# Nested superstatistics: When the fluctuation parameter fluctuates

Kamel Ourabah*
*Theoretical Physics Laboratory, Faculty of Physics, University of Bab-Ezzouar, USTHB, Boite Postale 32, El Alia, Algiers 16111, Algeria*

(Received 3 July 2026; accepted 19 October 2026; published 2 November 2026)

Many nonequilibrium systems exhibit a hierarchy of well-separated dynamical scales. Typically, microscopic variables such as energy or velocity evolve rapidly, whereas environmental parameters such as the inverse temperature $\beta$ fluctuate more slowly. This separation naturally leads to superstatistics, in which local equilibrium factors are averaged over fluctuations of $\beta$, whose relative strength can be characterized by $q = \langle\beta^2\rangle/\langle\beta\rangle^2$. Here, we introduce an additional hierarchical level by allowing $q$ itself to fluctuate on a still slower scale. Physically, such fluctuations may arise when different regions or time intervals are characterized by different fluctuation intensities, correlation structures, or effective numbers of microscopic degrees of freedom. The resulting nested construction produces effective distributions that depend on the full distribution $g(q)$ and, in the weak-fluctuation regime, reduce to a universal correction governed only by the mean and variance of $q$. Using a Langevin model, we show explicitly that the nested construction emerges dynamically when the three relevant timescales are sufficiently well separated, and we provide a method for reconstructing $g(q)$ from a time series of the fast variable. Finally, we apply the theory to solar-wind electron distributions and cosmic-ray electron spectra, and identify observational situations in which the systematic deviations from fixed-$q$ superstatistics are consistent with a fluctuating superstatistical parameter.

DOI: 10.1103/dtcy-db7k

## I. INTRODUCTION

"Complexity frequently takes the form of hierarchy," Herbert Simon observed, adding that hierarchic systems may possess common properties that are independent of their specific content [1]. This observation is particularly pertinent to nonequilibrium statistical mechanics. Many complex systems, whether physical, biological, or social, are organized across well-separated dynamical scales. A simple physical example is a Brownian particle in a slowly varying environment: The particle velocity changes rapidly, while the environmental parameters controlling its local statistics, such as the temperature of the surrounding medium, vary much more slowly. When such a separation of scales holds, the statistical description can be greatly simplified. The global statistics may then be constructed as a superposition of local equilibrium statistics, each associated with a different value of the slowly varying environmental parameter.

This idea has a long history in statistical mechanics, dating back to the works of Kubo *et al*. and Lavenda on fluctuating thermodynamic parameters [2,3]. Closely related constructions also have a long tradition in probability theory and mathematical statistics, where they are known as *compounding* or *probability mixtures* [4]. In the contemporary physics literature, this approach is known as *superstatistics*, so named because it consists of a superposition of statistics [5]. It has become a standard tool for describing nonequilibrium systems characterized by a clear separation of timescales, providing successful descriptions across a wide spectrum of phenomena, including turbulence [6–9], plasmas [10–14], gravity-dominated systems [15–17], and cold atoms [18,19].

One of the first successes of superstatistics, and in fact the historical motivation behind its introduction [20,21], was to provide a dynamical origin for $q$-exponential distributions. Such distributions had already been introduced in space-plasma physics by Olbert and Vasyliunas in the 1960s [22,23], well before the development of nonextensive statistical mechanics by Tsallis [24]. In nonextensive statistical mechanics, Tsallis distributions are usually derived by maximizing the generalized entropy $S_q$, with $q$ entering as an entropic index [25], but their broad empirical success in complex systems [26–30] makes the dynamical origin of this parameter a natural question. Superstatistics offers a different route: The same distributions arise in a particular microscopic scenario in which the inverse temperature $\beta$ fluctuates slowly according to a $\chi^2$ distribution. This case is natural when the fluctuations of $\beta$ result from the combined contribution of many independent microscopic degrees of freedom [31]. The parameter $q$ then acquires a transparent physical meaning: It measures the strength of the inverse-temperature fluctuations and is directly related to the effective number of degrees of freedom that generate them. This picture, however, relies on a clear dynamical hierarchy, with microscopic variables evolving rapidly and temperature fluctuations occurring on a slower scale.

*Contact author: kam.ourabah@gmail.com, kourabah@usthb.dz

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

This hierarchy, however, need not stop at the level of the fluctuating temperature. In realistic nonequilibrium systems, the intensity of the inverse-temperature fluctuations may itself vary on an even slower scale, as the large-scale conditions controlling the local dynamics evolve. This is natural, for instance, in intermittent turbulent media, where local dissipation rates and fluctuation amplitudes are strongly inhomogeneous, or in space plasmas, where particle distributions and turbulence properties evolve across different regions and scales [32–35]. A more direct indication is provided by collisional superstatistical relaxation, where the parameter $q$ may evolve in time [36]. In the $\chi^2$ case, this amounts to saying that the effective number $n$ of independent contributions generating the fluctuations of $\beta$ is no longer fixed, but becomes a slowly varying quantity. This naturally invites the central idea of the present paper: a nested-superstatistical description in which the parameter controlling the fluctuations of $\beta$ becomes a fluctuating quantity in its own right.

Related ideas have appeared in the literature, although the specific question of a fluctuating superstatistical parameter has received comparatively little attention. Sob'yanin introduced generalized superstatistics, in which superstatistical systems are themselves mixed over higher-level control parameters [37,38], while the $H$ theory describes hierarchical chains of fluctuating intensive variables derived from an entropy-maximization principle [39,40]. From a different angle, Tsekouras and Tsallis constructed generalized entropies from a distribution of $q$ indices [41]. The present work takes a more specific and dynamical route. We extend the usual $\chi^2$ superstatistical construction by promoting $q$, or equivalently the effective number $n$ of degrees of freedom, to a slow dynamical variable. This leads to a nested construction governed by the distribution $g(q)$ of the fluctuation parameter itself. We show that this construction has a universal weak-fluctuation regime, emerges dynamically under a three-level separation of timescales, and can be reconstructed from time series. We also show that simple choices of $g(q)$ can reproduce observational features that cannot be accounted for by a single fixed value of $q$.

The remainder of the paper is organized as follows. In Sec. II, we introduce the nested-superstatistical construction. In Sec. III, we test its dynamical emergence with a Langevin model and identify the required hierarchy of timescales. In Sec. IV, we derive the universal regime of weak $q$ fluctuations, while Sec. V considers bimodal fluctuations of $q$ and illustrates their observational relevance with solar-wind electron distributions and cosmic-ray electron spectra. In Sec. VI, we show how the distribution $g(q)$ can be reconstructed from a time series of the fast variable. Section VII concludes the paper with a summary and some perspectives. Three Appendixes provide technical details on the construction and on finite-time-scale corrections.

## II. A HIERARCHY OF SUPERSTATISTICAL FLUCTUATIONS

Many nonequilibrium systems display a separation of scales: While the system as a whole is not in thermal equilibrium, sufficiently small spatial or temporal regions may remain close to local equilibrium. Within each such region, the system is approximately thermalized and its statistical properties can be described by equilibrium statistical mechanics. At larger scales, however, the temperature, or equivalently the inverse temperature $\beta = 1/k_B T$, is no longer constant, but varies slowly from one region or observation interval to another. The statistics of the complete system are then obtained by superposing the local equilibrium statistics. This construction defines superstatistics, with an effective Boltzmann factor of the form

$$B(E) \propto \int_0^\infty f(\beta) e^{-\beta E}\, d\beta, \tag{1}$$

where $f(\beta)$ is the probability density of the fluctuating inverse temperature.

In principle, different physical mechanisms may lead to different forms of $f(\beta)$. A particularly relevant example is the $\chi^2$ (or Gamma) distribution

$$f_n(\beta) = \frac{1}{\Gamma(n/2)} \left(\frac{n}{2\beta_0}\right)^{n/2} \beta^{n/2-1} \exp\left(-\frac{n\beta}{2\beta_0}\right), \tag{2}$$

which satisfies

$$\langle \beta \rangle = \beta_0. \tag{3}$$

This distribution has a clear statistical origin. For integer $n$, it arises when the inverse temperature is proportional to the sum of the squares of $n$ independent Gaussian variables [31]. The parameter $n$ can therefore be interpreted as an *effective* number of degrees of freedom and controls the strength of the inverse-temperature fluctuations through

$$\frac{\mathrm{Var}(\beta)}{\beta_0^2} = \frac{2}{n}. \tag{4}$$

Substituting Eq. (2) into Eq. (1) gives

$$B_n(E) \propto \left(1 + \frac{2\beta_0 E}{n}\right)^{-n/2}. \tag{5}$$

Introducing

$$q := \frac{\langle \beta^2 \rangle}{\langle \beta \rangle^2} = 1 + \frac{2}{n}, \tag{6}$$

the effective factor takes the standard $q$-exponential form

$$B_q(E) = [1 + (q-1)\beta_0 E]^{-1/(q-1)}. \tag{7}$$

This power-law factor belongs to the family of $\kappa$ distributions introduced in space-plasma physics by Olbert and Vasyliunas [22,23] and can equivalently be expressed in the $q$-exponential form familiar from nonextensive statistical mechanics [25]. With $\kappa$ denoting the power-law exponent in Eq. (7), the corresponding parameters are related by $\kappa = 1/(q-1) = n/2$. The ordinary Boltzmann factor is recovered in the limit $n \to \infty$, or equivalently $q \to 1$. By construction, one has $q \geqslant 1$ in the present setting. When $q > 1$, the corresponding distributions develop heavy tails, making them well suited to the description of a broad range of systems, from astrophysical observations [42,43] to controlled laboratory experiments [18,19,44].

Equations (1)–(7) summarize the standard $\chi^2$ superstatistical construction. In ensemble language, superstatistics may be viewed as a *hyperensemble*, namely, an ensemble of canonical

ensembles characterized by different values of the inverse temperature $\beta$ [45]. The density $f_n(\beta)$ specifies how these canonical ensembles are weighted, while the parameter $n$ determines the structure of the hyperensemble by controlling the magnitude and shape of the inverse-temperature fluctuations.

In the conventional formulation, $n$, and hence $q$, remains fixed. In a heterogeneous nonequilibrium system, there is no general reason for this effective number to remain identical in all regions or over all observation windows. Different regions may exhibit different fluctuation intensities, correlation structures, degrees of intermittency, or effective numbers of active microscopic contributions. It is therefore natural to promote $n$ from a fixed parameter to a fluctuating quantity.

Let $h(n)$ denote the probability density of the effective number of degrees of freedom. The resulting effective factor is then

$$\overline{B}(E) = \int_0^\infty h(n) B_n(E)\, dn. \tag{8}$$

Although $n$ is introduced as a number of degrees of freedom, the $\chi^2$ distribution in Eq. (2) is mathematically well defined for any real $n > 0$. Indeed, Eq. (2) is a Gamma distribution with shape parameter $n/2$, which need not be an integer [5]. In this sense, $n$ is naturally understood as an effective continuous parameter controlling the strength of the inverse-temperature fluctuations, rather than as a literal count of microscopic degrees of freedom. The integral in Eq. (8) therefore represents a mixture over effective continuous parameters, rather than a sum over literal counts of Gaussian variables. Using Eq. (6), the same construction can be expressed in terms of the induced distribution $g(q)$,

$$\overline{B}(E) = \int_1^\infty g(q)[1 + (q-1)\beta_0 E]^{-1/(q-1)} dq. \tag{9}$$

Since

$$n = \frac{2}{q-1}, \tag{10}$$

the distributions $h(n)$ and $g(q)$ are related by

$$g(q) = h\left(\frac{2}{q-1}\right)\frac{2}{(q-1)^2}. \tag{11}$$

The end points $n = 0$ and $q = \infty$ in Eqs. (8) and (9) are understood as improper integration limits, rather than admissible parameter values. While the Gamma distribution in Eq. (2) is defined for any $n > 0$, additional restrictions arise when the effective Boltzmann factor is used to construct a normalized probability density. For the one-dimensional velocity distribution considered in Sec. III, normalization requires $n > 1$ ($q < 3$), while finite second and fourth moments require $n > 3$ ($q < 5/3$) and $n > 5$ ($q < 7/5$), respectively.

Equation (9) introduces a second statistical level beyond ordinary superstatistics. At the first level, the inverse temperature fluctuates according to $f_n(\beta)$ for a fixed value of $n$. At the second level, the parameter controlling these fluctuations is itself distributed according to $h(n)$, or equivalently $g(q)$. We refer to this two-level construction as *nested superstatistics*. The hyperensemble interpretation makes the additional level particularly transparent. At fixed $n$, superstatistics defines a hyperensemble of canonical ensembles with fluctuating inverse temperature. Allowing $n$ to vary therefore produces an *ensemble of hyperensembles*.

At the level of the final effective factor alone, Eq. (9) may be rewritten as an ordinary one-level superstatistics with an effective inverse-temperature distribution $f_{\rm eff}(\beta) = \int h(n) f_n(\beta)\, dn$. The distinction is therefore not that nested superstatistics produces a marginal distribution of $\beta$ that could not be cast in this form. Rather, it resolves the effective distribution into a conditional hierarchy: $\beta$ fluctuates at fixed $n$, while $n$ (or equivalently $q$) fluctuates on a slower level. This gives a physical meaning to $g(q)$ and makes the hierarchy dynamically testable.

For later use, we denote the mean and variance of the induced $q$ distribution by

$$q_0 = \langle q \rangle, \quad \sigma_q^2 = \langle q^2 \rangle - \langle q \rangle^2 \tag{12}$$

and introduce the dimensionless parameter

$$\alpha = \frac{\langle q^2 \rangle}{\langle q \rangle^2} = 1 + \frac{\sigma_q^2}{q_0^2}. \tag{13}$$

The conventional superstatistical result is recovered when

$$g(q) = \delta(q - q_0), \tag{14}$$

for which $\alpha = 1$. When $g(q)$ has a finite width, $\alpha > 1$, and the resulting effective factor is no longer described by a single $q$ exponential. Instead, it arises from a superposition of $q$-exponential factors with different indices. The observed statistics therefore reflect a hierarchy of fluctuations: first, fluctuations of the inverse temperature at fixed $n$, and second, fluctuations of the effective number of degrees of freedom itself.

At fixed $q$, the $q$-exponential form (7) can also arise from the maximization of Tsallis entropy under appropriate constraints [25], with the temperature parameter held fixed in the corresponding variational construction. Equation (9) may therefore be viewed, at the level of distribution functions, as a superposition of such factors with different indices, without implying that the resulting mixture itself maximizes Tsallis entropy. In the present work, in contrast, the same $q$-exponential factors emerge from fluctuations of the local inverse temperature $\beta$, while $q$ varies on a still slower timescale, providing the specific dynamical basis of nested superstatistics.

Note also that we adopt the type-A formulation of superstatistics, in which the unnormalized Boltzmann factor is averaged over $\beta$ before the resulting effective weight is normalized. This prescription is convenient here because, for the $\chi^2$ distribution, it directly yields the standard Tsallis $q$-exponential form. The relation with the alternative type-B formulation, where the local Boltzmann factor is normalized before averaging, is briefly discussed in Appendix A.

Finally, the statistical constructions above rely on a hierarchy of timescales. In ordinary superstatistics, the microscopic variables must relax toward local equilibrium on a timescale $\tau_{\rm rel}$ much shorter than the characteristic time $\tau_\beta$ over which the inverse temperature changes appreciably. The inverse temperature can then be regarded as approximately constant during each local relaxation process, which justifies Eq. (1). Nested

superstatistics requires a third, still slower timescale: The effective number of degrees of freedom $n$, or equivalently $q$, must vary on a timescale $\tau_n$ much longer than $\tau_\beta$. The complete hierarchy is therefore

$$\tau_{\rm rel} \ll \tau_\beta \ll \tau_n. \tag{15}$$

The dynamical validity of this hierarchy, as well as the deviations that arise when it is not satisfied, is examined in the next section.

## III. DYNAMICAL EMERGENCE OF NESTED SUPERSTATISTICS

The nested construction introduced in the previous section involves not only a hierarchy of statistical averages, but also an underlying hierarchy of timescales. To examine this dynamical requirement explicitly, we perform a controlled numerical experiment based on a minimal one-dimensional Langevin model. The purpose is to determine whether the nested-superstatistical prediction (8) is recovered when the environmental variables evolve sufficiently slowly, and how the resulting statistics are modified when the separation of timescales is reduced.

The observed variable is a velocity $v(t)$ governed by

$$\mathrm{d}v = -\gamma v\,\mathrm{d}t + \sqrt{\frac{2\gamma}{m\beta(t)}}\,\mathrm{d}W_t, \tag{16}$$

where $\gamma$ is the relaxation rate, $m$ is the particle mass, and $\mathrm{d}W_t$ denotes a Wiener increment. For a fixed value of $\beta$, Eq. (16) reduces to an Ornstein-Uhlenbeck process and relaxes toward the corresponding local Maxwellian distribution on the characteristic timescale

$$\tau_v = \gamma^{-1}. \tag{17}$$

The environmental variables are updated according to a two-level protocol. The inverse temperature $\beta(t)$ is kept constant over intervals of duration $\tau_\beta$, while the parameter $n$ is kept fixed over a longer interval $\tau_n$. The sampling protocol is chosen so that, in the adiabatic limit, the stationary velocity distribution reproduces the type-A nested-superstatistical construction introduced in Sec. II. The technical relation between this sampling protocol and the type-A weights is given in Appendix A.

In the numerical simulations, the prescribed weight $h(n)$ is chosen to be log-normal and may be generated as

$$n = n_0 \exp(\eta), \tag{18}$$

where $\eta$ is a Gaussian random variable with mean $-\sigma_\eta^2/2$ and variance $\sigma_\eta^2$. This choice guarantees positivity and gives $\langle n\rangle_h = n_0$, while $\sigma_\eta$ controls the amplitude of the fluctuations. Through the relation $q = 1 + 2/n$, the same prescription induces fluctuations of the parameter $q$. The log-normal distribution provides a convenient continuous realization of positive fluctuations in $n$, but is not a requirement of the nested construction. Other admissible choices are possible, including a rescaled beta distribution for $g(q)$, whose bounded support allows the range of $q$ to be restricted directly to physically admissible values. The log-normal parent distribution is truncated to $n > 5.05$, slightly above the threshold $n > 5$ required for a finite fourth moment of the one-dimensional velocity distribution.

For the numerical analysis, we introduce the dimensionless velocity

$$u = \frac{v}{v_0}, \quad v_0 = \sqrt{\frac{2}{m\beta_0}}. \tag{19}$$

In the adiabatic regime,

$$\tau_v \ll \tau_\beta \ll \tau_n, \tag{20}$$

each quantity has sufficient time to relax locally. The stationary velocity statistics are then expected to approach the (normalized) nested-superstatistical reference (8).

We test this expectation through two numerical experiments based on the same Langevin equation and the same environmental sampling protocol, but with different separations between the characteristic timescales. In the slow dynamics, the hierarchy in Eq. (20) is well satisfied, so that the velocity can relax before the environmental parameters are updated. In the fast dynamics, the environmental variables change before this local relaxation is completed. Since the distributions are identical in both simulations, the comparison isolates the effect of the timescale separation.

The comparison is performed through the survival probability

$$S(U) = \mathrm{Prob}(|u| > U), \tag{21}$$

which is particularly sensitive to differences in the tails of the velocity distribution.

Figure 1(a) shows that, in the slow regime, the numerical dynamics reproduces the adiabatic nested-superstatistical reference (8) with high accuracy, whereas reducing the separation of timescales leads to visible deviations. These differences are quantified in Fig. 1(b) through the ratios of the simulated survival probabilities to the reference curve. The ratio associated with the slow dynamics remains close to unity, while that of the fast dynamics departs systematically from it, confirming that the nested-superstatistical prediction is recovered only in the adiabatic regime.

These results show that the nested-superstatistical limit emerges only when the three dynamical levels remain sufficiently well separated. In that regime, the velocity equilibrates locally before $\beta$ changes, while the statistics of $\beta$ are sampled before $n$ evolves. Once this ordering is weakened, the stationary distribution no longer coincides with the adiabatic reference.

The departure from the adiabatic limit can be quantified perturbatively in the ratios $\tau_v/\tau_\beta$ and $\tau_v/\tau_n$. The leading finite-time-scale corrections are derived in Appendix B, where they are shown to reduce the excess kurtosis relative to its adiabatic value, consistently with the suppression of the tails observed in the fast dynamics.

## IV. UNIVERSAL REGIME OF WEAK $q$ FLUCTUATIONS

As in ordinary superstatistics, where the distribution $f(\beta)$ determines the departure from Boltzmann statistics, the distribution $g(q)$ controls how the nested construction departs from the $q$-exponential distribution (7). When $g(q)$ is sharply

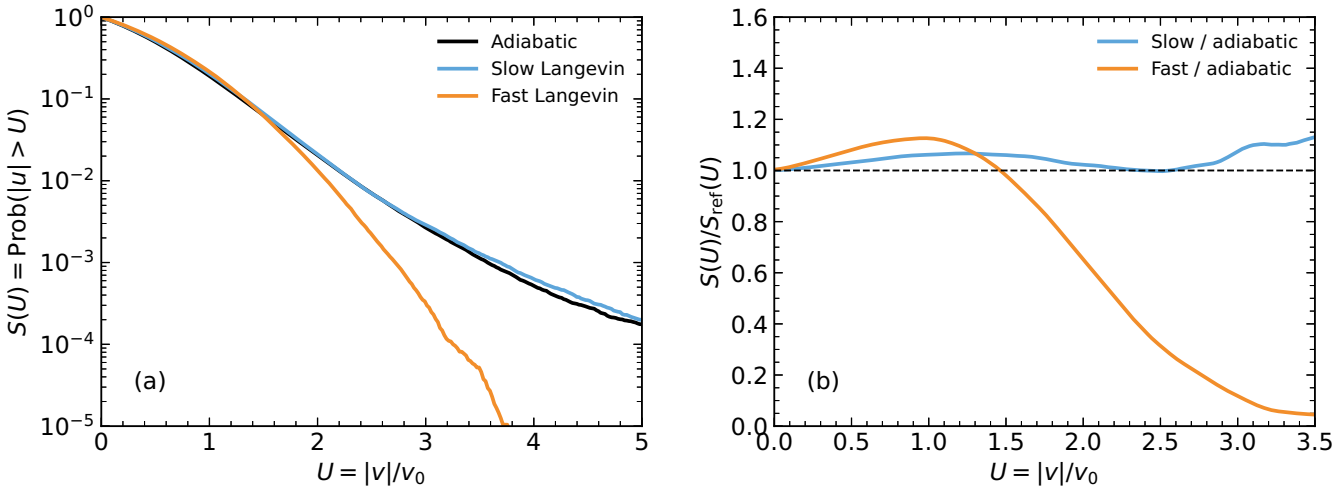



FIG. 1. Dynamical test of nested superstatistics through a controlled Langevin experiment. (a) Survival probability $S(U) = \mathrm{Prob}(|u| > U)$, with $u = v/v_0$, for the adiabatic nested-superstatistical reference and for slow and fast environmental dynamics. The slow dynamics closely follows the reference, whereas the fast dynamics shows clear deviations. (b) Corresponding ratios to the adiabatic reference. The slow dynamics remains close to unity, while the fast dynamics departs systematically from it. The ratios are shown only where the survival probabilities remain statistically resolved.

peaked around its mean value $q_0 = \langle q \rangle$, however, the leading correction becomes universal, depending only on the mean and variance of $q$ rather than on the detailed form of $g(q)$. It is therefore useful to examine this weak-fluctuation regime before turning to specific models for $g(q)$.

Expanding the $q$-exponential factor in Eq. (9) around $q = q_0$ gives

$$B_q(E) = B_{q_0}(E) + (q - q_0) \times \left.\frac{\partial B_q(E)}{\partial q}\right|_{q_0} + \frac{(q-q_0)^2}{2}\left.\frac{\partial^2 B_q(E)}{\partial q^2}\right|_{q_0} + \cdots . \tag{22}$$

Since $\langle q - q_0 \rangle = 0$, the linear contribution vanishes after averaging. To leading nontrivial order, one therefore obtains

$$\overline{B}(E) \simeq B_{q_0}(E) + \frac{\sigma_q^2}{2}\left.\frac{\partial^2 B_q(E)}{\partial q^2}\right|_{q_0}, \tag{23}$$

where $\sigma_q^2 = q_0^2(\alpha - 1)$. Equation (23) is universal in the sense that, at this order, the correction depends on $g(q)$ only through its mean and variance. The skewness and higher moments of $g(q)$ enter only at higher orders.

Introducing the dimensionless quantities

$$x = \beta_0 E, \quad s_0 = q_0 - 1 \tag{24}$$

and evaluating the second derivative of the $q$ exponential in Eq. (7), the weak-fluctuation correction can be written in the form

$$\overline{B}(E) \simeq B_{q_0}(E)\big[1 + q_0^2(\alpha - 1)\mathcal{F}(x, q_0)\big], \tag{25}$$

where

$$\mathcal{F}(x, q_0) = \tfrac{1}{2}\big[\mathcal{D}_1^2(x, s_0) + \mathcal{D}_2(x, s_0)\big], \tag{26}$$

with

$$\mathcal{D}_1(x, s) = \frac{\ln(1 + sx)}{s^2} - \frac{x}{s(1 + sx)} \tag{27}$$

and

$$\mathcal{D}_2(x, s) = -\frac{2\ln(1 + sx)}{s^3} + \frac{2x}{s^2(1 + sx)} + \frac{x^2}{s(1 + sx)^2}. \tag{28}$$

Note that, although Eqs. (27) and (28) contain inverse powers of $s$, their apparent singularities at $s = 0$ are removable. Their limiting values are $\mathcal{D}_1(x, 0) = x^2/2$ and $\mathcal{D}_2(x, 0) = -2x^3/3$.

If, in addition, the mean value $q_0$ is close to the Boltzmann limit $q_0 \approx 1$, the correction function simplifies to

$$\mathcal{F}(x, q_0) = \frac{x^4}{8} - \frac{x^3}{3} + \mathcal{O}(q_0 - 1). \tag{29}$$

In this regime, Eq. (25) reduces to the compact expression

$$\overline{B}(E) \simeq B_{q_0}(E)\left[1 + q_0^2(\alpha - 1)\left(\frac{(\beta_0 E)^4}{8} - \frac{(\beta_0 E)^3}{3}\right)\right]. \tag{30}$$

The behavior of the universal correction is illustrated in Fig. 2, where $\mathcal{F}(x, q_0)$ is shown for several values of $q_0$ and compared with its near-Boltzmann approximation. As $q_0$ approaches unity, the exact result converges to the polynomial form in Eq. (29). Since the correction in Eq. (25) is proportional to $\mathcal{F}(x, q_0)$, its sign determines whether fluctuations of $q$ produce a relative suppression or enhancement with respect to the fixed-$q_0$ factor. In the near-Boltzmann limit, the sign change occurs at $x = 8/3$, whereas for finite $q_0$ the zero is progressively shifted toward higher energies.

The expansion is valid provided that

$$q_0^2(\alpha - 1)|\mathcal{F}(\beta_0 E, q_0)| \ll 1, \tag{31}$$

over the energy range considered. Beyond this perturbative regime, the detailed structure of $g(q)$ becomes important. In

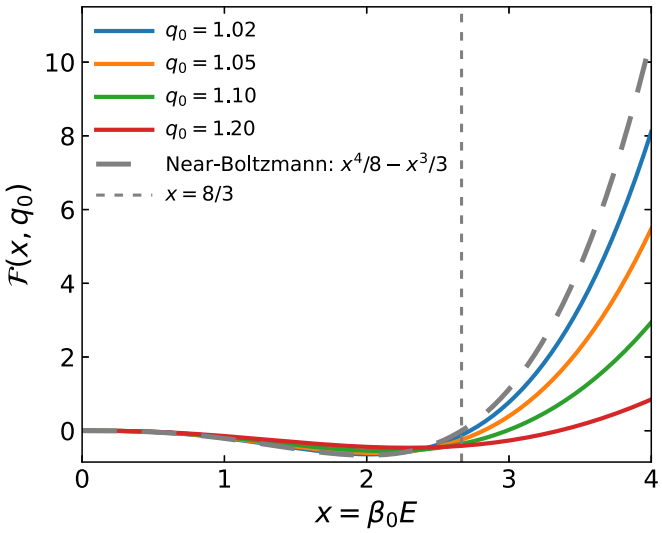



FIG. 2. Universal weak-fluctuation correction function $\mathcal{F}(x, q_0)$, defined in Eq. (26), as a function of $x = \beta_0 E$ for several values of $q_0 \equiv \langle q \rangle$. The long-dashed gray curve shows the near-Boltzmann approximation $\mathcal{F}(x, q_0) \simeq x^4/8 - x^3/3$, while the finely dashed vertical line marks its zero at $x = 8/3$. As $q_0 \to 1$, the exact correction approaches the near-Boltzmann result.

the following sections, we therefore turn to two complementary questions: first, how a simple nonperturbative choice of $g(q)$ can improve the description of observational data, and second, how such a distribution may be reconstructed from a time series.

## V. BIMODAL FLUCTUATIONS OF $q$

In the weak-fluctuation regime, fluctuations of $q$ induce the universal correction derived in the previous section. Beyond this regime, however, the resulting distributions depend explicitly on the detailed form of $g(q)$ and may exhibit features that cannot be captured by an ordinary superstatistical model with a fixed value of $q$. For suitable choices of $g(q)$, the nested construction can thus provide an improved description of observational or experimental data for which standard superstatistics is insufficient.

A simple yet physically meaningful realization is obtained when $q$ fluctuates between two preferred values. The corresponding distribution is

$$g(q) = w\,\delta(q - q_1) + (1 - w)\,\delta(q - q_2), \tag{32}$$

with

$$0 \leqslant w \leqslant 1, \quad 1 < q_1 < q_2. \tag{33}$$

The two delta peaks here need not represent strictly discrete microscopic states. They may instead be viewed as a coarse-grained representation of a bimodal distribution concentrated around two characteristic values of $q$. In this case, the nested construction (9) gives

$$\overline{B}(E) = w[1 + (q_1 - 1)\beta_0 E]^{-1/(q_1-1)} + (1 - w)[1 + (q_2 - 1)\beta_0 E]^{-1/(q_2-1)}. \tag{34}$$

The effective Boltzmann factor is therefore a weighted superposition of two $q$-exponential factors with different tail indices.

This is probably the simplest model of $g(q)$. It may also be viewed as the first step toward more general distributions, since an arbitrary $g(q)$ can be approximated, at least formally, by a weighted sum of narrow peaks or Dirac delta functions. Physically speaking, this construction may describe a system containing two coexisting statistical populations or a system alternating between two long-lived dynamical regimes characterized by different intensities of inverse-temperature fluctuations. In that interpretation, $w$ is the occupation probability of the first regime at the level of the fluctuating parameter $q$. In what follows, we illustrate the usefulness of this bimodal construction in two different settings: solar-wind electron distributions and the cosmic-ray electron spectrum.

### A. Solar-wind electron distributions

A particularly relevant limit of Eq. (34) is obtained by taking $q_1 \to 1$ and denoting $q_2 \equiv q > 1$, such that

$$\overline{B}(E) = w\,\exp(-\beta_0 E) + (1 - w)[1 + (q - 1)\beta_0 E]^{-\frac{1}{q-1}}. \tag{35}$$

In this case, the first contribution is Maxwellian, whereas the second decays algebraically. For sufficiently large $w$, the Maxwellian contribution controls the central part of the distribution, while the $q$-exponential component eventually dominates the high-energy tail. Such a structure can arise when a comparatively thermalized bulk population coexists with a dilute suprathermal component maintained by intermittent energization processes.

This situation is common in space plasmas, whose velocity distributions often display a quasi-Maxwellian core together with a pronounced suprathermal tail [46–49]. A standard single-$q$ distribution generally cannot adjust these two regions independently, because the same value of $q$ controls both the deformation of the core and the asymptotic decay. The bimodal nested model provides the additional flexibility required to describe the two regions simultaneously.

To illustrate this point, we apply Eq. (35) to an electron velocity-distribution measurement obtained by the *Ulysses* spacecraft in the slow solar wind. The observation was made on 15 January 2002, at 03:33:42 UT. The observational data were digitized from Fig. 1 of Lazar *et al.* [50], which was constructed from Ulysses/SWOOPS electron pitch-angle measurements available through the ESA Ulysses archive [51]. The one-dimensional velocity distribution used in the fit is constructed from the effective Boltzmann factor (35) by setting $E = mv^2/2$,

$$f(v) = \frac{1}{Z_v}\overline{B}\left(\frac{1}{2}mv^2\right), \tag{36}$$

where $m$ is the electron mass and $Z_v$ ensures normalization with respect to the measured velocity variable.

The results are shown in Fig. 3. The fit is performed in logarithmic space because the data span several orders of magnitude. The best bimodal fit is obtained for $w = 0.97$ and $q = 1.31$, whereas the best single-$q$ description

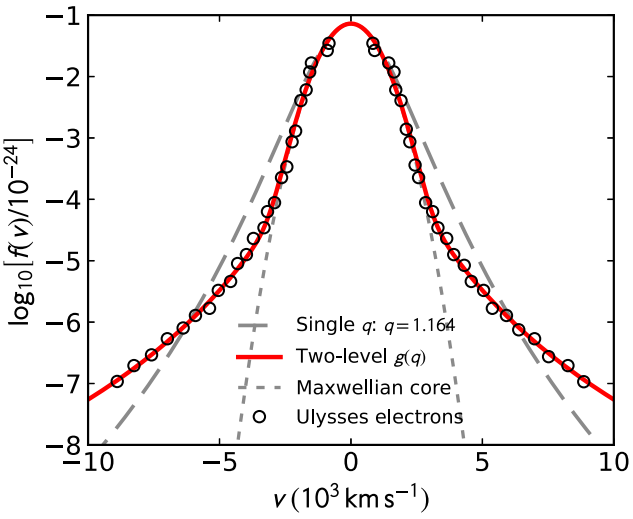



FIG. 3. Base-10 logarithm of the electron velocity-distribution function measured by *Ulysses* spacecraft in the slow solar wind on day 15 of 2002 at 03:33:42 UT. The observational data are shown by circles. The solid line is the bimodal fit obtained from Eq. (35), with $q = 1.31$ and $w = 0.97$. The long-dashed line shows the best single-$q$ fit, with $q = 1.16$, which cannot reproduce simultaneously the Maxwellian core and the suprathermal tail. The finely dashed gray line represents the Maxwellian component of the bimodal model. For all three curves, the width parameter is fixed by the core and corresponds to $\beta_0^{-1} \simeq 3.35\,\mathrm{eV}$.

corresponds to $q = 1.16$. One may see that the bimodal model reproduces both the quasi-Maxwellian core and the suprathermal tail, while the single-$q$ distribution cannot fit the two regions simultaneously. Quantitatively, the bimodal model reduces the residual sum of squares by a factor of approximately 28.5. The improvement remains strongly favored after accounting for the additional fitting parameter. Defining

$$\Delta\mathrm{AIC} = \mathrm{AIC}_{\text{single-}q} - \mathrm{AIC}_{\text{bimodal}},$$
$$\Delta\mathrm{BIC} = \mathrm{BIC}_{\text{single-}q} - \mathrm{BIC}_{\text{bimodal}}, \tag{37}$$

we obtain $\Delta\mathrm{AIC} \simeq 165.4$ and $\Delta\mathrm{BIC} \simeq 163.5$, so that both criteria strongly favor the bimodal model.

### B. Cosmic-ray electron spectrum

A second application of the bimodal construction arises in cosmic-ray physics. Cosmic rays are energetic charged particles propagating through interstellar space. They consist predominantly of atomic nuclei, but also contain electrons, positrons, and other subatomic particles. Their differential fluxes carry information about the astrophysical mechanisms governing their production, acceleration, and propagation.

More specifically, we consider the cosmic-ray electron flux measured by the Alpha Magnetic Spectrometer (AMS-02), a magnetic particle detector operating aboard the International Space Station. The measurement is based on $28.1 \times 10^6$ electron events collected between May 2011 and November 2017 and covers the energy range from 0.5 GeV to 1.4 TeV [52]. The numerical values used below are taken from the corresponding tabulation in the Cosmic-Ray DataBase [53].

Cosmic-ray spectra are commonly expressed in terms of the particle rigidity,

$$R = \frac{pc}{|Z|e}, \tag{38}$$

where $p$ is the momentum and $Ze$ is the particle charge. For ultrarelativistic electrons, $|Z| = 1$ and $pc \simeq E$, so that the numerical values of $R$ in GV and $E$ in GeV are approximately equal. The differential electron flux is defined as

$$\Phi_{e^-}(R) = \frac{dN}{dA\,dt\,d\Omega\,dR} \tag{39}$$

and gives the number of electrons detected per unit area, time, solid angle, and rigidity interval. AMS-02 found that the electron spectrum departs significantly from its lower-energy trend above approximately 42 GeV, indicating a progressive spectral hardening [52]. Such a gradual change of slope can naturally emerge from a superposition of components characterized by different values of $q$.

In the ultrarelativistic regime, the three-dimensional phase-space factor contributes a term proportional to $R^2$. The nested model therefore gives

$$\Phi_{e^-}(R) = AR^2 \int g(q)[1 + (q-1)\beta_0 R]^{-\frac{1}{q-1}}\,dq, \tag{40}$$

where $A$ is an overall normalization.

Before considering the bimodal case, it is useful to examine the full fixed-$q$ model,

$$\Phi_{e^-}^{(q)}(R) = AR^2[1 + (q-1)\beta_0 R]^{-\frac{1}{q-1}}. \tag{41}$$

Its local spectral index is

$$\gamma_{\mathrm{eff}}(R) = -\frac{d\ln\Phi_{e^-}^{(q)}}{d\ln R} = \frac{\beta_0 R}{1 + (q-1)\beta_0 R} - 2. \tag{42}$$

Moreover,

$$\frac{d\gamma_{\mathrm{eff}}}{d\ln R} = \frac{\beta_0 R}{[1 + (q-1)\beta_0 R]^2} > 0. \tag{43}$$

A single fixed-$q$ distribution therefore predicts a gradual steepening of the spectrum, rather than the hardening observed by AMS-02. At sufficiently large rigidity,

$$(q-1)\beta_0 R \gg 1, \tag{44}$$

Equation (41) approaches a single power law,

$$\Phi_{e^-}^{(q)}(R) \sim \widetilde{A} R^{-\gamma(q)}, \quad \gamma(q) = \frac{1}{q-1} - 2, \tag{45}$$

where the dependence on $\beta_0$ is absorbed into the normalization $\widetilde{A}$. A direct fit of the full fixed-$q$ expression (41) drives $\beta_0$ toward values for which the asymptotic condition is satisfied over the entire fitted range. In this limit, $\beta_0$ becomes unidentifiable and the fitted curve is numerically indistinguishable from the pure power law (45). For this reason, the fixed-$q$ reference used below is represented by its best-fit asymptotic form.

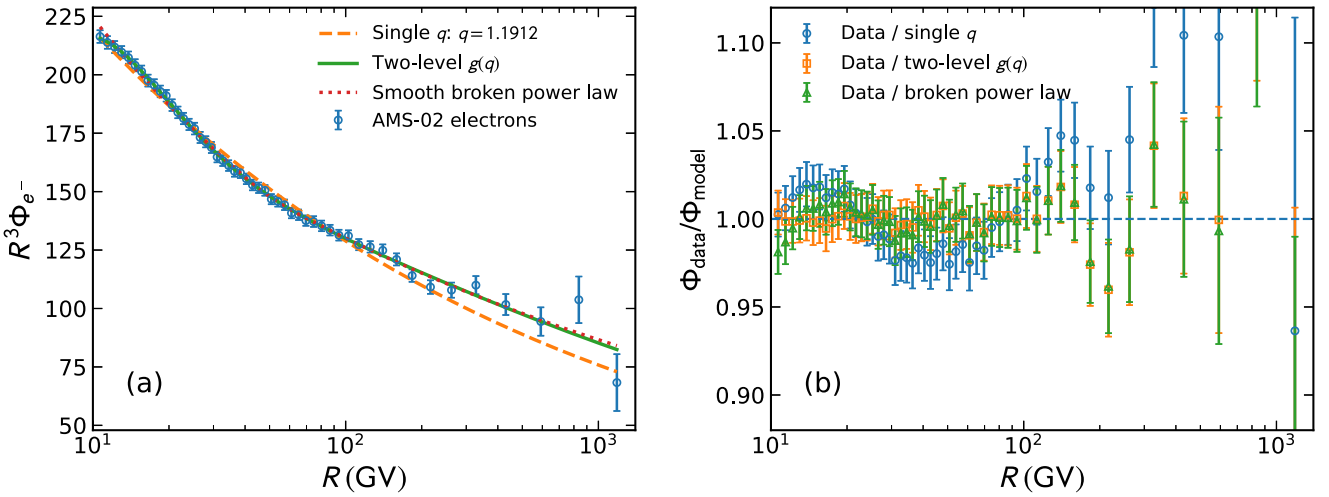



FIG. 4. AMS-02 cosmic-ray electron flux above $R = 10\,\mathrm{GV}$. Panel (a) shows $R^3\Phi_{e^-}$ as a function of rigidity. The observational data are represented by circles, with statistical and systematic uncertainties added in quadrature. The dashed line is the best asymptotic fixed-$q$ reference, corresponding to $q = 1.1912$. The solid line is the bimodal nested model (46), with $q_1 = 1.1406$, $q_2 = 1.1920$, $w = 0.8839$, and $\beta_0 = 2.397\,\mathrm{GV}^{-1}$. The dotted line shows the best smoothly broken power-law (47) fit. Panel (b) displays the corresponding data-to-model ratios. The fixed-$q$ reference exhibits systematic deviations associated with the spectral hardening, whereas the bimodal model reproduces the data over the full fitted rigidity range.

For the bimodal distribution (32), the electron flux (40) becomes

$$\Phi_{e^-}(R) = AR^2\Big\{w[1+(q_1-1)\beta_0 R]^{-\frac{1}{q_1-1}} + (1-w)[1+(q_2-1)\beta_0 R]^{-\frac{1}{q_2-1}}\Big\}. \quad (46)$$

Two-component $q$-generalized descriptions of cosmic-ray electron and positron spectra have been considered previously in related contexts, in particular, through combinations of escort and nonescort distributions [54]. In the present case, however, the two components have a different origin: They arise from a bimodal distribution $g(q)$ of the superstatistical parameter itself.

We fit the AMS-02 data above $R = 10\,\mathrm{GV}$, excluding the low-rigidity region where solar-modulation effects are strongest. Statistical and systematic uncertainties are added in quadrature. In addition to the asymptotic fixed-$q$ reference model (45), we also consider a smoothly broken power law, a widely used phenomenological parametrization in cosmic-ray studies to describe spectral hardenings or smooth transitions between different rigidity regimes [55,56],

$$\Phi_{\mathrm{SBPL}}(R) = C\left(\frac{R}{R_0}\right)^{-\gamma_1}\left[\frac{1+(R/R_b)^{1/s}}{1+(R_0/R_b)^{1/s}}\right]^{-s(\gamma_2-\gamma_1)}, \quad (47)$$

where $R_b$ is the break rigidity and $s$ controls the smoothness of the transition.

The results are displayed in Fig. 4. In panel (a), we plot $R^3\Phi_{e^-}$ rather than the flux itself. Multiplication by $R^3$ removes most of the steep overall decrease and makes the spectral curvature more visible. The fixed-$q$ model displays a clear systematic departure from the data, whereas the bimodal model follows the observed hardening over the full fitted range. Panel (b) shows the corresponding data-to-model ratios. The residuals of the fixed-$q$ reference exhibit a coherent rigidity dependence, while those of the bimodal model fluctuate around unity without an obvious systematic trend.

The best asymptotic fixed-$q$ reference gives

$$q = 1.1912, \quad \chi^2 = 85.80, \quad (48)$$

for 49 degrees of freedom. In contrast, the bimodal model gives

$$q_1 = 1.1406, \quad q_2 = 1.1920, \quad w = 0.8839, \quad \beta_0 = 2.397\,\mathrm{GV}^{-1}, \quad (49)$$

with

$$\chi^2 = 13.66 \quad (50)$$

for 46 degrees of freedom. The improvement is therefore

$$\Delta\chi^2 = \chi^2_{\text{fixed-}q} - \chi^2_{\text{bimodal}} = 72.14. \quad (51)$$

After accounting for the additional parameters, the information criteria remain strongly in favor of the bimodal construction,

$$\Delta\mathrm{AIC} = 66.14, \quad \Delta\mathrm{BIC} = 60.34. \quad (52)$$

A parametric-bootstrap analysis confirms that the two components are well separated, with nonoverlapping 95% confidence intervals $q_1 \in [1.1318, 1.1528]$ and $q_2 \in [1.1900, 1.1940]$.

Beyond the goodness of fit, the physical origin of the hardening is particularly transparent within this construction. The component associated with the smaller value $q_1$ has a steeper asymptotic decay and dominates at low rigidities, whereas the heavier-tailed $q_2$ component progressively takes over at higher rigidities despite its smaller weight. This crossover produces a smooth hardening without the need for an explicit spectral break. The AMS-02 spectrum therefore provides a concrete example in which fluctuations of $q$ generate a progressive

hardening that cannot be reproduced by the intrinsic curvature of a single fixed-$q$ distribution.

For comparison, the five-parameter phenomenological smoothly broken power law (47) gives

$$\chi^2_{\mathrm{SBPL}} = 20.92, \tag{53}$$

for 46 degrees of freedom, with a break rigidity close to $R_b = 51$ GV. Thus, for the same number of adjustable parameters, the bimodal nested model provides a better fit, while retaining a direct interpretation in terms of fluctuations of $q$.

## VI. RECONSTRUCTING THE DISTRIBUTION OF $q$ FROM TIME SERIES

The previous section considered the simplest nontrivial approximation to a fluctuating superstatistical parameter, namely, a bimodal distribution $g(q)$. More generally, a broad class of distributions can be represented, at least in a coarse-grained sense, by a weighted sum of narrow components,

$$g(q) \simeq \sum_{j=1}^{N} w_j\, \delta(q - q_j), \qquad \sum_{j=1}^{N} w_j = 1. \tag{54}$$

The bimodal case (33) corresponds to the first nontrivial member of this family. For sufficiently large $N$, such a discrete representation provides a systematic approximation to a continuous distribution $g(q)$. This raises the natural question of how the weights and characteristic values of $q$, or more generally the full distribution $g(q)$, can be inferred from data rather than postulated.

In an experimental setting, neither the slowly varying inverse temperature nor the fluctuations of $q$ are generally accessible directly. It is therefore important to determine whether the hierarchical structure underlying nested superstatistics can be inferred from a measured time series of the fast variable alone.

For ordinary superstatistics, a systematic procedure for extracting the slowly varying intensive parameter from a time series was introduced in Ref. [31]. In general terms, the idea is to identify time windows over which the fast variable has relaxed to a local equilibrium statistics characterized by an approximately constant inverse temperature $\beta$. For a velocity-like variable, as in the Langevin model, this local equilibrium distribution is Gaussian, and $\beta$ can be estimated from the local variance. The resulting sequence of local estimates then provides an effective time series $\beta(t)$, from which the distribution and correlation time of the $\beta$ process can be reconstructed. Here, we extend this procedure by adding a second level of inference: Once the intermediate process $\beta(t)$ has been reconstructed, its local statistics are used to identify the still slower process $q(t)$ and its distribution $g(q)$.

### A. Reconstruction of the intermediate process

We first describe the reconstruction of the intermediate inverse-temperature process from a velocitylike time series. Let $u(t)$ denote the measured fast variable. For any time-dependent quantity $X(t)$ constructed from this signal, we define the local average over a window of duration $\Delta$ by

$$\langle X\rangle_{t,\Delta} = \frac{1}{\Delta}\int_t^{t+\Delta} X(t')\,dt'. \tag{55}$$

The local mean and central moments of $u(t)$ in the same window are

$$\mu_{t,\Delta} = \langle u\rangle_{t,\Delta}, \quad M_k(t,\Delta) = \langle (u-\mu_{t,\Delta})^k\rangle_{t,\Delta}. \tag{56}$$

The average local flatness is then

$$\mathcal{K}_u(\Delta) = \left\langle \frac{M_4(t,\Delta)}{[M_2(t,\Delta)]^2}\right\rangle_t, \tag{57}$$

where the outer average is taken over the positions of the time window. Following the ordinary superstatistical analysis [31], the intermediate scale $T_\beta$ is determined from the condition

$$\mathcal{K}_u(T_\beta) = 3. \tag{58}$$

This condition identifies a scale at which the local distribution of $u$ is approximately Gaussian, while its variance still changes appreciably between different windows. A necessary consistency condition is

$$\tau_u \ll T_\beta, \tag{59}$$

where $\tau_u$ is the relaxation time extracted, for example, from the autocorrelation function of $u(t)$.

For the dimensionless velocity defined in Eq. (19), the local stationary distribution at fixed $\beta$ has zero mean and variance

$$\mathrm{Var}_\beta(u) = \langle u^2\rangle_\beta - \langle u\rangle_\beta^2 = \langle u^2\rangle_\beta = \frac{\beta_0}{2\beta}. \tag{60}$$

Assuming that $\beta$ is approximately constant over a window of duration $T_\beta$, the theoretical variance in Eq. (60) can be identified with the empirical local variance of the signal. The inverse-temperature process can therefore be estimated as

$$\widehat{\beta}(t) = \frac{\beta_0}{2\left[\langle u^2\rangle_{t,T_\beta} - \langle u\rangle^2_{t,T_\beta}\right]}. \tag{61}$$

Only the relative fluctuations of $\widehat{\beta}(t)$ are needed to infer $q$, since $q$ is determined from moment ratios such as $\langle\beta^2\rangle/\langle\beta\rangle^2$. These ratios are invariant under a constant rescaling of $\beta$, so that the absolute calibration of $\beta_0$ does not affect the reconstruction of $q$.

### B. Second-level inference

Once the intermediate process $\widehat{\beta}(t)$ has been reconstructed, it can be analyzed as the fluctuating variable of the second statistical level. The idea is to divide the $\widehat{\beta}$ time series into longer windows, within which the value of $q$ is approximately constant while the fluctuations of $\widehat{\beta}$ are sufficiently well sampled. Denoting the duration of these outer windows by $\Delta_q$, the required hierarchy is

$$T_\beta \ll \Delta_q \ll \tau_q, \tag{62}$$

where $\tau_q$ is the characteristic timescale of the slowest process.

Let the $\ell$th outer window be centered at time $t_\ell$. Within this window, we compute the local moment ratio

$$R_\ell = \frac{\langle\widehat{\beta}^2\rangle_{t_\ell,\Delta_q}}{\langle\widehat{\beta}\rangle^2_{t_\ell,\Delta_q}}. \tag{63}$$

For a $\chi^2$ distribution sampled directly, this ratio would itself provide the local estimate of $q := \langle\beta^2\rangle/\langle\beta\rangle^2$.

Within each outer window, the local fluctuations of $\widehat{\beta}$ are then used to estimate a local value of the superstatistical parameter $\widehat{q}_\ell$. Operationally, this can be done either from local moment ratios or from a local maximum-likelihood fit of the conditional $\widehat{\beta}$ statistics. Because the Langevin dynamics samples locally normalized Maxwellian states, the relation between these local statistics and the type-A parameter $q$ is slightly different from the standard $\chi^2$ moment relation. The explicit estimators and the consistency relations used in the reconstruction are given in Appendix C.

The choice of $\Delta_q$ is part of the reconstruction. If $\Delta_q$ is too short, the local fluctuations of $\widehat{\beta}$ are poorly sampled and the estimates $\widehat{q}_\ell$ are noisy. If $\Delta_q$ is too long, different values of $q$ are mixed within the same window and the reconstructed distribution is artificially smoothed. The appropriate range of $\Delta_q$ is therefore the range over which the conditional $\widehat{\beta}$ statistics remain compatible with the local $\chi^2$ form and the estimates of $\widehat{q}_\ell$ are stable. Once the sequence $\widehat{q}_\ell$ has been obtained, its autocorrelation function provides an independent estimate of the slow timescale $\tau_q$.

### C. Empirical reconstruction of $g(q)$

The sequence of reconstructed values $\widehat{q}_\ell$ provides an empirical sampling of the slow superstatistical fluctuations. It can therefore be used to construct a nonparametric approximation of the distribution $g(q)$. In its most general weighted form, this empirical measure is written as

$$\widehat{g}_M(q) = \sum_{\ell=1}^{M} w_\ell\, \delta(q - \widehat{q}_\ell), \qquad \sum_{\ell=1}^{M} w_\ell = 1. \tag{64}$$

For uniformly sampled values, one has $w_\ell = 1/M$, while more general weights can be used to account for the sampling procedure. The explicit expressions used in the reconstruction are given in Appendix C.

The corresponding effective factor reconstructed from the time series is then

$$\widehat{\overline{B}}_M(E) = \sum_{\ell=1}^{M} w_\ell [1 + (\widehat{q}_\ell - 1)\beta_0 E]^{-\frac{1}{\widehat{q}_\ell - 1}}. \tag{65}$$

Thus, a finite superposition of $q$ exponentials arises naturally as the empirical representation of an otherwise unknown distribution $g(q)$. The bimodal model considered in Sec. V corresponds to the minimal nontrivial case in which the empirical measure is concentrated around two well-separated values of $q$. A continuous distribution is recovered in the weak limit as the number of resolved values increases.

### D. Numerical reconstruction test

We now test the reconstruction procedure on the hierarchical Langevin dynamics introduced in Sec. III. In contrast with the direct comparison performed there, the environmental variables are not used in the analysis. Only the simulated trajectory of the fast variable $u(t)$ is retained, and the processes $\beta(t)$ and $q(t)$ are treated as hidden variables.

The slowest fluctuations are chosen so as to generate a smooth distribution of the superstatistical parameter $q$. In the present example, this distribution is obtained from a log-normal modulation of the slow variable controlling $q$, as specified in Appendix C. This choice provides a continuous target distribution rather than a finite number of prescribed states, and therefore tests whether a broad distribution $g(q)$ can be recovered from the temporal organization of the fast signal alone.

The analysis proceeds in three steps. First, the relaxation time of $u$ is estimated from its autocorrelation function, and the intermediate scale $T_\beta$ is determined from the local-flatness condition (58). Second, the intermediate process $\widehat{\beta}(t)$ is reconstructed from the local variance of $u(t)$ using Eq. (61). Third, the reconstructed $\widehat{\beta}(t)$ signal is divided into outer windows, from which the local values $\widehat{q}_\ell$ are inferred. These values define the empirical measure $\widehat{g}_M(q)$ and the corresponding reconstructed effective factor $\widehat{\overline{B}}_M(E)$.

The result is shown in Fig. 5. Panel (a) shows how the average local flatness $\mathcal{K}_u(\Delta)$ changes with the window duration. The intermediate scale $T_\beta$ is not imposed externally; it is selected operationally as the value of $\Delta$ for which $\mathcal{K}_u(\Delta) = 3$. This is the scale at which the local statistics of the fast variable become approximately Gaussian.

Panel (b) provides a consistency check on the temporal hierarchy. The autocorrelations of $u$ and of the reconstructed $\widehat{\beta}$ decay on scales much shorter than the slow decay of $\widehat{q}$. Together with the selection of $T_\beta$ in panel (a), this supports the ordering

$$\tau_u \ll T_\beta \ll \tau_q. \tag{66}$$

This separation is essential for the reconstruction: $T_\beta$ must be long enough for the fast variable to be locally Gaussian, while the outer windows used to infer $\widehat{q}_\ell$ must remain short compared with the evolution time of the slowest process.

Panel (c) compares the reconstructed empirical distribution $\widehat{g}_M(q)$ with the prescribed distribution $g(q)$. Although the reconstruction is necessarily coarse grained by the finite number of outer windows, it recovers both the location and the width of the underlying distribution. This shows that the slow fluctuations of the superstatistical parameter can be inferred without direct access to the hidden environmental variables.

Finally, panel (d) tests the reconstructed distribution at the level of the effective statistical factor. The ratio $\widehat{\overline{B}}_M(x)/\overline{B}_{\text{true}}(x)$ shows that the empirical reconstruction captures the main effect of the fluctuating superstatistical parameter, but also that finite-window and finite-sample errors are amplified at large $x$. This sensitivity is expected, since the large-$x$ behavior depends strongly on the tail of the reconstructed distribution $g(q)$. Nevertheless, the reconstructed factor remains substantially closer to the exact nested result than the fixed-$q$ approximation evaluated at the reconstructed mean value $\widehat{q}_0$. This shows that retaining the distribution of $q$ contains information that is lost when the nested fluctuations are reduced to a single effective value.

This numerical experiment gives an operational meaning to the finite-mixture representation introduced above. The discrete measure in Eq. (64) should not be interpreted

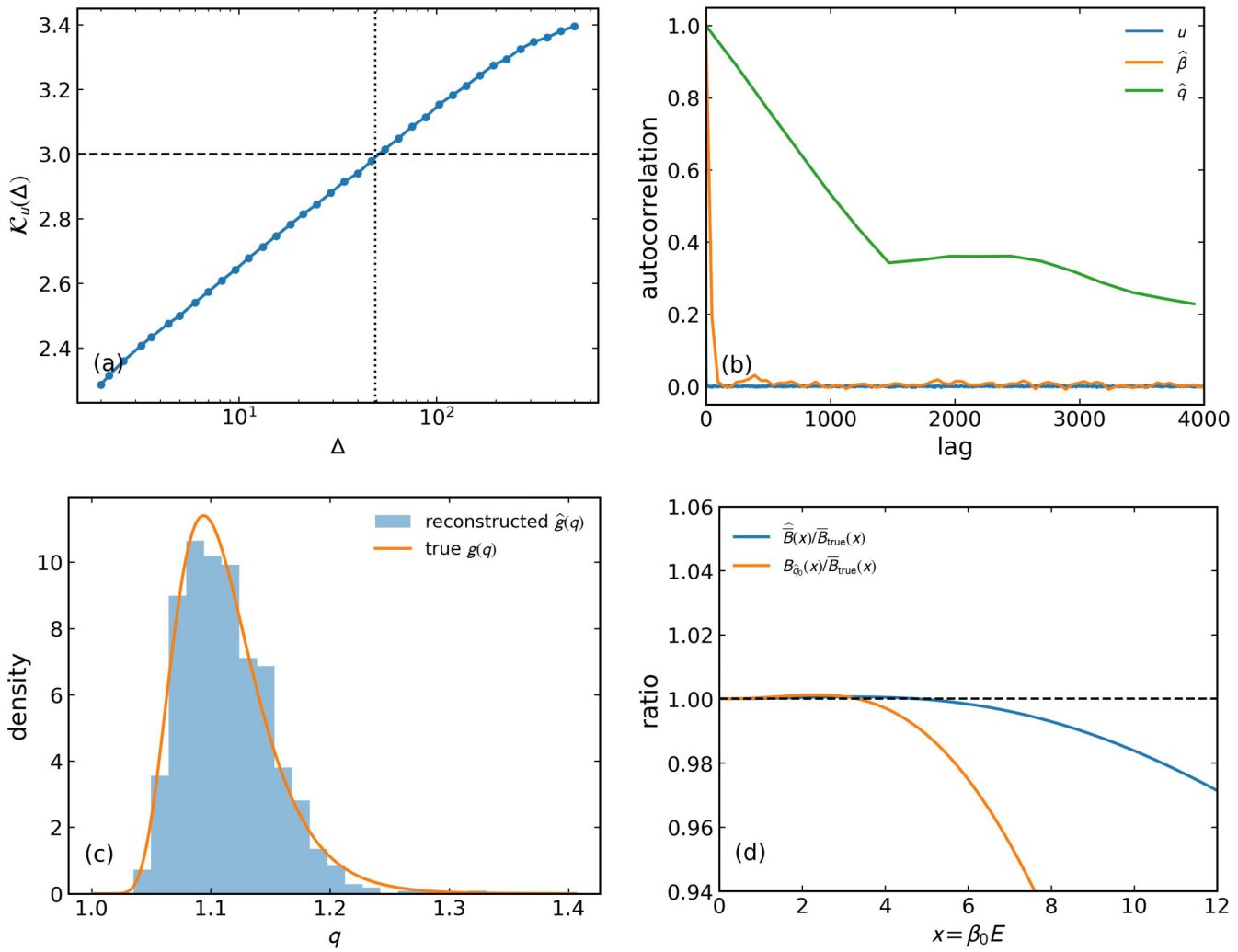



FIG. 5. Reconstruction of nested superstatistics from a single simulated time series. The environmental variables $\beta(t)$ and $q(t)$ are hidden during the analysis, and only the fast trajectory $u(t)$ is used. (a) Average local flatness $\mathcal{K}_u(\Delta)$ as a function of the window duration $\Delta$. The horizontal dashed line marks the Gaussian value $\mathcal{K}_u = 3$, and the vertical dotted line marks the value $\Delta = T_\beta$ selected by this criterion. (b) Normalized autocorrelation functions of the observed variable $u$ and of the reconstructed processes $\widehat{\beta}$ and $\widehat{q}$, plotted as functions of the time lag. The much slower decay of $\widehat{q}$ compared with $u$ and $\widehat{\beta}$ displays the separation between the slowest level and the faster reconstructed variables. (c) Comparison between the prescribed distribution $g(q)$ and the empirical distribution $\widehat{g}_M(q)$ reconstructed from the time series. (d) Ratios of the reconstructed effective factor $\widehat{\overline{B}}_M(x)$ and of the fixed-$q$ approximation $\overline{B}_{\widehat{q}_0}(x)$ to the exact nested factor, with $x = \beta_0 E$ and $\widehat{q}_0 = \sum_\ell w_\ell \widehat{q}_\ell$.

as requiring genuinely discrete physical states. Rather, it is the finite-resolution empirical representation of a continuous distribution of slow superstatistical fluctuations.

## VII. CONCLUSION

In this paper, we have introduced a nested extension of superstatistics in which the parameter controlling the strength of the fluctuations is itself allowed to fluctuate. We focused on the physically important case of a $\chi^2$-distributed inverse temperature, for which ordinary superstatistics leads to the family of Tsallis, or $q$-exponential, distributions [20,21,31]. These distributions appear in a wide range of nonequilibrium systems [26–30]. In the $\chi^2$ superstatistical setting, the parameter $q$ has a direct interpretation as a measure of the intensity of the inverse-temperature fluctuations [31]. The central idea developed here is that this intensity need not be fixed. In heterogeneous nonequilibrium systems, different regions or time intervals may be characterized by different fluctuation strengths, correlation structures, or effective numbers of active microscopic contributions. This naturally leads to a three-level hierarchy: a fast microscopic variable, a more slowly fluctuating inverse temperature $\beta$, and a still slower fluctuation of the superstatistical parameter $q$, or equivalently of the effective number of degrees of freedom.

We have shown explicitly that this nested-superstatistical construction can emerge dynamically in a Langevin model, provided that the three relevant timescales are sufficiently well separated. In the weak-fluctuation regime, the effect of the fluctuating parameter reduces to a universal correction governed only by the mean and variance of $q$. Beyond this perturbative regime, simple nontrivial choices of

$g(q)$ can generate observable features that are absent from fixed-$q$ superstatistics. This was illustrated with two examples: the electron velocity distribution in the solar wind, where a bimodal distribution of $q$ accounts simultaneously for the quasi-Maxwellian core and the suprathermal tail, and the cosmic-ray electron spectrum, where it naturally produces a progressive spectral hardening. Finally, we have provided a procedure for reconstructing $g(q)$ from a time series of the fast variable, thereby making the nested construction directly testable from dynamical data.

The present work also opens several directions, both at the fundamental level and in applications. At the fundamental level, a natural direction is to extend the nested construction to the other universality classes of superstatistics. Indeed, besides the $\chi^2$ class considered here, which is the most prominent because of its direct connection with Tsallis statistics, two other universality classes have been identified: inverse-$\chi^2$ superstatistics and log-normal superstatistics [31,57]. The latter is particularly relevant when the fluctuations of $\beta$ are not generated by additive microscopic random variables, but by multiplicative processes, as in turbulence and cascade dynamics [9,31–33]. It would therefore be interesting to examine whether the fluctuation parameter associated with these classes can also develop slow fluctuations, leading to analogous nested constructions.

At the applied level, an important direction is to confront the nested construction with high-precision data for which a single Tsallis or $q$-exponential distribution gives a good first description but leaves systematic residuals. The aim is not merely to introduce additional fitting flexibility. Rather, the question is whether these residuals can be organized in terms of a distribution $g(q)$, whose width would quantify slow heterogeneity in the local fluctuation intensity. In this sense, deviations from a fixed-$q$ description may provide direct evidence for a second level of superstatistical fluctuations. This possibility should be relevant in systems such as turbulent flows, space and astrophysical plasmas, and high-energy particle spectra, where approximate $q$-exponential behavior is common but a single value of $q$ is often too restrictive [6,7,9,11–14,54].


## ACKNOWLEDGMENT

I thank Roberto Onofrio for reading the manuscript and for his warm hospitality during my visit to Dartmouth College, where I had the idea for the present work.


## DATA AVAILABILITY

The PYTHON codes used to perform the Langevin simulations, compute the survival probabilities, and reconstruct the distribution $g(q)$ from simulated time-series data are publicly available in Ref. [58]. The repository also contains the numerical arrays and parameter files used to generate Figs. 1 and 5. The observational data used for the solar-wind electron distributions and cosmic-ray electron spectra were taken from the publicly available sources cited in the manuscript.

## APPENDIX A: TYPE-A VERSUS TYPE-B SUPERSTATISTICS AND THE LANGEVIN MODEL

This Appendix presents the difference between type-A and type-B superstatistics and collects the technical details connecting the type-A nested-superstatistical weights used in the main text to the normalized distributions sampled by the Langevin dynamics. It also gives the corresponding adiabatic reference distribution used in Sec. III.

In the type-A formulation adopted in the main text, the unnormalized Boltzmann factor is averaged over the inverse-temperature distribution and the resulting effective weight is normalized only afterward. Thus, for an energy variable $E$,

$$B_A(E) \propto \int_0^\infty f(\beta)\exp(-\beta E)\,d\beta. \tag{A1}$$

The alternative type-B formulation instead averages locally normalized canonical distributions,

$$B_B(E) = \int_0^\infty f(\beta)\frac{\exp(-\beta E)}{Z(\beta)}\,d\beta, \tag{A2}$$

where $Z(\beta)$ is the local partition function. The two prescriptions are related by a reweighting of the inverse-temperature density: A type-B average with $f(\beta)$ is equivalent to a type-A average with

$$\widetilde{f}(\beta) = \frac{f(\beta)/Z(\beta)}{\displaystyle\int_0^\infty f(\beta')/Z(\beta')\,d\beta'}. \tag{A3}$$

This point is relevant for the Langevin dynamics used in Sec. III. For the dimensionless velocity (19), the local stationary distribution at fixed $\beta$ is

$$p(u \mid \beta) = \frac{1}{Z_u(\beta)}\exp\left(-\frac{\beta u^2}{\beta_0}\right), \quad Z_u(\beta) = \sqrt{\frac{\pi\beta_0}{\beta}}. \tag{A4}$$

If the Langevin process sampled the type-A density $f_n(\beta)$ directly, the resulting marginal would contain the additional factor $1/Z_u(\beta)$ and would therefore correspond to a type-B average. To reproduce the type-A weight dynamically, the residence distribution of $\beta$ at fixed $n$ must instead be

$$\rho(\beta \mid n) = \frac{Z_u(\beta)f_n(\beta)}{\mathcal{Z}_n}, \tag{A5}$$

with

$$\mathcal{Z}_n = \int_0^\infty Z_u(\beta)f_n(\beta)\,d\beta. \tag{A6}$$

For the $\chi^2$ weight (2), this normalization is

$$\mathcal{Z}_n = \sqrt{\frac{\pi n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \tag{A7}$$

and the residence distribution becomes

$$\rho(\beta \mid n) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{n}{2\beta_0}\right)^{\frac{n-1}{2}}\beta^{\frac{n-3}{2}}\exp\left(-\frac{n\beta}{2\beta_0}\right), \tag{A8}$$

which is well defined for $n > 1$. Thus, the residence distribution is again a $\chi^2$ distribution, but with the shape parameter shifted from $n/2$ to $(n-1)/2$ because of the local Maxwellian normalization.

The same reweighting must be applied at the slower level. If $h(n)$ denotes the prescribed type-A distribution of the effective number of degrees of freedom, the values of $n$ visited by the Langevin protocol must be drawn from

$$\rho_n(n) = \frac{h(n)\mathcal{Z}_n}{\int_0^\infty h(n')\mathcal{Z}_{n'}\,dn'}. \tag{A9}$$

The factor $\mathcal{Z}_n$ appears because different values of $n$ have different intermediate normalizations after averaging over the locally sampled values of $\beta$.

The stationary marginal generated by the locally normalized Langevin dynamics is then

$$p_{\rm dyn}(u) = \int_0^\infty \rho_n(n) \int_0^\infty \rho(\beta \mid n) p(u \mid \beta)\, d\beta\, dn. \tag{A10}$$

Substituting Eqs. (A4)–(A9) gives

$$p_{\rm dyn}(u) = \frac{\int_0^\infty h(n) B_n\big(u^2/\beta_0\big)\, dn}{\int_0^\infty h(n)\mathcal{Z}_n\, dn}. \tag{A11}$$

This is the normalized type-A nested-superstatistical reference used in Sec. III.

The distinction between the type-A and type-B prescriptions affects the precise reference distribution, but not the dynamical conclusion tested in Sec. III. If the Langevin protocol is implemented directly with the unmodified densities $f_n(\beta)$ and $h(n)$, the corresponding adiabatic limit is the type-B nested distribution. In that case, the same numerical test shows the same qualitative behavior: The slow dynamics approaches its appropriate adiabatic reference, whereas the fast dynamics departs from it. The role of the reweighting used above is therefore not to create the timescale effect, but only to ensure that the adiabatic reference coincides with the type-A construction adopted throughout the main text.

## APPENDIX B: FINITE-TIME-SCALE CORRECTIONS TO THE ADIABATIC LIMIT

The adiabatic reference associated with the hierarchical Langevin dynamics is recovered when the relaxation of the velocity is much faster than the evolution of the environmental variables. Here, we derive the leading corrections produced by a finite separation of timescales for the piecewise-constant protocol used in Sec. III. We assume

$$\tau_v \ll \tau_\beta \ll \tau_n, \tag{B1}$$

and, for simplicity, that $\tau_n = N\tau_\beta$, with $N$ an integer. The velocity is sampled along the stationary trajectory, and the corresponding adiabatic reference is derived in Appendix A.

Using the dimensionless velocity (19), the Langevin dynamics at fixed $\beta$ has stationary variance

$$w(\beta) = \frac{\beta_0}{2\beta}. \tag{B2}$$

Over a time interval $s$ during which $\beta$ remains constant, the Ornstein-Uhlenbeck propagator gives

$$u(t+s) = a(s)u(t) + \sqrt{1-a^2(s)}\,\xi_\beta, \quad a(s) = e^{-s/\tau_v}, \tag{B3}$$

where $\xi_\beta$ is a centered Gaussian variable with variance $w(\beta)$. At fixed $n$, let

$$\Phi_n(k) = \int_0^\infty \rho(\beta \mid n) \exp\left[-\frac{k^2}{2} w(\beta)\right] d\beta \tag{B4}$$

denote the characteristic function of the adiabatic velocity distribution. Averaging over the fluctuations of $n$ gives

$$\Phi_{\rm ad}(k) = \int_0^\infty \rho_n(n)\Phi_n(k)\, dn. \tag{B5}$$

After a renewal of $\beta$ at fixed $n$, Eq. (B3) yields the transient characteristic function

$$\Phi^{\rm tr}_\beta(k, s \mid n) = \Phi_n(a(s)k)\Phi_n(\sqrt{1-a^2(s)}\,k). \tag{B6}$$

If the renewal also changes $n$, averaging independently over the old and new environments gives

$$\Phi^{\rm tr}_n(k, s) = \Phi_{\rm ad}(a(s)k)\Phi_{\rm ad}(\sqrt{1-a^2(s)}\,k). \tag{B7}$$

Since each transient lasts for a time of order $\tau_v$, their leading contributions can be added independently in the adiabatic regime. Because $\tau_v \ll \tau_\beta$, the upper integration limit may be extended to infinity. The stationary characteristic function is then

$$\Phi(k) \simeq \Phi_{\rm ad}(k) + \left(\frac{1}{\tau_\beta} - \frac{1}{\tau_n}\right)\mathcal{C}_\beta(k) + \frac{1}{\tau_n}\mathcal{C}_n(k), \tag{B8}$$

where

$$\mathcal{C}_\beta(k) = \int_0^\infty ds\left[\int_0^\infty \rho_n(n)\Phi^{\rm tr}_\beta(k, s \mid n)\, dn - \Phi_{\rm ad}(k)\right] \tag{B9}$$

and

$$\mathcal{C}_n(k) = \int_0^\infty ds\big[\Phi^{\rm tr}_n(k, s) - \Phi_{\rm ad}(k)\big]. \tag{B10}$$

Equations (B8)–(B10) show that the two finite-time-scale corrections are controlled by $\tau_v/\tau_\beta$ and $\tau_v/\tau_n$, respectively.

A simple consequence follows for the low-order moments. Define

$$A_1(n) = \langle w(\beta)\rangle_{\beta|n}, \quad A_2(n) = \langle w^2(\beta)\rangle_{\beta|n}, \tag{B11}$$

where the conditional averages are taken with respect to $\rho(\beta \mid n)$, and angular brackets with subscript $\rho_n$ denote averaging over $\rho_n(n)$. The second moment is unaffected by the renewal transients and therefore remains

$$\langle u^2\rangle = \langle A_1\rangle_{\rho_n}. \tag{B12}$$

At fixed $n$, the transient fourth moment following a renewal of $\beta$ is

$$\langle u^4\rangle_{\rm tr} = 3A_2 - 6a^2(s)[1-a^2(s)](A_2 - A_1^2). \tag{B13}$$

Using

$$\int_0^\infty a^2(s)[1-a^2(s)]\, ds = \frac{\tau_v}{4} \tag{B14}$$

and combining the contributions from renewals of $\beta$ and $n$, one obtains

$$\langle u^4\rangle \simeq 3\langle A_2\rangle_{\rho_n} - \frac{3\tau_v}{2\tau_\beta}\langle A_2 - A_1^2\rangle_{\rho_n} - \frac{3\tau_v}{2\tau_n}\,\mathrm{Var}_{\rho_n}(A_1). \quad \text{(B15)}$$

The corresponding excess kurtosis is therefore

$$\kappa \simeq \kappa_{\mathrm{ad}} - \frac{3\tau_v}{2\langle A_1\rangle_{\rho_n}^2}\left[\frac{\langle A_2 - A_1^2\rangle_{\rho_n}}{\tau_\beta} + \frac{\mathrm{Var}_{\rho_n}(A_1)}{\tau_n}\right], \quad \text{(B16)}$$

with

$$\kappa_{\mathrm{ad}} = \frac{3\langle A_2\rangle_{\rho_n}}{\langle A_1\rangle_{\rho_n}^2} - 3. \quad \text{(B17)}$$

Both corrections in Eq. (B16) are nonpositive. A finite separation of timescales therefore reduces the excess kurtosis as compared to the adiabatic value, consistently with the suppression of the tails observed for the fast dynamics in Fig. 1.

For the $\chi^2$ weight (2), the sampling distribution is the $\chi^2$ law given in Eq. (A8). One then finds at fixed $n$,

$$A_1(n) = \frac{n}{2(n-3)}, \quad A_2(n) = \frac{n^2}{4(n-3)(n-5)} \quad (n > 5). \quad \text{(B18)}$$

The adiabatic excess kurtosis is then $\kappa_{\mathrm{ad}} = 6/(n-5)$, and Eq. (B16) reduces to

$$\kappa \simeq \frac{6}{n-5}\left(1 - \frac{\tau_v}{2\tau_\beta}\right). \quad \text{(B19)}$$

The moment formulas require the corresponding moments to exist. The second-moment result requires $\rho_n(n)$ to be supported on $n > 3$, while the fourth-moment and kurtosis formulas require support on $n > 5$. A distribution of $n$ extending below these thresholds must therefore be suitably truncated if the corresponding ensemble moments are to remain finite. In contrast, the characteristic-function correction in Eq. (B8) does not rely on the existence of these moments and remains the more general result.

## APPENDIX C: RECONSTRUCTION DETAILS FOR $g(q)$

This Appendix gives the technical details behind the reconstruction procedure used in Sec. VI. In the main text, the measured signal $u(t)$ is first converted into an intermediate reconstructed process $\widehat{\beta}(t)$ through Eq. (61). The second level of the reconstruction then requires three additional ingredients: the estimator of the local values $\widehat{q}_\ell$, the weights $w_\ell$ entering the empirical measure (64), and the reference distribution $g_{\mathrm{true}}(q)$ used in the numerical test.

We first derive the estimator of $\widehat{q}_\ell$ from the local statistics of $\widehat{\beta}(t)$. At fixed $n$, the conditional law of $\beta$ sampled by the locally normalized Langevin dynamics is given in Eq. (A8). Using

$$q = 1 + \frac{2}{n}, \quad \text{(C1)}$$

this conditional law can be written in terms of $q$ as

$$\rho(\beta \mid q) = \frac{1}{\Gamma(\nu(q)/2)[(q-1)\beta_0]^{\nu(q)/2}} \times \beta^{\nu(q)/2-1}\exp\left[-\frac{\beta}{(q-1)\beta_0}\right], \quad \text{(C2)}$$

where

$$\nu(q) = \frac{2}{q-1} - 1 = n - 1. \quad \text{(C3)}$$

Equivalently,

$$\frac{2\beta}{(q-1)\beta_0} \sim \chi^2_{\nu(q)}. \quad \text{(C4)}$$

Thus, the reconstructed values of $\widehat{\beta}$ inside one outer window are expected to follow a scaled $\chi^2$ law whose effective number of degrees of freedom is determined by the local value of $q$.

The moment-ratio estimator used in Sec. VI follows directly from this conditional law. Its moments satisfy

$$\frac{\langle\beta^p\rangle_q}{\beta_0^p} = \prod_{j=0}^{p-1}\left[1 + \left(j - \frac{1}{2}\right)(q-1)\right], \quad \text{(C5)}$$

for positive integers $p$. In particular,

$$R_q := \frac{\langle\beta^2\rangle_q}{\langle\beta\rangle_q^2} = 1 + \frac{2}{\nu(q)} = 1 + \frac{2}{n-1}. \quad \text{(C6)}$$

This is the theoretical counterpart of the empirical ratio $R_\ell$ defined in Eq. (63). Since $R_q$ is not identical to the parameter $q$, Eq. (C6) must be inverted. This gives

$$q = \frac{3R_q - 1}{R_q + 1}. \quad \text{(C7)}$$

Therefore, in the $\ell$th outer window, one may estimate

$$\widehat{R}_\ell = \frac{\langle\widehat{\beta}^2\rangle_{t_\ell,\Delta_q}}{\langle\widehat{\beta}\rangle^2_{t_\ell,\Delta_q}} \quad \text{(C8)}$$

and then set

$$\widehat{q}_\ell = \frac{3\widehat{R}_\ell - 1}{\widehat{R}_\ell + 1}. \quad \text{(C9)}$$

This is the explicit moment estimator referred to in the second-level inference step of Sec. VI.

For finite samples, a maximum-likelihood estimate may be more stable. Treating the scale of the scaled $\chi^2$ law as a nuisance parameter, the estimated effective number of degrees of freedom $\widehat{\nu}_\ell$ in the $\ell$th outer window satisfies

$$\ln\left(\frac{\widehat{\nu}_\ell}{2}\right) - \psi\left(\frac{\widehat{\nu}_\ell}{2}\right) = \ln\overline{\widehat{\beta}}_\ell - \overline{\ln\widehat{\beta}}_\ell, \quad \text{(C10)}$$

where $\psi$ is the digamma function and the averages are evaluated within the same outer window. The corresponding local estimate of the superstatistical parameter is

$$\widehat{q}_\ell = 1 + \frac{2}{\widehat{\nu}_\ell + 1}. \quad \text{(C11)}$$

Equations (C9) and (C11) provide the two operational ways of obtaining the sequence $\widehat{q}_\ell$ used in the empirical reconstruction.

The local $\chi^2$ assumption can be checked through moment-closure relations. If

$$M_r(t,\Delta) = \langle \widehat{\beta}^r \rangle_{t,\Delta}, \quad R(t,\Delta) = \frac{M_2(t,\Delta)}{M_1(t,\Delta)^2}, \tag{C12}$$

then an exact scaled $\chi^2$ law at fixed $q$ satisfies

$$\frac{M_3}{M_1^3} = R(2R-1) \tag{C13}$$

and

$$\frac{M_4}{M_1^4} = R(2R-1)(3R-2). \tag{C14}$$

These relations provide the consistency checks mentioned in Sec. VI. Significant deviations indicate either insufficient statistics inside the outer window or a window duration large enough to mix different values of $q$.

We now explain the weights $w_\ell$ appearing in the empirical measure (64). The values $\widehat{q}_\ell$ reconstructed from the time-series sample the temporal distribution generated by the slow dynamics. The distribution entering the nested statistical factor is obtained by correcting this temporal sampling with the normalization factor $\mathcal{Z}_n$ defined in Appendix A. With

$$n(q) = \frac{2}{q-1}, \tag{C15}$$

the empirical distribution is written as

$$\widehat{g}_M(q) = \sum_{\ell=1}^{M} w_\ell \, \delta(q - \widehat{q}_\ell), \tag{C16}$$

where

$$w_\ell = \frac{d_\ell / \mathcal{Z}_{\widehat{n}_\ell}}{\sum_{j=1}^{M} d_j / \mathcal{Z}_{\widehat{n}_j}}, \quad \widehat{n}_\ell = \frac{2}{\widehat{q}_\ell - 1}. \tag{C17}$$

Here, $d_\ell$ is the duration represented by the $\ell$th outer window. For equally spaced windows of equal duration, the factors $d_\ell$ cancel. These are the weights used in Eq. (64) and in the reconstructed effective factor (65).

The moments of the reconstructed distribution are then estimated from the same weighted empirical measure:

$$\widehat{q}_0 = \sum_{\ell=1}^{M} w_\ell \widehat{q}_\ell, \quad \widehat{\langle q^2 \rangle} = \sum_{\ell=1}^{M} w_\ell \widehat{q}_\ell^2, \tag{C18}$$

and

$$\widehat{\alpha} = \frac{\sum_{\ell=1}^{M} w_\ell \widehat{q}_\ell^2}{\left( \sum_{\ell=1}^{M} w_\ell \widehat{q}_\ell \right)^2}. \tag{C19}$$

These quantities cannot be obtained directly from global moment ratios of the full reconstructed process $\widehat{\beta}(t)$, because the conditional mean of $\widehat{\beta}$ itself depends on the local value of $q$.

Finally, we give the reference distribution used in the numerical reconstruction test. In that test, the slow modulation is prescribed through

$$n = n_0 \exp(\eta), \quad \eta \sim \mathcal{N}\left( -\frac{\sigma_\eta^2}{2}, \sigma_\eta^2 \right). \tag{C20}$$

Using $q = 1 + 2/n$, the induced distribution of $q$ is

$$g_{\text{true}}(q) = \frac{1}{\mathcal{N}_q} \frac{1}{(q-1)\sigma_\eta \sqrt{2\pi}} \times \exp\left\{ -\frac{\left[ \ln\left( \frac{2}{n_0(q-1)} \right) + \frac{\sigma_\eta^2}{2} \right]^2}{2\sigma_\eta^2} \right\}, \tag{C21}$$

where $\mathcal{N}_q$ is the normalization constant over the domain of $q$ used in the numerical reconstruction. This is the distribution denoted $g_{\text{true}}(q)$ in Fig. 5.


[1] H. A. Simon, The architecture of complexity, Proc. Am. Philos. Soc. **106**, 467 (1962).

[2] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics* (Springer, Berlin, 1991).

[3] B. H. Lavenda, *Statistical Physics: A Probabilistic Approach* (Wiley, New York, 1991).

[4] O. Barndorff-Nielsen, J. Kent, M. Sørensen, and M. Sorensen, Normal variance-mean mixtures and $z$ distributions, Int. Stat. Rev. **50**, 145 (1982).

[5] C. Beck and E. G. D. Cohen, Superstatistics, Physica A **322**, 267 (2003).

[6] C. Beck, Statistics of three-dimensional Lagrangian turbulence, Phys. Rev. Lett. **98**, 064502 (2007).

[7] A. Reynolds, Superstatistical mechanics of tracer-particle motions in turbulence, Phys. Rev. Lett. **91**, 084503 (2003).

[8] M. Mehrafarin, Superstatistics as the statistics of quasiequilibrium states: Application to fully developed turbulence, Phys. Rev. E **84**, 022102 (2011).

[9] S. Jung and H. L. Swinney, Velocity difference statistics in turbulence, Phys. Rev. E **72**, 026304 (2005).

[10] F. Sattin and L. Salasnich, Multiparameter generalization of nonextensive statistical mechanics, Phys. Rev. E **65**, 035106(R) (2002).

[11] K. Ourabah, L. Ait Gougam, and M. Tribeche, Nonthermal and suprathermal distributions as a consequence of superstatistics, Phys. Rev. E **91**, 012133 (2015).

[12] S. Davis, G. Avaria, B. Bora, J. Jain, J. Moreno, C. Pavez, and L. Soto, Single-particle velocity distributions of collisionless, steady-state plasmas must follow superstatistics, Phys. Rev. E **100**, 023205 (2019).

[13] K. Ourabah, Demystifying the success of empirical distributions in space plasmas, Phys. Rev. Res. **2**, 023121 (2020).

[14] S. Davis, G. Avaria, B. Bora, J. Jain, J. Moreno, C. Pavez, and L. Soto, Kappa distribution from particle correlations in nonequilibrium, steady-state plasmas, Phys. Rev. E **108**, 065207 (2023).

[15] K. Ourabah, Quasiequilibrium self-gravitating systems, Phys. Rev. D **102**, 043017 (2020).

[16] K. Ourabah, Generalized statistical mechanics of stellar systems, Phys. Rev. E **105**, 064108 (2022).

[17] M. Iranmanesh, H. Arjomand Kermani, and K. Ourabah, Superstatistics and stellar rotation: Modeling velocity distributions in six stellar groups, New Astron. **119**, 102408 (2025).

[18] I. Rouse and S. Willitsch, Superstatistical velocity distributions of cold trapped ions in molecular-dynamics simulations, Phys. Rev. A **92**, 053420 (2015).

[19] I. Rouse and S. Willitsch, Superstatistical energy distributions of an ion in an ultracold buffer gas, Phys. Rev. Lett. **118**, 143401 (2017).

[20] G. Wilk and Z. Włodarczyk, Interpretation of the nonextensivity parameter $q$ in some applications of Tsallis statistics and Lévy distributions, Phys. Rev. Lett. **84**, 2770 (2000).

[21] C. Beck, Dynamical foundations of nonextensive statistical mechanics, Phys. Rev. Lett. **87**, 180601 (2001).

[22] S. Olbert, Summary of experimental results from M. I. T. detector on IMP-1, in *Physics of the Magnetosphere*, edited by R. L. Carovillano, J. F. McClay, and H. R. Radoski (D. Reidel, Dordrecht, 1968), pp. 641–659.

[23] V. M. Vasyliunas, A survey of low-energy electrons in the evening sector of the magnetosphere with OGO 1 and OGO 3, J. Geophys. Res. **73**, 2839 (1968).

[24] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys. **52**, 479 (1988).

[25] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World* (Springer, New York, 2009).

[26] J. Jersblad, H. Ellmann, K. Støchkel, A. Kastberg, L. Sanchez-Palencia, and R. Kaiser, Non-Gaussian velocity distributions in optical lattices, Phys. Rev. A **69**, 013410 (2004).

[27] P. Douglas, S. Bergamini, and F. Renzoni, Tunable Tsallis distributions in dissipative optical lattices, Phys. Rev. Lett. **96**, 110601 (2006).

[28] B. Liu and J. Goree, Superdiffusion and non-Gaussian statistics in a driven-dissipative 2D dusty plasma, Phys. Rev. Lett. **100**, 055003 (2008).

[29] G. Livadiotis and D. J. McComas, Beyond kappa distributions: Exploiting Tsallis statistical mechanics in space plasmas, J. Geophys. Res. **114**, A11105 (2009).

[30] J. Cleymans and D. Worku, Relativistic thermodynamics: Transverse momentum distributions in high-energy physics, Eur. Phys. J. A **48**, 160 (2012).

[31] C. Beck, E. G. D. Cohen, and H. L. Swinney, From time series to superstatistics, Phys. Rev. E **72**, 056133 (2005).

[32] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).

[33] B. Castaing, Y. Gagne, and E. J. Hopfinger, Velocity probability density functions of high Reynolds number turbulence, Physica D **46**, 177 (1990).

[34] R. Bruno and V. Carbone, The solar wind as a turbulence laboratory, Living Rev. Solar Phys. **10**, 2 (2013).

[35] G. Livadiotis and D. J. McComas, Understanding kappa distributions: A toolbox for space science and astrophysics, Space Sci. Rev. **175**, 183 (2013).

[36] K. Ourabah, Superstatistics from a dynamical perspective: Entropy and relaxation, Phys. Rev. E **109**, 014127 (2024).

[37] D. N. Sob'yanin, Generalization of the Beck-Cohen superstatistics, Phys. Rev. E **84**, 051128 (2011).

[38] D. N. Sob'yanin, Hierarchical maximum entropy principle for generalized superstatistical systems and Bose-Einstein condensation of light, Phys. Rev. E **85**, 061120 (2012).

[39] A. M. S. Macêdo, I. R. R. González, D. S. P. Salazar, and G. L. Vasconcelos, Universality classes of fluctuation dynamics in hierarchical complex systems, Phys. Rev. E **95**, 032315 (2017).

[40] G. L. Vasconcelos, D. S. P. Salazar, and A. M. S. Macêdo, Maximum entropy approach to $H$-theory: Statistical mechanics of hierarchical systems, Phys. Rev. E **97**, 022104 (2018).

[41] G.-A. Tsekouras and C. Tsallis, Generalized entropy arising from a distribution of $q$ indices, Phys. Rev. E **71**, 046144 (2005).

[42] J. C. Carvalho, J. D. do Nascimento, Jr., R. Silva, and J. R. de Medeiros, Non-Gaussian statistics and stellar rotational velocities of main-sequence field stars, Astrophys. J. Lett. **696**, L48 (2009).

[43] M. P. da Silva, M. M. F. de Lima, E. N. Velloso, and J.-D. do Nascimento, Jr., Nonextensive behavior of stellar rotation in the Galactic disk components, Astrophys. J. **958**, 32 (2023).

[44] K. Ourabah, Fingerprints of nonequilibrium stationary distributions in dispersion relations, Sci. Rep. **11**, 12103 (2021).

[45] G. E. Crooks, Beyond Boltzmann-Gibbs statistics: Maximum entropy hyperensembles out of equilibrium, Phys. Rev. E **75**, 041119 (2007).

[46] R. P. Lin, Wind observations of suprathermal electrons in the interplanetary medium, Space Sci. Rev. **86**, 61 (1998).

[47] M. Maksimovic, S. P. Gary, and R. M. Skoug, Solar wind electron suprathermal strength and temperature gradients: Ulysses observations, J. Geophys. Res. **105**, 18337 (2000).

[48] M. Maksimovic, I. Zouganelis, J.-Y. Chaufray, K. Issautier, E. E. Scime, J. E. Littleton, E. Marsch, D. J. McComas, C. Salem, R. P. Lin, and H. Elliott, Radial evolution of the electron distribution functions in the fast solar wind between 0.3 and 1.5 AU, J. Geophys. Res. **110**, A09104 (2005).

[49] Š. Štverák, P. Trávníček, M. Maksimovic, E. Marsch, A. N. Fazakerley, and E. E. Scime, Electron temperature anisotropy constraints in the solar wind, J. Geophys. Res. **113**, A03103 (2008).

[50] M. Lazar, V. Pierrard, S. M. Shaaban, H. Fichtner, and S. Poedts, Dual Maxwellian-Kappa modeling of the solar wind electrons: New clues on the temperature of Kappa populations, Astron. Astrophys. **602**, A44 (2017).

[51] S. J. Bame *et al.*, The ULYSSES solar wind plasma experiment, Astron. Astrophys. Suppl. Ser. **92**, 237 (1992).

[52] M. Aguilar *et al.* (AMS Collaboration), Towards understanding the origin of cosmic-ray electrons, Phys. Rev. Lett. **122**, 101101 (2019).

[53] D. Maurin, M. Ahlers, H. Dembinski, A. Haungs, P.-S. Mangeard, F. Melot, P. Mertsch, D. Wochele, and J. Wochele, A cosmic-ray database update: CRDB v4.1, Eur. Phys. J. C **83**, 971 (2023).

[54] G. C. Yalcin and C. Beck, Generalized statistical mechanics of cosmic rays: Application to positron-electron spectral indices, Sci. Rep. **8**, 1764 (2018).

[55] M. Korsmeier and A. Cuoco, Implications of lithium to oxygen AMS-02 spectra on our understanding of cosmic-ray diffusion, Phys. Rev. D **103**, 103016 (2021).

[56] M. Di Mauro, F. Donato, and S. Manconi, Novel interpretation of the latest AMS-02 cosmic-ray electron spectrum, Phys. Rev. D **104**, 083012 (2021).

[57] C. Beck, Recent developments in superstatistics, Braz. J. Phys. **39**, 357 (2009).

[58] K. Ourabah, kamourabah-oss/nested-superstatistics: Reproducibility package for nested superstatistics (Version v1.0.1) [Computer software], Zenodo, 2026, https://zenodo.org/records/21170588.