

Memory beyond memory in heart beating, a sign of a healthy physiological condition

P. Allegrini,¹ P. Grigolini,^{2,3,4} P. Hamilton,⁵ L. Palatella,³ and G. Raffaelli³

¹*Istituto di Linguistica Computazionale del Consiglio Nazionale delle Ricerche, Area della Ricerca di Pisa-S. Cataldo, Via Moruzzi 1, 56124, Ghezzano-Pisa, Italy*

²*Center for Nonlinear Science, University of North Texas, P.O. Box 311427, Denton, Texas, 76203-1427*

³*Dipartimento di Fisica dell'Università di Pisa and INFN Piazza Torricelli 2, 56127 Pisa, Italy*

⁴*Istituto di Biofisica del Consiglio Nazionale delle Ricerche, Area della Ricerca di Pisa-S. Cataldo, Via Moruzzi 1, 56124, Ghezzano-Pisa, Italy*

⁵*Center for Nonlinear Science, Texas Woman's University, P.O. Box 425498, Denton, Texas 76204*

(Received 25 July 2001; revised manuscript received 10 January 2002; published 11 April 2002)

We describe two types of memory and illustrate each using artificial and actual heartbeat data sets. The first type of memory, yielding anomalous diffusion, implies the inverse power-law nature of the waiting time distribution and the second the correlation among distinct times, and consequently also the occurrence of many pseudoevents, namely, not genuinely random events. Using the method of diffusion entropy analysis, we establish the scaling that would be determined by the real events alone. We prove that the heart beating of healthy patients reveals the existence of many more pseudoevents than in the patients with congestive heart failure.

DOI: 10.1103/PhysRevE.65.041926

PACS number(s): 87.19.Hh, 05.45.Tp, 05.40.Fb

The analysis of time series of physiological significance is currently done by many research groups using the paradigm of anomalous scaling [1]. This means that a time series is converted into a diffusion process described by the probability distribution $p(x,t)$ of the diffusing variable x , which is expected to fit the scaling property

$$p(x,t) = \frac{1}{t^\delta} F\left(\frac{x}{t^\delta}\right) \quad (1)$$

with the “degree of anomaly” being measured by the distance of the scaling parameter δ from the standard value 0.5. It is straightforward to prove that the Shannon entropy

$$S(t) = - \int_{-\infty}^{\infty} p(x,t) \ln[p(x,t)] dx \quad (2)$$

of a process fitting the scaling condition of Eq. (1) yields

$$S(t) = A + \delta \ln(t), \quad (3)$$

where A is a constant, whose explicit form is not relevant for the ensuing discussion. This result is immediately obtained by plugging Eq. (1) into Eq. (2). We thus find a method to evaluate the scaling parameter δ , more efficient than the calculation of the second moment of the probability distribution. Note that when the distribution density under study departs from the ordinary Gaussian case and the function $F(y)$ has slow tails with an inverse power-law nature [2,3] the second moment is a divergent quantity. This diverging quantity is made finite by the unavoidable statistical limitation. In this case, the second moment analysis would yield misleading results, determined by the statistical inaccuracy, while the method based on Eq. (3) yields correct results [2,3]. This method is denoted as diffusion entropy (DE) method.

The aim of this paper is to show that the entropy of a diffusion process generated by a physiological time series

according to the prescriptions of Refs. [2,3] yields a scaling exponent that depends only on genuinely random events. The time distances τ 's between nearest-neighbor events can be evaluated numerically and can be associated to a density distribution $\psi(\tau)$. In the case of physiological processes, the waiting time distribution is expected to be an inverse power law, with index μ . According to the theory of Ref. [3] there exists a simple relation between δ and μ . Thus, the experimental determination of $\psi(\tau)$ should yield the same information as the DE method. This is true when the events are genuinely random events. If the events are not genuinely random, and a memory, or time correlation exists, the DE method and the direct evaluation of $\psi(\tau)$ do not yield equivalent results, and the conflict between them is an important information on the physiological process under study.

Prior to the physiological applications, we will show this important property of the DE analysis with a dynamic model that generates genuinely random events. From now on for simplicity we shall denote the genuinely random events as *events*, while we shall refer to those that are not genuinely random as *pseudoevents*. The dynamic model, generating events, is given by

$$\dot{x} = \Phi(x) > 0, \quad (4)$$

where x denotes the coordinate of a particle, moving within the interval $I \equiv [0,1]$, from the left to the right, with times of arrival at $x=1$ determined by Eq. (4) and by the initial condition. When the particle reaches the right border of I , it is injected back to a new initial condition selected with uniform probability on I . Consequently, the times of arrival at $x=1$, t_1, \dots, t_i, \dots , represent events. It is straightforward to prove that the choice $\Phi(x) = kx^z$, with $z > 1$ and $k > 0$, yields for the waiting times $\tau_i \equiv t_i - t_{i-1}$ the following distribution density:

$$\psi(\tau) = (\mu - 1) \frac{T^{\mu-1}}{(T + \tau)^\mu}, \quad (5)$$

with $\mu = z/(z-1)$ and $T = (\mu - 1)/k$. Note that the mean waiting time $\langle \tau \rangle$ is determined by T through $\langle \tau \rangle = T/(\mu - 2)$.

Next, we convert the time series $\{\tau_i\}$ into a random walk. We select a rule [3] that makes the random walker move, always in the same direction and by a step of constant intensity, only when an event occurs. This means that the sequence $\{\tau_i\}$ is converted into a sequence of 0's and 1's as follows. We take the integer part of τ_1 , say n_1 , and we fill the first $n_1 - 1$ sites of the sequence with 0's. We assign the symbol 1 to the n_1 th site and we move to τ_2 , and so on. The resulting sequence, formed by attaching patches together in the same order as the sequence of τ_i , has a number of sites, say N , given by the sum of the integer parts of τ_i 's. Then, this sequence is converted into many trajectories of a given length l . A window of size l moves along the sequence and for any window position, the portion of the whole sequence spanned by the window is regarded as a single trajectory of length l . Thus, we obtain $N + 1 - l$ trajectories that are assumed to start from the origin, and are used to create a diffusion distribution, at time l . If there is scaling, the DE method detects it by means of Eq. (3) (for more details see Refs. [2,3]).

In short, this mathematical description corresponds to a random-walk process where the random walker remains in a state of rest, with a distribution of sojourn times given by Eq. (5). At the end of any sojourn the random walker makes an abrupt jump ahead. This kind of random walker has been studied in Ref. [3] by means of the continuous time random walk [4] and by means of the generalized central limit theorem [5]. The resulting diffusion process is shown by both methods to be an asymmetric Lévy distribution, fitting Eq. (1) and yielding for the scaling parameter δ , for $2 < \mu < 3$, the following value:

$$\delta = \frac{1}{\mu - 1} \quad (\delta = 0.5 \text{ if } \mu > 3). \quad (6)$$

This prediction is known [3,6] to be correct, only when an ordinary form of memory exists [7].

Let us discuss, first of all, what do we mean by ordinary memory. A Markov master equation, namely, a stochastic process without memory, is characterized by a waiting time distribution $\psi(\tau)$ with an exponential form. This implies that a marked deviation from the exponential condition, and thus from $\delta = 0.5$, is a signature of the presence of memory [7]. We refer to this memory as Type 1 memory. From an intuitive point of view we can say that this kind of memory means that the occurrence of the symbol 0 at a given time implies that at the next time, with a large probability, the same symbol will appear again. In a sense the system at later times keeps memory of earlier times. The exponential decay of the function $\psi(\tau)$ would imply a fast process of memory erasure, and the system would rapidly become memoryless. In the specific case where only Type 1 memory is present,

the direct evaluation of the waiting function $\psi(\tau)$ and the DE method provide equivalent results. In this case the times τ_i are uncorrelated and the deviation from ordinary statistical mechanics is only due to the inverse power-law nature of the waiting time distribution $\psi(\tau)$.

What about Type 2 memory? This rests on the correlation among the times τ_i , consequently implying that the sequence $\{\tau_i\}$ is a mixture of events and pseudoevents. In fact, to illustrate Type 2 memory, we discuss a dynamic model generating both events and pseudoevents. For this purpose let us consider a two-variable model. The equation referring to the first variable, is given by Eq. (4), and the one concerning the variable y , is given by

$$\dot{y} = \chi(y) > 0. \quad (7)$$

The variables x and y are the coordinates of two particles, both moving in the interval I , always from the left to the right. The initial conditions of the variable y are always chosen randomly, thereby corresponding to events. The initial conditions of x , on the contrary, are not always chosen randomly, but they are only when the variable y reaches the border at least once, during the sojourn of x within the range $[0, 1]$. Let us consider the sojourn time interval $[t_i, t_{i+1}]$. If in this time interval the variable y remains within the interval, without touching the right border, then we set $x(t_{i+1}) = x(t_i)$. This means that the next waiting time is equal to the preceding one, and consequently the time t_{i+2} , which might be predicted, represents a pseudoevent. Thus, the sequence $\{t_i\}$ reflects a mixture of events and pseudoevents. Let us consider the case where $\chi(y) = k'y^{z'}$ with $z' > 1$ and $k' > 0$, so as to produce the power index $\mu' = z'/(z' - 1)$, with $\mu' > 2$, a property of real events. Let us set the condition $\langle \tau \rangle_x \ll \langle \tau \rangle_y$. In this case, it is possible to prove with intuitive arguments that the waiting time distribution of x of Eq. (5) is changed into one much sharper than the original. The intuitive arguments run as follows. Let us consider the sojourn times of the particle with coordinate y . They will have a time duration of the order of $\langle \tau \rangle_y$. The unperturbed waiting times of the particle with coordinate y have a much shorter time duration. However, among these, those with a shorter length will be repeated for a larger number of times than those of a larger length. Let us consider a sojourn time with length comparable to $\langle \tau \rangle_y$. In this case there is a significant probability for the fast particle, with coordinate x , to reach the border after the arrival of the slow particle, with coordinate y . In this case, the initial condition for the fast particle will be selected randomly, and consequently the same sojourn time will not be repeated, if not by a fortuitous result of mere randomness. The same property will certainly apply, with much larger probability, to the sojourn times of length larger than $\langle \tau \rangle_y$. It is thus evident that in the perturbed distribution of waiting times for the particle of coordinate x , the statistical weight of the shorter times is increased, while the statistical weight of the larger times is decreased. As a result, the decay of the perturbed waiting time distribution will be much faster than the decay of the unperturbed distribution.

Let us consider the case where the unperturbed waiting time distribution $\psi(\tau)$ is characterized by $\mu > 3$ ($\mu = 5$, in

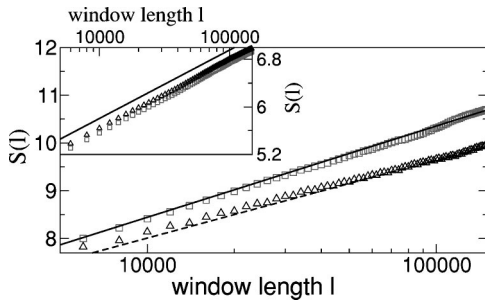


FIG. 1. DE for two-variables model as a function of the window length l . The squares correspond to $k'=0.018$, $z'=1.83$, and to a slope equal to 0.83, while the diamonds to $k'=0.011$, $z'=1.71$, and to a slope equal to 0.71. For both curves $k=0.4$, $z=1.25$. In the inset: the same curves after shuffling, the straight line slope is 0.5.

the case of Fig. 1). The perturbed waiting time distribution, to be identified with that experimentally observed, is even faster. This is so because of the earlier described enhancement of the statistical weight of the short waiting times and reduction of that of long waiting times. On the other hand, as we have earlier observed [see Eq. (6)] when there is no memory, the distribution with $\mu > 3$ falls in the Gauss basin of attraction. Let us adopt now an experimental point of view. Let us look at the perturbed distribution pretending that we do not know that it is affected by the hidden action of the variable y . We perceive a waiting time distribution whose decay is even faster than the unperturbed one, which sits in the Gauss basin of attraction. We pretend that we do not know anything about the presence of Type 2 memory. Thus, we make the prediction that Eq. (6) applies, thereby yielding the scaling $\delta=0.5$. If we make this plausible conjecture, then we find that the experimental result of Fig. 1 is striking and apparently incomprehensible. In fact, we see that the DE now yields $\delta=1/(\mu'-1)$, quite different from the prescription of Eq. (6) that would produce $\delta=0.5$. This is so because the experimental $\psi(\tau)$ is not simply a reflection of real events but is determined by pseudoevents as well. The breakdown of Eq. (6) is a manifestation of Type 2 memory, referred to by us as *memory beyond memory*. In fact, the existence of pseudoevents implies correlation among different times of the series $\{\tau_i\}$, and thus a memory of earlier events. This result is so surprising that we need to support it with compelling arguments. The Type 2 memory is a consequence of correlation among the waiting times of the particle x . We note that shuffling the sequence $\{\tau_i\}$ does not have any effect on the shape of the waiting time distribution. Shuffling has the effect of annihilating Type 2 memory. At the end of this shuffling process, we can correctly state that, being in the Gauss basin of attraction, the system is forced to obey the prescription of Eq. (6), thereby resulting in ordinary scaling. The inset of Fig. 1 shows that, as expected, shuffling the order of the corresponding patches has the effect of yielding $\delta=0.5$, as the experimental $\psi(\tau)$ implies. The scaling detected by the DE method does not depend on the pseudoevents, but only on the hidden events, and thus on a time distribution, which cannot be experimentally detected, slower than $\psi(\tau)$. This hidden waiting time distribution of waiting

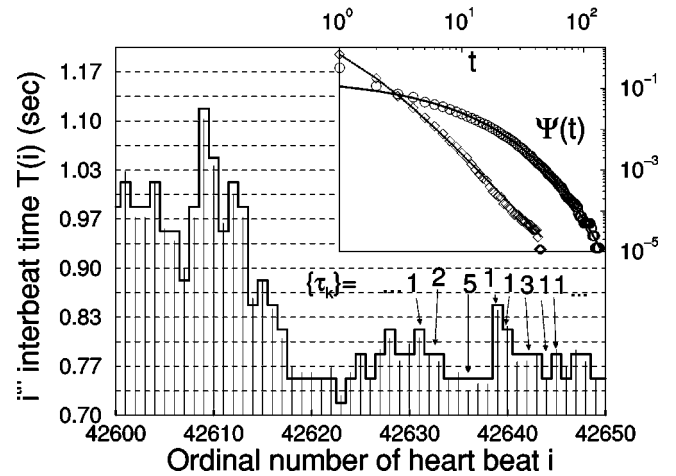


FIG. 2. The interbeat time $T(i)$ as a function of the number of beats, i . The thick line denotes the trajectory corresponding to the coarse graining given by $\Delta T=1/30$ sec. The vertical lines denote the height of the original data. The arrows and the integer labels illustrates how the sequence of τ_k 's is generated. Inset: Survival probabilities. The circles denote the c.h.f. patients and the corresponding fitting function is $\Psi(t)=0.19 \exp[-(t/3.1)^{0.6}]$. The diamonds denote the healthy patient and the corresponding fitting function is $\Psi(t)=5.71/(0.93+t)^{3.25}$.

times is responsible for the anomalous scaling, $\delta > 0.5$, of the two unshuffled curves of Fig. 1,

Let us show now that the heart beating is a process where the memory beyond memory effect is important. Let us apply the earlier perspective to the real data taken from [8]. We apply our technique to 33 long-time electrocardiogram (ECG) records (about 20 hours each), 18 healthy and 15 with congestive heart failure (c.h.f.). Following Ref. [8], we refer to all the ECG records of the *MIT-BIH Normal Sinus Rhythm Database* and of the *BIDMC Congestive Heart Failure Database*, the healthy and the c.h.f. patients, respectively.

The data under study are time series of the kind of that illustrated in Fig. 2, where the length of the vertical lines expresses $T(i)=t_i-t_{i-1}$ as a function of the integer number i . The letter i denotes the i th heartbeating of an electrocardiogram, and t_i is the time at which the R wave of this heartbeat occurs.

We make these data suitable for the illustration of the *memory beyond memory* effect as follows. We adopt a procedure illustrated with the help of Fig. 2. The vertical axis, concerning the variable $T(i)$, is divided into many cells of a given size ΔT . Thus the $[T(i), i]$ plane is divided into many horizontal strips with a constant width equal to ΔT . This coarse-graining prescription yields the thick line of Fig. 2. The curve corresponds to many horizontal intervals separated by vertical up and down jumps. The widths of these horizontal intervals define a sequence of numbers τ_i that is the object of our statistical analysis. To make this analysis as efficient as possible we have to make a proper choice of the value of ΔT , since an excessively small value would produce too many pseudoevents and an excessively large would yield poor statistics. The results of our statistical analysis were proven to be insensitive to changing ΔT over the relatively

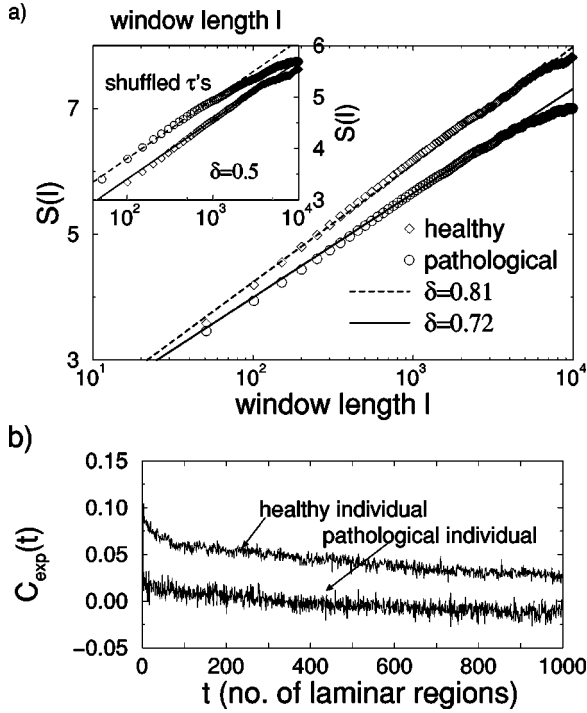


FIG. 3. (a) The DE as a function of the window length l . The inset illustrates the effect of shuffling, the two lines correspond to $\delta=0.5$. (b) The correlation function $C_{exp}(t)$ as a function of t (number of laminar regions) on a healthy and on a c.h.f. individual.

large range from 0.1 to 0.02 sec. We assign to ΔT an intermediate value, namely, $\Delta T=1/30$ sec.

The events under study refer to the jumps from one strip to another. To assess whether these jumps are events or pseudoevents, we have to compare the waiting time distribution $\psi(\tau)$ to the scaling detected by means of the DE method. For the sake of statistical accuracy we decided to evaluate the probability of finding waiting times larger than a given value t . This is the function $\Psi(t)$ defined by $\Psi(t) \equiv \int_t^\infty d\tau \psi(\tau)$. The results illustrated in the inset of Fig. 2 imply the Brownian scaling $\delta=0.5$. In fact, the function $\Psi(t)$ of the heart failure subjects is a stretched exponential and the healthy subjects yield the power index 3.25 that corresponds to the waiting time distribution density $\psi(\tau)$ with $\mu=4.25$. The same Brownian condition applies to all the subjects. On the contrary, the DE method yields for the healthy subjects a mean value $\delta=0.82 \pm 0.04$ and for the heart failure subjects $\delta=0.71 \pm 0.06$. It is interesting to notice that Fig. 3 refers to the same subjects as those of the inset of Fig. 2, and yields for the healthy $\mu'=2.17$ and for the heart failure subjects $\mu'=2.4$. If we shuffle the numbers of the sequence τ_i we recover $\delta=0.5$, a fact proving that the *memory beyond memory* effect is a genuine property of heartbeat.

The additional memory is confirmed by the numerical evaluation of the normalized correlation function of the variable $\tau_i - \langle \tau_i \rangle$, denoted by $C_{exp}(t)$, where the symbol t is the continuous approximation of the discrete patch label i . The two-variable model that we are using to explain the *memory beyond memory* effect would yield

$$C(t) \propto 1/t^\beta. \quad (8)$$

The index β in that case would be a complicated function of the four parameters involved by the two-variable model. This means that the DE is a more efficient memory detector and much less ambiguous than the correlation function. The DE selects the really random events from the distribution of times described by the arbitrary $\psi(\tau)$, yielding a nonarbitrary distribution with a unique μ' . The correlation function $C(t)$, on the contrary, depends on the details of the model, but does not afford an easy way to define them. For the main purpose of this paper it is enough to point out that the form of the correlation function $C_{exp}(t)$ is

$$C_{exp}(t) = (1 - \epsilon^2)W(t) + \epsilon^2 C(t). \quad (9)$$

Here $W(t)$ denotes a function dropping from 1 to 0 in one time step, while the function $C(t)$, with the asymptotic form of Eq. (8), is continuous for $t \rightarrow 0$. We account for the structure of Eq. (9) as follows. The sequence $\{\tau_i\}$ is generated by the joint use of two models.

The first is the model of Eq. (4) with no additional variables and no *memory beyond memory* property, the second is the model with two variables. These two models generate two independent sequences $\{\tau_i\}$. To any index i we assign, with probability ϵ , the value provided by the model with additional memory, and, with probability $1 - \epsilon$, the value provided by the model with only one variable. This model is reminiscent of one adopted to account for the statistical properties of DNA sequences [9]. The function $C_{exp}(t)$ in one step drops from the value $C_{exp}(0)=1$ to the value $C_{exp}(1) = \epsilon^2 C(1) \approx \epsilon^2$, thereby allowing us to derive ϵ from the experimental correlation function at $t=1$.

In conclusion, the meaning of the parameter ϵ is as follows. The value $\epsilon=1$ would imply that the heart beating is described entirely by a model with two variables, x and y . In other words, the larger ϵ the larger the weight of the *memory beyond memory* effect. The parameter δ is connected to the time distance between two nearest-neighbor real events. If this time distribution is exponential, there is no Type 1 memory, as earlier observed. As μ' becomes closer to $\mu'=2$, this conventional memory becomes stronger. Thus, to establish a more intuitive understanding of what happens to memory, regardless of whether it is Type 1 or Type 2, let us adopt the following perspective. The condition of highest memory corresponds to $\epsilon=1$ and $\delta=1$. This would mean that the heart beating depends only on the *memory beyond memory* model, and, that, at the same time, $\mu'=2$. The opposite case, of complete absence of memory, implies $\epsilon=0$ and $\mu' \rightarrow \infty$. This would mean that the heart beating is very well modeled by the one-variable model of Eq. (4), with an exponential distribution of waiting times, in other words, without either type of memory. This leads us to express the distribution of patients in the (δ, ϵ^2) plan of Fig. 4. We note the surprising result that all the healthy subjects and all the heart-failure subjects are contained in the top-right region and in the bottom-left one, respectively. We also notice that all the healthy subjects, except two, are localized within a small portion of the top-right region of the graph, not far from the border with the heart-failure region. We have the impression that this reflects the fact that the healthy function

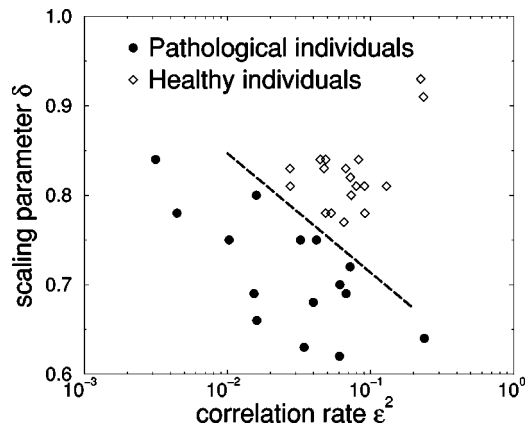


FIG. 4. Values of the scaling parameter δ and of ϵ^2 for the healthy and c.h.f. individuals of this analysis.

of the heartbeating system depends on a proper balance of memory and randomness that the analysis of this paper makes apparent. The distribution of the heart-failure subjects within the bottom-left region is much broader. It would be desirable to have at our disposal the patient survival prob-

ability as, for example, in Ref. [10], to assess whether a physiological reaction to the c.h.f. pathology, responsible for the bottom-left broadened distribution, plays a negative or a positive role. The advocates of the second possibility might argue that higher randomness and broader distribution reflect an effort of the perturbed heartbeating system to explore all possible states to recover the lost function. An interesting issue to discuss would be the connection between these results and those of Refs. [11,12]. This is left as a subject for future research work.

As a final conclusion, we want to recognize that the main result of this paper does not have yet any importance from the diagnostic point of view. The paper does not yield a criterion to distinguish pathological from healthy condition of medical interest. Rather, it establishes that the main difference between healthy and c.h.f. patients is that in the latter group of patients the extent of memory beyond memory is much reduced compared to the former group. Further research work is necessary to establish the physiological significance of this result, and this, in turn, might help the diagnostic purpose. The discovery of the memory beyond effect is, on the other hand, by itself very interesting and very surprising.

-
- [1] M. F. Shlesinger, Ann. N.Y. Acad. Sci. **504**, 214 (1987); J. B. Bassingthwaite, L. S. Liebovitch, and B. J. West, *Fractal Physiology* (Oxford University Press, New York, 1994).
- [2] N. Scafetta, P. Hamilton, and P. Grigolini, *Fractals* **9**, 193 (2001).
- [3] P. Grigolini, L. Palatella, and G. Raffaelli, *Fractals* **9**, 439 (2001).
- [4] E. W. Montroll and G. H. Weiss, *J. Math. Phys.* **6**, 178 (1965).
- [5] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Random Variables* (Addison-Wesley, Reading, MA, 1954).
- [6] N. Scafetta, V. Latora, and P. Grigolini, cond-mat/0105041.
- [7] D. Bedeaux, K. Lakatos Lindenber, and K. E. Shuler, *J. Math. Phys.* **12**, 2116 (1971).
- [8] A. L. Goldberger, L. A. Nunes Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, *Circulation* **101(23)**, e215 (2000).
- [9] P. Allegrini, M. Buiatti, P. Grigolini, and B. J. West, *Phys. Rev. E* **57**, 4558 (1998).
- [10] K. Saermark, M. Moeller, U. Hintze, H. Moelgaard, P. E. Bloch Thomsen, H. Huikuri, T. Makikallio, J. Levitan, and M. Lewkowicz, *Fractals* **8**, 315 (2000).
- [11] P. Ch. Ivanov, L. A. Nunes Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, *Nature (London)* **399**, 461 (1999).
- [12] Y. Ashkenazy, P. Ch. Ivanov, S. Havlin, C.-K. Peng, A. L. Goldberger, and H. E. Stanley, *Phys. Rev. Lett.* **86**, 1900 (2001).