Long Timescales, Individual Differences, and Scale Invariance in Animal Behavior

William Bialek[1,2] and Joshua W. Shaevitz[1]
[1]Joseph Henry Laboratories of Physics and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, USA
[2]Center for Studies in Physics and Biology, Rockefeller University, New York, New York 10065, USA

(Received 22 April 2023; accepted 27 November 2023; published 22 January 2024)

The explosion of data on animal behavior in more natural contexts highlights the fact that these behaviors exhibit correlations across many timescales. However, there are major challenges in analyzing these data: records of behavior in single animals have fewer independent samples than one might expect. In pooling data from multiple animals, individual differences can mimic long-ranged temporal correlations; conversely, long-ranged correlations can lead to an overestimate of individual differences. We suggest an analysis scheme that addresses these problems directly, apply this approach to data on the spontaneous behavior of walking flies, and find evidence for scale-invariant correlations over nearly three decades in time, from seconds to one hour. Three different measures of correlation are consistent with a single underlying scaling field of dimension $\Delta = 0.180 \pm 0.005$.

DOI: 10.1103/PhysRevLett.132.048401

Animals, including humans, exhibit behaviors with structure on many timescales. In one view, the many timescales result from distinct processes, perhaps organized hierarchically [1,2]. In another view, the wide range of timescales emerges from interactions among many underlying degrees of freedom, perhaps approaching a nearly scale-invariant continuum [3]. Scale invariance is especially tantalizing because of possible links to the renormalization group and critical phenomena [4–6].

The literature on scale invariance in living systems is dominated by theoretical arguments: why it might be advantageous for organisms to operate in this regime or how apparent signatures of criticality and scale invariance might have more prosaic explanations. This is an opportune moment to revisit the experiments because we have seen an explosion of quantitative data on animal behavior under more naturalistic conditions, in systems ranging from worms to primates [7–20]. High resolution video imaging and efficient AI tools combine to make these approaches more generally applicable [21–23], and these data have brought renewed attention to the wide range of timescales in behavior [24–30]. Although we focus here on the behavior of individual organisms, there is strong evidence for scale invariance in collective behaviors of flocks and swarms [31].

We can characterize a system either by measuring the mean behavior in response to external perturbations or by analyzing the correlations in spontaneous fluctuations of behavior. Here we take the latter approach, across a wide range of timescales. However, if correlations are sufficiently long ranged, then in a single behavioral trajectory we never have truly independent samples and many statistical intuitions break down. Also, if we are interested in the longest timescales that we can access in a given experiment, then by definition we do not have many samples, independent or not. In order to increase statistical power, quantitative studies of animal behavior often average over multiple individual organisms, but this makes sense only if the different organisms behave in the same way. In trying to measure correlations over long times, individual differences are an important confounding factor, essentially because each individual has an infinite memory of its own identity. The goal of this Letter is to disentangle the nonindependence of samples, individual differences, and genuinely long-ranged correlations.

Many different notions of scale invariance have been considered in the analysis of behavior, including language. Maybe the first example is Zipf's law [32], a kind of scale invariance in our vocabulary. The signature here is a power-law distribution for individual words or for the states in a network of neurons at a single moment in time [33,34]. Closely related is the possibility that in foraging a single step of exploration spans a distance drawn from a power-law distribution, as in a Lévy flight [35]. Another idea is that time for transitions from one state to the next, especially from a quiescent to an active state, may have a power-law distribution [36]. Still another idea is that behavior or its underlying neural dynamics may consist of an intermittent sequence of "avalanches," and the size or duration of these avalanches may have a power-law distribution [37–39], as in the original sandpile models for self-organized criticality [3,40]. Although often grouped together under some umbrella of scale invariance or criticality, none of these notions require correlations in the states of a system at widely separated times. A familiar example of scale-invariant temporal correlations is provided by anomalous diffusion [41].

We ground our discussion in the analysis of experiments on the behavior of walking flies [11,12], and we will see that these data provide evidence for scale-invariant correlations over nearly three decades, from timescales of seconds to one hour. The raw data are high resolution videos of fruit flies, from an inbred laboratory stock, walking in a featureless arena. Video frames are of duration $\Delta t = 0.01$ s, single trajectories are of length $T_{\rm max} = 3600$ s, and we count time $t = 1, 2, \ldots$ in units of Δt; there are $N_f = 59$ individual flies in the dataset [11,12]. Through a combination of linear and nonlinear dimensionality reduction, these data can be embedded in a low-dimensional space. In this space, the fly repeatedly visits small neighborhoods and then jumps to another, defining 122 discrete behavioral states, which also can be seen as peaks in the probability distribution over the continuous space; in addition, there is a null state. Some of these states have names (e.g., different forms of grooming) and some do not. The same states can be identified across related species of flies, and the distribution over these states varies systematically with evolutionary distance [42]. Trajectories through the discrete state space are strongly non-Markovian [12] and compressed versions of these state sequences are described by models with nearly scale-invariant interactions [30].

We define $n_{\rm i}(t) = 1$ if an individual fly is in state i in the small bin of duration Δt surrounding time t. We can characterize the correlations, summed over the individual states, by the (connected) two-point function

$$C(t,t') = \sum_{\rm i=1}^{N_s} \left[\langle n_{\rm i}(t) n_{\rm i}(t')\rangle - \langle n_{\rm i}\rangle^2\right] \tag{1}$$

$$= P_c(t,t') - P_c({\rm ind}), \tag{2}$$

where $P_c(t,t')$ is the coincidence probability of finding a fly in the same state at times t and t', and $P_c({\rm ind})$ is the coincidence probability if we draw two independent samples from the distribution over states. If the behavior is statistically stationary, then $C(t,t') = C(t-t')$.

Perhaps surprisingly, a serious problem in estimating these correlations from data is subtracting the mean. If we estimate $\langle n_{\rm i}\rangle$ as a time average over the behavioral trajectory of a single fly, then the correlation function must integrate to zero over the finite duration of our observations [43]. If trajectories are very long compared with the timescales of correlation this does not matter, but long-ranged correlations will be significantly distorted. An alternative is to estimate $\langle n_{\rm i}\rangle$ as an average both over time and over an ensemble of flies. However, if individuals have even slightly different mean behaviors, this also distorts the correlation function, even violating the condition that $C(\tau)$ should vanish at large $|\tau|$.

To disentangle long-ranged correlations and individual differences, we take a direct approach. We estimate the probability that fly α occupies state i by averaging over a window of duration T,

$$\hat{P}^\alpha_{\rm i}(T) = \frac{\Delta t}{T}\sum_{t=1}^{T/\Delta t} n^\alpha_{\rm i}(t). \tag{3}$$

With an ensemble of N_f flies, we can estimate the variance across individuals, sampled in independent experiments and then summed over states,

$$\Phi_2(T) = \sum_{\rm i=1}^{N_s}\frac{1}{N_f}\sum_{\alpha=1}^{N_f}\left[\hat{P}^\alpha_{\rm i}(T) - \frac{1}{N_f}\sum_{\beta=1}^{N_f}\hat{P}^\beta_{\rm i}(T)\right]^2. \tag{4}$$

This is an *apparent* variance across individuals. It includes both real differences in the probabilities with which different individuals visit the behavioral states and the fact that our estimates of these probabilities are based on samples of duration T.

If there are no real individual differences, then the only reason we see any variance across the ensemble of flies is because of statistical errors; that is, because our estimates $\hat{P}^\alpha_{\rm i}(T)$ are based on a finite number of samples and thus differ from the true probabilities $P^\alpha_{\rm i}$. In this case, $\Phi_2(T)$ should get smaller at larger T and eventually vanish. However, if there are true individual differences, then $\Phi_2(T\to\infty)$ will measure the variance of these differences. Formally,

$$\langle\Phi_2(T)\rangle = \Phi_{2c}(T) + \Phi_{2,{\rm ind}}. \tag{5}$$

Here the first term comes from the (connected) correlations in the behaviors of individuals in Eq. (1),

$$\Phi_{2c}(T) = \frac{1}{N_f}\sum_{\alpha=1}^{N_f}\left(\frac{\Delta t}{T}\right)^2\sum_{t=1}^{T}\sum_{t'=1}^{T} C^\alpha(t-t'), \tag{6}$$

where we assume stationarity, while the second term comes from individual differences,

$$\Phi_{2,{\rm ind}} = \frac{1}{N_f}\sum_{\alpha=1}^{N_f}\sum_{\rm i=1}^{N_s}\left[P^\alpha_{\rm i} - \frac{1}{N_f}\sum_{\beta=1}^{N_f}P^\beta_{\rm i}\right]^2. \tag{7}$$

We notice that $\Phi_{2c}(T)$ is the double integral of an underlying correlation function, in the same way that the mean-square displacement of a diffusing particle is the double integral of the velocity (auto)correlation function [41]. As with diffusion, long-ranged behavior in $C(\tau)$ should appear as an anomalous behavior of the $\Phi_{2c}(T)$.

If correlations in the behavioral trajectories of individuals are short-ranged, then at value of T larger than the correlation time we will see the connected part $\Phi_{2c}(T) \sim 1/T$. This corresponds to the intuition that

variances should decay as the inverse of the number of independent samples. On the other hand, if correlations are long-ranged, with a power-law decay $C^{\alpha}(\tau \to \infty) \sim 1/|\tau|^{2\Delta}$, then $\Phi_{2c}(T) \sim 1/T^{2\Delta}$ at large T. Power-law decays are referred to colloquially as scale invariant because there is no characteristic time over which the correlations decay. However, scale invariance is more than the observation of a single power law. If behavior is determined by an internal variable that undergoes scale-invariant fluctuations, then we should see related power laws in different moments of these fluctuations [5]. Concretely, we define

$$\Phi_n(T) = \sum_{i=1}^{N_s} \frac{1}{N_f} \sum_{\alpha=1}^{N_f} \left[\hat{P}_i^{\alpha}(T) - \frac{1}{N_f} \sum_{\beta=1}^{N_f} \hat{P}_i^{\beta}(T) \right]^n, \quad (8)$$

and the prediction of scale invariance is that

$$\langle \Phi_n(T \to \infty) \rangle = A_n \left(\frac{\Delta t}{T} \right)^{\gamma_n} + B_n, \quad (9)$$

with $\gamma_n = n\Delta$, and Δ is the "scaling dimension".

We use these ideas to analyze the behavioral state sequences in walking flies described above [11,12]. To estimate $\Phi_2(T)$, we follow these steps: (1) Choose a random window of duration T from the recording of each individual and estimate the state probabilities. (2) Compute the summed variance of these probabilities across 1000 random halves of the individuals. (3) Use these many random halves to compute a mean and standard error for $\Phi_2(T)$. Results are in Fig. 1.

On a linear scale (Fig. 1, left), we see a rapid, subsecond decay of $\Phi_2(T)$. On a semilogarithmic plot (Fig. 1, center) we see gradual decay out to 1 min, but to reveal the full behavior we need a doubly logarithmic plot (Fig. 1, right). This spans five decades in time, which is equivalent to measuring correlations between letters from neighboring letters out to the length of a short story. Beyond ~ 1 s there is no sign of a characteristic timescale and no clear sign of a plateau at long times, suggesting that genuine individual differences are small. The decay is much slower than $\Phi_2 \sim 1/T$, suggesting that the system has long-ranged correlations. Indeed, the data for $T > 7$ s are an excellent fit to the prediction of Eq. (9).

The choice to fit only $T > 7$ s is motivated both by the appearance of $\Phi_2(T)$ and by the fact that mean residence times in individual states can be as long as ~ 7 s [11]. Thus, for $T < 7$ s we are probing, in part, the structure of waiting time distributions for transitions out of single states, while for $T > 7$ s we are sensitive mostly to recurrences where the fly returns to a previously visited state. We emphasize that, in problems we understand, scaling is asymptotic [4,6], and so we expect that scaling behaviors are clearer at larger T. Indeed, it is surprising that scaling becomes a good description almost at the smallest possible values of T.

If our description of $\Phi_2(T)$ is correct, then we can subtract our estimate of the variance across individuals [B_2 in Eq. (9)] to reveal a "clean" power-law decay, as shown at left in Fig. 2. We see that, over three decades in time, all of the data points are within errors of a power law with exponent $\gamma_2 = 2\Delta$, $\Delta = 0.180 \pm 0.004$. Since our definition of states uses features on the ~ 1 s timescale [11], and our recordings are ~ 1 h in duration, it is impossible to see a "better" power-law in these data.

In the same way that individual differences can mimic long-ranged correlations, these correlations can mimic individual differences. Naively, if we estimate state probabilities over the hour long experiment, there is a variance across individuals that is $\sim 5\times$ larger than our best estimate of $\Phi_{2,\mathrm{ind}}$. It would be tempting to interpret this as biological variability, but this assumes that averaging over one hour is enough to push the statistical fluctuations below the individual variations. Because of long-ranged correlations, this turns out not to be true. Our best estimate is that individual differences contribute less than 1% of the total variance in behavior.

Following the discussion above, we can do the same analysis for $\Phi_3(T)$, with the results shown at right in Fig. 2 (blue), and again the fit to Eq. (9) is excellent. We have tried

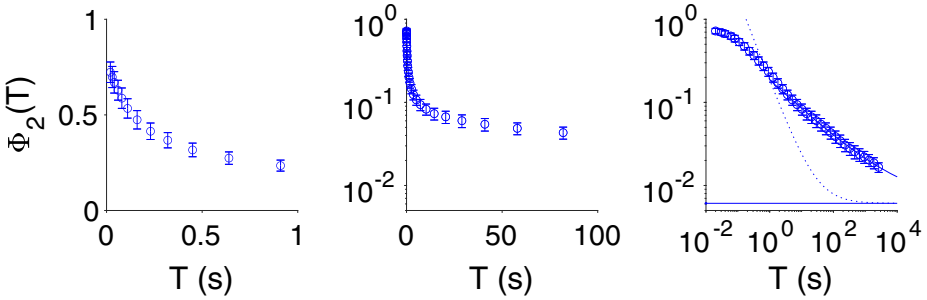

FIG. 1. Estimates of the summed variance in state probabilities for walking flies as a function of averaging time; raw data from Refs. [11,12]. Means and standard errors computed from random halves of the data. Left: linear plot. Center: semilog plot. Right: log-log plot; solid line is Eq. (9) for $n = 2$ and $T > 7$ s, with parameters in Table I. Horizontal line shows the estimated plateau $B_2 = \Phi_{2,\mathrm{ind}}$, and dotted line is a decay $\sim 1/T$ to the same plateau, as expected if correlation times were short.

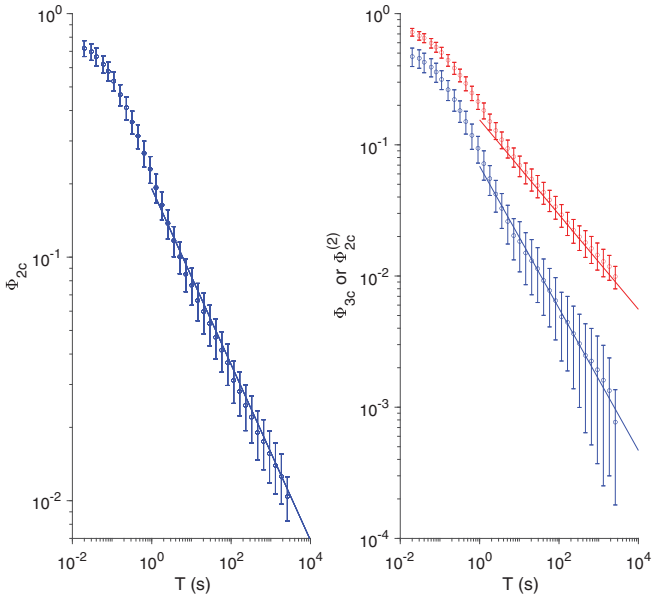

FIG. 2. Summed variance in state probabilities as a function of averaging time, with background subtracted. Left: $\Phi_{2c}(T)$. Right: $\Phi_{3c}(T)$ (blue) and $\Phi_{2c}^{(2)}(T)$ (red). Error bars include the standard error computed from random halves of the data, as in Fig. 1, and uncertainty in the background. Lines are best fits as in Fig. 1, parameters from Table I.

to analyze $\Phi_4(T)$, but the error bars are too large for this to be informative. If there is a single underlying scale-invariant process, then the exponents for $\Phi_2(T)$ and $\Phi_3(T)$ should be $\gamma_2 = 2\Delta$ and $\gamma_3 = 3\Delta$. Thus, the different moments provide independent estimates of the underlying scaling dimension Δ, and Table I shows that these estimates agree to the third decimal place.

An important lesson from statistical physics is that structures on long time or length scales are robust to a range of choices in defining variables on smaller scales. To test this idea, we redefine the "state" of the system to be the combined states at two successive moments in time. Now there are $N_s^2 \sim 15\,000$ possible states, of which ~1200 occur in our sample of 2.1×10^7 frames. We can estimate the probability for each of these states, as before, from data in windows of duration T and define the variance across individuals for these "two-frame states," $\Phi_2^{(2)}(T)$. We expect that

TABLE I. Parameter estimates for different measures of variation across individuals. The underlying scaling dimensions Δ should agree, while amplitudes A and backgrounds B differ.

	Φ_2	Φ_3	$\Phi_2^{(2)}$
Δ	0.180 ± 0.004	0.181 ± 0.005	0.179 ± 0.004
A	0.936 ± 0.054	0.807 ± 0.047	0.839 ± 0.048
B	0.0060 ± 0.0003	0.0010 ± 0.0001	0.0050 ± 0.0003

$$\langle \Phi_2^{(2)}(T) \rangle \sim A_2^{(2)} \left(\frac{\Delta t}{T} \right)^{2\Delta} + B_2^{(2)}, \tag{10}$$

where Δ is the same scaling dimension as in $\Phi_2(T)$ and $\Phi_3(T)$, and at right in Fig. 2 (red) this is confirmed.

In summary, behavioral correlations exhibit precise power-law scaling over three decades in time. To attach errors to our estimates of exponents, we use the variance across random halves of the data to construct χ^2 between the predictions of Eq. (9) and the data, then take Monte Carlo samples from the distribution of fitting parameters $\propto \exp(-\chi^2/2)$. The result is that measurements of three different correlation functions are consistent with a single underlying scaling field, and estimates of the scaling dimension all agree within the 2%–3% experimental errors (Table I).

Temporal correlations that decay as a power law are mathematically equivalent to a mixture of many independent processes happening on different timescales. These could be trivially independent, as with the electron trapping processes that generate $1/f$ noise in metals [44], or they could be the "normal modes" of an interacting network so that the spectrum of timescales is an emergent property of the network dynamics [45]. However, true scale invariance is more than a single power-law decay. In particular, the consistency of scaling dimensions in three different correlation measures is much more difficult to account for with simple mixtures of timescales.

A complementary view is to think about the evolution operator for the probability distribution over states, what would be the diffusion or Fokker-Planck equation in simple cases [46]. Then the relevant modes are the eigenfunctions of this operator, and the power-law decay of the correlation function tells us that the corresponding eigenvalue spectrum has a near continuum extending toward zero. In general, the connection between correlation functions and eigenvalue spectra depends on the structure of the eigenfunctions, and again the constraints required for different correlation functions to have consistent exponents seem highly nontrivial.

Early evidence for scale-invariant temporal correlations in human behavior was hidden in Shannon's 1951 experiments on the prediction of text by human subjects [47,48]. What is surprising here is the precision with which scaling emerges from the data, at least in this one example. We see clean power-law behavior across three decades, from seconds to one hour; errors on exponents are in the third decimal place, and different correlation functions are described by exponents that are consistent with one another, within these small errors, on the hypothesis that there is a single scale-invariant process controlling behavior. Still, we are doing this analysis on data which now are nearly a decade old [11]. The next generation of experiments will make continuous measurements of behavior at the same resolution for weeks rather than for one hour [49].

While it will be necessary to separate circadian and secular variations, these experiments have the potential to test for scaling over six decades and to reduce errors in the seconds to hours range by more than an order of magnitude, getting close to the precise tests of scaling in equilibrium critical phenomena [50].

We thank G. J. Berman, M. P. Brenner, A. Cavagna, I. Giardina, O. Kimchi, R. Pang, and F. C. Pereira for helpful discussions. This work was supported in part by the National Science Foundation through the Center for the Physics of Biological Function (PHY–1734030), the National Institutes of Health (R01EB026943 and R01NS104899), the John Simon Guggenheim Memorial Foundation, and the Simons Foundation.

[1] R. Dawkins, Hierarchical organization: A candidate principle for ethology, in *Growing Points in Ethology*, edited by P. Bateson and R. Hinde (Cambridge University Press, Cambridge, England, 1976), pp. 7–54.
[2] H. A. Simon, The architecture of complexity, Proceedings of the American Philosophical Society **106**, 467 (1982), https://www.jstor.org/stable/985254.
[3] It is hard to find a single origin for this idea in the context of human or animal behavior. Interest was stimulated by the discovery of self-organized criticality in the mid 1980s, leading to provocative discussions, e.g., P. Bak, *How Nature Works* (Springer, New York, 1996).
[4] K. G. Wilson, Problems in physics with many scales of length, Sci. Am. **241**, No. 2, 158 (1979).
[5] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, England, 1996).
[6] J. Sethna, *Statistical Mechanics: Entropy, Order Parameters and Complexity* (Oxford University Press, New York, 2006).
[7] L. C. Osborne, S. G. Lisberger, and W. Bialek, A sensory source for motor variation, Nature (London) **437**, 412 (2005).
[8] G. J. Stephens, B. Johnson-Kerner, W. Bialek, and W. S. Ryu, Dimensionality and dynamics in the behavior of C. Elegans, PLoS Comput. Biol. **4**, e1000028 (2008).
[9] T. Ahamed, A. C. Costa, and G. J. Stephens, Capturing the continuous complexity of behaviour in *Caenorhabditis elegans*, Nat. Phys. **17**, 275 (2021).
[10] K. Branson, A. A. Robie, J. Bender, P. Perona, and M. H. Dickinson, High-throughput ethomics in large groups of *Drosophila*, Nat. Methods **6**, 451 (2009).
[11] G. J. Berman, D. M. Choi, W. Bialek, and J. W. Shaevitz, Mapping the stereotyped behaviour of freely moving fruit flies, J. R. Soc. Interface **11**, 20146072 (2014).
[12] G. J. Berman, W. Bialek, and J. W. Shaevitz, Predictability and hierarchy in *Drosophila* behavior, Proc. Natl. Acad. Sci. U.S.A. **113**, 11943 (2016).
[13] A. B. Wiltschko, M. J. Johnson, G. Iurilli, R. E. Peterson, J. M. Katon, S. L. Pashkovski, V. E. Abraira, R. P. Adams, and S. R. Datta, Mapping sub-second structure in mouse behavior, Neuron **88**, 1 (2015).
[14] J. D. Marshall, D. E. Aldarondo, T. W. Dunn, W. L. Wang, G. J. Berman, and B. P. Ölveczky, Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire, Neuron **109**, 420 (2021).
[15] G. Reddy, L. Desban, H. Tanaka, J. Rouseel, O. Mirat, and C. Wyart, A lexical approach for identifying behavioural action sequences, PLoS Comput. Biol. **18**, e1009672 (2020).
[16] R. E. Johnson, S. Linderman, T. Panier, C. L. Wee, E. Song, K. J. Herrera, A. Miller, and F. Engert, Probabilistic models of larval zebrafish behavior reveal structure on many scales, Curr. Biol. **30**, 70 (2020).
[17] M. Ghosh and J. Rihel, Hierarchical compression reveals sub-second to day-long structure in larval zebrafish behavior, eNeuro **7**, 0408 (2020).
[18] J. E. Markowitz, E. Ivie, L. Kligler, and T. J. Gardner, Long-range order in canary song, PLoS Comput. Biol. **9**, e1003052 (2013).
[19] A. Kershenbaum, A. E. Bowles, T. M. Freeberg, D. Z. Jin, A. R. Lameira, and K. Bohn, Animal vocal sequences: Not the Markov chains we thought they were, Proc. R. Soc. B **281**, 20141370 (2014).
[20] A. Kershenbaum *et al.*, Acoustic sequences in non-human animals: A tutorial review and prospectus, Biol. Rev. Camb. Philos. Soc. **91**, 13 (2016).
[21] A. Mathis, P Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, and M. Bethge, DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning, Nat. Neurosci. **21**, 1281 (2018).
[22] T. Pereira, D. Aldarondo, L. Willmore, M. Kislin, S. S. Wang, M. Murthy, and J. W. Shaevitz, Fast animal pose estimation using deep neural networks, Nat. Methods **16**, 117 (2019).
[23] T. D. Pereira *et al.*, SLEAP: A deep learning system for multi-animal pose tracking, Nat. Methods **19**, 486 (2022).
[24] A. E. X. Brown and B. de Bivort, Ethology as a physical science, Nat. Phys. **14**, 653 (2017).
[25] G. J. Berman, Measuring behavior across scales, BMC Biol. **16**, 23 (2018).
[26] S. R. Datta, D. J. Anderson, K. Branson, P. Perona, and A. Leifer, Computational neuroethology: A call to action, Neuron **104**, 11 (2019).
[27] T. D. Pereira, J. W. Shaevitz, and M. Murthy, Quantifying behavior to understand the brain, Nat. Neurosci. **23**, 1537 (2020).
[28] M. W. Mathis and A. Mathis, Deep learning tools for the measurement of animal behavior in neuroscience, Curr. Opin. Neurobiol. **60**, 1 (2020).
[29] W. Bialek, On the dimensionality of behavior, Proc. Natl. Acad. Sci. U.S.A. **119**, e2021860119 (2022); arXiv:2008.09574.
[30] V. Alba, G. J. Berman, W. Bialek, and J. W. Shaevitz, Exploring a strongly non-Markovian animal behavior, arXiv:2012.15681.
[31] A. Cavagna, I. Giardina, and T. S. Grigera, The physics of flocking: Correlation as a compass from experiments to theory, Phys. Rep. **728**, 1 (2018).
[32] G. K. Zipf, *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Reading, MA, 1949).

[33] T. Mora and W. Bialek, Are biological systems poised at criticality?, J. Stat. Phys. **144**, 268 (2011).
[34] G. Tkačik, T. Mora, O. Marre, D. Amodei, M. J. Berry II, and W. Bialek, Thermodynamics and signatures of criticality in a network of neurons, Proc. Natl. Acad. Sci. U.S.A. **112**, 11508 (2015).
[35] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. G. E. da Luz, E. P. Raposo, and H. E. Stanley, Optimizing the success of random searches, Nature (London) **401**, 911 (1999).
[36] A. Proekt, J. R. Banavar, A. Maritan, and D. W. Pfaff, Scale invariance in the dynamics of spontaneous behavior, Proc. Natl. Acad. Sci. U.S.A. **109**, 10564 (2012).
[37] J. M. Beggs and D. Plenz, Neuronal avalanches in neocortical circuits, J. Neurosci. **23**, 167 (2003).
[38] J. M. Beggs and D. Plenz, Neuronal avalanches are diverse and precise patterns of activity that are stable for many hours in cortical slice cultures, J. Neurosci. **24**, 5216 (2004).
[39] N. Friedman, S. Ito, B. Brinkman, M. Shimono, R. L. DeVille, K. Dahmen, J. M. Beggs, and T. C. Butler, Universal critical dynamics in high resolution neuronal avalanche data, Phys. Rev. Lett. **108**, 208102 (2012).
[40] P. Bak, C. Tang, and K. Wiesenfeld, Self-organized criticality: An explanation of the $1/f$ noise, Phys. Rev. Lett. **59**, 381 (1987).
[41] J. S. Lucas, Y. Zhang, O. K. Dudko, and C. Murre, 3D trajectories adopted by coding and regulatory DNA elements: First-passage times for genomic interactions, Cell **158**, 339 (2014).
[42] D. G. Hernández, C. Rivera, J. Cande, B. Zhou, D. L. Stern, and G. J. Berman, A framework for studying behavioral evolution by reconstructing ancestral repertoires, eLife **10**, e61806 (2021).
[43] The same issue arises in analyzing the spatial correlations of velocity fluctuations in a flock of birds [31].
[44] P. Dutta and P. H. Horn, Low-frequency fluctuations in solids: $1/f$ noise, Rev. Mod. Phys. **53**, 497 (1981).
[45] For modes in neural networks, and especially the emergence of modes that can carry correlations over long times, see, e.g., H. S. Seung, How the brain keeps the eyes still, Proc. Natl. Acad. Sci. U.S.A. **93**, 13339 (1996).
[46] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer-Verlag, Berlin, 1984).
[47] C. E. Shannon, Prediction and entropy of written English, Bell Syst. Tech. J. **30**, 50 (1951).
[48] W. Hilberg, Der bekannte Grenzwert der redundanzfreien Information in Texten: Eine Fehlinterpretation der Shannonschen Experimente?, Frequenz **44**, 243 (1990).
[49] G. McKenzie-Smith, S. W. Wolf, J. F. Ayroles, and J. W. Shaevitz, Capturing continuous, long timescale behavioral changes, in Drosophila melanogaster postural data, arXiv: 2309.04044.
[50] J. A. Lipa, D. R. Swanson, J. A. Nissen, T. C. P. Chui, and U. S. Israelsson, Heat capacity and thermal relaxation of bulk helium very near the lambda point, Phys. Rev. Lett. **76**, 944 (1996).