Effects of correlated variability on information entropies in nonextensive systems

Hideo Hasegawa*
Department of Physics, Tokyo Gakugei University, Koganei, Tokyo 184-8501, Japan
(Received 1 April 2008; published 28 August 2008)

We have calculated the Tsallis entropy and Fisher information matrix (entropy) of spatially correlated nonextensive systems, by using an analytic non-Gaussian distribution obtained by the maximum entropy method. The effects of the correlated variability on the Fisher information matrix are shown to be different from those on the Tsallis entropy. The Fisher information is increased (decreased) by a positive (negative) correlation, whereas the Tsallis entropy is decreased with increasing absolute magnitude of the correlation, independently of its sign. This fact arises from the difference in their characteristics. It implies from the Cramér-Rao inequality that the accuracy of an unbiased estimate of fluctuation is improved by a negative correlation. A critical comparison is made between the present study and previous ones employing the Gaussian approximation for the correlated variability due to multiplicative noise.

DOI: 10.1103/PhysRevE.78.021141 PACS number(s): 05.70.−a, 05.10.Gg, 05.45.−a

I. INTRODUCTION

It is well known that the Tsallis entropy and Fisher information entropy (matrix) are very important quantities expressing information measures in nonextensive systems. The Tsallis entropy for an N-unit nonextensive system is defined by [1–3]

$$S_q^{(N)} = \frac{(1 - c_q^{(N)})}{(q-1)}, \tag{1}$$

with

$$c_q^{(N)} = \int [p^{(N)}(\{x_i\})]^q \prod_i dx_i, \tag{2}$$

where q is the entropic index $(0<q<3)$, and $p^{(N)}(\{x_i\})$ denotes the probability distribution of N variables $\{x_i\}$. In the limit of $q \to 1$, the Tsallis entropy reduces to the Boltzmann-Gibbs-Shannon entropy given by

$$S_1^{(N)} = -\int p^{(N)}(\{x_i\}) \ln p^{(N)}(\{x_i\}) \prod_i dx_i. \tag{3}$$

The Boltzmann-Gibbs-Shannon entropy is extensive in the sense that, for a system consisting of N independent but equivalent subsystems, the total entropy is the sum of the constituent subsystems: $S_1^{(N)} = N S_1^{(1)}$. In contrast, the Tsallis entropy is nonextensive: $S_q^{(N)} \neq N S_q^{(1)}$ for $q \neq 1.0$, and $|q-1|$ expresses the degree of nonextensivity of a given system. The Tsallis entropy is the basis of nonextensive statistical mechanics, which has been successfully applied to a wide class of systems including physics, chemistry, mathematics, biology, and others [3].

The Fisher information matrix provides us with an important measure of information [4]. Its inverse expresses the lower bound of decoding errors for an unbiased estimator in the Cramér-Rao inequality. It denotes also the distance between neighboring points in the Riemann space spanned by probability distributions in the information geometry. The Fisher information matrix expresses a local measure of a positive amount of information whereas the Boltzmann-Gibbs-Shannon-Tsallis entropy represents a global measure of ignorance [4]. In recent years, many authors have investigated the Fisher information in nonextensive systems [5–17]. In a previous paper [17], we pointed out that two types of *generalized* and *extended* Fisher information matrices are necessary for nonextensive systems [17]. The generalized Fisher information matrix $g_{ij}^{(N)}$ obtained from the generalized Kullback-Leibler divergence in conformity with the Tsallis entropy is expressed by

$$g_{ij}^{(N)} = qE\left[\left(\frac{\partial \ln p^{(N)}(\{x_i\})}{\partial \theta_i}\right)\left(\frac{\partial \ln p^{(N)}(\{x_i\})}{\partial \theta_j}\right)\right], \tag{4}$$

where $E[\cdots]$ denotes the average over $p^{(N)}(\{x_i\})$ $[=p^{(N)}(\{x_i\};\{\theta_k\})]$ characterized by a set of parameters $\{\theta_k\}$. On the contrary, the extended Fisher information matrix $\tilde{g}_{ij}^{(N)}$ derived from the Cramér-Rao inequality in nonextensive systems is expressed by [17]

$$\tilde{g}_{ij}^{(N)} = E_q\left[\left(\frac{\partial \ln P_q^{(N)}(\{x_i\})}{\partial \theta_i}\right)\left(\frac{\partial \ln P_q^{(N)}(\{x_i\})}{\partial \theta_j}\right)\right], \tag{5}$$

where $E_q[\cdots]$ expresses the average over the escort probability $P_q^{(N)}(\{x_i\})$ given by

$$P_q^{(N)}(\{x_i\}) = \frac{[p^{(N)}(\{x_i\})]^q}{c_q^{(N)}}, \tag{6}$$

$c_q^{(N)}$ being given by Eq. (2). In the limit of $q=1.0$, both the generalized and extended Fisher information matrices reduce to the conventional Fisher information matrix.

Studies of the information entropies have been made mainly for independent (uncorrelated) systems. The effects of correlated noise and inputs on the Fisher information matrix and Shannon's mutual information have been extensively studied in neuronal ensembles (for a recent review, see Ref. [18], and related references therein). It is a fundamental problem in neuroscience to determine whether correlations in neural activity are important for decoding, and what is the impact of correlations on information transmission. When neurons fire independently, the Fisher information increases

*hideohasegawa@goo.jp

proportionally to the population size. In ensembles with limited-range correlations, however, the Fisher information is shown to saturate as a function of population size [19–21]. In recent years the interplay between fluctuations and correlations in nonextensive systems has been investigated [22–24]. It has been demonstrated that, in some globally correlated systems, the Tsallis entropy becomes extensive while the Boltzmann-Gibbs-Shannon entropy is nonextensive [22]. Thus the correlation plays an important role in discussing the properties of information entropies in nonextensive systems.

It is the purpose of the present paper to study the effects of spatially correlated variability on the Tsallis entropy and Fisher information in nonextensive systems. In Sec. II, we will discuss information entropies of correlated nonextensive systems, by using probability distributions derived by the maximum entropy method (MEM). In Sec. III, we discuss the marginal distribution to study the properties of probability distributions obtained by the MEM. Previous related studies are critically discussed also. The final Sec. IV is devoted to our conclusion. In Appendix A, results of the MEM for uncorrelated, nonextensive systems are briefly summarized [6,9,10,17,25].

II. CORRELATED NONEXTENSIVE SYSTEMS

A. The case of $N=2$

We consider correlated N-unit nonextensive systems, for which the probability distribution is derived with the use of the MEM under the constraints given by

$$1 = \int p^{(N)}(\{x_i\})\prod_i dx_i, \tag{7}$$

$$\mu = \frac{1}{N}\sum_i E_q[x_i], \tag{8}$$

$$\sigma^2 = \frac{1}{N}\sum_i E_q[(x_i-\mu)^2], \tag{9}$$

$$s\sigma^2 = \frac{1}{N(N-1)}\sum_i \sum_{j(\neq i)} E_q[(x_i-\mu)(x_j-\mu)], \tag{10}$$

μ, σ^2, and s expressing the mean, variance, and degree of the correlated variability, respectively. Cases with $N=2$ and arbitrary N will be separately discussed in Secs. II A and II B, respectively.

For a given correlated nonextensive system with $N=2$, the MEM with constraints given by Eqs. (7)–(10) yields (details being explained in Appendix B)

$$p^{(2)}(x_1,x_2) = \frac{1}{Z_q^{(2)}} \exp_q\left[-\left(\frac{1}{2}\right)\sum_{i=1}^{2}\sum_{j=1}^{2} A_{ij}(x_i-\mu)(x_j-\mu)\right], \tag{11}$$

with

$$A_{ij} = a\delta_{ij} + b(1-\delta_{ij}), \tag{12}$$

$$a = \frac{1}{\nu_q^{(2)}\sigma^2(1-s^2)}, \tag{13}$$

$$b = -\frac{s}{\nu_q^{(2)}\sigma^2(1-s^2)}, \tag{14}$$

$$Z_q^{(2)} = \begin{cases} \dfrac{2\nu_q^{(2)}\sigma^2 r_q^{(2)}}{(q-1)} B\left(\dfrac{1}{2},\dfrac{1}{q-1}-\dfrac{1}{2}\right) B\left(\dfrac{1}{2},\dfrac{1}{q-1}-1\right) & \text{for } 1<q<3, \quad (15) \\ 2\pi\sigma^2 r_q^{(2)} & \text{for } q=1, \quad (16) \\ \dfrac{2\nu_q^{(2)}\sigma^2 r_q^{(2)}}{(1-q)} B\left(\dfrac{1}{2},\dfrac{1}{1-q}+1\right) B\left(\dfrac{1}{2},\dfrac{1}{1-q}+\dfrac{3}{2}\right) & \text{for } 0<q<1, \quad (17) \end{cases}$$

$$r_q^{(2)} = \sqrt{1-s^2}, \tag{18}$$

$$\nu_q^{(2)} = (2-q), \tag{19}$$

where $B(x,y)$ denotes the Beta function and $\exp_q(x)$ expresses the q-exponential function defined by

$$\exp_q(x) \equiv [1+(1-q)x]^{1/(1-q)}. \tag{20}$$

The matrix A with elements A_{ij} is expressed by the inverse of the covariant matrix Q given by

$$\mathsf{A} = \mathsf{Q}^{-1}, \tag{21}$$

with

$$Q_{ij} = \nu_q^{(2)}\sigma^2[\delta_{ij} + s(1-\delta_{ij})] \quad \text{for } i,j=1,2. \tag{22}$$

In the limit of $q=1.0$, the distribution $p^{(2)}(x_1,x_2)$ reduces to

$$p^{(2)}(x_1,x_2) = \frac{1}{2\pi\sigma^2\sqrt{1-s^2}} \times \exp\left[-\left(\frac{1}{2}\right)\sum_{i=1}^{2}\sum_{j=1}^{2}(x_i-\mu)(\mathsf{Q}^{-1})_{ij}(x_j-\mu)\right], \tag{23}$$

which is nothing but the Gaussian distribution for $N=2$.

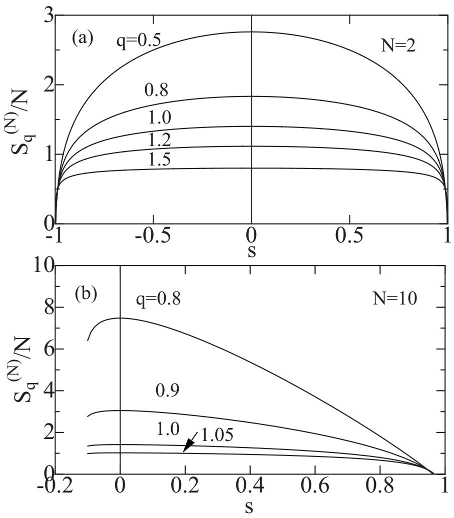

FIG. 1. s dependence of the Tsallis entropy per element, $S_q^{(N)}/N$, with (a) $N=2$ and (b) 10 for various q values with $\mu=0.0$ and $\sigma^2=1.0$; the s value is allowed to be $-1.0<s<1.0$ for $N=2$, and $-0.11<s<1.0$ for $N=10$ [Eq. (43)].

We have calculated information entropies by using the distribution given by Eq. (11).

1. Tsallis entropy

We obtain

$$S_q^{(2)} = \begin{cases} [1+\ln(2\pi\sigma^2)] + \ln(r_q^{(2)}) & \text{for } q=1, \quad (24) \\ \dfrac{1-c_q^{(2)}}{q-1} & \text{for } q \neq 1, \quad (25) \end{cases}$$

with

$$c_q^{(2)} = \nu_q^{(2)} (Z_q^{(2)})^{1-q}, \tag{26}$$

where $Z_q^{(2)}$ is given by Eqs. (15)–(17). From $r_q^{(2)}$ given by Eqs. (18), we may obtain the s dependence of $c_q^{(2)}$ as given by

$$c_q^{(2)}(s) = c_q^{(2)}(0)(1-s^2)^{(1-q)/2} \tag{27}$$

$$\simeq c_q^{(2)}(0)\left(1 + \frac{(q-1)}{2}s^2\right) \quad \text{for } |s| \ll 1, \tag{28}$$

which yields

$$S_q^{(2)}(s) \simeq S_q^{(2)}(0) - \frac{c_q^{(2)}(0)}{2}s^2 \quad \text{for } |s| \ll 1. \tag{29}$$

Figure 1(a) shows $S_q^{(N)}/N$ as a function of the correlation s for $N=2$ (values of $\mu=0.0$ and $\sigma^2=1.0$ are hereafter adopted in the model calculations shown in Figs. 1–7). We note that the Tsallis entropy is decreased with increasing absolute value of s, independently of its sign.

2. Fisher information

By using Eqs. (4) and (5) for $\theta_i=\theta_j=\mu$, we obtain the Fisher information matrices given by

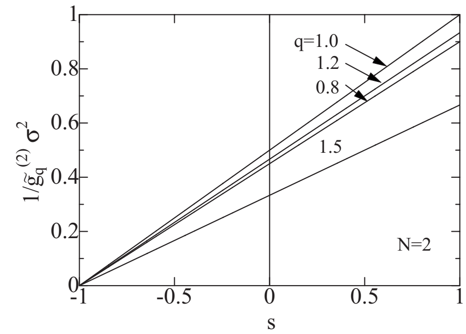

FIG. 2. s dependence of the inverse of the extended Fisher information $\tilde{g}_q^{(2)}$ for various q values with $N=2$ ($\mu=0.0$, $\sigma^2=1.0$).

$$g_q^{(2)} = \frac{2}{\sigma^2(1+s)}, \tag{30}$$

$$\tilde{g}_q^{(2)} = \frac{2q}{(2q-1)(2-q)\sigma^2(1+s)}, \tag{31}$$

which show that $g_q^{(2)}$ is independent of q and that the inverses of both matrices are proportional to $\sigma^2(1+s)$. Figure 2 shows the s dependence of the extended Fisher information for $N=2$, whose inverse is increased (decreased) for a positive (negative) s, depending on the sign of s, in contrast to S_q.

B. The case of arbitrary N

It is possible to extend our approach to the case of arbitrary N, for which the MEM with the constraints given by Eqs. (7)–(10) leads to the distribution given by (details being given in Appendix B)

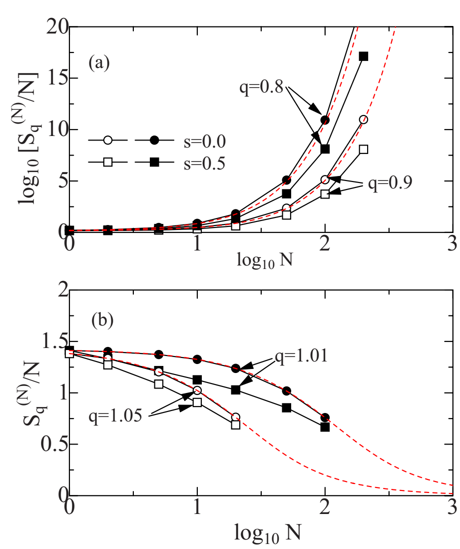

FIG. 3. (Color online) (a) N dependence of the Tsallis entropy per element, $S_q^{(N)}/N$, for $q \leqslant 1.0$: $(q,s)=(0.8,0.0)$ (filled circles), (0.8,0.5) (filled squares), (0.9,0.0) (open circles), and (0.9,0.5) (open squares). (b) $S_q^{(N)}/N$ for $q \geqslant 1.0$: $(q,s)=(1.01,0.0)$ (filled circles), (1.01,0.5) (filled squares), (1.05,0.0) (open circles), and (1.05,0.5) (open squares). Dashed curves denote exact results given by Eq. (A13) [Figs. 7(a) and 7(b)] in Appendix A. Note the logarithmic and linear vertical scales in (a) and (b), respectively.

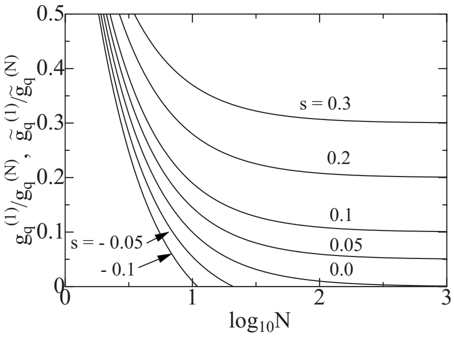

FIG. 4. N dependences of inverses of the Fisher information matrices, $g_q^{(1)}/g_q^{(N)}$ and $\tilde{g}_q^{(1)}/\tilde{g}_q^{(N)}$, for various s values given by Eq. (51); results for $s=-0.05$ and $s=-0.1$ are valid for $N \leqslant 21$ and $N \leqslant 11$, respectively.

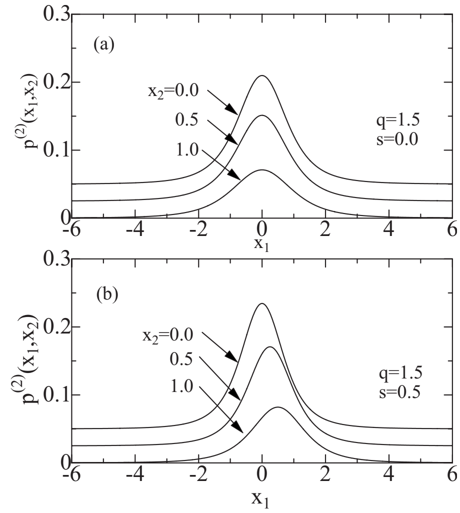

FIG. 5. Probability distribution of $N=2$ systems, $p^{(2)}(x_1,x_2)$ [Eq. (11)], for (a) $s=0.0$ and (b) 0.5 with $q=1.5$ as a function of x_1 for $x_2=0.0$, 0.5, and 1.0.

$$p^{(N)}(\{x_i\}) = \frac{1}{Z_q^{(N)}} \exp_q\left[-\left(\frac{1}{2}\right)\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}(x_i-\mu)(x_j-\mu)\right], \tag{32}$$

with

$$A_{ij} = a\delta_{ij} + b(1-\delta_{ij}), \tag{33}$$

$$a = \frac{[1+(N-2)s]}{\nu_q^{(N)}\sigma^2(1-s)[1+(N-1)s]}, \tag{34}$$

$$b = -\frac{s}{\nu_q^{(N)}\sigma^2(1-s)[1+(N-1)s]}, \tag{35}$$

$$Z_q^{(N)} = \begin{cases} \dfrac{(2\nu_q^{(N)}\sigma^2)^{N/2} r_q^{(N)}}{(q-1)^{N/2}} \displaystyle\prod_{i=1}^{N} B\left(\frac{1}{2}, \frac{1}{q-1} - \frac{i}{2}\right) & \text{for } 1 < q < 3, \quad (36) \\ (2\pi\sigma^2)^{N/2} r_q^{(N)} & \text{for } q = 1, \quad (37) \\ \dfrac{(2\nu_q^{(N)}\sigma^2)^{N/2} r_q^{(N)}}{(1-q)^{N/2}} \displaystyle\prod_{i=1}^{N} B\left(\frac{1}{2}, \frac{1}{1-q} + \frac{(i+1)}{2}\right) & \text{for } 0 < q < 1, \quad (38) \end{cases}$$

$$r_q^{(N)} = \{(1-s)^{N-1}[1+(N-1)s]\}^{1/2}, \tag{39}$$

$$\nu_q^{(N)} = \frac{[(N+2)-Nq]}{2}. \tag{40}$$

The matrix A is expressed by the inverse of the covariant matrix Q whose elements are given by

$$Q_{ij} = \nu_q^{(N)}\sigma^2[\delta_{ij} + s(1-\delta_{ij})]. \tag{41}$$

In the limit of $q=1.0$, the distribution given by Eq. (32) becomes the multivariate Gaussian distribution given by

$$p(\{x_i\}) \propto \exp\left[-\left(\frac{1}{2}\right)\sum_{ij}(x_i-\mu)(\mathsf{Q}^{-1})_{ij}(x_j-\mu)\right]. \tag{42}$$

It is necessary to note that there is a condition for a physically conceivable s value given by [see Eq. (C8), details being discussed in Appendix C]

$$s_L \leqslant s \leqslant s_U, \tag{43}$$

where the lower and upper critical s values are given by $s_L = -1/(N-1)$ and $s_U=1.0$, respectively. In the cases of $N=2$ and 10, for example, we obtain $s_L=-1.0$ and $s_L=-0.11$, respectively.

By using the probability distribution given by Eq. (32), we have calculated information entropies whose s dependences are given as follows.

1. Tsallis entropy

We obtain

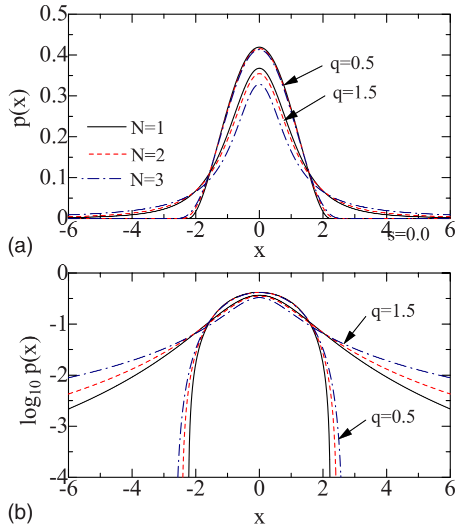

FIG. 6. (Color online) Uncorrelated distribution for $N=1$, $p^{(1)}(x_1)$ [Eq. (62), solid curves], and marginal distributions of $p_m^{(2)}(x_1)$ for $N=2$ [Eq. (61), dashed curves], and $p_m^{(3)}(x_1)$ for $N=3$ [Eq. (63), chain curves] with $q=0.5$ and $q=1.5$ in (a) linear and (b) logarithmic vertical scales.

$$S_q^{(N)} = \begin{cases} \frac{N}{2}[1+\ln(2\pi\sigma^2)] + \ln(r_q^{(N)}) & \text{for } q=1, \quad (44) \\ \frac{1-c_q^{(N)}}{q-1} & \text{for } q \neq 1, \quad (45) \end{cases}$$

with

$$c_q^{(N)} = \nu_q^{(N)}(Z_q^{(N)})^{1-q}, \tag{46}$$

$$\simeq c_q^{(N)}(0)\left(1 + \frac{(q-1)N(N-1)}{4}s^2\right)$$

$$\text{for } |s| \ll 2/\sqrt{N(N-1)}, \tag{47}$$

where the s dependence of $c_q^{(N)}$ arises from the factor of $r_q^{(N)}$ in Eq. (39), and $c_q^{(N)}(0)$ expresses the $s=0.0$ value of $c_q^{(N)}$. Equation (46) yields the s-dependent $S_q^{(N)}$ given by

$$S_q^{(N)}(s) \simeq S_q^{(N)}(0) - \left(\frac{N(N-1)c_q^{(N)}(0)}{4}\right)s^2$$

$$\text{for } |s| \ll 2/\sqrt{N(N-1)}, \tag{48}$$

where $S_q^{(N)}(0)$ stands for the Tsallis entropy for $s=0.0$. The region where Eqs. (47) and (48) hold becomes narrower for larger N.

The s dependence of $S_q^{(N)}/N$ for $N=10$ is shown in Fig. 1(b), where $S_q^{(N)}/N$ has a peak at $s=0.0$ and it is decreased with increasing $|s|$. Comparing Fig. 1(b) with Fig. 1(a), we notice that the s dependence of $S_q^{(N)}/N$ for $N=10$ is more significant than that for $N=2$ [Eq. (48)].

The circles in Figs. 3(a) and 3(b) show $S_q^{(N)}/N$ with $s=0.0$ for $q<1.0$ and $q>1.0$, respectively, calculated with the use of the expressions given by Eqs. (45) and (46). They are in good agreement with the dashed curves showing the exact results which are given by Eq. (A13) and shown in Figs. 7(a) and 7(b) in Appendix A. The squares show $S_q^{(N)}/N$ with $s=0.5$ calculated by using Eqs. (45) and (46). The Tsallis entropy is decreased by an introduced correlation. Because of a computational difficulty [26], calculations using Eqs. (45) and (46) cannot be performed for larger N than those shown in Figs. 3(a) and 3(b).

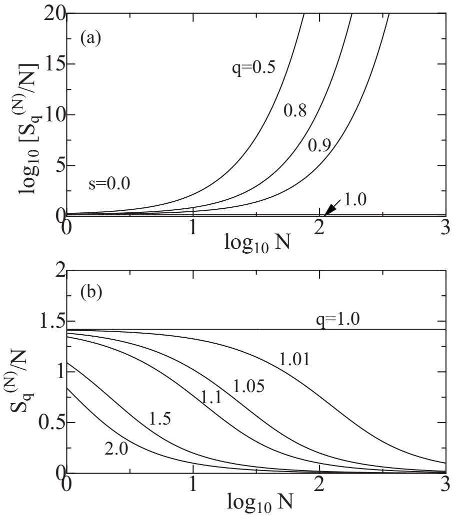

FIG. 7. N dependence of the Tsallis entropy per element, $S_q^{(N)}/N$, in uncorrelated systems for (a) $q\leqslant 1.0$ and (b) $q\geqslant 1.0$; note the logarithmic and linear vertical scales in (a) and (b), respectively.

2. Fisher information

The generalized and extended Fisher information matrices are given by

$$g_q^{(N)} = \frac{N}{\sigma^2[1+(N-1)s]}, \tag{49}$$

$$\tilde{g}_q^{(N)} = \frac{Nq(q+1)}{\sigma^2(3-q)(2q-1)[1+(N-1)s]}. \tag{50}$$

The results for $q=1.0$ given by Eqs. (49) and (50) are consistent with those derived with the use of the multivariate Gaussian distribution [19]. By using the Fisher information matrices for $N=1$, $g_q^{(1)}$ and $\tilde{g}_q^{(1)}$, given by Eqs. (A17) and (A18), we obtain

$$\frac{g_q^{(1)}}{g_q^{(N)}} = \frac{\tilde{g}_q^{(1)}}{\tilde{g}_q^{(N)}} = \frac{1}{N} + \left(1 - \frac{1}{N}\right)s, \tag{51}$$

$$= \begin{cases} s & \text{for } N\to\infty, \quad (52) \\ \frac{1}{N} & \text{for } s=0, \quad (53) \\ 1 & \text{for } s=s_U, \quad (54) \\ 0 & \text{for } s=s_L, \quad (55) \end{cases}$$

which holds independently of q. The inverses of the Fisher information matrices approach the value of s for $N\to\infty$, and are proportional to $1/N$ for $s=0.0$. In particular, they vanish at $s=s_L$. These features are clearly seen in Fig. 4, where the inverses of the Fisher information matrices, $g_q^{(1)}/g_q^{(N)}$ and $\tilde{g}_q^{(1)}/\tilde{g}_q^{(N)}$, are plotted as functions of N for various s values.

III. DISCUSSION

A. Marginal distributions

In the present study, we have obtained the probability distributions, by applying the MEM to spatially correlated nonextensive systems. We will examine our probability distributions in more detail. The x_1 dependences of $p^{(2)}(x_1,x_2)$ for $N=2$ given by Eq. (11) with $s=0.0$ and 0.5 are plotted in Figs. 5(a) and 5(b), respectively, where x_2 is treated as a parameter. When $s=0.0$, the distribution is symmetric with respect to x_1 for all x_2 values. When the correlated variability of $s=0.5$ is introduced, peak positions of the distribution appear at finite x_1 for $x_2=0.5$ and 1.0.

In the limit of $s=0.0$ (i.e., no correlated variability), $p^{(2)}(x_1,x_2)$ given by Eq. (11) becomes

$$p^{(2)}(x_1,x_2) \propto \left(1 - \frac{(1-q)(x_1^2+x_2^2)}{2\nu_q^{(2)}\sigma^2}\right)^{1/(1-q)}, \quad (56)$$

which does not agree with the exact result (except for $q=1.0$), as given by

$$p^{(1)}(x_1)p^{(1)}(x_2) \propto \left(1 - \frac{(1-q)(x_1^2+x_2^2)}{2\nu_q^{(1)}\sigma^2} + \frac{(1-q)^2}{4(\nu_q^{(1)})^2\sigma^4}x_1^2x_2^2\right)^{1/(1-q)} \quad (57)$$

$$\neq p^{(2)}(x_1,x_2), \quad (58)$$

because of the properties of the q-exponential function defined by Eq. (20): $\exp_q(x+y) \neq \exp_q(x)\exp_q(y)$. By using the q-product $\otimes_q$ defined by [27]

$$x\otimes_q y \equiv (x^{1-q}+y^{1-q}-1)^{1/(1-q)}, \quad (59)$$

we may obtain the expression given by

$$p^{(1)}(x_1)\otimes_q p^{(1)}(x_2) \propto \left(1 - \frac{(1-q)(x_1^2+x_2^2)}{2\nu_q^{(1)}\sigma^2}\right)^{1/(1-q)}, \quad (60)$$

which coincides with $p^{(2)}(x_1,x_2)$ given by Eq. (56) apart from the difference between $\nu_q^{(1)}$ and $\nu_q^{(2)}$. In deriving Eq.(60), however, we have not included normalization factors of $p^{(1)}(x_1)$ and $p^{(1)}(x_2)$.

In order to study the properties of the probability distribution of $p^{(2)}(x_1,x_2)$ in more detail, we have calculated its marginal probability (with $s=0.0$) given by

$$p_m^{(2)}(x_1) = \int p^{(2)}(x_1,x_2)dx_2 \propto \left(1 - \frac{(1-q)x_1^2}{2\nu_q^{(2)}\sigma^2}\right)^{1/(1-q)+1/2}. \quad (61)$$

The dashed curves in Figs. 6(a) and 6(b) show $p_m^{(2)}(x_1)$ in linear and logarithmic scales, respectively. The marginal distributions are in good agreement with the solid curves showing $p^{(1)}(x_1)$ [Eq. (A5)],

$$p^{(1)}(x_1) \propto \left(1 - \frac{(1-q)x_1^2}{2\nu_q^{(1)}\sigma^2}\right)^{1/(1-q)}. \quad (62)$$

In the case of $N=3$, the distribution given by Eq. (32) yields its marginal distribution (with $s=0.0$) given by

$$p_m^{(3)}(x_1) = \int\int p^{(3)}(x_1,x_2,x_3)dx_2dx_3 \propto \left(1 - \frac{(1-q)x_1^2}{2\nu_q^{(3)}\sigma^2}\right)^{1/(1-q)+1}. \quad (63)$$

The chain curves in Figs. 6(a) and 6(b) represent $p_m^{(3)}(x_1)$, which is again in good agreement with the solid curves showing $p^{(1)}(x_1)$. These results justify, to some extent, the probability distribution adopted in our calculation.

The marginal distribution for an arbitrary N (with $s=0.0$) is given by

$$p_m^{(N)}(x_1) = \int\int p^{(N)}(x_1,\ldots,x_N)dx_2\cdots dx_N \quad (64)$$

$$\propto \left(1 - \frac{(1-q)x_1^2}{2\nu_q^{(N)}\sigma^2}\right)^{1/(1-q)+(N-1)/2}, \quad (65)$$

$$\propto \left(1 - \frac{(1-q_N)x_1^2}{2\nu_N\sigma^2}\right)^{1/(1-q_N)}, \quad (66)$$

with

$$q_N = \frac{(N-1)-(N-3)q}{(N+1)-(N-1)q}, \quad (67)$$

$$\nu_N = \frac{(N+2)-Nq}{(N+1)-(N-1)q}. \quad (68)$$

Equations (66)–(68) show that in the limit of $N\to\infty$ we obtain $q_N=1.0$ and $\nu_N=1.0$, and $p_m^{(N)}(x_1)$ reduces to the Gaussian distribution.

B. Comparison with related studies

One typical microscopic nonextensive system is the Langevin model subjected to multiplicative noise, as given by [28–30]

$$\frac{dx_i}{dt} = -\lambda x_i + \beta\xi_i(t) + \alpha x_i\eta_i(t) + H(I) \quad (i=1-N). \quad (69)$$

Here λ expresses the relaxation rate, $H(I)$ denotes a function of an external input I, and α and β stand for the magnitudes of multiplicative and additive noise, respectively, with zero-mean white noise given by $\eta_i(t)$ and $\xi_i(t)$ with the correlated variability

$$\langle\eta_i(t)\eta_j(t')\rangle = \alpha^2[\delta_{ij}+c_M(1-\delta_{ij})]\delta(t-t'), \quad (70)$$

$$\langle\xi_i(t)\xi_j(t')\rangle = \beta^2[\delta_{ij}+c_A(1-\delta_{ij})]\delta(t-t'), \quad (71)$$

$$\langle \eta_i(t)\xi_j(t')\rangle = 0, \quad (72)$$

where c_A and c_M express the degrees of correlated variabilities of the additive and multiplicative noise, respectively. The Fokker-Planck equation (FPE) for the probability distribution $p(\{x_k\},t)$ $(=p)$ is given by

$$\frac{\partial}{\partial t}p = -\sum_i \frac{\partial}{\partial x_i}[(-\lambda x_i + H)p] + \frac{\beta^2}{2}\sum_i\sum_j [\delta_{ij} + c_A(1-\delta_{ij})]\frac{\partial^2}{\partial x_i \partial x_j}p + \frac{\alpha^2}{2}\sum_i\sum_j [\delta_{ij} + c_M(1-\delta_{ij})]\frac{\partial}{\partial x_i}x_i\frac{\partial}{\partial x_j}(x_j p) \quad (73)$$

in the Stratonovich representation.

For additive noise only ($\alpha=0$), the stationary distribution is given by

$$p(\{x_i\}) \propto \exp\left(-\frac{1}{2}\sum_{ij}(x_i-\mu_i)(\mathsf{Q}^{-1})_{ij}(x_j-\mu_j)\right), \quad (74)$$

where $\mu_i = H/\lambda$ and Q expresses the covariance matrix given by

$$Q_{ij} = (\beta^2/2\lambda)[\delta_{ij} + c_A(1-\delta_{ij})]. \quad (75)$$

When multiplicative noise exists ($\alpha \neq 0.0$), the calculation of even stationary distributions becomes difficult, and it is generally not given by the Gaussian. Indeed, the stationary distribution for noncorrelated multiplicative noise with $\alpha \neq 0.0$, $\beta \neq 0.0$, and $c_A = c_M = 0.0$ is given by [17,28–30]

$$p(\{x_i\}) \propto \prod_i \left(1-(1-q)\left(\frac{x_i^2}{2\phi^2}\right)\right)^{1/(1-q)} e^{Y(x_i)}, \quad (76)$$

with

$$q = 1 + \frac{2\alpha^2}{2\lambda + \alpha^2}, \quad (77)$$

$$\phi^2 = \frac{\beta^2}{2\lambda + \alpha^2}, \quad (78)$$

$$Y(x_i) = \left(\frac{2H}{\alpha\beta}\right)\tan^{-1}\left(\frac{\alpha x_i}{\beta}\right). \quad (79)$$

The probability distribution given by Eq. (76) for $H=0$ ($c_A = c_M = 0$) agrees with that derived by the MEM for $\phi^2 = \nu_q^{(1)}\sigma^2$ [Eq. (A5)]. For $\alpha \neq 0.0$, $\beta = 0.0$, and $H>0$ ($c_A = c_M = 0$), Eq. (76) becomes [17]

$$p(x) \propto |x|^{-2/(q-1)} e^{-2H/\alpha^2 x}\Theta(x), \quad (80)$$

yielding the Fisher information given by

$$g_q^{(N)} = \frac{Nq^4}{\sigma_q^2} = \frac{2N\lambda q^4}{\alpha^2\mu^2}, \quad (81)$$

where $\sigma_q^2 = \alpha^2\mu^2/2\lambda$ and $\Theta(x)$ is the Heaviside function.

The probability distribution for correlated multiplicative noise ($\alpha \neq 0.0$, $c_M \neq 0.0$) is also a non-Gaussian, which is easily confirmed by direct simulations of the Langevin model with $N=2$ [31]. In some previous studies [19–21], the stationary distribution of the Langevin model subjected to correlated multiplicative noise with $c_M \neq 0.0$, $\beta = 0.0$, and $H = \mu$ is assumed to be expressed by the Gaussian distribution with the covariance matrix given by

$$Q_{ij} = \alpha^2\mu^2[\delta_{ij} + c_M(1-\delta_{ij})]. \quad (82)$$

This is equivalent to assuming that

$$\frac{\partial}{\partial x_i}\left(x_i \frac{\partial}{\partial x_j}(x_j p)\right) \simeq \frac{\partial}{\partial x_i}\left(\langle x_i\rangle \frac{\partial}{\partial x_j}(\langle x_j\rangle p)\right) = \mu^2 \frac{\partial^2 p}{\partial x_i \partial x_j} \quad (83)$$

in the FPE given by Eq. (73). By using such an approximation, Abbott and Dayan (AD) [19] calculated the Fisher information matrix of a neuronal ensemble with correlated variability, which is given by

$$g_{\rm AD}^{(N)} = \frac{NK}{\alpha^2[1+(N-1)c_M]} + 2NK = \frac{N}{\alpha^2\mu^2[1+(N-1)c_M]} + \frac{2N}{\mu^2}, \quad (84)$$

with a spurious second term ($2NK$), where $K = N^{-1}\Sigma_i[d \ln H_i(\mu)/d\mu]^2 = 1/\mu^2$ [Eq. (4.7) of Ref. [19] in our notation]. Equation (84) is not in agreement with either Eq. (49) or Eq. (50) derived by the MEM. Furthermore, the result of AD in the limit of $c_M = 0$, $g_{\rm AD}^{(N)} = (N/\alpha^2\mu^2 + 2N/\mu^2)$, does not agree with the exact result given by Eq. (81) for the Langevin model. This fact casts some doubt on the results of Refs. [19–21] based on the Gaussian approximation given by Eq. (82) or (83), which has no physical or mathematical justification. The Fisher information matrix depends on the detailed structure of the probability distribution because it is expressed by the derivative of the distribution with respect to x, as given by

$$g_q^{(N)} = -NqE\left(\frac{\partial^2 \ln p^{(1)}(x)}{\partial x^2}\right). \quad (85)$$

We must take into account the non-Gaussian structure of the probability distribution in discussing the Fisher information of nonextensive systems.

IV. CONCLUSION

We have discussed the effects of spatially correlated variability on the Tsallis entropy and Fisher information matrix in nonextensive systems, by using the probability distribution derived by the MEM. Although the analytical distribution obtained in the limit of $s=0.0$ does not hold the relation given by $p^{(N)}(\{x_k\}) = \Pi_{i=1}^N p^{(1)}(x_i)$ for $q \neq 1.0$, it numerically yields good results (Fig. 3), reducing to the multivariate Gaussian distribution in the limit of $q=1.0$. Our calculations have shown that (i) the Tsallis entropy is decreased by both positive and negative correlations, and (ii) the inverses of the Fisher information matrices are increased (decreased) by a positive (negative) correlation.

The difference between the s dependences of S_q and g_q arises from the difference in their characteristics: S_q provides us with a global measure of ignorance and g_q a local measure of a positive amount of information [4]. The item (ii) implies that the accuracy of an unbiased estimate of fluctuation is improved by negative correlation. If there is known, and strong, negative correlation between successive pairs of data points, estimating the unknown parameter as their simple average must reduce the error, as negatively correlated errors tend to cancel in taking the difference.

In connection with the discussion presented in Sec. III, it is interesting to make a detailed study of the properties of information entropies in the Langevin model subjected to correlated as well as uncorrelated multiplicative noise. Such a calculation is in progress and will be reported elsewhere [31].

ACKNOWLEDGMENTS

This work is partly supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

APPENDIX A: MEM FOR NONCORRELATED NONEXTENSIVE SYSTEMS

We summarize the results of the MEM for nonextensive systems [6,9,10,17,25]. In order to apply the MEM to N-unit nonextensive systems, we impose three constraints given by

$$1 = \int p^{(N)}(\{x_i\}) \prod_i dx_i, \tag{A1}$$

$$\mu = \frac{1}{N}\sum_i E_q[x_i], \tag{A2}$$

$$\sigma^2 = \frac{1}{N}\sum_i E_q[(x_i-\mu)^2], \tag{A3}$$

where the q-dependent μ and σ^2 correspond to μ_q and σ_q^2, respectively, in [17]. For a given nonextensive system, the variational condition for the Tsallis entropy given by Eq. (1) with the constraints (A1)–(A3) yields the q-Gaussian distribution given by [6,9,10,17,25]

$$p^{(N)}(\{x_k\}) = \prod_{i=1}^{N} p^{(1)}(x_i), \tag{A4}$$

with

$$p^{(1)}(x) = \frac{1}{Z_q^{(1)}} \exp_q\left[-\left(\frac{(x-\mu)^2}{2\nu_q^{(1)}\sigma^2}\right)\right], \tag{A5}$$

$$Z_q^{(1)} = \int \exp_q\left(-\frac{(x-\mu)^2}{2\nu_q^{(1)}\sigma^2}\right)dx \tag{A6}$$

$$= \begin{cases} \left(\dfrac{2\nu_q^{(1)}\sigma^2}{q-1}\right)^{1/2} B\left(\dfrac{1}{2},\dfrac{1}{q-1}-\dfrac{1}{2}\right) & \text{for } 1<q<3, \quad \text{(A7)} \\ \sqrt{2\pi}\sigma & \text{for } q=1.0, \quad \text{(A8)} \\ \left(\dfrac{2\nu_q^{(1)}\sigma^2}{1-q}\right)^{1/2} B\left(\dfrac{1}{2},\dfrac{1}{1-q}+1\right) & \text{for } 0<q<1, \quad \text{(A9)} \end{cases}$$

$$\nu_q^{(1)} = \frac{3-q}{2}, \tag{A10}$$

where $B(x,y)$ denotes the Beta function.

1. Tsallis entropy

Substituting the probability distribution given by Eq. (A4) into Eqs. (1) and (2), we obtain the Tsallis entropy expressed by

$$S_q^{(N)} = \begin{cases} \dfrac{N}{2}[1+\ln(2\pi\sigma^2)] & \text{for } q=1, \quad \text{(A11)} \\ \dfrac{[1-(c_q^{(1)})^N]}{(q-1)} & \text{for } q\neq 1, \quad \text{(A12)} \end{cases}$$

where $c_q^{(1)} = \nu_q^{(1)}(Z_q^{(1)})^{1-q}$. We may express $S_q^{(N)}$ in terms of $S_q^{(1)}$ by [17]

$$S_q^{(N)} = \sum_{k=1}^{N} C_k^N (-1)^{k-1}(q-1)^{k-1}(S_q^{(1)})^k \tag{A13}$$

$$= NS_q^{(1)} - \frac{N(N-1)(q-1)}{2}(S_q^{(1)})^2 + \cdots, \tag{A14}$$

where $C_k^N = N!/(N-k)!k!$. Equation (A14) clearly shows that the Tsallis entropy is generally nonextensive except for $q=1.0$ for which $S_q^{(N)} = NS_q^{(1)}$. Figures 7(a) and 7(b) show the N dependence of the Tsallis entropy per element, $S_q^{(N)}/N$, of uncorrelated systems ($s=0.0$), which are calculated with the use of Eq. (A13). With increasing N, $S_q^{(N)}/N$ is decreased for $q>1.0$, whereas it is significantly increased for $q<1.0$.

2. The Fisher information

With the use of Eqs. (4) and (5) for $\theta_i=\theta_j=\mu$, the generalized and extended Fisher information matrices are given by [17]

$$g_q^{(N)} = N g_q^{(1)}, \tag{A15}$$

$$\tilde{g}_q^{(N)} = N \tilde{g}_q^{(1)}, \tag{A16}$$

with

$$g_q^{(1)} = \frac{1}{\sigma^2}, \tag{A17}$$

$$\tilde{g}_q^{(1)} = \frac{q(q+1)}{(3-q)(2q-1)\sigma^2}, \tag{A18}$$

which show that Fisher information matrices are extensive.

APPENDIX B: MEM FOR CORRELATED NONEXTENSIVE SYSTES

In the case of $N=2$, the probability distribution $p^{(2)}(x_1,x_2)$ given by Eqs. (11) and (12) is rewritten as

$$p^{(2)}(x_1,x_2) \propto [1-(1-q)\Phi^{(2)}(x_1,x_2)]^{1/(1-q)}, \tag{B1}$$

with

$$\Phi^{(2)}(x_1,x_2) = \frac{1}{2}[a(x_1^2+x_2^2)+2bx_1x_2] \tag{B2}$$

$$=\lambda_1 y_1^2+\lambda_2 y_2^2, \tag{B3}$$

where λ_i and y_i $(i=1,2)$ are eigenvalues and eigenvectors of $\Phi(x_1,x_2)$ given by

$$\lambda_1 = \frac{1}{2}(a+b), \tag{B4}$$

$$\lambda_2 = \frac{1}{2}(a-b), \tag{B5}$$

$$y_1 = \frac{1}{\sqrt{2}}(x_1+x_2), \tag{B6}$$

$$y_2 = \frac{1}{\sqrt{2}}(x_1-x_2). \tag{B7}$$

The averages of $E_q[y_1^2]$ and $E_q[y_1^2]$ are given by

$$E_q[y_1^2] = \frac{1}{\nu_q^{(2)}\lambda_1}, \tag{B8}$$

$$E_q[y_2^2] = \frac{1}{\nu_q^{(2)}\lambda_2}, \tag{B9}$$

from which we obtain σ^2 and $s\ \sigma^2$ as

$$\sigma^2 = \frac{1}{2}E_q[x_1^2+x_2^2] = \frac{1}{2}E_q[y_1^2+y_2^2] = \frac{a}{\nu_q^{(2)}(a^2-b^2)}, \tag{B10}$$

$$s\sigma^2 = E_q[x_1x_2] = \frac{1}{2}E_q[y_1^2-y_2^2] = -\frac{b}{\nu_q^{(2)}(a^2-b^2)}. \tag{B11}$$

By using Eqs. (B10) and (B11), a and b are expressed in terms of σ^2 and s as

$$a = \frac{1}{\nu_q^{(2)}\sigma^2(1-s^2)}, \tag{B12}$$

$$b = -\frac{s}{\nu_q^{(2)}\sigma^2(1-s^2)}, \tag{B13}$$

which yield the matrix of A given by the inverse of the covariance matrix of Q [Eq. (22)].

A calculation for the case of arbitrary N may be similarly performed as follows. The distribution given by Eqs. (32) and (33) is rewritten as

$$p^{(N)}(\{x_i\}) \propto [1-(1-q)\Phi^{(N)}(\{x_i\})]^{1/(1-q)}, \tag{B14}$$

with

$$\Phi^{(N)}(\{x_i\}) = \frac{1}{2}\left(a\sum_i x_i^2 + 2b\sum_{i<j} x_ix_j\right) \tag{B15}$$

$$=\sum_i \lambda_i y_i^2, \tag{B16}$$

where λ_i and y_i are eigenvalues and eigenvectors, respectively. With the use of eigenvalues given by

$$\lambda_i = \begin{cases} \frac{1}{2}[a+(N-1)b] & \text{for } i=1, \quad \text{(B17)} \\ \frac{1}{2}(a-b) & \text{for } 1<i\le N, \quad \text{(B18)} \end{cases}$$

Eqs. (9) and (10) lead to

$$\sigma^2 = \frac{1}{N}\sum_i E_q[y_i^2] = \frac{1}{\nu_q^{(N)}N}\sum_i\left(\frac{1}{\lambda_i}\right) = \frac{[a+b(N-2)]}{\nu_q^{(N)}(a-b)[a+(N-1)b]}, \tag{B19}$$

$$s\sigma^2 = \frac{1}{N(N-1)}\sum_{i<j}E_q[y_i^2-y_j^2] = \left(\frac{1}{\nu_q^{(N)}N(N-1)}\right)\sum_{i<j}\left(\frac{1}{\lambda_i}-\frac{1}{\lambda_j}\right) = -\frac{b}{\nu_q^{(N)}(a-b)[a+(N-1)b]}. \tag{B20}$$

From Eqs. (B19) and (B20), a and b are expressed in terms of σ^2 and s, as given by

$$a = \frac{[1+(N-2)s]}{\nu_q^{(N)}\sigma^2(1-s)[1+(N-1)s]}, \tag{B21}$$

$$b = -\frac{s}{\nu_q^{(N)}\sigma^2(1-s)[1+(N-1)s]}. \tag{B22}$$

APPENDIX C: THE CONDITION FOR A CONCEIVABLE s VALUE

In order to discuss the condition for a physically conceivable s value, we consider the global variable $X(t)$ defined by

$$X(t) = \frac{1}{N}\sum_i x_i(t). \tag{C1}$$

The first and second q moments of $X(t)$ are given by

$$E_q[X(t)] = \frac{1}{N}\sum_i E_q[x_i(t)] = \mu(t), \tag{C2}$$

$$E_q[\{\delta X(t)\}^2] = \frac{1}{N^2}\sum_i\sum_j E_q[\delta x_i(t)\delta x_j(t)] \tag{C3}$$

$$=\frac{1}{N^2}\sum_i E_q[\{\delta x_i(t)\}^2] + \frac{1}{N^2}\sum_i\sum_{j(\neq i)} E_q[\delta x_i(t)\delta x_j(t)] \tag{C4}$$

$$=\frac{\sigma(t)^2}{N}[1+(N-1)s(t)], \tag{C5}$$

where $\delta x_i = x_i(t) - \mu(t)$ and $\delta X(t) = X(t) - \mu(t)$. Since the global fluctuation in X is smaller than the average of the local fluctuation in $\{x_i\}$, we obtain

$$0 \leqslant E_q[\{\delta X(t)\}^2] \leqslant \frac{1}{N}\sum_i E[\{\delta x_i(t)\}^2] = \sigma(t)^2. \tag{C6}$$

Equations (C5) and (C6) yield

$$0 \leqslant \frac{[1+(N-1)s(t)]}{N} \leqslant 1.0, \tag{C7}$$

which leads to

$$s_L \leqslant s(t) \leqslant s_U, \tag{C8}$$

with $s_L = -1/(N-1)$ and $s_U = 1.0$.

[1] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[2] C. Tsallis, R. S. Mendes, and A. R. Plastino, Physica A **261**, 534 (1998).
[3] C. Tsallis, Physica D **193**, 3 (2004).
[4] B. R. Frieden, *Physics from Fisher information: A Unification* (Cambridge University Press, Cambridge, U.K., 1998).
[5] A. Plastino and A. R. Plastino, Physica A **222**, 347 (1995).
[6] C. Tsallis and D. J. Bukman, Phys. Rev. E **54**, R2197 (1996).
[7] A. Plastino, A. R. Plastino, and H. G. Miller, Physica A **235**, 577 (1997).
[8] F. Pennini, A. R. Plastino, and A. Plastino, Physica A **258**, 446 (1998).
[9] L. Borland, F. Pennini, A. R. Plastino, and A. Plastino, Eur. Phys. J. B **12**, 285 (1999).
[10] A. R. Plastino, M. Casas, and A. Plastino, Physica A **280**, 289 (2000).
[11] S. Abe, Phys. Rev. E **68**, 031101 (2003).
[12] J. Naudts, J. Inequal. Pure Appl. Math. **5**, 102 (2004).
[13] F. Pennini and A. Plastino, Physica A **334**, 132 (2004).
[14] M. Portesi, A. Plastino, and F. Pennini, Physica A **365**, 173 (2006).
[15] M. Portesi, F. Pennini, and A. Plastino, Physica A **373**, 273 (2007).
[16] M. Masi, e-print arXiv:cond-mat/0611300.
[17] H. Hasegawa, Phys. Rev. E **77**, 031133 (2008).
[18] B. B. Averbeck, P. E. Latham, and A. Pouget, Nat. Rev. Neurosci. **358**, 358 (2006).
[19] L. F. Abbott and P. Dayan, Neural Comput. **11**, 91 (1999).
[20] S. D. Wilke and C. W. Eurich, Neurocomputing **44-46**, 1023 (2002).
[21] S. Wu, S. Amari, and H. Nakamura, Neural Networks **17**, 205 (2004).
[22] C. Tsallis, M. Gell-mann, and Y. Sato, Proc. Natl. Acad. Sci. U.S.A. **102**, 15377 (2005).
[23] C. Tsallis, Physica A **375**, 7 (2006).
[24] G. Wilk and Z. Wlodarczyk, Physica A **376**, 279 (2007).
[25] H. Hasegawa, Physica A **365**, 383 (2006).
[26] Calculations of $Z_q^{(N)}$ for large N with the use of Eqs. (45) and (46) need a computer program of the Γ function with a negative (real) argument, which is not available in our facility.
[27] E. P. Borges, Physica A **340**, 95 (2004).
[28] H. Hasegawa, Physica A **384**, 241 (2007).
[29] H. Sakaguchi, J. Phys. Soc. Jpn. **70**, 3247 (2001).
[30] C. Anteneodo and C. Tsallis, J. Math. Phys. **44**, 5194 (2003).
[31] H. Hasegawa (unpublished).