Information entropy as a measure of nonexponentiality of waiting-time distributions

Sergei F. Chekmarev
Institute of Thermophysics, 630090 Novosibirsk, Russia
and Department of Physics, Novosibirsk State University, 630090 Novosibirsk, Russia
(Received 10 May 2008; revised manuscript received 18 September 2008; published 23 December 2008)

It is shown that the information entropy based on waiting-time distributions (WTDs) offers a natural and robust measure of nonexponentiality of the distributions in the form of the Schrödinger-Brillouin negentropy, or equivalently the Kullback-Leibler divergence, and has a straightforward interpretation in terms of transition state theory. Other measures of nonexponentiality of WTDs, based on comparison of the standard deviation and the median with the mean waiting time, are also discussed. The theoretical analysis is illustrated with results from protein folding studies.

DOI: 10.1103/PhysRevE.78.066113 PACS number(s): 82.20.Uv, 87.15.hm

I. INTRODUCTION

The exponential decay distributions represent a specific class of waiting-time distributions (WTDs) corresponding to a Poisson process of transitions from one state of the system to another, e.g., from a reactant to the product. They are applicable to various phenomena of discrete nature, where the events occur randomly with a given mean frequency. At the same time, in many cases the WTDs reveal a nonexponential behavior. Some examples are given by dynamics of recombination processes [1] and relaxation phenomena [2] in crystals and glasses, single-molecule [3–5] and fluorescence resonance energy transfer [6] experiments, motor protein dynamics [7], self-organized criticality dynamics inherent to many natural phenomena [8–10], protein folding [11–17], and isomerization processes [18–20] and diffusion in clusters [21]. In such cases it may be useful to know how far the process under consideration deviates from the Poisson process, or which of the processes deviates more. For this, a quantitative measure of nonexponentiality of the WTD is needed.

Such a measure can be introduced in different ways. One possibility is to utilize the fact that for the exponential WTD both the median (divided by factor ln 2) and the standard deviation are equal to the mean waiting time [22] (see, e.g., Refs. [13,14], respectively). Another possibility is to use the information entropy approach, which allows not only quantification of the degree of deviation from the Poisson process, but also some insight into the nature of deviation. Two well-known examples of the "entropic" measure are the negative entropy (negentropy), due to Schrödinger [23] and Brillouin [24], and its generalization to arbitrary distributions, the Kullback-Leibler divergence [25]. In application to the WTD analysis, they are shown to be equivalent and thus present the same measure. Meaning mostly physical aspects of the problem, we will refer to this measure as "negative entropy." One essential feature of the negative entropy is that the information entropy, on which it is based, presents a functional that is maximized by the exponential distribution, so that any deviation from the exponential distribution results in an increase of the negative entropy. This contrasts the negative entropy with the above "nonentropic" measures, which do not possess this property. Also, the negative entropy has a straightforward interpretation in terms of transition state theory, which may be useful in a broad range of applications, to which the concepts of information entropy and transition state theory are relevant (chemical reactions, conformation transitions in polyatomic systems, etc.).

The paper is organized as follows. Section II describes different measures of nonexponentiality of WTDs: In Sec. II A the negative entropy in the form of the Schrödinger-Brillouin negentropy and its interpretation in terms of transition state theory are considered, in Sec. II B the equivalence of the Kullback-Leibler divergence and the negative entropy, and in Sec. II C the nonentropic measures. In Sec. II D, it is analyzed how the shape of a WTD affects the measures of its nonexponentiality. Section III illustrates the application of the entropic and nonentropic measures to the WTD analysis in the protein folding problem. Section IV summarizes the results and presents some concluding remarks.

II. NEGATIVE ENTROPY AND OTHER MEASURES OF NONEXPONENTIALITY OF WAITING-TIME DISTRIBUTIONS

A. Negative entropy

In the differential form, the information (Shannon) entropy is defined by the functional

$$S = -\int_0^\infty p(t)\ln p(t)dt, \tag{1}$$

where t is the time, and $p(t)$ is the WTD, i.e., the probability density of waiting times. The WTD satisfies two constraints: $\int_0^\infty p(t)dt=1$ and $\int_0^\infty tp(t)dt=\langle t\rangle$, where $\langle t\rangle$ is the mean waiting time. Applying the method of Lagrange multipliers, one finds that the functional is extremized by the exponential distribution (see, e.g., [26])

$$p_{\mathrm{P}}(t) = k\exp(-kt), \tag{2}$$

which is the WTD for the Poisson process ($k=1/\langle t\rangle$ is the mean frequency of events). Correspondingly, the maximum (Poisson) entropy is

$$S_{\mathrm{P}} = 1 - \ln k = 1 + \ln\langle t\rangle. \tag{3}$$

Since the probability density $p(t)$ has the dimension of reciprocal time, the term $\ln p(t)$ under the integral in Eq. (1) requires the time t to be dimensionless. Therefore both Eqs. (1) and (3) determine the value of the entropy up to a constant, which depends on the time scale. More specifically, if the time scale changes from τ to $\tilde{\tau}$, the entropy shifts from S to

$$\tilde{S} = S + \ln(\tau/\tilde{\tau}). \tag{4}$$

The same is valid for the maximum entropy S_P, Eq. (3).

It is instructive to compare Eq. (3) with the corresponding expressions from the transition state theory for unimolecular reactions. For this, we rewrite Eq. (3) as

$$\langle t\rangle = \exp(S_{\mathrm{P}} - 1). \tag{5}$$

In a microcanonical ensemble, according to the Rice-Ramsperger-Kassel-Marcus (RRKM) theory [27], the mean waiting time is written as $\langle t'\rangle = h\rho(E)/N^{\star}(E^{\star})$, where h is the Planck constant, E the energy, $\rho(E)$ the density of states of the reactant, $E^{\star}$ the available energy at the transition state, $N^{\star}(E^{\star})$ the number of possible states at the transition state, and the prime at t denotes that the time is dimensional. We introduce a time scale as $\tau = h/\epsilon$, where ϵ is a characteristic energy, and define the entropies of the reactant and transition states as $S(E) = k_{\mathrm{B}} \ln \rho(E)$ and $S^{\star}(E^{\star}) = k_{\mathrm{B}} \ln[N^{\star}(E^{\star})/\epsilon]$, respectively ($k_{\mathrm{B}}$ is the Boltzmann constant). Then the RRKM equation rewrites as $\langle t\rangle = \exp\{[S(E) - S^{\star}(E^{\star})]/k_{\mathrm{B}}\}$, where $\langle t\rangle = \langle t'\rangle/\tau$. Comparison of this equation with Eq. (5) yields

$$S_P = [S(E) - S^{\star}(E^{\star})]/k_{\mathrm{B}} + 1, \tag{6}$$

i.e., S_P is essentially a loss of physical entropy in the transition from the reactant to the transition state.

A similar result is obtained for a canonical ensemble, except that the partition function plays a role of the density of states and, correspondingly, the free energy appears instead of the entropy. Specifically, in this case $\langle t'\rangle = (h/k_{\mathrm{B}}T)Z(T)/Z(T)^{\star}$, where T is the temperature, and $Z(T)$ and $Z^{\star}(T)$ are the partition functions of the reactant and the transition state, respectively [28]. Introducing the time scale as $\tau = h/k_{\mathrm{B}}T$, and the free energies of the reactant and the transition state as $F(T) = -k_{\mathrm{B}}T \ln Z(T)$ and $F^{\star}(T) = -k_{\mathrm{B}}T \ln Z^{\star}(T)$, respectively, one finds $\langle t\rangle = \exp\{-[F(T) - F^{\star}(T)]/k_{\mathrm{B}}T\}$. Correspondingly, comparison of this equation with Eq. (5) gives

$$S_P = -[F(T) - F^{\star}(T)]/k_{\mathrm{B}}T + 1. \tag{7}$$

We define the measure of dissimilarity between a given WTD and the Poisson distribution as

$$\Delta S = S_{\mathrm{P}} - S = 1 + \ln\langle t\rangle + \int_0^{\infty} p(t)\ln p(t)dt \tag{8}$$

provided that both distributions have the same value of the mean waiting time $\langle t\rangle$. By formal definition, this quantity is essentially the Brillouin negentropy [24], generalizing the Schrödinger negative entropy for living systems [23]. The only difference is that in Eq. (8) the probability for the system to escape from a certain state at a given time is considered instead of the probability for the system to be found in a certain state [24]. Although, by definition, ΔS [Eq. (8)] is closer to Brillouin's negentropy, we will use for it a more general term, i.e., "negative entropy." According to Eqs. (1)–(3), ΔS is always non-negative and vanishes as the WTD approaches the Poisson distribution. It is essential that ΔS is determined by the degree of deviation of the WTD from the exponential distribution, i.e., by the shape of the WTD, and it does not depend on the mean waiting time. This simply follows from the fact that both the entropies, Eqs. (1) and (3), depend on the time scale in the same way [Eq. (4)].

The utility of the negative entropy defined by Eq. (8) is twofold. First, if $\Delta S > 0$, it follows that $p(t)$ is less random than the Poisson distribution, so that the process under consideration should contain, by implication, some deterministic elements, which break the Poisson process. Second, according to Eqs. (6) and (7), ΔS can be expressed in terms of physical entropy and free energy, which offers a possibility to estimate the contribution of these deterministic elements in the k_{B} and $k_{\mathrm{B}}T$ units, respectively. Potentially, the negative entropy approach can be applied to any problem to which the concepts of information entropy and transition state theory are relevant.

B. Kullback-Leibler divergence

Since $p(t)$ and $p_{\mathrm{P}}(t)$ in Eq. (8) are assumed to have the same value of the mean waiting time $\langle t\rangle$, $S_{\mathrm{P}} = -\int_0^{\infty} p_{\mathrm{P}}(t)\ln p_{\mathrm{P}}(t)dt = -\int_0^{\infty} p(t)\ln p_{\mathrm{P}}(t)dt$. Then, the right-hand side of Eq. (8) can be rewritten as $\int_0^{\infty} p(t)\ln[p(t)/p_{\mathrm{P}}(t)]dt$, which presents a specific form of the Kullback-Leibler divergence [25], i.e., the divergence of $p(t)$ from $p_{\mathrm{P}}(t)$ [Eq. (2)] with the logarithm taken to the natural base. It follows that the Kullback-Leibler divergence has all the properties of the negative entropy ΔS, defined by Eq. (8), and, correspondingly, ΔS could equivalently be referred to as the Kullback-Leibler divergence.

C. "Nonentropic" measures

Another way to measure nonexponentiality of WTDs is to compare the standard deviation of waiting times, $\sigma[p(t)]$, or the median of waiting times, $\mu_{1/2}[p(t)]$, with the mean waiting time. The corresponding measures can be, respectively, defined as

$$D_{\mathrm{sd}} = \sigma[p(t)]/\langle t\rangle \tag{9}$$

and

$$D_{\mathrm{med}} = \{\mu_{1/2}[p(t)]/\ln 2\}/\langle t\rangle. \tag{10}$$

For an exponential distribution both D_{sd} and D_{med} should be equal to unity [22]. A shortcoming of these measures is that they do not present functionals that are extremized by the exponential distribution (in contrast to the information entropy). Therefore, although in certain cases D_{sd} and D_{med} can be used for the WTD analysis (see, e.g., Refs. [14,13], respectively), the equality of these quantities to unity does not generally guarantee that the WTD is exponential (see also the discussion of this issue in Secs. II D and III).

D. Model waiting-time distributions

To see how the shape of a WTD affects the measure of its nonexponentiality, consider the WTDs of three characteristic types, specifically, double-exponential, three-exponential, and nonexponential distributions; the latter are taken in the form of the function $p(t) \sim \exp[(t/\tau)^{\beta}]$, where β varies from $\beta<1$ (stretched-exponential distributions) to $\beta>1$ (squeezed-exponential). A double-exponential WTD generally represents the process of transitions from a reactant to the product in the presence of an off-pathway intermediate, which is separated from the reactant by a barrier not considerably lower than the barrier between the reactant and the product. Similarly, a three-exponential WTD represents the process of transitions that is complicated by the presence of two such intermediates (see, e.g., Refs. [15,16]). A stretched (squeezed) exponential WTD can be considered as a limit case of these processes when an array of intermediates, which are separated from the reactant by relatively low barriers, is involved (see, e.g., Refs. [29,30], respectively); correspondingly, the WTD is characterized by a broad (quasi-) continuous spectrum of decay times and associated with so-called "strange kinetics" [31]. Figures 1(a)–1(c) present series of these distributions; in each case the WTD varies from the exponential distribution to a distribution which considerably deviates from it. The parameters of the WTDs are chosen to be such that the series of the curves would be more or less similar in shape. Figure 2 shows the values of the entropy for the distributions of Fig. 1, which are calculated according to Eq. (1), and compares these values with the values of the maximum entropy, Eq. (3), corresponding to the Poisson process [32].

Experimental and simulated WTDs are often noisy (see, e.g., Fig. 6). To see how this circumstance may affect the accuracy of calculation of the entropy, the WTDs of Fig. 1 were subjected to a statistic noise. For this, the probability density $p(t)$ was shifted according to the Gaussian distribution with the root-mean-square deviation $\sim 1/\sqrt{p(t)N}$ ($N=5 \times 10^3$ is the number of trials), which mimics the inherent statistic broadening of exponential-like distributions at small $p(t)$ [15]. As an example of the noisy WTDs, Fig. 3 presents such distributions for the three-exponential WTDs shown in Fig. 1(b). The values of the entropy for the noisy WTDs are shown in Fig. 2 by open symbols. It is seen that the statistical noise of a reasonable amplitude effects the value of the entropy very slightly. It should be noted, however, that a poor sampling may pose a problem for accurate estimation of entropy (for a discussion of this problem, see, e.g., Ref. [33] and the papers cited therein).

Comparison of Figs. 1 and 2 shows that ΔS [Eq. (8)] increases as the WDT deviates from the exponential distribution. Moreover, the less regular the behavior of the WTD, the higher ΔS. This is illustrated by Fig. 4, which shows two-exponential, three-exponential, and stretched-exponential distributions characterized by approximately the same value of the mean waiting time. As one can see, comparing Figs. 4 and 2, ΔS drastically increases as the slope of the curve changes less monotonically, specifically, from ≈ 0.08 for the stretched distribution to ≈ 0.25 for the three-exponential distribution, and then to ≈ 0.5 for the double-exponential distribution.

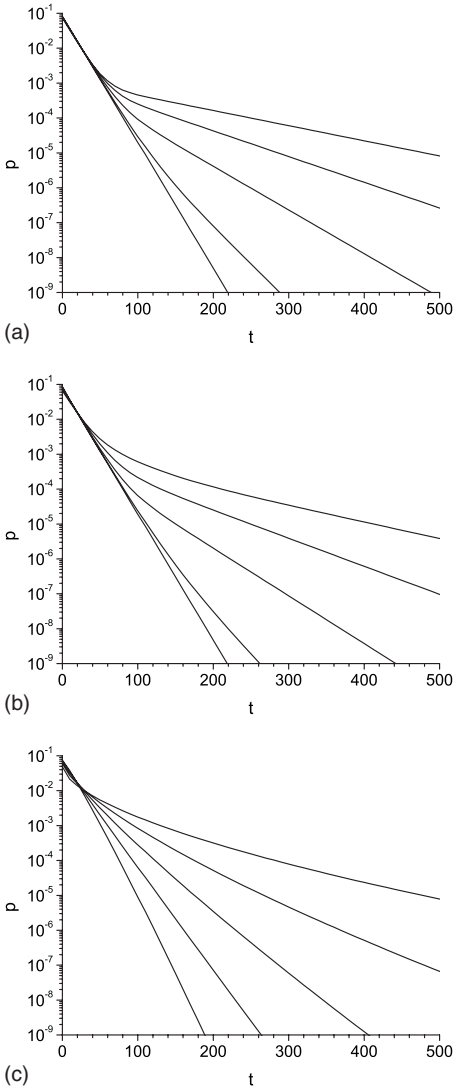

FIG. 1. Model waiting-time distributions: double-exponential (a), three-exponential (b), and nonexponential (c). The *double-exponential* distributions are determined as $p(t)=c_1 \exp(t/\tau_1)+c_2 \exp(t/\tau_2)$, where $c_1:c_2=60:1$ and $\tau_1=\tau_2=12$ in the lowest curve, and τ_2 increases successively by a factor of 1.7 in each of the higher curves; the *three-exponential* distributions are determined as $p(t)=c_1 \exp(t/\tau_1)+c_2 \exp(t/\tau_2)+c_3 \exp(t/\tau_3)$, where $c_1:c_2:c_3=67:6.7:1$ and $\tau_1=12$, $\tau_2=4$, and $\tau_3=11$ in the lowest curve, and the latter two increase successively by a factor of 1.7 in each of the higher curves; and the *nonexponential* distributions are determined as $p(t) \sim \exp[(t/\tau)^{\beta}]$, where $\tau=13.5$ and $\beta=0.6$, and β increases successively by 0.125. The increase of the characteristic times in the double- and three-exponential distributions [(a) and (b)], as well as the decrease of β in the nonexponential distribution [(c)], is accompanied by a monotonic decrease of the slope of the decay curves at longer times.

Also, consider the WTDs which differ from the exponential distribution in that in the semilog representation one is concave and the other is convex. In Fig. 5 the concave distribution is represented by the stretched-exponential function ($\beta=0.7$, dashed curve), and the convex distribution by the squeezed-exponential function ($\beta=1.3$, dotted curve). For the exponential distribution ($\beta=1$, solid curve in Fig. 5) ΔS

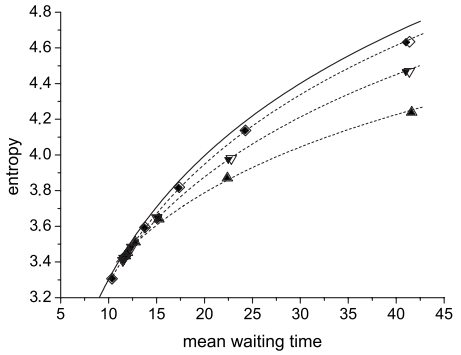

FIG. 2. Information entropy vs mean waiting time for the model distributions of Fig. 1. Solid up triangles, down triangles, and diamonds are for the double-exponential, three-exponential, and nonexponential distributions, respectively. The corresponding empty symbols, which are larger in size for illustrative purposes, are for the corresponding noisy distributions, such as in Fig. 3. The thin dashed curves depict the exponential fits to the data, and the thick solid curve presents the maximum entropy, Eq. (3).

[Eq. (8)], $\log_{10} D_{\mathrm{sd}}$ [Eq. (9)] and $\log_{10} D_{\mathrm{med}}$ [Eq. (10)] are essentially zero, whereas for the concave and convex distributions they are, respectively, as follows: $\Delta S = 2 \times 10^{-2}$ and 9×10^{-3}, $\log_{10} D_{\mathrm{sd}} = 9 \times 10^{-2}$ and -5×10^{-2}, and $\log_{10} D_{\mathrm{med}} = -7 \times 10^{-2}$ and 4×10^{-2}. It shows that if ΔS, as expected, remains positive in both cases of deviation from the exponential distribution, $\log_{10} D_{\mathrm{sd}}$ and $\log_{10} D_{\mathrm{med}}$ change their signs. The latter is a result of the above-mentioned property of D_{sd} and D_{med} (Sec. I C), i.e., that neither of them presents a functional which is extremized by the exponential distribution. One possible consequence of this is that the opposite deviations from the exponential distribution, if they exist, can cancel each other, reducing $\log_{10} D_{\mathrm{sd}}$ and $\log_{10} D_{\mathrm{med}}$ to zero. Such effect has been observed in Ref. [14], where $\langle t \rangle / \sigma$ in Fig. 4(g) was found close to unity "even though the distribution of folding times is nonexponential."

III. APPLICATION TO PROTEIN FOLDING

To illustrate one possible application of the negative entropy approach, consider several examples from protein folding. In this case, the first-passage time from the denatured state of a protein to its native state plays the role of the

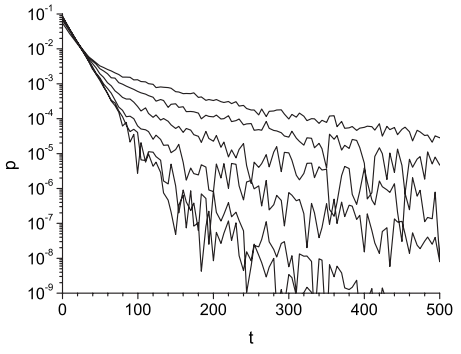

FIG. 3. Noisy three-exponential waiting-time distributions corresponding to (b) of Fig. 1.

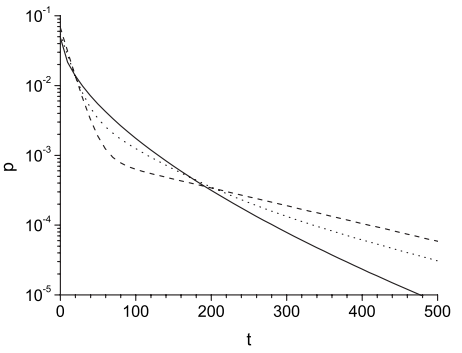

FIG. 4. Waiting-time distributions of Fig. 1, corresponding to approximately the same value of the mean waiting time ($\langle t \rangle \approx 41$): double-exponential (dashed line), three-exponential (dotted), and stretched-exponential (solid).

waiting time. Typically, the folding process consists of a fast collapse of the protein into a semicompact (globular) state, which is followed by a slow search for a pathway to the native state. The deviation of the WTD from an exponential distribution indicates that the folding kinetics are more complex than the simple two-state kinetics, which are associated with a Poisson process of transitions from the globule to the native state. It can thus be expected that some on/off-pathway intermediates are involved into the folding process in this case. From the information entropy viewpoint, the presence of intermediates can be considered as the appearance of some deterministic elements in the folding process due to the transitions between the globule and/or intermediates. These transitions break the Poisson process of the escape from the free energy basin corresponding to the globule, and should thus lead to an increase of the negative entropy ΔS.

At low temperatures, the mean folding time decreases with temperature because the system spends less time in local minima on the energy surface, and at high temperatures it increases with temperature because the system explores a larger portion of semicompact states. Due to this, the mean folding time exhibits a U-shaped behavior with the temperature [34,35], with the minimum achieved at some $T = T_{\mathrm{f}}$. Figure 6 shows the WTDs for a 27-residue lattice protein studied in Ref. [15], $T_{\mathrm{f}} \approx 1.9$ (for simulation results, here and below, the temperature is measured in units of the interaction

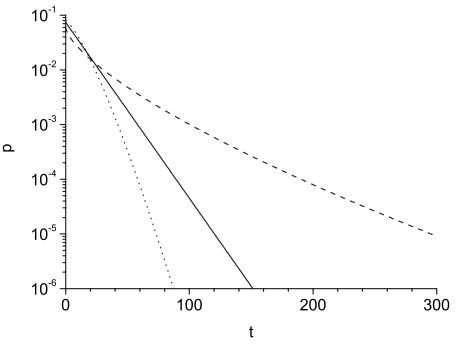

FIG. 5. The distributions $p(t) \sim \exp[(t/\tau)^{\beta}]$ with $\tau = 13.5$ and $\beta = 0.7$ (stretched-exponential, dashed line), $\beta = 1.0$ (exponential, solid line), and $\beta = 1.3$ (squeezed-exponential, dotted line).

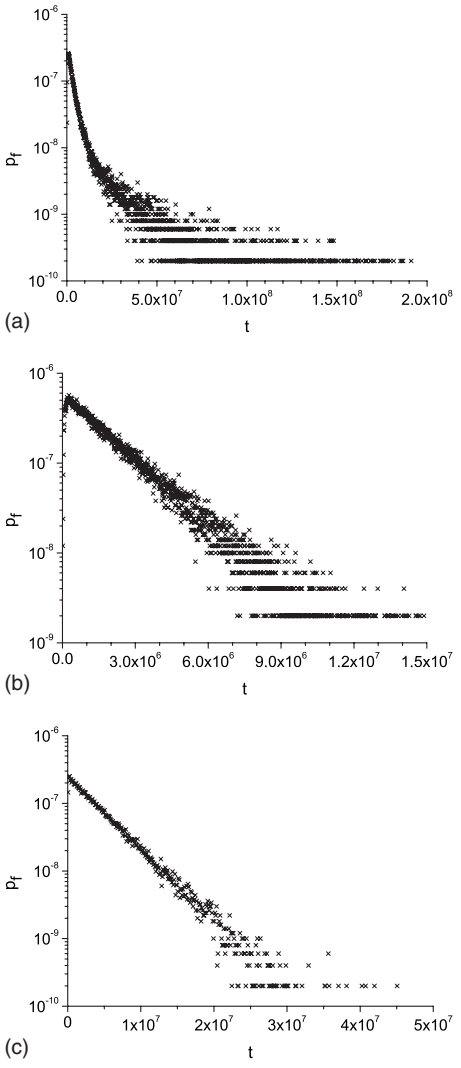

FIG. 6. Folding time distributions for a 27-residue lattice protein [15]: $T=1.5$ (a), $T=1.9$ (b), and $T=2.3$ (c).

energy between the monomers, which represent residues in the protein). Each of the distributions was obtained on the basis of 5×10^4 folding trajectories. Figure 6(a) presents a characteristic WTD below T_f, Fig. 6(b) the WTD at $T=T_f$, and Fig. 6(c) a characteristic WTD above this temperature. At and above T_f the WTDs are close to exponential, and below T_f they are nonexponential. The reason behind this difference is that along with the free energy minimum corresponding to the globule the system has off-pathway minima associated with kinetic intermediates. At $T\geqslant T_f$ these intermediates are in equilibrium with the globule, forming an "extended globule," and at $T<T_f$ they are not [15].

Figure 7 plots the negative entropy $\Delta S=S_P-S$ and two nonentropic measures, $\log_{10} D_{sd}$ and $\log_{10} D_{med}$, against the temperature. They were calculated on the basis of Eqs. (8)–(10), respectively, using the WTDs obtained in the simulations in Ref. [15]. It is seen that the behavior of the negative entropy and nonentropic measures is, in general, similar: At low temperatures, where the WTDs are nonexponential [as in Fig. 6(a)], all ΔS, $\log_{10} D_{sd}$, and $\log_{10} D_{med}$ are of considerable value. Then, as the temperature approaches T_f,

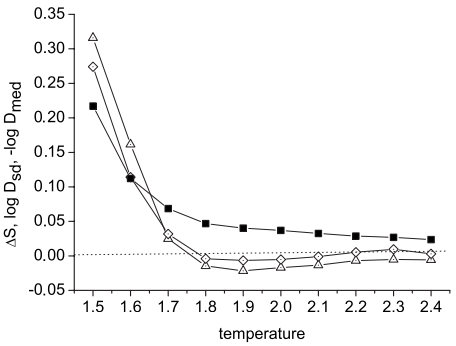

FIG. 7. The temperature dependence of the negative entropy ΔS [Eq. (8), cubes], and two "nonentropic" measures, $\log_{10} D_{sd}$ [Eq. (9), triangles] and $-\log_{10} D_{med}$ [Eq. (10), diamonds], for folding of a 27-residue lattice protein [15]. The solid lines are to guide the eye, and the dotted line is to show the zero deviation from the exponential distribution.

where the WTD acquires an exponential form [Fig. 6(b)], they become small. At the higher temperatures, where the WTDs are exponential [Fig. 6(c)], these quantities change slightly. Figure 7 clearly indicates that the deviation from two-state kinetics can be expected below $T=1.9$, which agrees well with the results of a detailed study of the folding kinetics of the given system in Ref. [15]. At the same time, it is seen that none of ΔS, $\log_{10} D_{sd}$, and $\log_{10} D_{med}$ is essentially close to zero at $T\geqslant T_f$. The reason is that the WTDs at these temperatures [Figs. 6(b) and 6(c)] are not truly exponential. When the protein folds from a fully extended state, as it is often assumed in the simulations (including Ref. [15]), some time is required for the protein to collapse into a semicompact globule. This leads to a sharp increase of the WTD at short times [15], which is particularly well seen in Fig. 6(b). Due to this deviation from the exponential distribution, ΔS retains a finite value at high temperatures. For D_{sd} and D_{med} the consequence is more dramatic, i.e., because the WTDs are convex at short times, $\log_{10} D_{sd}$ and $-\log_{10} D_{med}$ become negative (see the discussion of Fig. 5 in Sec. II D). Correspondingly, when T approaches T_f from the side of low temperatures, $\log_{10} D_{sd}$ and $-\log_{10} D_{med}$ decrease excessively fast. A similar irregular behavior of D_{med} and D_{sd} is observed for the model proteins studied in Refs. [13,14]; for the 27-mer lattice protein [13] $\log_{10} D_{med}$ changes sign as the transition midpoint in the chevron plot is approached (Fig. 3 of Ref. [13], folding, open, and closed squares), and for the off-lattice SH3 domain [14] $\log_{10} D_{sd}$ behaves similar to that in Fig. 7 [Fig. 4(g) of Ref. [14]]. Thus although in principle every above measure, i.e., ΔS, D_{sd}, and D_{med}, can be used to quantify nonexponentiality of a WTD, the negative entropy has a more regular behavior.

Simulations for a lattice protein with an amyloidogenic latent state [17] show the same character of variation of the negative entropy with temperature as in Fig. 7. Here at $T=0.65$, $\Delta S\approx0.2$ (a double-exponential WTD, multistate kinetics), at $T=T_f=0.6725$, $\Delta S\approx0.03$ (the WTD is close to an exponential WTD, and the kinetics are close to two-state kinetics), and at $T=0.695$, $\Delta S\approx0.01$ (an exponential WTD, two-state kinetics). Similar results are obtained for an off-lattice ubiquitinlike protein [12], where $\Delta S\approx0.1$ at $T=0.44$

and $\Delta S \sim 0.01$ at $T=0.48-0.57$ [to calculate ΔS, Eq. (8), the fitting functions for the WTDs, presented in Ref. [12], were used].

The negative entropy can also increase with temperature, as, e.g., in the case of "strange kinetics" observed by Sabelko *et al.* [11] for a two-domain enzyme yeast phosphoglycerate kinase (PGK) and double mutant (F45W, V26G) of human ubiquitin (Ub*G), which unfold on cold denaturation. Provided that a nonspecific relaxation of the polypeptide chain in the unfolded state that occurs at small times is excluded, both the proteins have the exponential WTDs at lower temperatures and the stretched-exponential WTDs at higher temperatures [11]. What is of interest is that the values of ΔS are found to be of the same order as in the previously discussed simulations. Calculations for PGK give $\Delta S \approx 0.04$ at $T=5C$ and $\Delta S \approx 0.16$ at $T=19C$, and for Ub*G $\Delta S \approx 0.016$ at $T=2C$ and $\Delta S \approx 0.5$ at $T=8C$. It should be noted, however, that taking into account the previously mentioned relaxation results in more complex kinetics at low temperatures (see also [16]). In particular, for Ub*G at $T=2C$ the WTD transforms into a double-exponential distribution with vastly different decay times (22 μs and 5 ms) [11], which makes ΔS as large as ≈ 2.3.

In the case of canonical ensemble, according to Eqs. (7) and (8), ΔS represents a change in the free energy. Therefore, provided that both the folding time distribution and the free energy barrier separating the unfolded and folded states (ΔF_{UF}) are known, ΔS can be compared with $\Delta F_{\mathrm{UF}}/k_{\mathrm{B}}T$. Such information is available for the 27-residue lattice protein of Ref. [15] at $T=1.5$. Here $\Delta S \approx 0.22$ (Fig. 7), and $\Delta F_{\mathrm{UF}}/k_{\mathrm{B}}T = \ln Z_d/Z_{\mathrm{d,n}} \approx 5.9$, where Z_{d} and $Z_{\mathrm{d,n}}$ are the partition functions for the denatured state and the transition state ensemble, respectively (they were estimated in Ref. [15] with the use of the transition disconnectivity graph [36]). It follows that although the folding time distribution apparently deviates from the exponential distribution [Fig. 6(a)], the increase in the free energy does not exceed $\approx 4\%$ of the free energy barrier, i.e., folding kinetics can be considered as two-state kinetics in the leading order.

IV. CONCLUSION

It has been shown that the information entropy based on waiting-time distributions (WTDs) offers a natural and robust measure of nonexponentiality of the distributions. To quantify the deviation of a WTD from the exponential distribution, corresponding to a Poisson process, the difference between the information entropies for the exponential distribution and a given WTD has been employed. This difference is essentially the Schrödinger-Brillouin [23,24] negative entropy (negentropy), except that the probability for the system to escape from a certain state at a given time is considered instead of the probability for the system to be found in a certain state. The utility of the negative entropy in application to the WTD analysis is twofold. First, a positive value of the negative entropy indicates that the WTD is less random than the Poisson distribution, so that the process under consideration should involve some deterministic element(s), which breaks the Poisson process. Second, the negative entropy has a straightforward interpretation in terms of transition state theory, so that it can be expressed in terms of physical entropy and free energy, and, correspondingly, be measured in the k_{B} and $k_{\mathrm{B}}T$ units, respectively. Potentially, the negative entropy approach has a broad range of application for the WTD analysis because it is applicable to any problem to which the concepts of information entropy and transition state theory are relevant.

The negative entropy thus introduced presents a specific case of the well-known Kullback-Leibler divergence [25], when the Poisson distribution is used as the reference distribution and the logarithm is taken to the natural base. Correspondingly, it can equivalently be referred to as the Kullback-Leibler divergence.

Along with the entropic measures, the nonentropic measures D_{sd} [Eq. (9)] and D_{med} [Eq. (10)], which are based on the comparison of the standard deviation and median of waiting times with the mean waiting time, can, in principle, be used. However, since they do not present functionals that are extremized by the exponential distribution (in contrast to the information entropy), the equality of D_{sd} and D_{med} to unity does not generally guarantee that the WTD is exponential. Therefore care should be taken in the application of these measures.

The theoretical analyses have been illustrated with simulation and experimental results from protein folding. Considering a limited but not specific set of proteins, it has been found that the negative entropy typically varies in the range of several hundredths of $k_{\mathrm{B}}T$ (two-state kinetics) to several tenths of $k_{\mathrm{B}}T$ (multistate kinetics). The knowledge of the negative entropy and the free energy barrier between the unfolded and folded states of the protein allows estimation of the relative deviation of the folding process from two-state kinetics.

ACKNOWLEDGMENTS

I thank Professor M. Karplus and Dr. S. Krivov for useful discussions. This work was supported in part by the grants from the Russian Foundation for Basic Research (Grants No. 06-04-48587 and No. 08-04-91104) and the Civilian Research and Development Foundation (Grant No. RUB2-2913-NO-07).

[1] N. Jaba, A. Kanoun, H. Mejri, H. Maaref, and A. Brenier, J. Phys.: Condens. Matter **12**, 7303 (2000).
[2] K. Trachenko and M. T. Dove, Phys. Rev. B **70**, 132202 (2004).
[3] X. S. Xie, Annu. Rev. Phys. Chem. **49**, 441 (1998).
[4] D. Chakrabarti and B. Bagchi, J. Chem. Phys. **118**, 7965 (2003).
[5] O. Flomenbom, K. Velonia, D. Loos, S. Masuo, M. Cotlet, Y. Engelborghs, J. Hofkens, A. E. Rowan, R. J. M. Nolte, M. Van der Auweraer, F. C. de Schryver, and J. Klafter, Proc. Natl. Acad. Sci. U.S.A. **102**, 2368 (2005).
[6] G. Shrinivas, A. Yethiraj, and B. Bagchi, J. Chem. Phys. **114**, 9170 (2001).
[7] M. E. Fisher and A. B. Kolomeisky, Proc. Natl. Acad. Sci. U.S.A. **98**, 7748 (2001).
[8] H. J. Jensen, *Self-Organized Criticality* (Cambridge University Press, Cambridge, England, 1998).
[9] C. Adami, Phys. Lett. A **203**, 23 (1995).
[10] R. Sánchez, D. E. Newman, and B. A. Carreras, Phys. Rev. Lett. **88**, 068302 (2002).
[11] J. Sabelko, J. Ervin, and M. Gruebele, Proc. Natl. Acad. Sci. U.S.A. **96**, 6031 (1999).
[12] J. M. Sorenson and T. Head-Gordon, Proteins: Struct., Funct., Genet. **46**, 368 (2002).
[13] H. Kaya and H. S. Chan, Proteins: Struct., Funct., Genet. **52**, 524 (2003).
[14] J. M. Borreguero, F. Ding, S. V. Buldyrev, H. E. Stanley, and N. V. Dokholyan, Biophys. J. **87**, 521 (2004).
[15] S. F. Chekmarev, S. V. Krivov, and M. Karplus, J. Phys. Chem. B **109**, 5312 (2005).
[16] S. F. Chekmarev, S. V. Krivov, and M. Karplus, J. Phys. Chem. B **110**, 8865 (2006).
[17] A. Yu. Palyanov, S. V. Krivov, M. Karplus, and S. F. Chekmarev, J. Phys. Chem. B **111**, 2675 (2007).
[18] F. A. Amar and R. S. Berry, J. Chem. Phys. **85**, 5943 (1986).
[19] C. Seko and K. Takatsuka, J. Chem. Phys. **104**, 8613 (1996).
[20] M. A. Miller and D. J. Wales, J. Chem. Phys. **107**, 8568 (1997).
[21] S. V. Krivov and S. F. Chekmarev, Eur. Phys. J. D **9**, 205 (1999).
[22] G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* 2nd ed. (McGraw-Hill, New York, 1968).
[23] E. Schrödinger, *What is Life? The Physical Aspect of the Living Cell* (Cambridge University Press, Cambridge, England, 1945).
[24] L. Brillouin, J. Appl. Phys. **24**, 1152 (1953).
[25] S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).
[26] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[27] W. Forst, *Theory of Unimolecular Reactions* (Academic Press, New York, 1973).
[28] H. Eyring, S. H. Lin, and S. M. Lin, *Basic Chemical Kinetics* (Wiley, New York, 1980).
[29] R. Metzler, J. Klafter, J. Jortner, and M. Volk, Chem. Phys. Lett. **293**, 477 (1998).
[30] H. K. Nakamura, M. Sasai, and M. Takano, Proteins: Struct., Funct., Bioinf. **55**, 99 (2004).
[31] M. F. Shlesinger, G. M. Zaslavsky, and J. Klafter, Nature (London) **363**, 31 (1993).
[32] Although in Fig. 1 the time scale is restricted by the value of $t=5\times10^2$, the entropy was calculated using the interval with the upper bound $t=2.5\times10^3$, where all $p(t)$ are as small as 1×10^{-9} at least.
[33] M. Montemurro, R. Senatore, and S. Panzeri, Neural Comput. **19**, 2913 (2007).
[34] O. V. Galzitskaya and A. V. Finkelstein, Protein Eng. **8**, 883 (1995).
[35] M. Karplus, Fold. Des. **2**, 569 (1997).
[36] S. V. Krivov and M. Karplus, J. Chem. Phys. **117**, 10894 (2002).