Letter

Logarithmic dispersion at marginal stability in collective attention

Keisuke Okamura*
Ministry of Education, Culture, Sports, Science and Technology, Tokyo, Japan

(Received 18 April 2026; accepted 23 October 2026; published 12 November 2026)

How cross-sectional heterogeneity evolves in nonequilibrium populations of competing, aging units is a collective question that cannot be resolved from a representative trajectory or the ensemble mean alone. Analysis of more than two million age-aligned English Wikipedia pageview trajectories from six yearly cohorts spanning 2015–2020 shows that the variance of logarithmic attention grows linearly with $\ln t$, while the ensemble mean relaxes on the same logarithmic clock. This dispersion law recurs across all six cohorts and is systematically favored over the tested power-law alternative; temporal increments also exhibit consistent antipersistent scaling. A complementary trajectory-level analysis reveals broad, non-Gaussian variation in effective relaxation rates, showing that this collective regularity coexists with heterogeneous observed trajectories. A minimal scale-free stochastic model combines mean relaxation with self-affine temporal fluctuations of age-dependent amplitude. Within this model class, logarithmic spreading emerges when fluctuation accumulation balances aging attenuation, defining the marginal boundary between algebraic broadening and saturation. Together, age-resolved cross-sectional dispersion and temporal-increment scaling provide a transferable ensemble-level criterion for distinguishing amplification, marginal spreading, and saturation in other populations of competing, aging units.

DOI: 10.1103/c9ny-57bd

Introduction. In many nonequilibrium systems, individual trajectories and their ensemble average provide the conventional starting point for describing populations of competing units. They characterize unit-level evolution and mean behavior, whereas ensemble dispersion—quantified here by the variance—captures how heterogeneity develops across the population. Yet the scaling of this dispersion with age has received comparatively little attention, despite its potential to distinguish qualitatively different collective regimes of amplification, marginal spreading, and saturation.

Collective attention provides a particularly well-resolved empirical setting for this question: the activity of many items competing for visibility can be followed simultaneously over time. Across information environments, collective attention exhibits bursts, cascade structures, and rapid turnover [1–3]. At the level of individual items, existing descriptions include stochastic relaxation and novelty decay [4,5], self-exciting or epidemic cascades [6,7], and preferential or multiplicative growth mechanisms [8,9]. Models with explicit competition among multiple items further explain heterogeneity and critical popularity statistics under limited attention [10,11]. These approaches, however, do not determine how cross-sectional dispersion evolves with age. Consequently, trajectory-level relaxation laws or static popularity statistics alone do not locate the marginal boundary between amplification and saturation considered here.

Here we address this problem using English Wikipedia page views, a large-scale observational record of collective attention across a vast information space. Across six yearly cohorts from 2015 to 2020, we find that the ensemble mean and variance consistently evolve on the same logarithmic clock: the variance of logarithmic attention grows linearly with $\ln t$, while its mean relaxes proportionally to $-\ln t$. Temporal increments in all six cohorts also exhibit consistent antipersistent scaling. For the 2020 cohort examined at the trajectory level, finite-trajectory estimates of effective decay rates form a broad, non-Gaussian distribution. Together, these observations show that reproducible ensemble-level scaling coexists with pronounced heterogeneity among observed trajectories.

To account for these ensemble laws, we formulate a minimal scale-free stochastic model for logarithmic attention that combines self-affine temporal correlations with an age-dependent fluctuation amplitude. Using fractional Brownian motion (fBM) as a parsimonious Gaussian reference for the measured second-order correlations [12], the model introduces aging attenuation into the constant-amplitude setting previously considered for citation dynamics [13]. Within this model class, logarithmic dispersion emerges when self-affine fluctuation accumulation balances aging attenuation, marking the boundary between algebraic broadening and saturation. The resulting description locates marginality at the level of ensemble dispersion rather than representative article dynamics and provides a transferable diagnostic, based on age-resolved dispersion and temporal-increment scaling, for other populations in which activity is allocated competitively.

Data and empirical scaling. To construct the age-aligned ensembles analyzed here, we combine the English Wikipedia "Topviews" rankings with pageview records for 2015–2020 [14,15]. For each reference day d, we select the 1000 most-viewed articles and follow their activity over the subsequent $T = 365$ days. Because the Pageviews API record begins on July 1, reference days cover July–December in 2015 and the

*Contact author: okamura@alumni.lse.ac.uk

 ©2026 American Physical Society

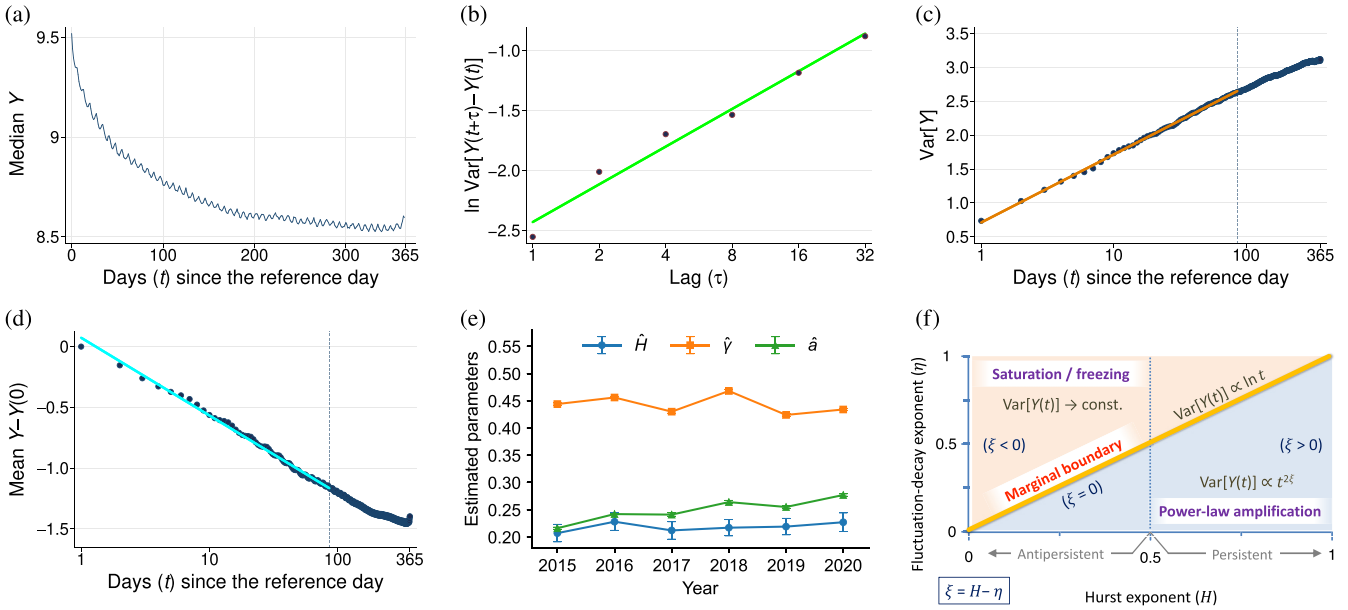


FIG. 1. Empirical scaling and the scale-free phase portrait of collective attention. (a) Median postselection trajectory across the 2020 trajectory episodes, showing slow decay with weekly modulation. (b) Pooled increment-variance scaling $\mathrm{Var}[Y_{t+\tau} - Y_t] \propto \tau^{2H}$, yielding $H \approx 0.22$. (c) Ensemble variance as a function of $\ln t$; the fitted linear regime gives $\mathrm{Var}[Y_t] = \gamma \ln t + \mathrm{const}$. (d) Ensemble mean relative to the selection-day value; its linear dependence on $\ln t$ gives $\langle Y_t - Y_0 \rangle = -a \ln t + \mathrm{const}$. (e) Estimated H, γ, and a across the 2015–2020 cohorts. (f) Phase diagram of the scale-free model introduced below. For a fluctuation amplitude $\beta(t) \propto t^{-\eta}$, the coordinate $\xi = H - \eta$ separates algebraic amplification ($\xi > 0$), marginal logarithmic spreading ($\xi = 0$), and saturation ($\xi < 0$). Dotted vertical lines in (c) and (d) mark the fitted endpoint $t = t_c$.

full calendar year thereafter. Further data-construction details and cohort sizes are provided in Sec. S1A and Table S1 of the Supplemental Material (SM) [16].

For a trajectory episode, let V_t denote the page views at elapsed day t, with V_0 denoting the selection-day count, and define

$$X_t = V_t + 1, \quad Y_t = \ln X_t. \tag{1}$$

The unit shift retains zero-view days while having negligible effect when pageview counts are large. Each reference-day–article pair is treated as a distinct trajectory episode, so repeated appearances of the same article represent distinct episodes of salience. All moment curves are calculated across these trajectory episodes at fixed elapsed time. The resulting ensemble laws therefore characterize the postselection evolution of salient articles rather than preselection dynamics or the full population of Wikipedia pages.

Figure 1(a) shows the typical postselection relaxation together with weekly modulation. We estimate temporal roughness from increments, which reduce the influence of the slowly varying mean. For each dyadic lag $\tau \in \{1, 2, 4, 8, 16, 32\}$, increments are pooled across trajectory episodes and all admissible postselection elapsed times within each yearly cohort. Their empirical variance is well described by the cohort-level lag scaling

$$\mathrm{Var}[Y_{t+\tau} - Y_t] \propto \tau^{2H}. \tag{2}$$

For 2020, the fit gives $\hat{H} \approx 0.22$ [Fig. 1(b)]; across all six cohorts, the estimates lie in the narrow range $\hat{H} \approx 0.21$–0.23 (Fig. 1(e) and Sec. S1B and Fig. S1 of the SM [16]). Within an fBM reference description, the sublinear exponent $2\hat{H} < 1$ corresponds to antipersistent fluctuations and sample paths rougher than Brownian motion. Equation (2) should be read as an effective finite-window lag scaling; its geometric interpretation and relation to the local increment law under an age-dependent fluctuation amplitude are detailed in Secs. S2D and S2E of the SM [16].

The principal empirical observable is the cross-sectional variance of Y_t. As shown for 2020 in Fig. 1(c), it grows linearly with logarithmic age,

$$\mathrm{Var}[Y_t] = \gamma \ln t + \mathrm{const}, \tag{3}$$

over the data-adaptive range $1 \leqslant t \leqslant t_c$. The endpoint t_c is determined by a single-breakpoint least-squares analysis of the variance curve and then held fixed for the mean fit and the power-law comparison (Sec. S1C of the SM [16]). A direct comparison with a two-parameter power law over the same fitted interval favors the logarithmic form in every yearly cohort (Sec. S1D and Table S2 of the SM [16]). The corresponding cohortwise variance curves are shown in Fig. S2 of the SM [16].

The mean evolves on the same logarithmic time axis [Fig. 1(d)]:

$$\langle Y_t - Y_0 \rangle = -a \ln t + \mathrm{const}. \tag{4}$$

The additive constant absorbs the difference between the selection-day value and the first postselection observation used in the scaling fit. Equation (4) corresponds to an effective cohort-level drift proportional to $-1/t$. Across years, $\hat{H}$ and $\hat{\gamma}$ remain comparatively stable, whereas $\hat{a}$ shows a modest overall upward tendency, increasing from 0.216 in 2015 to 0.277 in 2020 despite nonmonotonic year-to-year variation

(Fig. 1(e) and Sec. S1E, Fig. S3, and Table S3 of the SM [16]). Larger values of a correspond to faster mean relaxation, so this overall tendency indicates a steeper decline in the later cohorts.

The coefficients a, γ, and H describe complementary aspects of the dynamics: a governs relaxation of the distribution center, γ measures the logarithmic growth rate of cross-sectional variance, and H characterizes temporal-increment scaling. Together, the common logarithmic clock of the first two ensemble moments and the cross-year stability of H and γ define a reproducible ensemble structure even as the mean-relaxation rate varies.

Scale-free stochastic model. The logarithmic variance growth in Eq. (3) is the central empirical clue. The observed growth is slower than the algebraic broadening generated by unattenuated self-affine fluctuations, yet shows no flattening toward saturation over the fitted window. This intermediate behavior motivates a scale-free stochastic model for logarithmic attention.

Let $X(t) > 0$ denote the continuous-time model counterpart of the empirical positive variable $X_t = V_t + 1$ defined in Eq. (1), and set $Y(t) := \ln X(t)$. We write its evolution as

$$Y(t) - Y(t_0) = \int_{t_0}^{t} \alpha(s)\,\mathrm{d}s + \int_{t_0}^{t} \beta(s)\,\mathrm{d}B_H(s), \quad t \geqslant t_0 > 0. \tag{5}$$

Here B_H is fBM with Hurst exponent $H \in (0, 1)$ [12], $\alpha(t)$ is a deterministic drift, and $\beta(t)$ is an age-dependent fluctuation amplitude. We identify t_0 with the first postselection day entering the scaling analysis; although $t_0 = 1$ day empirically, we retain it symbolically to keep dimensionless time ratios and scale transformations explicit. The stochastic construction and its relation to the positive process $X(t)$ are detailed in Sec. S2A of the SM [16].

The empirical logarithmic relaxation of the mean corresponds to $\alpha(t) = -a/t$, while the effective increment scaling in Eq. (2) provides the estimate of the self-affine exponent H used below. To allow the fluctuations to weaken with age without introducing a characteristic time, we use the minimal scale-free ansatz

$$\beta(t) = \frac{b}{t_0^H}\left(\frac{t}{t_0}\right)^{-\eta}, \quad \eta \geqslant 0, \tag{6}$$

where $b > 0$ sets the overall fluctuation scale; the factor t_0^{-H} fixes its normalization relative to the reference time. Equation (6) provides a phenomenological interpolation between unattenuated fBM fluctuations at $\eta = 0$ and progressively stronger attenuation at larger η. Writing

$$I(t; t_0) := \int_{t_0}^{t} \beta(s)\,\mathrm{d}B_H(s), \tag{7}$$

the deterministic drift gives

$$\mathrm{Var}[Y(t) - Y(t_0)] = \mathrm{Var}[I(t; t_0)]. \tag{8}$$

Under the simplifying ensemble assumption that $Y(t_0)$ is uncorrelated with the subsequent stochastic accumulation, the initial cross-sectional variance contributes an additive baseline to $\mathrm{Var}[Y(t)]$; the general covariance relation is given in Sec. S2A of the SM [16].

The competition between fluctuation accumulation and aging attenuation becomes transparent by comparing intervals of fixed logarithmic length $[t, \rho t]$, with $\rho > 1$. The stochastic contribution over such an interval is the increment

$$I_\rho(t) := I(\rho t; t_0) - I(t; t_0) = \frac{b}{t_0^H}\int_t^{\rho t} \left(\frac{s}{t_0}\right)^{-\eta} \mathrm{d}B_H(s). \tag{9}$$

Self-similarity of B_H supplies a factor t^{2H} to the variance, whereas the homogeneous amplitude supplies a factor $t^{-2\eta}$. Consequently,

$$\mathrm{Var}[I_\rho(t)] = \left(\frac{t}{t_0}\right)^{2(H-\eta)} \mathrm{Var}[I_\rho(t_0)]. \tag{10}$$

The net exponent governing the magnitude of fluctuations accumulated per logarithmic time interval is therefore

$$\xi := H - \eta. \tag{11}$$

Equivalently, under the dilation $t \mapsto \rho t$ the root-mean-square fluctuation accumulated over a fixed logarithmic interval rescales by ρ^ξ; $\xi = 0$ is marginal in the precise sense that this fluctuation scale is invariant under dilations of age. Thus, ξ serves as the phase coordinate distinguishing the accumulation regimes.

A logarithmic-time decomposition, including correlations between successive intervals, yields (Sec. S2B of the SM [16])

$$\mathrm{Var}[I(t; t_0)] \sim \begin{cases} C_+ (t/t_0)^{2\xi}, & \xi > 0, \\ C_0 \ln(t/t_0), & \xi = 0, \\ C_-, & \xi < 0, \end{cases} \tag{12}$$

where C_+, C_0, and C_- are positive constants. For $\xi > 0$, successive logarithmic time intervals contribute fluctuations of increasing magnitude, producing algebraic broadening. For $\xi < 0$, their contributions weaken with age and the accumulated variance saturates. At $\xi = 0$, intervals related by a fixed rescaling contribute on the same scale. The corresponding increments then form a stationary sequence in logarithmic time with summable correlations, and their cumulative variance gives the ultraslow $\ln(t/t_0)$ law (Sec. S2B of the SM [16]). The logarithmic law is therefore not postulated separately, but emerges as the marginal member of the same scale-free family that produces algebraic broadening for $\xi > 0$ and saturation for $\xi < 0$. Under the ensemble assumption stated above, these are also the asymptotic regimes of the age-dependent part of $\mathrm{Var}[Y(t)]$, with the initial variance absorbed into the additive constant in Eq. (3).

The two empirical measurements play complementary roles: temporal-increment scaling estimates H, whereas the evolution of cross-sectional variance identifies the accumulation regime [Figs. 1(b) and 1(c)]. Within this ansatz, the logarithmic variance places the ensembles near $\xi = 0$; combining this result with the measured H then gives $\eta \simeq H$ without imposing the balance *a priori*. Across cohorts, the variance curves are consistent with marginality over the fitted windows, showing neither the systematic upward curvature predicted for $\xi > 0$ nor the flattening predicted for $\xi < 0$ (Sec. S2C and Table S4 of the SM [16]).

The resulting phase portrait [Fig. 1(f)] is dynamically reminiscent of spreading and arrest in glassy systems [17]:

"liquidlike" amplification for $\xi > 0$ and "frozen" saturation for $\xi < 0$ are separated by ultraslow spreading at the marginal boundary $\xi = 0$. These labels refer to contrasting regimes of ensemble-dispersion dynamics rather than thermodynamic phases and provide a structural language for comparing amplification, marginal spreading, and saturation across nonequilibrium aging systems.

At marginality, setting $\eta = H$ cancels the reference-time factors in Eq. (6) and gives the compact model

$$dY(t) = -\frac{a}{t}\,dt + \frac{b}{t^H}\,dB_H(t). \tag{13}$$

Equation (13) is shorthand for the canonical Wiener-integral construction in Eq. (5), as detailed in Sec. S2A of the SM [16]. The same empirically estimated exponent H thus characterizes the self-affine driver and the aging of its amplitude. Within this marginal ensemble description, the empirical slope γ corresponds to the coefficient C_0 in Eq. (12). At fixed H and fBM normalization, $C_0 \propto b^2$; hence, once H is determined from increment scaling, the measured γ calibrates the fluctuation scale b (Sec. S2B of the SM [16]).

Summary and discussion. The central result is a reproducible logarithmic law for cross-sectional dispersion in collective attention, observed across all six yearly cohorts alongside logarithmic mean relaxation and stable antipersistent increment scaling. Within the scale-free model, this dispersion law marks the marginal balance between self-affine fluctuation accumulation and the age-dependent attenuation of fluctuation amplitude. At this balance, each fixed multiplicative interval in age contributes fluctuations of the same order, so that the variance grows in proportion to the number of such intervals and hence logarithmically with age; weaker or stronger attenuation instead produces algebraic broadening or saturation, respectively. Because both the mean drift and the fluctuation amplitude decrease algebraically with elapsed age, the effective pace of evolution slows as an article ages. This explicit age dependence provides a concrete connection to aging dynamics and anomalous transport [18].

Although Eq. (13) is written at the trajectory level, it is calibrated to the observed first- and second-order ensemble laws and should not be read as a complete generative model for every article. The distribution of fitted trajectory-level relaxation coefficients is broad and markedly non-Gaussian (Secs. S3A–S3C and Fig. S4 of the SM [16]). This pronounced heterogeneity coexists with cross-year stable values of H and γ, showing that reproducible second-order scaling does not require trajectory-level uniformity. The logarithmic law thus expresses collective statistical organization across heterogeneous attention histories: marginality characterizes the evolution of ensemble dispersion, not individual articles.

The cross-cohort comparison further separates mean relaxation from dispersion dynamics. The estimates of H and γ remain comparatively stable, whereas the coefficient a exhibits a modest overall upward tendency. Because a enters only through the cohort-level drift $\alpha(t) = -a/t$, its variation changes the relaxation of the distribution center without, by itself, changing the accumulation coordinate ξ. The larger values of a in the later cohorts are consistent with the acceleration of collective attention dynamics reported across information environments [3], although the nonmonotonic sequence of six cohorts is insufficient to establish a secular trend or identify its cause.

The comparison with citation dynamics clarifies the role of the age-dependent fluctuation amplitude. In Ref. [13], a constant amplitude allowed self-affine temporal correlations to produce algebraic growth of the cross-sectional variance of logarithmic citation counts. Here, scale-free attenuation balances self-affine accumulation, yielding logarithmic growth of the cross-sectional variance of log-transformed page views. Thus, the temporal correlation exponent alone does not determine whether dispersion broadens algebraically, grows logarithmically, or saturates; the age dependence of the fluctuation amplitude is equally decisive.

What selects the apparent balance $\eta \simeq H$ remains open. Competition for finite collective attention may contribute, but the present observations do not distinguish external resource constraints from platform-mediated exposure, endogenous feedback, or other mechanisms. A microscopic derivation of Eq. (13) would therefore need to account for both the correlated effective noise and the aging of its amplitude.

Several qualifications delimit the present conclusion. The analysis concerns postselection episodes of salient articles in English Wikipedia rather than the full population of pages, and the six cohorts establish temporal reproducibility within one information environment rather than replication across independent systems. The finite fitted windows support a robust logarithmic regime but do not establish an exact asymptotic law. Moreover, fBM is used as a parsimonious Gaussian reference for the measured second-order correlations; neither the event-level origin of this effective process nor the microscopic source of trajectory-level heterogeneity is uniquely identified. In particular, the distribution of fitted relaxation rates cannot be equated directly with a distribution of latent article parameters.

These limitations define the natural empirical tests of the framework. The analysis requires age-resolved activity trajectories but not user-level microdata. Longer observation windows, other Wikipedia language editions, other information platforms, and other populations of competing, aging units can determine whether logarithmic spreading persists or crosses over to algebraic broadening or saturation. Joint measurement of temporal-increment scaling and cross-sectional dispersion provides a common basis for these comparisons even when individual trajectory shapes differ substantially. The broader implication is that the scaling of ensemble dispersion supplies a collective coordinate for locating a stability boundary whose position is not determined by a representative trajectory or the ensemble mean alone.

Acknowledgments. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

The views and conclusions expressed herein are solely those of the author and should not be construed as necessarily reflecting the official policies or endorsements—whether explicit or implicit—of any organizations with which the author is presently or has previously been affiliated.

Data availability. The data that support the findings of this article are openly available [19].

[1] A.-L. Barabási, The origin of bursts and heavy tails in human dynamics, Nature (London) **435**, 207 (2005).
[2] S. Goel, A. Anderson, J. Hofman, and D. J. Watts, The structural virality of online diffusion, Manag. Sci. **62**, 180 (2016).
[3] P. Lorenz-Spreen, B. M. Mønsted, P. Hövel, and S. Lehmann, Accelerating dynamics of collective attention, Nat. Commun. **10**, 1759 (2019).
[4] Z. Dezsö, E. Almaas, A. Lukács, B. Rácz, I. Szakadát, and A.-L. Barabási, Dynamics of information access on the web, Phys. Rev. E **73**, 066132 (2006).
[5] F. Wu and B. A. Huberman, Novelty and collective attention, Proc. Natl. Acad. Sci. USA **104**, 17599 (2007).
[6] A. G. Hawkes, Spectra of some self-exciting and mutually exciting point processes, Biometrika **58**, 83 (1971).
[7] R. Crane and D. Sornette, Robust dynamic classes revealed by measuring the response function of a social system, Proc. Natl. Acad. Sci. USA **105**, 15649 (2008).
[8] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).
[9] J. Ratkiewicz, S. Fortunato, A. Flammini, F. Menczer, and A. Vespignani, Characterizing and modeling the dynamics of online popularity, Phys. Rev. Lett. **105**, 158701 (2010).
[10] J. P. Gleeson, J. A. Ward, K. P. O'Sullivan, and W. T. Lee, Competition-induced criticality in a model of meme popularity, Phys. Rev. Lett. **112**, 048701 (2014).
[11] L. Weng, A. Flammini, A. Vespignani, and F. Menczer, Competition among memes in a world with limited attention, Sci. Rep. **2**, 335 (2012).
[12] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. **10**, 422 (1968).
[13] K. Okamura, Fractional stochastic model of citation dynamics with memory and volatility, Phys. Rev. E **112**, 044304 (2025).
[14] Wikimedia Foundation, Wikimedia Analytics API: Project metrics, https://doc.wikimedia.org/generated-data-platform/aqs/analytics-api/examples/project-metrics.html (accessed November 2025).
[15] Wikimedia Foundation, Wikimedia Analytics API: Page metrics, https://doc.wikimedia.org/generated-data-platform/aqs/analytics-api/examples/page-metrics.html (accessed November 2025).
[16] See Supplemental Material at http://link.aps.org/supplemental/10.1103/c9ny-57bd for methodological details, theoretical derivations and diagnostics, and trajectory-level heterogeneity analysis, which includes Refs. [20–24].
[17] J.-P. Bouchaud, Weak ergodicity breaking and aging in disordered systems, J. Phys. I **2**, 1705 (1992).
[18] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: A fractional dynamics approach, Phys. Rep. **339**, 1 (2000).
[19] K. Okamura, Dataset for "Logarithmic dispersion at marginal stability in collective attention" [Dataset], Zenodo, 2026, https://doi.org/10.5281/zenodo.18277388.
[20] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, Mosaic organization of DNA nucleotides, Phys. Rev. E **49**, 1685 (1994).
[21] K. Hu, P. C. Ivanov, Z. Chen, P. Carpena, and H. E. Stanley, Effect of trends on detrended fluctuation analysis, Phys. Rev. E **64**, 011114 (2001).
[22] J.-M. Bardet and I. Kammoun, Asymptotic properties of the detrended fluctuation analysis of long-range-dependent processes, IEEE Trans. Inf. Theory **54**, 2041 (2008).
[23] D. Nualart, Stochastic calculus with respect to fractional Brownian motion, Ann. Fac. Sci. Toulouse **15**, 63 (2006).
[24] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 3rd ed. (John Wiley & Sons, Ltd, Chichester, UK, 2014).