First-principles derivation of the $q$-canonical ensemble in Bayesian superstatistics

Sergio Davis$^{1,2}$

$^1$Research Center on the Intersection of Plasma Physics, Matter and Complexity (P$^2$mc), Comisión Chilena de Energía Nuclear, Casilla 188-D, Santiago, Chile
$^2$Departamento de Física, Facultad de Ciencias Exactas, Universidad Andres Bello. Sazié 2212, piso 7, 8370136, Santiago, Chile

E-mail: sergio.davis@cchen.cl

Abstract. In this work it is shown that the main ensemble of Tsallis nonextensive statistics, namely the $\chi^2$-superstatistics, can be obtained as a noninformative prior distribution that is singled out by the mathematical structure of superstatistics itself. There are no assumptions involved about the physics of the systems of interest, or regarding their complexity or range of interactions, which supports the thesis that the success of the $\chi^2$-superstatistics is information-theoretical in its origin, and explainable in terms of Jaynes’ original maximum entropy principle.

1. Introduction

Understanding the origin of non-canonical probability distributions occurring in Nature, particularly the $q$-canonical distributions, sometimes referred to as Tsallis distributions, is one of the remaining fundamental problems of Nonequilibrium Statistical Mechanics. Regarding this question, the issue of whether a generalization of the Boltzmann-Gibbs entropy, such as the one proposed by Tsallis [1, 2] is required to explain such distributions is still an open question, given the rise of alternative proposals such as Superstatistics [3, 4] among others [5, 6]. Superstatistics in particular is strongly founded on the rules of probability theory while still being compatible with Jaynes’ original maximum entropy principle [7] and is therefore an appealing choice of framework. Moreover, superstatistics is especially suited to the language of Bayesian probability, as has been shown in several different works [8–13].

In this work we show that the $\chi^2$-family of superstatistics, from which the Tsallis distributions are obtained, arise naturally from the superstatistical framework as a noninformative prior distribution of inverse temperature, that we recover from a standard application of the principle of maximum entropy and requiring only compatibility with the structure of superstatistics itself. We therefore provide a strong argument for the ubiquity of the $q$-canonical ensembles solely based on the standard principles of statistical mechanics and Bayesian/information theoretical reasoning, possibly enabling future insights regarding the meaning of the entropic index $q$. 
2. A Bayesian retelling of the canonical ensemble

As is well known, in standard Boltzmann-Gibbs statistics the probability distribution of microstates $\Gamma$ is the canonical ensemble, given by

$$P(\Gamma | \beta) = \frac{\exp(-\beta H(\Gamma))}{Z(\beta)},$$

where the normalization factor $Z(\beta)$ is the partition function,

$$Z(\beta) = \int_U d\Gamma \exp(-\beta H(\Gamma)).$$

The corresponding distribution of energies is, in turn, given by

$$P(E | \beta) = \frac{\exp(-\beta E)\Omega(E)}{Z(\beta)},$$

with $\Omega(E)$ the density of states

$$\Omega(E) = \int_U d\Gamma \delta(E - H(\Gamma)).$$

Upon simple inspection we notice an asymmetry between equations (1) and (3), in that the energy distribution is weighted by an extra factor $\Omega(E)$, and in order to put both distributions in the same footing we will introduce the idea of Bayesian ensembles. These are probability distributions of the form

$$P(\zeta | \mathcal{R}, I_0) = P(\zeta | I_0) \frac{P(\mathcal{R} | \zeta, I_0)}{P(\mathcal{R} | I_0)},$$

where $\zeta$ is a generic set of variables and $\mathcal{R}$ is some evidence or piece of information that updates the prior distribution $P(\zeta | I_0)$ to a posterior distribution $P(\zeta | \mathcal{R}, I_0)$ that incorporates $\mathcal{R}$. We see then that in (1) we have omitted the prior distribution of microstates $P(\Gamma | I_0)$, which was originally taken as a constant by virtue of the postulate of a priori equiprobability of microstates and thus cancelled out. Including this prior we should write

$$P(\Gamma | \beta, I_0) = P(\Gamma | I_0) \frac{\exp(-\beta H(\Gamma))}{Z_0(\beta)}.$$

as a replacement for (1) where now

$$Z_0(\beta) = \int_U d\Gamma P(\Gamma | I_0) \exp(-\beta H(\Gamma))$$

is the partition function. Following this argument, the density of states has to be replaced by the prior distribution of energy

$$P(E | I_0) := \int_U d\Gamma P(\Gamma | I_0) \delta(E - H(\Gamma)).$$
so that we can write

\[ P(\Gamma | \beta, I_0) = P(\Gamma | I_0) \frac{\exp(-\beta H(\Gamma))}{Z_0(\beta)}, \quad (9a) \]

\[ P(E | \beta, I_0) = P(E | I_0) \frac{\exp(-\beta E)}{Z_0(\beta)}. \quad (9b) \]

We can see that the definition of \( P(E | I_0) \) is consistent with the proper normalization of (9b), because

\[ Z_0(\beta) = \int dE P(E | I_0) \exp(-\beta E) \]

\[ = \int dE \left\{ \int_U d\Gamma P(\Gamma | I_0) \delta(E - H(\Gamma)) \right\} \exp(-\beta E) \]

\[ = \int_U d\Gamma P(\Gamma | I_0) \exp(-\beta H(\Gamma)). \quad (10) \]

Moreover, now the equations (9) can both be derived using the maximization of a single form of entropy, namely the parameterization-invariant Shannon-Jaynes entropy

\[ S(I \rightarrow I_0) = - \int d\zeta P(\zeta | I) \ln \left[ \frac{P(\zeta | I)}{P(\zeta | I_0)} \right] \quad (11) \]

under the constraint of fixed expected energy, where the generic variables \( \zeta \) can be replaced by either the microstate variables \( \Gamma \) or the energy \( E \). We notice also that both (8) and (9b) can be unified as

\[ P(E | S) = \left\langle \delta(E - H) \right\rangle_S = \int_U d\Gamma P(\Gamma | S) \delta(E - H(\Gamma)) \quad (12) \]

for \( S \) either \( I_0 \) or \( (\beta, I_0) \), or in general for any state of knowledge.

3. Generalized ensembles and superstatistics the Bayesian way

Now, instead of the constraint of the expected energy we will determine the effect of a constraint on the form of the energy distribution \( P(E | S) \), which we will assume is known and given by the function \( p(E) \). That is, we must impose

\[ P(E | S) = \left\langle \delta(H - E) \right\rangle_S = p(E) \quad \forall E \geq 0. \quad (13) \]

By maximizing the Shannon-Jaynes entropy in (11) from the prior \( P(\Gamma | I_0) \) we obtain

\[ P(\Gamma | S) = \frac{1}{\eta} P(\Gamma | I_0) \exp \left( - \int_0^\infty dE \mu(E) \delta(H(\Gamma) - E) \right) \]

\[ = \frac{1}{\eta} P(\Gamma | I_0) \exp(-\mu(H(\Gamma))), \quad (14) \]
where $\mu(E)$ is the Lagrange multiplier function conjugate to $p(E)$. In other words, we obtain the ensemble $S$ as

$$P(\Gamma|S) = P(\Gamma|I_0)\rho(H(\Gamma); S),$$

(15)

with $\rho(E; S)$ a non-negative function of the energy known as the ensemble function. This result is already consistent with Bayes’ theorem in equation (5) if we use $\zeta = \Gamma$ and let

$$\rho(H(\Gamma); S) = \frac{P(R|\Gamma, I_0)}{P(R|I_0)}.$$

(16)

The distribution of energy $P(E|S)$ that generalizes equation (3) is then

$$P(E|S) = \left< \delta(H - E) \right>_S = \int_U d\Gamma P(\Gamma|S)\delta(H(\Gamma) - E) = \int_U d\Gamma P(\Gamma|I_0)\rho(H(\Gamma))\delta(H(\Gamma) - E) = P(E|I_0)\rho(E; S),$$

(17)

again in agreement with Bayes’ theorem in (5) if $\zeta = E$ and we let

$$\rho(E; S) = \frac{P(R|E, I_0)}{P(R|I_0)}.$$

(18)

Using equations (14), (15) and (17) we have

$$\rho(E; S) = \frac{1}{\eta} \exp(-\mu(E)) = \frac{P(E|S)}{P(E|I_0)}$$

(19)

and then, if we let $p_0(E) = P(E|I_0)$ the maximum entropy ensemble is simply

$$P(\Gamma|S) = P(\Gamma|I_0) \left[ \frac{1}{\eta} \frac{p(H(\Gamma))}{p_0(H(\Gamma))} \right].$$

(20)

While, in principle, any pair of energy distributions $p(E)$ and $p_0(E)$ produces a valid steady-state ensemble in this way, we are interested in those compatible with the idea of temperature in the framework of superstatistics. This is a particular class of models defined by a distribution of inverse temperature $P(\beta|S)$. The central equation of superstatistics, as presented in recent works [14,15], is the joint distribution of microstates $\Gamma$ and inverse temperature $\beta$, which is given by the product rule of probability,

$$P(\Gamma, \beta|S) = P(\Gamma|\beta)P(\beta|S),$$

(21)

with $P(\Gamma|\beta)$ the canonical ensemble of equation (1). By the marginalization rule we can extract the distribution of microstates,

$$P(\Gamma|S) = \int_0^\infty d\beta P(\Gamma, \beta|S) = \int_0^\infty d\beta P(\Gamma|\beta)P(\beta|S),$$

(22)
First-principles derivation of the q-canonical ensemble in Bayesian superstatistics

which upon replacing equation (6) reads

\[ P(\Gamma|S) = P(\Gamma|I_0) \int_0^\infty d\beta \left[ \frac{\exp(-\beta H(\Gamma))}{Z_0(\beta)} \right] P(\beta|S). \] (23)

Direct comparison with equation (15) gives the ensemble function \( \rho(E; S) \) as

\[ \rho(E; S) = \int_0^\infty d\beta \left[ \frac{P(\beta|S)}{Z_0(\beta)} \right] \exp(-\beta E) \] (24)

which, by defining the superstatistical weight function

\[ f(\beta; S) := \frac{P(\beta|S)}{Z_0(\beta)} \] (25)

simply tells us that \( \rho(E; S) \) and \( f(\beta; S) \) are connected through a Laplace transform,

\[ \rho(E; S) = \int_0^\infty d\beta f(\beta; S) \exp(-\beta E). \] (26)

From this point forward, we will omit the label \( S \) on the functions \( \rho(E) \) and \( f(\beta) \) for simplicity.

The joint distribution of energy and inverse temperature \( P(E, \beta|S) \) can be factorized using the product rule,

\[ P(E, \beta|S) = P(E|\beta)P(\beta|S), \] (27)

and by replacing equation (3) it can be written compactly as

\[ P(E, \beta|S) = \exp(-\beta E)P(E|I_0)f(\beta). \] (28)

Note that this structure of a superstatistical joint distribution of \( E \) and \( \beta \) implies that the quantity

\[ P(E, \beta|S) \exp(\beta E) \]

is always separable as a product of a function of energy and a function of (inverse) temperature.

4. The fundamental inverse temperature

For steady-state ensembles of the form in equation (15) we define the fundamental inverse temperature as the function

\[ \beta_F(E) := -\frac{\partial}{\partial E} \ln \rho(E). \] (29)

This function contains all the information about the ensemble, because integration of equation (29) readily gives

\[ \rho(E) = \rho(0) \exp \left( -\int_0^E dE' \beta_F(E') \right), \] (30)
and thus, more importantly for us, recovering the correct form of $\beta_F(E)$ is equivalent to recovering the correct ensemble function $\rho(E)$. The fundamental inverse temperature function is also usually easier to read and manipulate than the ensemble function, and can even be used in situations beyond superstatistics [16]. As the simplest example, the canonical ensemble at $\beta = \beta_0$ is represented with a constant function $\beta_F(E) = \beta_0$.

In the case of superstatistics, we can show a connection between the moments of $P(\beta|E, S)$ and $\beta_F$. First, we use equations (28) and (17) as

$$P(\beta|E, S) = \frac{P(E, \beta|S)}{P(E|S)} = \frac{\exp(-\beta E)f(\beta)P(E|I_0)}{\rho(E)P(E|I_0)} = \frac{\exp(-\beta E)f(\beta)}{\rho(E)},$$

where we have cancelled the energy prior $P(E|I_0)$. Now we use the identity

$$(-1)^n \frac{\partial^n}{\partial E^n} \left\{ \exp(-\beta E)f(\beta) \right\} = \exp(-\beta E)f(\beta)\beta^n$$

(32)

to write the $n$-th moment of $\beta$ given $E$ as

$$\langle \beta^n \rangle_{E,S} = \int_0^\infty d\beta \left[ \frac{\exp(-\beta E)f(\beta)}{\rho(E)} \right] \beta^n = (-1)^n \frac{\partial^n}{\partial E^n} \frac{\rho(E)}{\rho(E)} \left\{ \int_0^\infty d\beta \exp(-\beta E)f(\beta) \right\}$$

$$= (-1)^n \frac{\partial^n}{\partial E^n} \left\{ \int_0^\infty d\beta \exp(-\beta E)f(\beta) \right\}$$

$$= (-1)^n \frac{\partial^n}{\partial E^n} \rho(E)$$

(33)

Therefore we have that the moments of $\beta$ at a fixed energy $E$ are given by

$$\langle \beta^n \rangle_{E,S} = \frac{(-1)^n}{\rho(E)} \frac{\partial^n}{\partial E^n} \rho(E),$$

(34)

and for $n = 1$ this yields

$$\langle \beta \rangle_{E,S} = -\frac{\partial}{\partial E} \ln \rho(E),$$

that is,

$$\langle \beta \rangle_{E,S} = \beta_F(E).$$

(35)

This equality gives meaning to the fundamental inverse temperature in superstatistics [17], as the conditional mean inverse temperature for a fixed energy $E$. Using $n = 2$ we obtain

$$\langle \beta^2 \rangle_{E,S} = \frac{1}{\rho(E)} \frac{\partial^2}{\partial E^2} \rho(E) = \left( \frac{\partial}{\partial E} \ln \rho(E) \right)^2 + \frac{\partial^2}{\partial E^2} \ln \rho(E) = \beta_F(E)^2 - \beta_F'(E)$$

(36)

from which it follows that

$$\langle (\delta \beta)^2 \rangle_{E,S} = \langle \beta^2 \rangle_{E,S} - \langle \beta \rangle^2_{E,S} = -\beta_F'(E),$$

(37)

and therefore, for every valid superstatistics it must hold that

$$\beta_F'(E) \leq 0.$$  

(38)
5. The $q$-canonical ensemble

We will define the $q$-canonical ensemble by its fundamental inverse temperature function

$$\beta_F(E; \beta_0, q) = \frac{\beta_0}{1 - (1 - q)\beta_0 E},$$  

(39)

where $q$ is traditionally known as the entropic index, and $\beta_0$ an inverse temperature that serves as a scale parameter. Here it is immediate clear that $q = 1$ recovers the canonical ensemble with inverse temperature $\beta_0$. Integration according to equation (30) gives the usual $[2]$ ensemble function

$$\rho(E; \beta_0, q) = \rho(0) \exp\left(\int_0^E \frac{\beta_0 dE'}{1 - (1 - q)\beta_0 E'}\right) = \rho(0) [1 - (1 - q)\beta_0 E]^{\frac{1}{1-q}}$$  

(40)

By defining the $q$-exponential function

$$\exp(x; q) := \left[1 + (1 - q)x\right]_+^{\frac{1}{1-q}},$$  

(41)

with $[x]_+ = \max(x, 0)$ such that $\exp(x; q) \to \exp(x)$ when $q \to 1$, we can write

$$\rho(E; \beta_0, q) = \frac{\exp(-\beta_0 E; q)}{Z_q(\beta_0)}.$$  

(42)

As first noted by Beck and Cohen [3], this ensemble can be obtained in superstatistics from equation (26) with a weight function $f(\beta)$ of the form

$$f(\beta) = \frac{1}{\eta(\beta_0, q)} \exp\left(-\frac{\beta}{\beta_0(q-1)}\right) \beta^\frac{1}{q-1},$$  

(43)

referred to as the $\chi^2$-distribution. We can see this by considering the integral

$$\int_0^\infty d\beta \exp\left(-\frac{\beta}{\beta_0(q-1)}\right) \exp(-\beta E)\beta^{\frac{1}{q-1}} = \Gamma\left(\frac{1}{q-1}\right) \left[\frac{1 + \beta_0 E(q-1)}{\beta_0(q-1)}\right]^{\frac{1}{1-q}} = \Gamma\left(\frac{1}{q-1}\right) (\beta_0(q-1))^{\frac{1}{q-1}} Z_q(\beta_0) \rho(E; q, \beta_0),$$  

(44)

from which it follows that

$$\eta(q, \beta_0) = \Gamma\left(\frac{1}{q-1}\right) \left[\beta_0(q-1)\right]^{\frac{1}{q-1}} Z_q(\beta_0).$$  

(45)

Neglecting a normalization constant, $f(\beta)$ is also a Gamma distribution with shape parameter $k = 1/(q-1)$ and scale parameter $\theta = \beta_0(q-1)$, having a mean equal to $k\theta = \beta_0$ and a variance $k\theta^2 = \beta_0^2(q-1)$. In the limit $q \to 1$ the variance vanishes, and $f(\beta)$ becomes proportional to $\delta(\beta - \beta_0)$, recovering the canonical ensemble.
6. A first-principles derivation of the $q$-canonical ensemble

Now we have all the elements needed to recover the $q$-canonical ensemble in the form of its fundamental inverse temperature (39), by maximizing the Shannon-Jaynes entropy under constraints that impose consistency with the superstatistical framework. More specifically, we will search for the maximum entropy joint distribution $P(E, \beta|S)$ under the constraint that equation (35) holds for the whole range of allowed energies,

$$\langle \beta \rangle_{E,S} = \beta_F(E) \quad \forall \ E \geq 0. \quad (46)$$

Additionally, because of the structure noted in equation (28), we will require that the joint prior distribution of $E$ and $\beta$ is separable, i.e.,

$$P(E, \beta|I_0) = P(E|I_0)P(\beta|I_0). \quad (47)$$

These are “bare-bones” constraints with almost no content. In fact, we have not included any information about a particular system, only asking that the resulting model must be compatible with superstatistics with a fundamental inverse temperature. However, in the following we will show that the least-biased pair of functions $\beta_F(E)$ and $P(\beta|I_0)$ that are consistent with these constraints and the structure of superstatistics are simply

$$\beta_F(E) = \frac{\beta_0}{1 - (1 - q)\beta_0 E}, \quad (48)$$

that is, the $q$-canonical ensemble, and the inverse temperature prior

$$P(\beta|I_0) \propto \beta^{-\frac{1}{q-1}}. \quad (49)$$

Equations (48) and (49) constitute the main results of this work, and we proceed to prove them now. Recalling the Bayesian expectation identity (equation 14 of Ref. [7]),

$$\langle A\delta(B - b) \rangle_I = P(B = b|I)\langle A \rangle_{B=b,I}, \quad (50)$$

valid for any pair of quantities $A$ and $B$, we have

$$\langle \beta\delta(H - E) \rangle_S = \langle \beta \rangle_{E,S} P(E|S) = \beta_F(E)\langle \delta(H - E) \rangle_S. \quad (51)$$

Therefore, the constraint in equation (46) can be expressed as

$$\langle \left[ \beta - \beta_F(E) \right] \delta(H - E) \rangle_S = 0 \quad \forall E \geq 0. \quad (52)$$

Maximizing the Shannon-Jaynes entropy in the $(E, \beta)$ space with the constraint in equation (52) and the prior $P(E, \beta|I_0)$ we obtain

$$P(E, \beta|S) = \frac{1}{\zeta[\lambda]} P(E, \beta|I_0) \exp \left( - \int_0^\infty dE' \lambda(E') \left[ \beta - \beta_F(E') \right] \delta(E - E') \right)$$

$$= \frac{P(E|I_0)P(\beta|I_0)}{\zeta[\lambda]} \exp \left( - \lambda(E) \left[ \beta - \beta_F(E) \right] \right), \quad (53)$$
where $\lambda(E)$ is a Lagrange multiplier function to be determined, precisely the conjugate to the fundamental inverse temperature $\beta_F(E)$ that fixes the constraint in equation (46). Note that, at this point, the functions $\lambda(E)$, $\beta_F(E)$ and $P(\beta|I_0)$ are all unknown, and in order to determine them, we impose consistency between equations (28) and (53).

On the one hand, from equation (28) we obtain that

$$\frac{\partial}{\partial E} \ln P(E, \beta|S) = -\beta + \frac{\partial}{\partial E} \ln P(E|I_0).$$

while equation (53) produces

$$\frac{\partial}{\partial E} \ln P(E, \beta|S) = -\beta \lambda'(E) + \frac{\partial}{\partial E} \left( \lambda(E) \beta_F(E) \right) + \frac{\partial}{\partial E} \ln P(E|I_0).$$

As both must be true for all $\beta$ and $E$, we have that

$$-\beta = -\beta \lambda'(E) + \frac{\partial}{\partial E} \left( \lambda(E) \beta_F(E) \right),$$

and by comparing powers of $\beta$ on both sides it follows that $\lambda'(E) = 1$, therefore

$$\lambda(E) = E + \lambda_0$$

where $\lambda_0$ is an integration constant, and

$$\frac{\partial}{\partial E} \left( \lambda(E) \beta_F(E) \right) = 0$$

from which it follows that

$$\beta_F(E) = \frac{C}{E + \lambda_0}$$

where $C$ is an additional integration constant. By introducing new parameters

$$\beta_0 := \frac{C}{\lambda_0}, \quad q := 1 + \frac{1}{C}$$

we recognize $\beta_F(E)$ in (59) as the fundamental inverse temperature of the $q$-canonical ensemble,

$$\beta_F(E) = \frac{\beta_0}{1 - (1 - q)\beta_0 E}.$$  

Now all that remains is to show that there is a prior $P(\beta|I_0)$ compatible with $\lambda(E)$ and $\beta_F(E)$. Replacing $\lambda(E)$ and $\beta_F(E)$ in equation (53) we obtain

$$P(E, \beta|S) = \left[ \frac{\exp(1/(q - 1))}{\zeta(q, \beta_0)} P(\beta|I_0) \exp \left( -\frac{\beta}{\beta_0 (q - 1)} \right) \right] \exp(-\beta E) P(E|I_0),$$

but agreement with equation (28) requires the quantity in square brackets to be equal to $f(\beta)$, therefore we have

$$f(\beta) = \frac{P(\beta|S)}{Z_0(\beta)} = \frac{\exp(1/(q - 1))}{\zeta(q, \beta_0)} P(\beta|I_0) \exp \left( -\frac{\beta}{\beta_0 (q - 1)} \right).$$

Finally, because our fundamental inverse temperature $\beta_F(E)$ corresponds to the $q$-canonical ensemble, the function $f(\beta)$ in (63) must agree (up to a constant factor) with $f(\beta)$ in equation (43) for the $\chi^2$-distribution. This forces our prior for $\beta$ to be

$$P(\beta|I_0) \propto \beta_{q-1}^{-1},$$

(64)

and at this point we can state our result as the fact that the constraints (46) and (47) lead to

$$f(\beta) = \frac{1}{\eta(\beta_0, q)} \exp \left( - \frac{\beta}{\beta_0(q - 1)} \right) \beta_{q-1}^{-1}$$

as the least-biased model, which is $\chi^2$-superstatistics. We can also interpret the result in (64) as a noninformative prior [18] for $\beta$ that does not refer to any particular system. Because $q \geq 1$, this prior is improper, i.e. not normalizable by itself, but this is a common situation for noninformative priors in Bayesian inference, as is the case for instance for the Jeffreys’ prior [19, 20]. The entropic index $q$ then can be understood as the shape parameter of this noninformative prior.

7. Concluding remarks

A Bayesian view of statistical mechanics that includes superstatistics at its core, and uses the concept of fundamental inverse temperature, is by itself capable of producing the main ensemble of Tsallis’ nonextensive statistics. The entropic index $q$ appears to be the shape parameter of the noninformative prior distribution of inverse temperature, and its origin does not invoke the use of any entropy other than the Boltzmann-Gibbs entropy, or rather, the parameterization-invariant Shannon-Jaynes entropy. These results suggest that the $q$-canonical ensemble is not the posterior distribution of a generalized inference procedure using constraints but, rather, a (noninformative) prior distribution which is built into superstatistics and should be its default. This is consistent with the derivation in Ref. [17] that produces the $q$-canonical ensemble from invariance requirements on the fundamental inverse temperature $\beta_F(E)$, and in both cases the parameters $q$ and $\beta_0$ are not fixed a priori.

Acknowledgements

SD thankfully acknowledges financial support from ANID FONDECYT 1171127 and ANID PIA ACT172101 grants.

References

[1] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. J. Stat. Phys., 52:479–487, 1988.
[2] C. Tsallis. Introduction to nonextensive statistical mechanics: approaching a complex world. Springer Science & Business Media, 2009.
[3] C. Beck and E.G.D. Cohen. Superstatistics. Phys. A, 322:267–275, 2003.
First-principles derivation of the $q$-canonical ensemble in Bayesian superstatistics

[4] C. Beck. Superstatistics: theory and applications. Continuum Mech. Thermodyn., 16:293–304, 2004.

[5] G. Kaniadakis. Non-linear kinetics underlying generalized statistics. Phys. A, 296:405–425, 2001.

[6] Q. A. Wang. Nonextensive statistics and incomplete information. Eur. Phys. J. B, 26:357–368, 2002.

[7] S. Davis. Conditional maximum entropy and superstatistics. J. Phys. A: Math. Theor., 53:445006, 2020.

[8] F. Sattin. Bayesian approach to superstatistics. Eur. Phys. J. B, 49:219–224, 2006.

[9] N. A. Alves and R. B. Frigori. Superstatistics and the quest of generalized ensembles equivalence in a system with long-range interactions. Phys. A, 446:195–203, 2016.

[10] F. Sattin. Superstatistics and temperature fluctuations. Phys. Lett. A, 382:2551–2554, 2018.

[11] S. Davis and G. Gutiérrez. Temperature is not an observable in superstatistics. Phys. A, 505:864–870, 2018.

[12] E. Sánchez C. and P. Vega-Jorquera. New Bayesian frequency-magnitude distribution model for earthquakes applied in Chile. Phys. A, 508:305–312, 2018.

[13] H. Umpierrez and S. Davis. Fluctuation theorems in $q$-canonical ensembles. Phys. A, 563:125337, 2021.

[14] P. D. Dixit. Detecting temperature fluctuations at equilibrium. Phys. Chem. Chem. Phys., 17:13000–13005, 2015.

[15] S. Davis. On the possible distributions of temperature in nonequilibrium steady states. J. Phys. A: Math. Theor., 53:045004, 2020.

[16] S. Davis. Fluctuating temperature outside superstatistics: Thermodynamics of small systems. Phys. A, 589:126665, 2022.

[17] S. Davis and G. Gutiérrez. Emergence of Tsallis statistics as a consequence of invariance. Phys. A, 533:122031, 2019.

[18] J.O. Berger. Statistical Decision Theory and Bayesian Analysis. Springer Series in Statistics. Springer New York, 2013.

[19] E. T. Jaynes. Probability theory: The logic of science. Cambridge university press, 2003.

[20] D. Sivia and J. Skilling. Data analysis: a Bayesian tutorial. OUP Oxford, 2006.