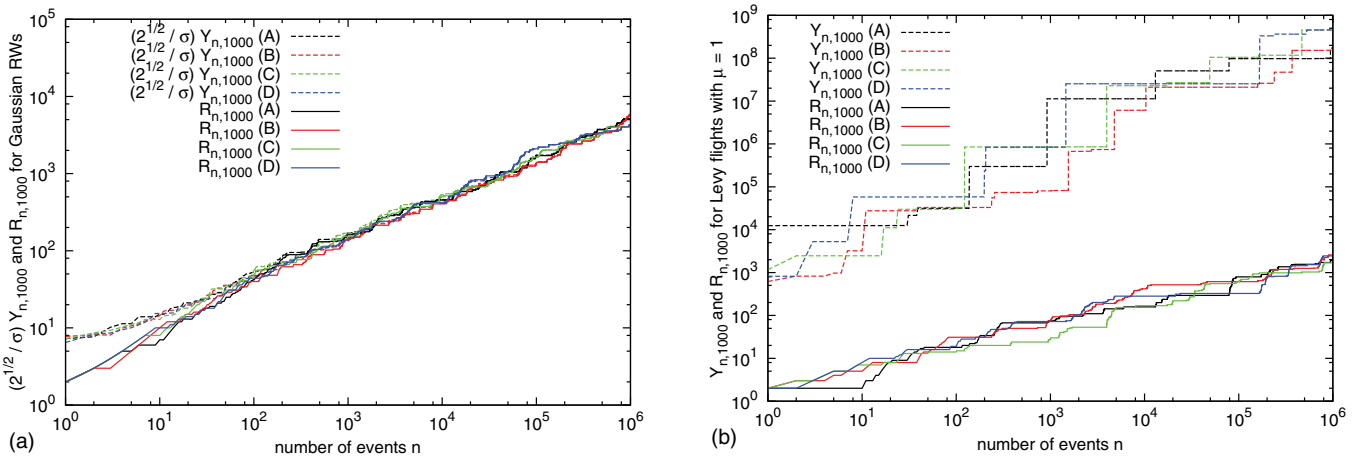

FIG. 6. (Color online) (a) Time evolution of the record number $R_{n,N}$ and the rescaled maximum value $(\sqrt{2}/\sigma)\, Y_{n,N}$ reached up to the nth step for four different realizations of $N = 1000$ random walks (labeled A, B, C, D) with Gaussian jump distribution (zero mean and $\sigma = 1$) (case I). Here, the results look rather deterministic and for $n \to \infty$, we find $(\sqrt{2}/\sigma)\, Y_{n,N} \approx R_{n,N}$ for every realization. (b) $R_{n,N}$ and $Y_{n,N}$ for four different realizations of $N = 1000$ Lévy flights (labeled A, B, C, D) with $\mu = 1$. The behavior of $Y_{n,N}$ for the Lévy flight is completely different from $R_{n,N}$: While $R_{n,N}$ is self-averaging, $Y_{n,N}$ fluctuates widely from one sample to another and is not self-averaging.

Therefore, it is very simple to give the pairwise correlations between record events. In fact, from the above argument, we have

$$\begin{aligned}&\text{Prob[rec. at } n-k \text{ and } n]\\ &\quad = \text{Prob[rec. at } n-k]\text{Prob[rec. at } k] = r_{n-k,1}r_{k,1}. \end{aligned} \tag{66}$$

Using the results from [28] this gives the following (exact) result for Prob [rec. at $n-k$ and n]:

$$\text{Prob[rec. at } n-k \text{ and } n] = \binom{2(n-k)}{n-k}\binom{2k}{k}2^{-2n}. \tag{67}$$

In the special case of $k=1$ we find Prob [rec. at $n-1$ and n] $=\frac{1}{2}r_{n-1}$. With this we find that the conditional probability of a second record directly following a record that just occurred is always given by

$$\text{Prob[rec. at } n|\text{rec. at } n-1] = \tfrac{1}{2}. \tag{68}$$

In our efforts to understand and compute the distribution of the record number $R_{n,N}$ for Lévy flights (case II), we considered the correlations between successive record events also for $N \gg 1$ random walks. If the correlations between successive record events in the large N limit would vanish, one could assume that the asymptotic distribution of $R_{n,N}$ approaches a Gaussian. However, we found that this is not the case. Figure 7 gives the behavior of Prob [rec. at $n-k$ and n] for the $N=1$ case, as well as for $N=10^3$ for random walks of the two cases I and II with Lévy indices $\mu=2$ and $\mu=1$. In all three cases Prob [rec. at $n-1$ and n] approaches $\frac{1}{2}$Prob [rec. at $n-1$] $=\frac{1}{2}r_{n,N}$ for large n, proving that for $n\to\infty$ the probability for a second record after an occurred one is just $1/2$. The inset of Fig. 7 also shows this behavior. Here, while for $N=1$ the conditional probability for a second record is always $1/2$, this value is only approached for larger n

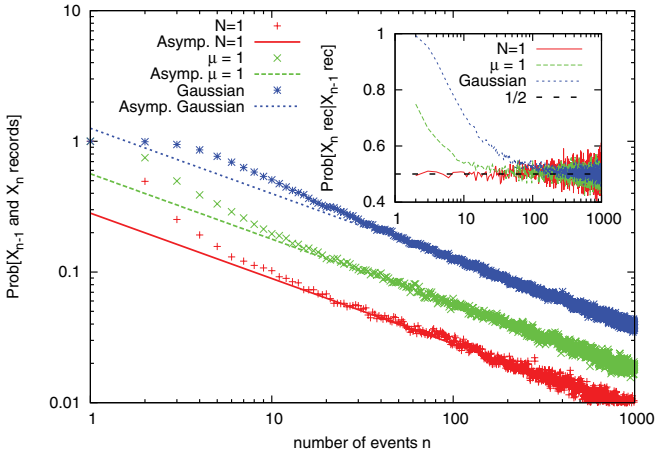

FIG. 7. (Color online) Probability for two successive records at times $n-1$ and n for a single random walker as well as N random walks with jump distributions of tail exponents $\mu=2$ (case 1) and $\mu=1$ (case 2). In all three cases this probability asymptotically approaches a value of $1/2$ times the probability for a record in the nth step. Therefore, for large n, the correlations between the record events in the $N\gg 1$ regime are the same as in the $N=1$ case. This is also shown in the inset, where we plotted the probability for a record in the nth step conditioned on a record in the previous one, which approaches $1/2$ in all three cases.

in the case of $N\gg 1$. For small n the conditional probability is larger.

VI. COMPARISON TO STOCK PRICES

The oldest application of the random walk model to analyze stock prices in markets goes back to Louis Bachelier (1900) who proposed a simple model based on arithmetic Brownian motion [46]. Later, Osborne (1959) proposed the geometric Brownian motion as a more accurate model for stock price evolutions [47]. This model was revisited and reanalyzed with modern stock data in Refs. [48,49]. In this model the stock prices perform a so-called geometric random walk and trends in the stocks are modeled by a linear drift in the logarithms of the stock prices. In [30] the record statistics of stocks in the Standard and Poor's 500 [34] (S&P 500) index were compared to the records in a random walk with a drift. The authors could show that on average the statistics of upper records in individually detrended stocks are in good agreement with the same statistics of a random walk with a symmetric jump distribution. In our notation they found that for a single stock from the S&P 500:

$$\langle R_{n,1}\rangle \approx \langle R_{n,1}\rangle^{(\text{GaussianRW})} \approx \frac{2}{\sqrt{\pi}}\sqrt{n}. \tag{69}$$

The lower records, however, showed a significant deviation from this model. These results were also found for daily data from other stock markets and seem to be relatively independent of the length of the time series n [50,51].

Here, we want to extend this analysis to the record statistics of N stocks. The question is to what degree the record statistics of N normalized and randomly chosen stocks from the S&P 500 can be compared to the record statistics of N independent random walks. As in [30], the observational data we used consisted of 366 stocks that remained in the S&P 500 index for the entire time-span from January 1990 to March 2009. Overall, we had data for 5000 consecutive trading days for each stock at our disposal. In [30] we found that it is useful to analyze these data over smaller intervals, on which one can then detrend the measurements. We decided to split up the 5000 trading days into 20 consecutive intervals of each 250 trading days, which is roughly one calendar year. In each of these intervals we considered the logarithms $X_i=\ln S_i/S_0$ of the stock prices S_i, where S_0 is the first trading day. The random walk model then suggests that these logarithms X_i perform a biased random walk that starts at the origin ($X_0=0$). Since our analytical theory presented in this paper only works for symmetric random walks we had to detrend the stocks. We subtracted an index-averaged linear trend obtained by linear regression from the X_i's in order to obtain symmetric random walkers. Finally, in order to make the stocks comparable to our model of N random walkers of the same jump distribution, we had to normalize the X_i's by dividing through the standard deviations of the respective individual jump distributions. After this detrending and normalization we can assume that the jump distributions in the individual time series have at least the same mean and the same variance, which should then be given by 0 and 1.

For a fixed N, we randomly selected subsets of size N from the 366 detrended and normalized stocks for each of the

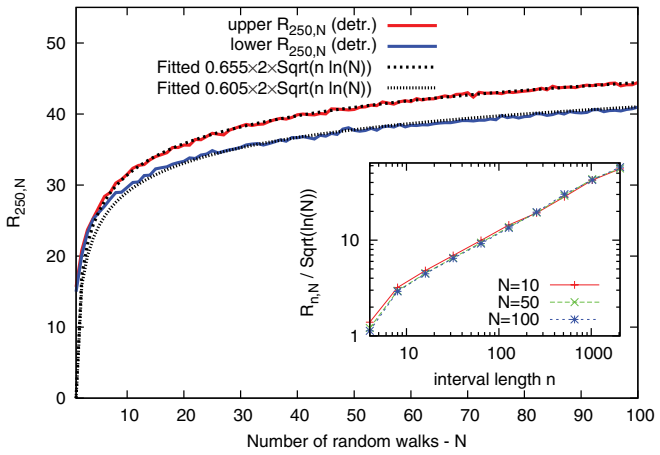

FIG. 8. (Color online) The averaged upper and lower record number after $n = 250$ trading days in the S&P 500 stock data. The 5000 trading days in [34] were subdivided into 20 intervals of 250 days and then linearly detrended in these intervals using the average linear trend of the index. Then we chose N stocks randomly out of the total number of 366 stocks and analyzed the evolution of the record number in this set. This random picking was repeated 10^4 times and the results were averaged to obtain the figure. The dashed lines give our analytical prediction for N Gaussian random walks multiplied by fitted prefactors. The inset gives the behavior of the $\langle R_{n,N}\rangle/\sqrt{\ln N}$ for different N plotted against the interval length n, confirming the proportionality $\langle R_{n,N}\rangle \propto \sqrt{\ln N}$.

20 intervals of length $n = 250$ and computed the evolution of the record number $R_{n,N}$ in these subsets. To get reliable statistics we average $R_{n,N}$ over 10^4 different subsets with a fixed N and also averaged over the 20 consecutive intervals. The resulting $\langle R_{n,N}\rangle$'s for the upper and the lower record number and N between 1 and 100 are given in Fig. 8. We find that both the curve for the upper and the curve for the lower mean record number are not in agreement with our theoretical prediction for the case of N Gaussian random walks given by $\langle R_{n,N}\rangle = 2\sqrt{n \ln N}$. However, Fig. 8 shows, that the $\langle R_{n,N}\rangle$ for the stocks increase with N. We also considered subsets of size $N > 100$ and found that for N closer to the maximal value of 366, $\langle R_{n,N}\rangle$ gets almost constant in N. While the increase of $\langle R_{n,N}\rangle$ for smaller N indicates that the statistics behave like N independent Gaussian random walks, the fact that it saturates for large N could indicate that they behave like a Lévy flight with Lévy index $\mu < 2$. We know, however, that the tail exponent of the daily returns $\ln S_i/S_{i-1}$ in the stock data is much larger than $\mu = 2$ and that they definitely do not perform a Lévy flight [50,52,53]. Much more likely is that the correlations [48,49,54,55] between the individual stocks play an important role. In addition, we observed that at least for $N < 100$, $\langle R_{n,N}\rangle$ grows proportional to $2\sqrt{n \ln N}$. One way to interpret this finding is the following: Since we know that the statistics of upper records of an individual stock is well described by a single Gaussian random walk and the jump distribution in the stock data has a finite variance, the origin of the inconsistency between the N stocks and our model of N independent random walks can only be in the interstock correlations in the market. When we assume that in N stocks only a smaller number of N^γ (with $\gamma \in \mathbb{R}^+$ and $\gamma < 1$) is effectively independent and that only these N^γ stocks contribute to the record statistics, the mean record number should be given by

$$\langle R_n, N\rangle = \langle R_n, N^\gamma\rangle^{(\text{Gaussian})} = 2\sqrt{\gamma\, n \ln N}, \tag{70}$$

and saturates if the value of $N^\gamma_{\max}$ is achieved, where $N_{\max}$ is the total number of stocks. In Fig. 8 we fitted curves of the form $\sqrt{\gamma_\pm}\, 2\sqrt{n \ln N}$ with $\sqrt{\gamma_+} \approx 0.655$ and $\sqrt{\gamma_-} \approx 0.605$ to the development of the upper and lower $\langle R_{n,N}\rangle$. The good agreement with the fitted curves and the data confirms our assumption. Apparently, the record statistics of N detrended and normalized stocks is the same as the one of N^γ independent Gaussian random walks. This finding is also confirmed by the inset in Fig. 8. There we plotted $\langle R_{n,N}\rangle/\sqrt{\ln N}$ for different interval length n and some different subset sizes N. The fact that all the lines collapse tells us that $\langle R_{n,N}\rangle/\sqrt{\ln N}$ is independent of N and therefore

$$\langle R_{n,N}\rangle \propto \sqrt{n \ln N}. \tag{71}$$

It is important to note that this finding is primarily an observation and the proposed "effective number of stocks" is just a possible interpretation. Unfortunately, up until now, we were not aware of a method to compute the record statistics of N correlated random walks directly. Even though it is well established that the prices of individual stocks are correlated [48,49,54,55], an effective number of stocks as introduced above was not computed before. Our model can therefore be seen as a method to compute such a number.

VII. CONCLUSION

In conclusion, we have presented a thorough analysis of the record statistics of N independent random walkers with continuous and symmetric jump distributions. For $N > 1$, we have found two distinct cases: the case where the variance of the jump distribution σ^2 is finite and the case where σ^2 does not exist (case II), as in the case of Lévy random walkers with index $0 < \mu < 2$. In the first case we have found that the mean record number behaves like $\langle R_{n,N}\rangle \approx 2\sqrt{\ln N}\sqrt{n}$ for $n, N \gg 1$ while in case II, $\langle R_{n,N}\rangle \approx \sqrt{4/\pi}\sqrt{n}$ for $n, N \gg 1$.

We have then argued that, in the first case, the full distribution of the scaled number of records $R_{n,N}/\sqrt{n}$ is given by the distribution of the maximum of N independent Brownian motions with diffusion coefficient $D = 1$. This statement was suggested by an exact result for lattice random walks and it was corroborated (i) by our exact calculation of the first moment $\langle R_{n,N}\rangle/\sqrt{n}$ valid for any value of N and (ii) by our numerical simulations. Of course it would be very interesting to obtain a proof of this result. From this connection with extreme value statistics, one thus expects that the distribution of $R_{n,N}/\sqrt{n}$ converges, for $N \to \infty$, to a Gumbel form (7). This connection between record statistics and extreme value statistics could also be useful to compute the record statistics in other models discussed in the Introduction, for instance, in the linear drift model [10,16,30]. In the case of Lévy random walkers, we have shown numerically that the full distribution of $\langle R_{n,N}\rangle/\sqrt{n}$ converges, when $N \to \infty$, to a limiting distribution $F_2(x)$ which is independent of μ. The exact computation of this universal distribution remains a challenging problem. Other interesting questions concern the extension of these results to include a linear drift [29] or to

the case of constrained Lévy walks, like Lévy bridges which were recently studied in the context of real space condensation phenomena [56]. Finally, the applications of our results to the record statistics of stock prices from the S&P 500 index suggest that, among a set of N stocks, only a smaller number, which scales like N^γ, with $0 < \gamma < 1$, are effectively independent. The record statistics of these N^γ stocks is then very similar to the statistics of N^γ independent random walkers. This idea might be useful for future investigations of the fluctuations of such ensemble of stock prices.

ACKNOWLEDGMENTS

We acknowledge support by ANR Grant No. 2011-BS04-013-01 WALKMAT and the Department of Business Administration and Finance at the University of Cologne for providing access to the stock data from the S&P 500. G.W. is grateful for the kind hospitality of the Laboratoire de Physique Théorique et Modèles Statistiques during the completion of this work and for the financial support provided by DFG within the Bonn Cologne Graduate School of Physics and Astronomy.

APPENDIX A: SCALING BEHAVIOR OF $p_m(x)$ AND $q_m(x)$ FOR FINITE σ^2

We start from Eqs. (23) and (24). When σ^2 is finite, by central limit theorem, the typical position of a walker after m steps scales as $m^{1/2}$ for large m. Hence, the natural scaling variable is $z = x/m^{1/2}$. Consider first Eq. (23) satisfied by $p_m(x)$. To extract the leading scaling function in the scaling limit $x \to \infty, m \to \infty$ but keeping $z = x/m^{1/2}$ fixed, we need to investigate $\phi(s,\lambda)$, given explicitly in Eq. (22), in the limit when $\lambda \to 0$, $s \to 1$ but keeping the ratio $\lambda/\sqrt{1-s}$ fixed. To extract the behavior of $\phi(s,\lambda)$ in this scaling limit, it is advantageous to work with an alternative expression of $\phi(s,\lambda)$ derived in Ref. [24] for finite σ^2

$$\phi(s,\lambda) = \frac{1}{[\sqrt{1-s} + \sigma\lambda\sqrt{s/2}]} \times \exp\left\{-\frac{\lambda}{\pi}\int_0^\infty \frac{dk}{\lambda^2+k^2}\,\ln\left[\frac{1-s\,\hat{f}(k)}{1-s+s\sigma^2k^2/2}\right]\right\}. \tag{A1}$$

This expression is more suitable for extracting the scaling limit. In the limit $\lambda \to 0$ and $s \to 1$, the expression inside the exponential in Eq. (A1) tends to 0 and hence, to leading order, we have

$$\phi(s,\lambda) \approx \frac{1}{[\sqrt{1-s} + \sigma\lambda\sqrt{s/2}]}. \tag{A2}$$

Inverting the Laplace transform with respect to λ [it has a simple pole at $\lambda = -\sqrt{2(1-s)}/\sigma$], one gets from Eq. (23)

$$\sum_{m=0}^\infty s^m\, p_m(x) \approx \frac{\sqrt{2}}{\sigma}\, e^{-\sqrt{2(1-s)}\,x/\sigma}. \tag{A3}$$

Setting $s = 1-p$ with $p \to 0$ in the scaling limit, the sum on the lhs of Eq. (A3) can be approximated, to leading order, by a continuous integral: $\sum_{m=0}^\infty s^m\, p_m(x) \approx \int_0^\infty p_m(x)\, e^{-pm} dm$ and we have

$$\int_0^\infty p_m(x)\, e^{-p\,m}\, dm \approx \frac{\sqrt{2}}{\sigma}\, e^{-\sqrt{2p}\,x/\sigma}. \tag{A4}$$

Next we need to invert the Laplace transform with respect to p. We use the explicit inversion formula, $\mathcal{L}^{-1}_{p\to m}[e^{-b\sqrt{p}}] = \frac{b}{2\sqrt{\pi}\,m^{3/2}}\,\exp[-b^2/4m]$. Applying this to Eq. (A4) gives, to leading order, in the scaling limit

$$p_m(x) \approx \frac{x}{\sigma^2\sqrt{\pi}\,m^{3/2}}\,\exp\left[-\frac{x^2}{2\sigma^2 m}\right], \tag{A5}$$

which can be reorganized in the scaling form

$$p_m(x) \to \frac{1}{\sqrt{2\sigma^2\,m}}\, g_1\left(\frac{x}{\sqrt{2\,\sigma^2\,n}}\right),$$
$$\text{where} \quad g_1(z) = \frac{2}{\sqrt{\pi}}\, z\, e^{-z^2}. \tag{A6}$$

Next we consider $q_m(x)$ given in Eq. (24). Following exactly the same procedure as in the case of $p_m(x)$ we find, in the scaling limit,

$$\int_0^\infty q_m(x)\, e^{-pm}\, dm \approx \frac{1}{p}[1 - e^{-\sqrt{2p}\,x/\sigma}]. \tag{A7}$$

Inverting the Laplace transform with respect to p upon using the explicit inversion formula, $\mathcal{L}^{-1}_{p\to m}[e^{-b\sqrt{p}}/p] = \text{erfc}(b/\sqrt{4m})$, we get, to leading order in the scaling limit

$$q_m(x) \approx 1 - \text{erfc}\left(\frac{x}{\sqrt{2\sigma^2\,m}}\right) = \text{erf}\left(\frac{x}{\sqrt{2\sigma^2\,m}}\right), \tag{A8}$$

which proves the result in Eq. (28).

APPENDIX B: SCALING BEHAVIOR OF $p_m(x)$ AND $q_m(x)$ FOR DIVERGENT σ^2

We consider jump distribution $f(\eta)$ such that its Fourier transform behaves, for small k, as $\hat{f}(k) \approx 1 - |ak|^\mu$, with $0 < \mu < 2$. In this case, the position of the walker after m steps, grows as $m^{1/\mu}$ for large m [23]. Hence, the natural scaling limit is $x \to \infty$, $m \to \infty$ with the ratio $x/m^{1/\mu}$ fixed. For $p_m(x)$, we expect a scaling form $p_m(x) \approx m^{-1/2-1/\mu} g_2(x/m^{1/\mu})$. The power of m outside the scaling function is chosen to ensure that $\int_0^\infty p_m(x)dx \sim m^{-1/2}$. This is needed since we know from Eq. (19) and the Sparre Andersen theorem in Eq. (25) that $\int_0^\infty p_m(x)dx = q_m(0) \sim 1/\sqrt{\pi m}$ for large m. Similarly, for $q_m(x)$, we expect a scaling form $q_m(x) \approx h_2(x/m^{1/\mu})$ in the scaling limit.

To extract the leading scaling functions $g_2(z)$ and $h_2(z)$, respectively, from Eqs. (23) and (24), we need to investigate the function $\phi(s,\lambda)$ in Eq. (22) in the corresponding scaling limit $\lambda \to 0$, $s \to 0$ but keeping the ratio $\lambda/(1-s)^{1/\mu}$ fixed. Fortunately, this was already done in Ref. [23] in a different context. Setting $s = 1-p$ with $p \to 0$, the leading behavior of $\phi(s,\lambda)$ in the scaling limit is given by (see Eqs. (43)–(47) of Ref. [23])

$$\phi(s,\lambda) \approx \frac{1}{\sqrt{p}}\,\exp\left\{-\frac{1}{\pi}\int_0^\infty \frac{du}{1+u^2}\,\ln\left[1 + \frac{1}{p}\,(a\,\lambda\,u)^\mu\right]\right\}. \tag{B1}$$

Let us first consider the function $p_m(x)$ in Eq. (23). We substitute the anticipated scaling form $p_m(x) = m^{-1/2-1/\mu} g_2(x/m^{1/\mu})$ on the lhs of Eq. (23). As before, setting $p = 1 - s$, we can replace, in the scaling limit, the sum over m by a continuous integral over m

$$\sum_{m=0}^{\infty} s^m \int_0^{\infty} p_m(x)\, e^{-\lambda x}\, dx \approx \int_0^{\infty}\int_0^{\infty} dx\, dm e^{-\lambda x - p\, m} m^{-1/2-1/\mu}\, g_2(x m^{-1/\mu}). \quad \text{(B2)}$$

We then make a change of variable $x\, m^{-1/\mu} = z$ and $p\, m = y$ to get

$$\sum_{m=0}^{\infty} s^m \int_0^{\infty} p_m(x)\, e^{-\lambda x}\, dx \approx \frac{1}{\sqrt{p}} \int_0^{\infty}\int_0^{\infty} dz\, dy\, g_2(z)\, y^{-1/2}\, e^{-(\lambda\, p^{-1/\mu})\, y^{1/\mu}\, z - y}. \quad \text{(B3)}$$

We next substitute Eq. (B3) on the lhs of Eq. (23) and Eq. (B1) on the rhs of Eq. (23). Writing the scaled variable as $\lambda\, p^{-1/\mu} = w$ and comparing lhs with the rhs, we see that the $1/\sqrt{p}$ cancels from both sides leaving us with

$$\int_0^{\infty} dz\, g_2(z) \int_0^{\infty} dy\, y^{-1/2}\, e^{-y}\, e^{-w\, y^{1/\mu}\, z} = \exp\left\{-\frac{1}{\pi}\int_0^{\infty} \frac{du}{1+u^2} \ln[1 + a^{\mu} w^{\mu} u^{\mu}]\right\} \equiv J_{\mu}(w)\,. \quad \text{(B4)}$$

Similarly, by substituting the anticipated scaling form $q_m(x) = h_2(x/m^{1/\mu})$ on the lhs of Eq. (24) and doing exactly the same series of manipulations, we get

$$\int_0^{\infty} dz\, h_2(z) \int_0^{\infty} dy\, y^{1/\mu}\, e^{-y}\, e^{-w\, y^{1/\mu}\, z} = \frac{1}{w}\, J_{\mu}(w), \quad \text{(B5)}$$

where $J_{\mu}(w)$ is defined in Eq. (B4).

For later purposes, it is further convenient to define a pair of Laplace transforms

$$\tilde{g}_2(\rho) = \int_0^{\infty} g_2(z)\, e^{-\rho\, z}\, dz, \quad \text{(B6)}$$

$$\tilde{h}_2(\rho) = \int_0^{\infty} h_2(z)\, e^{-\rho\, z}\, dz, \quad \text{(B7)}$$

in terms of which Eqs. (B4) and (B5) read

$$\int_0^{\infty} dy\, y^{-1/2}\, e^{-y}\, \tilde{g}_2(w\, y^{1/\mu}) = J_{\mu}(w), \quad \text{(B8)}$$

$$\int_0^{\infty} dy\, y^{1/\mu}\, e^{-y}\, \tilde{h}_2(w\, y^{1/\mu}) = \frac{1}{w}\, J_{\mu}(w). \quad \text{(B9)}$$

Equations (B4) and (B5) determine, in principle, the two scalings functions $g_2(z)$ and $h_2(z)$ for all z. In practice, it is hard to invert these two equations to obtain $g_2(z)$ and $h_2(z)$ for all z. However, it is possible to extract the large z asymptotics of these two functions by analyzing the leading singular behavior of $J_{\mu}(w)$ in Eq. (B4) as $w \to 0$. Clearly, it follows from the definition of $J_{\mu}(w)$ in Eq. (B4) that $J_{\mu}(0) = 1$. We are, however, interested in the leading singular correction term in $J_{\mu}(w)$ as $w \to 0$ which, as it turns out, depends on whether $0 < \mu < 1, 1 < \mu < 2$, or $\mu = 1$. Below, we consider the three cases separately.

1. The case $0 < \mu < 1$

We consider $J_{\mu}(w)$ in Eq. (B4) and compute the derivative $J'_{\mu}(w)$ as $w \to 0$. Simple computation shows that

$$J'_{\mu}(w) \xrightarrow{w \to 0} -\mu\, b_{\mu}\, w^{\mu-1}, \quad \text{where} \quad b_{\mu} = \frac{a^{\mu}}{\pi}\int_0^{\infty} \frac{u^{\mu}\, du}{1+u^2}. \quad \text{(B10)}$$

Note that the integral defining b_{μ} is convergent as $u \to \infty$ for $0 < \mu < 1$. Integrating over w and using $J_{\mu}(0) = 1$ we get the leading correction term as $w \to 0$,

$$J_{\mu}(w) \approx 1 - b_{\mu} w^{\mu} + \cdots, \quad \text{(B11)}$$

where b_{μ} is given in Eq. (B10).

Substituting this small w behavior of $J_{\mu}(w)$ on the rhs of Eq. (B8), it follows that to match the powers of w on both sides, the Laplace transform $\tilde{g}_2(\rho)$ must have the following small ρ behavior

$$\tilde{g}_2(\rho) \underset{\rho \to 0}{\sim} \frac{1}{\sqrt{\pi}} - \frac{2}{\sqrt{\pi}}\, b_{\mu}\, \rho^{\mu}. \quad \text{(B12)}$$

Using the classical Tauberian theorem (for a simple derivation see the Appendix A.2 of Ref. [57]), one immediately gets the following large z behavior of $g_2(z)$:

$$g_2(z) \underset{z \to \infty}{\sim} \frac{A_{\mu}}{z^{1+\mu}}, \quad \text{(B13)}$$

with the amplitude

$$A_{\mu} = \frac{2\mu}{\sqrt{\pi}}\, \frac{b_{\mu}}{\Gamma(1-\mu)} = \frac{2\mu}{\sqrt{\pi}}\, \beta_{\mu}, \quad \text{where}$$

$$\beta_{\mu} = \frac{b_{\mu}}{\Gamma(1-\mu)} = \frac{a^{\mu}}{\pi\Gamma(1-\mu)} \int_0^{\infty} \frac{u^{\mu}}{1+u^2}\, du. \quad \text{(B14)}$$

Similarly, substituting the small w behavior of $J_{\mu}(w)$ on the rhs of Eq. (B9) and matching powers of w on both sides, we get

$$\tilde{h}_2(\rho) \underset{\rho \to 0}{\sim} \frac{1}{\rho} - b_{\mu}\, \rho^{\mu-1}. \quad \text{(B15)}$$

Once again, using the Tauberian theorem of inversion, we get

$$h_2(z) \underset{z \to \infty}{\sim} 1 - \frac{B_{\mu}}{z^{\mu}}, \quad \text{(B16)}$$

with the amplitude

$$B_{\mu} = \frac{b_{\mu}}{\Gamma(1-\mu)} = \beta_{\mu}, \quad \text{(B17)}$$

where β_{μ} is given in Eq. (B14).

Finally, note that the ratio

$$\frac{A_{\mu}}{\mu\, B_{\mu}} = \frac{2}{\sqrt{\pi}} \quad \text{(B18)}$$

is universal in the sense that it is independent of $\mu \in (0,1)$ as well as on the scale factor a.

2. The case $1 < \mu < 2$

Unlike in the previous case, one finds that the first derivative of $J_\mu(w)$ at $w = 0$ is finite when $\mu \in [0,2]$ and is given by

$$\alpha_\mu = J'_\mu(0) = -\frac{a\,\mu}{\pi}\int_0^\infty \frac{z^{\mu-2}\,dz}{1+z^\mu}. \tag{B19}$$

Note that for $1 < \mu < 2$, the integral in Eq. (B19) is convergent as $z \to 0$. Thus, as $w \to 0$, $J_\mu(w) \to 1 - \alpha_\mu\, w$. To obtain the leading nonanalytic singular term, we need to compute the next term. By taking two derivatives with respect to w near $w = 0$ and then reintegrating back, we find the following leading singular behavior of $J_\mu(w)$ near $w = 0$:

$$\begin{aligned} &J_\mu(w) \approx 1 - \alpha_\mu\, w + c_\mu\, w^\mu + \cdots, \\ &\text{where} \quad c_\mu = \frac{2a^\mu}{\pi(\mu-1)}\int_0^\infty \frac{u^\mu\,du}{(1+u^2)^2}. \end{aligned} \tag{B20}$$

Substituting this small w behavior of $J_\mu(w)$ on the rhs of Eq. (B8) and matching powers of w on both sides we get

$$\tilde{g}_2(\rho) \underset{\rho\to 0}{\sim} \frac{1}{\sqrt{\pi}} - \frac{\alpha_\mu}{\Gamma(1/2+1/\mu)}\,\rho + \frac{2}{\sqrt{\pi}}\,c_\mu\,\rho^\mu, \tag{B21}$$

where α_μ and c_μ are given, respectively, in Eqs. (B19) and (B20). Again, inverting via the Tauberian theorem (see Ref. [57]), we get

$$g_2(z) \underset{z\to\infty}{\sim} \frac{A_\mu}{z^{1+\mu}}, \tag{B22}$$

with the amplitude

$$\begin{aligned} &A_\mu = \frac{2}{\sqrt{\pi}}\,\frac{\mu(\mu-1)c_\mu}{\Gamma(2-\mu)} = \frac{2\mu}{\sqrt{\pi}}\,\beta_\mu, \\ &\text{where} \quad \beta_\mu = \frac{2a^\mu}{\pi\Gamma(2-\mu)}\int_0^\infty \frac{u^\mu}{(1+u^2)^2}\,du\,. \end{aligned} \tag{B23}$$

Exactly in a similar way, we substitute the small w behavior of $J_\mu(w)$ on the rhs of Eq. (B9), match powers of w on both sides and find that

$$\tilde{h}_2(\rho) \underset{\rho\to 0}{\sim} \frac{1}{\rho} - \frac{\alpha_\mu}{\Gamma(1+2/\mu)} + c_\mu\,\rho^{\mu-1}, \tag{B24}$$

where α_μ and c_μ are defined in Eqs. (B19) and (B20). Inverting via the Tauberian theorem gives the desired result

$$h_2(z) \underset{z\to\infty}{\sim} 1 - \frac{B_\mu}{z^\mu}, \tag{B25}$$

with the amplitude

$$B_\mu = \frac{(\mu-1)c_\mu}{\Gamma(2-\mu)} = \beta_\mu, \tag{B26}$$

where β_μ is given in Eq. (B23).

In this case, also we note that the ratio

$$\frac{A_\mu}{\mu\,B_\mu} = \frac{2}{\sqrt{\pi}} \tag{B27}$$

is universal and does not depend explicitly on $1 < \mu < 2$ and a.

3. The case $\mu = 1$

In this case, from Eq. (B4),

$$\begin{aligned} &J_1(w) = \exp[-I_1(w)], \\ &\text{where} \quad I_1(w) = \frac{1}{\pi}\int_0^\infty \frac{du}{1+u^2}\,\ln(1+a\,w\,u). \end{aligned} \tag{B28}$$

Let us first derive the leading singular behavior of $I_1(w)$ as $w \to 0$. Making a change of variable $x = a\,w\,u$ in the integral we get

$$I_1(w) = \frac{aw}{\pi}\int_0^\infty \frac{dx}{x^2+a^2w^2}\,\ln(1+x). \tag{B29}$$

Next, we divide the integration range $[0,\infty)$ into two parts $[0,1]$ and $[1,\infty)$ and write $I_1(w) = Z_1(w) + Z_2(w)$. The second part $Z_2(w)$, that is, the integral over $[1,\infty]$ is a completely analytic function as $w \to 0$. Thus, the leading singular behavior of $I_1(w)$ as $w \to 0$ is contained only in the first part,

$$Z_1(w) = \frac{aw}{\pi}\int_0^1 \frac{dx}{x^2+a^2w^2}\,\ln(1+x). \tag{B30}$$

In this integral, we can now safely expand $\ln(1+x) = x - x^2/2 + x^3/3 + \cdots$ and perform the integral term by term. The leading singularity comes from the first term of this expansion,

$$\begin{aligned} Z_1(w) &\approx \frac{aw}{\pi}\int_0^1 \frac{x}{x^2+a^2w^2}\,dx \\ &= \frac{aw}{\pi}\ln\left[\frac{\sqrt{1+a^2w^2}}{aw}\right] \underset{w\to 0}{\sim} -\frac{a}{\pi}\,w\,\ln w, \end{aligned} \tag{B31}$$

which indicates, from Eq. (B28), that

$$J_1(w) \underset{w\to 0}{\sim} 1 + \frac{a}{\pi}\,w\,\ln w. \tag{B32}$$

Substituting this small w behavior of $J_1(w)$ on the rhs of Eq. (B8) and matching the leading behavior of w on both sides indicates that

$$\tilde{g}_2(\rho) \underset{\rho\to 0}{\sim} \frac{1}{\sqrt{\pi}} + \frac{2}{\sqrt{\pi}}\,\frac{a}{\pi}\,\rho\,\ln\rho. \tag{B33}$$

This indicates, using Tauberian inversion theorem (see Ref. [57]),

$$g_2(z) \underset{z\to\infty}{\sim} \frac{A_1}{z^2}, \quad \text{where} \quad A_1 = \frac{2}{\sqrt{\pi}}\,\frac{a}{\pi}. \tag{B34}$$

Similarly, substituting the small w behavior of $J_\mu(w)$ on the rhs of Eq. (B9) and matching leading behavior of w on both sides we get

$$\tilde{h}_2(\rho) \underset{\rho\to 0}{\sim} \frac{1}{\rho} + \frac{a}{\pi}\,\ln\rho, \tag{B35}$$

which, when inverted, provides the following large z behavior:

$$h_2(z) \underset{z\to\infty}{\sim} 1 - \frac{B_1}{z}, \quad \text{where} \quad B_1 = \frac{a}{\pi}. \tag{B36}$$

Finally, we notice that even for this marginal $\mu = 1$ case, the ratio $A_1/B_1 = 2/\sqrt{\pi}$ has the same value as in the other two cases, namely for $0 < \mu < 1$ and $1 < \mu < 2$.

Let us remark that if one puts $\mu = 1$ in the expression of β_μ in Eq. (36), we get $\beta_1 = a/\pi$. Correspondingly $A_1 = 2a/\pi^{3/2}$

from Eq. (33) and $B_1 = a/\pi$ from Eq. (34), we find that they are consistent respectively with A_1 in (B34) and B_1 in (B36). In other words, the final asymptotic results for $g_2(z)$ and $h_2(z)$ in the marginal case $\mu = 1$ are included in the range $\mu \in [1,2]$, even though the details for $\mu = 1$ are quite different, as it has logarithmic singularities.

APPENDIX C: DISTRIBUTION OF THE MAXIMUM AND THE NUMBER OF RECORDS OF A LATTICE RANDOM WALK

In this appendix we show how to compute the distribution of $Y_{n,N}$, the global maximum of N independent lattice random walks, all starting simultaneously at the origin. Once this is obtained, we can then easily compute the distribution of the record number $R_{n,N}$ by exploiting the exact relation (58).

More precisely, we consider N lattice random walks (RWs) starting at $x_i(0) = 0$, for $i = 1,2,\ldots,N$ and evolving as

$$x_i(m) = x_i(m-1) + \eta_i(m), \tag{C1}$$

where the noise $\eta_i(m)$'s are i.i.d. RVs with a distribution $f(\eta) = \frac{1}{2}\delta(\eta - 1) + \frac{1}{2}\delta(\eta + 1)$.

We first consider a single random walk, $N = 1$, and denote by $W(j,n)$ the number of lattice RW starting at $x_1(0) = 0$ and ending at j after n steps. One has

$$W(j,n) = \begin{cases} \binom{n}{k}, & 2k = n + j,\ n + j \text{ even}, \\ 0, & n + j \text{ odd}. \end{cases} \tag{C2}$$

To compute the cdf of the maximal displacement of N walkers we need to compute the number of walks, for a single walker $N = 1$, which stay strictly below a given value M. We thus denote, for $N = 1$, by $W_M(j,n)$ the number of walks which stay strictly below an integer M and end up in j after n steps. To do this we use the reflection principle, for example, the method of images: $W_M(j,n)$ can be obtained by subtracting from $W(j,n)$ the number of free walks that start at $x(0) = 2M$ and end at j after n steps. This yields

$$W_M(j,n) = \begin{cases} \binom{n}{k} - \binom{n}{k-m}, & 2k = n + j,\ n + j \text{ even}, \\ 0, & n + j \text{ odd}. \end{cases} \tag{C3}$$

The total number of walks $W_M(n)$ which start at $x_1(0) = 0$ and stay strictly below M after n steps are obtained by summing $W(j,n)$ in Eq. (C3) over the end point $j < M$. This yields

$$W_M(n) = \sum_{k=0}^{\lfloor \frac{n+M}{2} \rfloor} \left[\binom{n}{k} - \binom{n}{k-m} \right], \tag{C4}$$

where $\lfloor x \rfloor$ is the largest integer not greater than x. Therefore, one has

$$\text{Prob}\left[m_1(n) = \max_{0 \leqslant m \leqslant n} x_1(m) < M \right] = \frac{W_M(n)}{2^n} = \frac{1}{2^n} \sum_{k=0}^{\lfloor \frac{n+M}{2} \rfloor} \left[\binom{n}{k} - \binom{n}{k-m} \right]. \tag{C5}$$

Knowing the distribution of the maximum of one walker, one can then use the independence of the walkers to compute the cumulative distribution of $Y_{n,N} = \max_{1 \leqslant i \leqslant N}[m_i(n)]$ as in Eq. (59). We get

$$\text{Prob}\left[Y_{n,N} < M \right] = \left(\frac{W_M(n)}{2^n} \right)^N = \left\{ \frac{1}{2^n} \sum_{k=0}^{\lfloor \frac{n+M}{2} \rfloor} \left[\binom{n}{k} - \binom{n}{k-M} \right] \right\}^N, \tag{C6}$$

from which one gets

$$\text{Prob}[Y_{n,N} = M] = \frac{1}{2^{nN}} \{ [W_{M+1}(n)]^N - [W_M(n)]^N \}. \tag{C7}$$

Finally, using the identity (58), one obtains the record number distribution

$$P(R_{n,N} = M, n) = \frac{1}{2^{nN}} \left\{ \sum_{k=0}^{\lfloor \frac{n+M}{2} \rfloor} \left[\binom{n}{k} - \binom{n}{k-M} \right] \right\}^N - \frac{1}{2^{nN}} \left\{ \sum_{k=0}^{\lfloor \frac{n+M-1}{2} \rfloor} \left[\binom{n}{k} - \binom{n}{k-m+1} \right] \right\}^N, \tag{C8}$$

TABLE I. First values of $\langle R_{n,N} \rangle$ obtained from Eq. (C8).

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$
$N = 1$	1	$\frac{3}{2} = 1.5$	$\frac{7}{4} = 1.75$	2	$\frac{35}{16} = 2.187...$
$N = 2$	1	$\frac{7}{4} = 1.75$	$\frac{35}{16} = 2.187...$	$\frac{81}{32} = 2.531...$	$\frac{723}{256} = 2.824...$
$N = 3$	1	$\frac{15}{8} = 1.875$	$\frac{157}{64} = 2.453...$	$\frac{731}{256} = 2.855...$	$\frac{13145}{4096} = 3.209...$
$N = 4$	1	$\frac{31}{16} = 1.937...$	$\frac{671}{256} = 2.621...$	$\frac{6303}{2048} = 3.077...$	$\frac{227343}{65536} = 3.468...$

where $\lfloor x \rfloor$ is the largest integer not greater than x. For instance, for $N = 1$ one gets from Eq. (C8)

$$P(R_{n,1} = M, n) = \frac{1}{2^n} \binom{n}{\left(\lceil \frac{n+M-1}{2} \rceil\right)}, \tag{C9}$$

where $\lceil x \rceil$ is the smallest integer not less than x. We have checked that this formula for $N = 1$ (C9) yields back the result for the first moment $\langle R_{n,1} \rangle$ as obtained in Ref. [28]. From the above formula (C8) one can also compute $\langle R_{n,N} \rangle$, for instance with MATHEMATICA, although obtaining a simple closed form formula for it for $N > 1$ seems rather difficult. In Table I we have reported the first few values of $\langle R_{n,N} \rangle$ for $N = 1$ to $N = 4$.

Using the identity (58), one can also obtain the large n behavior of $R_{n,N}$. Indeed, in this limit, each rescaled ordinary random walk $x_i(\tau n)/\sqrt{n}$ converges, when $n \to \infty$, to a Brownian motion $B_{D=\frac{1}{2},i}(\tau)$ with a diffusion coefficient $D = 1/2$, on the unit time interval, $\tau \in [0,1]$. Therefore, from the above identity (58) one gets, in the limit $n \to \infty$

$$\frac{R_{n,N}}{\sqrt{n}} \to \tilde{M}_N = \max_{1 \leqslant i \leqslant N} \max_{0 \leqslant \tau \leqslant 1} [B_{D=\frac{1}{2},i}(\tau)]. \tag{C10}$$

Now, the distribution of $\max_{0 \leqslant \tau \leqslant 1} B_{D=\frac{1}{2},i}(\tau)$, that is, the maximum of a single Brownian motion (with diffusion constant $D = 1/2$) over a unit interval, is well known (see, e.g., in [21]),

$$\text{Prob}\,[\tilde{M}_1 \leqslant m] = \sqrt{\frac{2}{\pi}} \int_0^m \exp\left(-\frac{x^2}{2}\right) dx = \text{erf}\left(\frac{m}{\sqrt{2}}\right). \tag{C11}$$

Eq. (C10) demonstrates that in this case, $R_{n,N}/\sqrt{n}$ is distributed like the maximum of a collection of N i.i.d. positive RVs $\{z_1, z_2, \ldots, z_N\}$, each drawn from the distribution: $p(z) = \sqrt{\frac{2}{\pi}}\, e^{-z^2/2}$ for $z \geqslant 0$ and $p(z) = 0$ for $z < 0$. From Eq. (C11) one obtains also that $\langle R_{1,n} \rangle \approx \sqrt{2n/\pi}$, for $n \gg 1$, as obtained in Ref. [28], using a different method. More generally for any N one has

$$\lim_{n \to \infty} \text{Prob}\left[\frac{R_{n,N}}{\sqrt{n}} \leqslant m\right] = (\text{Prob}\,[\tilde{M}_1 \leqslant m])^N = \left[\text{erf}\left(\frac{m}{\sqrt{2}}\right)\right]^N. \tag{C12}$$

In the limit $N \to \infty$, the distribution of $R_{n,N}/\sqrt{n}$ (C12), properly shifted and scaled, will thus converge to the Gumbel distribution [58].

[1] D. Gembris, J. G. Taylor, and D. Suter, Nature (London) **417**, 506 (2002).
[2] D. Gembris, J. G. Taylor, and D. Suter, J. Appl. Stat. **34**, 529 (2007).
[3] J. Krug and K. Jain, Physica A **358**, 1 (2005).
[4] J. Krug, J. Stat. Mech. (2007) P07001.
[5] L. P. Oliveira, H. J. Jensen, M. Nicodemi, and P. Sibani, Phys. Rev. B **71**, 104526 (2005).
[6] P. Sibani, G. F. Rodriguez, and G. G. Kenning, Phys. Rev. B **74**, 224407 (2006).
[7] C. Godrèche and J. M. Luck, J. Stat. Mech. (2008) P11006.
[8] S. Redner and M. R. Petersen, Phys. Rev. E **74**, 061114 (2006).
[9] G. A. Meehl, C. Tebaldi, G. Walton, D. Easterling, and L. McDaniel, Geophys. Res. Lett. **36**, L23701 (2009).
[10] G. Wergen and J. Krug, Europhys. Lett. **92**, 30008 (2010).
[11] A. Anderson and A. Kostinski, J. Appl. Meteor. Climatol. **50**, 1859 (2011).
[12] F. G. Foster and A. Stuart, J. R. Stat. Soc. **16**, 1 (1954).
[13] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *Records* (Wiley, New York, 1998).
[14] V. B. Nevzorov, *Records: Mathematical Theory* (American Mathematical Society, Providence, RI, 2001).
[15] R. Ballerini and S. Resnick, J. Appl. Probab. **22**, 487 (1985).
[16] J. Franke, G. Wergen, and J. Krug, J. Stat. Mech. (2010) P10013.
[17] G. Wergen, J. Franke, and J. Krug, J. Stat. Phys. **144**, 1206 (2011).
[18] J. Franke, G. Wergen, and J. Krug, Phys. Rev. Lett. **108**, 064101 (2012).
[19] G. H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam, 1994).
[20] S. Redner *A Guide to First-passage Processes* (Cambridge University Press, Cambridge, 2001).
[21] S. N. Majumdar, Physica A **389**, 4299 (2010).
[22] E. G. Coffman and P. W. Shor, Algorithmica **9**, 253 (1993); E. G. Coffman, P. Flajolet, L. Flato, and M. Hofri, Prob. Eng. Inf. Sci. **12**, 373 (1998).
[23] A. Comtet and S. N. Majumdar, J. Stat. Mech. (2005) P06013.
[24] S. N. Majumdar, A. Comtet, and R. M. Ziff, J. Stat. Phys. **122**, 833 (2006).
[25] J. Franke and S. N. Majumdar, J. Stat. Mech. (2012) P05024.
[26] N. R. Moloney, K. Ozogany, and Z. Racz, Phys. Rev. E **84**, 061101 (2011).
[27] G. Schehr and S. N. Majumdar, Phys. Rev. Lett. **108**, 040601 (2012).
[28] S. N. Majumdar and R. M. Ziff, Phys. Rev. Lett. **101**, 050601 (2008).
[29] P. Le Doussal and K. J. Wiese, Phys. Rev. E **79**, 051105 (2009).
[30] G. Wergen, M. Bogner, and J. Krug, Phys. Rev. E **83**, 051109 (2011).
[31] S. Sabhapandit, Europhys. Lett. **94**, 20003 (2011).
[32] Y. Edery, A. Kostinski, and B. Berkowitz, Geophys. Res. Lett. **38**, L16403 (2011).
[33] E. Sparre Andersen, Math. Scand. **2**, 195 (1954).
[34] *Thomson Datastream Advance 4.0 SP4* (Thomson Reuters, New York, 2003).
[35] V. V. Ivanov, Astron. Astrophys. **286**, 328, (1994).
[36] F. Pollaczek, Comptes Rendus **234**, 2334 (1952).
[37] F. Spitzer, Duke Math. J. **24**, 327 (1957).

[38] D. A. Darling, Trans. Am. Math. Soc. **83**, 164 (1956).
[39] M. Kwasnicki, J. Malecki, and M. Ryznar, arXiv:1103.0935 [Annals of Probability (to be published)].
[40] R. García-García, A. Rosso, and G. Schehr, Phys. Rev. E **86**, 011101 (2012).
[41] R. M. Ziff, S. N. Majumdar, and A. Comtet, J. Phys. C: Condens. Matter **19**, 065102 (2007).
[42] R. M. Ziff, S. N. Majumdar, and A. Comtet, J. Chem. Phys. **130**, 204104 (2009).
[43] G. Zumofen and J. Klafter, Phys. Rev. E **51**, 2805 (1995).
[44] A. Zoia, A. Rosso, and S. N. Majumdar, Phys. Rev. Lett. **102**, 120602 (2009).
[45] E. J. Gumbel, *Statistics of Extremes* (Dover, Mineola, NY, 1958).
[46] L. Bachelier, Ann. Sci. Ec. Norm. Super. **3**, 17 (1900).
[47] M. F. M. Osborne, Oper. Res. **7**, 145 (1959).
[48] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics, Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2000).
[49] J. Voit, *The Statistical Mechanics of Financial Markets* (Springer, Berlin, 2001).
[50] G. Wergen (unpublished).
[51] M. Bogner, Rekordstatistik in Finanzdaten, unpublished thesis (2009).
[52] X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, Nature (London) **423**, 267 (2003).
[53] V. Plerou, P. Gopikrishnan, L. A. N. Amaral, M. Meyer, and H. E. Stanley, Phys. Rev. E **60**, 6519 (1999).
[54] J. P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge University Press, Cambridge, 2009).
[55] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley, Phys. Rev. Lett. **83**, 7 (1999).
[56] G. Schehr and S. N. Majumdar, J. Stat. Mech. (2010) P08005.
[57] M. R. Evans, S. N. Majumdar, and R. K. P. Zia, J. Stat. Phys. **123**, 357 (2006).
[58] H. Larralde, P. Trunfio, S. Havlin, H. E. Stanley, and G. H. Weiss, Phys. Rev. A **45**, 7128 (1992).