Maximum likelihood estimators for generalized Cauchy processes

Hidetoshi Konno[a)]
Department of Risk Engineering, Faculty of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

Fumitoshi Watanabe
R & D Center, Tokyo Electric Power Company, Inc. 4-1 Egasaki-cho, Tsurumi-ku, Yokohama, 230-8510, Japan

(Received 25 July 2006; accepted 27 September 2007; published online 26 October 2007)

Maximum likelihood estimator (MLE) for a generalized Cauchy process (GCP) is studied with the aid of the method of information geometry in statistics. Our GCP is described by the Langevin equation with multiplicative and additive noises. The exact expressions of MLEs are given for the two cases that the two types of noises are uncorrelated and mutually correlated. It is shown that the MLEs for these two GCPs are free from divergence even in the parameter region wherein the ordinary moments diverge. The MLE relations can be regarded as a generalized fluctuation-dissipation theorem for the present Langevin equation. Availability of them and of some other higher order statistics is demonstrated theoretically and numerically.
 [DOI: 10.1063/1.2800162]

I. INTRODUCTION

Stochastic processes with multiplicative noise have been studied in conjunction with the appearance of power-law distribution in probability density and that of the long-time tail in time-correlation function.[1–7] In many practical situations, stochastic variables are often be driven by both multiplicative and additive noises. Some stochastic processes associated with the two types of noises have been studied by Deutsch,[8] Venkataramani *et al.*,[9] Takayasu *et al.*,[10] Levy and Soloman,[11] Nakao,[12] Konno *et al.*,[13] and others[14–22] in the fields of finance, physics, chemistry, biology, fluid mechanics, and so on.

Recently, the q-deformation technique has been utilized in various ways. It has been used to characterize the features of temporal correlation in time series and of structure of attractors in the phase space in complex chaotic/turbulent time series analysis.[23–27] On the other hand, Tsallis *et al.*[14] proposed a new theory of q-entropy to derive a power-law-type distribution. Their studies on the q-averaging method help us to avoid mathematical divergence of the variance. However, the validity of their entropy concept is not proven completely although there are some successful examples in the real world to support their theory. Also, they proposed[28] a special nonlinear Fokker-Planck (FP) equation associated with the power-law-type distributions and with the anomalous super-or subdiffusion. The method of estimating the parameters of the FP equation from the probability density function has not been given.

For practical applications, it is worth to give a method of inferring the parameters of the system only from time series data in such a realistic situation that a multiplicative noise is correlated with an additive noise. In the case of vector linear Langevin equation, the fluctuation-dissipation theorem[29,30] (FDT) can be utilized to infer the system parameters even when noise sources are mutually correlated. In spite of the vector linear Langevin equation, the general theory

[a)]Electronic mail: hkonno@sakura.cc.tsukuba.ac.jp

of FDT for vector nonlinear Langevin equations is not known. In statistics, it is recognized that the maximum likelihood estimator (MLE) is one of the most reliable tool to infer parameters of non-Gaussian probability distributions.

Here, we take the method of information geometry in statistics for generalized Cauchy processes (GCPs) to give the MLEs to infer the parameters only from time series data for the GCPs. The paper is organized as follows. Section II considers a GCP with two independent random forces. We show (i) how the parameters of a CGP is connected to the Langevin equation; (ii) how the universal statistics, free from the divergence, is derived by the method of information geometry; (iii) how the values of MLEs changes. In Sec. III, it is shown that a GCP with two dependent random forces will lead us to the probability density with broken symmetry. Then the expressions of MLEs are shown. Numerical examples are also given. Section IV discusses (A) physical significance of dual coordinate, (B) physical relation to FDT, (C) procedure to estimate parameters, (D) relation to Tsallis statistics, (E) connection to analysis on intermittent chaos, and (F) the relevance of the situation of dependent noises to the phase-velocity fluctuation in a stochastic complex Ginzburg-Landau equation. The final section is devoted to conclusions.

II. GENERALIZED CAUCHY PROCESS 1

A. Model with two independent noises

Let us consider a GCP of state variable x with a multiplicative and an additive noise,

$$\frac{d}{dt}x = -\alpha x + xF_p(t) + F_a(t), \tag{1}$$

where α is a positive constant, both $F_p(t)$ and $F_a(t)$ are independent Gaussian-white noises with null mean, i.e., $\langle F_p(t)\rangle = 0$ and $\langle F_a(t)\rangle = 0$,

$$\langle F_j(t)F_k(t')\rangle = 2D_j\delta_{j,k}\delta(t-t') \quad (j,k = p \text{ and } a). \tag{2}$$

B. Probability density

The FP equation for the stochastic differential equation (SDE) (1) becomes

$$\frac{\partial}{\partial t}P(x,t) = \frac{\partial}{\partial x}[K(x)P(x,t)] + \frac{\partial^2}{\partial x^2}[D(x)P(x,t)], \tag{3}$$

where $K(x) = (\alpha - D_p)x$ and $D(x) = D_p x^2 + D_a$. The term $D_p x$ in the regression term $K(x)$ come from the Wong-Zakai noise correction[31] due to the Storatonovich interpretation of Eq. (1). At the steady state, one obtains the probability density function in the form:[32,33]

$$P_s(x) = \frac{a^{2b-1}}{B(b-1/2,1/2)}\frac{1}{(a^2+x^2)^b}, \tag{4}$$

where $B(b-1/2,1/2)$ is the Beta function,

$$a^2 = \frac{D_a}{D_p}$$

and

$$b = \frac{\alpha}{2D_p} + \frac{1}{2}. \tag{5}$$

It is readily seen that $b - 1/2 = \alpha/2D_p > 0$, since α is assumed to be positive.

C. Moments

The equation of moments (for $n=1,2,\ldots$) can be derived from the FP equation in Eq. (3) as

$$\frac{d}{dt}\langle x^n\rangle = -n\langle x^{n-1}K(x)\rangle + n(n-1)\langle x^{n-2}D(x)\rangle, \tag{6}$$

where $K(x)=(\alpha-D_p)x$ and $D(x)=D_a+D_px^2$. It seems that the set of even order moments ($n=2,4,\ldots$) are regarded as a kind of FDT (or generalized Einstein relation): $-n\langle x^{n-1}K(x)\rangle+n(n-1)\langle x^{n-2}D(x)\rangle=0$. Putting $n=2$ and $n=4$, one obtains

$$\langle x^2\rangle = \frac{D_a}{\alpha-2D_p} = \frac{1}{2}\frac{a^2}{b-3/2} \tag{7}$$

and

$$\langle x^4\rangle = \frac{3D_a}{\alpha-4D_p}\langle x^2\rangle = \frac{3}{4}\frac{a^4}{(b-3/2)(b-5/2)}. \tag{8}$$

Since the second and the fourth moment should have positive values, the different constraints, i.e., $b>3/2$ for $n=2$ and $b>5/2$ for $n=4$ [i.e., $b>(2m+1)/2$ for the $2m$th moment, in general] must be imposed.

It is clear from the expression of the moments (7) and (8) that they cannot be utilized to infer the system parameters a and b for general circumstances even when $b>1/2$. Thus, it is not suitable to call them as the fluctuation-dissipation relations for the GCP in Eq. (1).

D. Maximum likelihood estimator

According to the method of information geometry in statistics by Amari,[34] Amari and Nagaoka,[35] Murrey and Rice,[36] and Kass and Vos,[37] the information geometrical potential (IGP) for the probability density function (pdf) in Eq. (4) is given by

$$\Psi(a,b) = \ln B(b-1/2,1/2) - (2b-1)\ln a. \tag{9}$$

The dual coordinate η_1 for the natural (canonical) parameter b is given by

$$\eta_1 = -\frac{\partial}{\partial b}\Psi(a,b) = \langle \ln(x^2+a^2)\rangle = \psi(b) - \psi(b-1/2) + 2\ln a, \tag{10}$$

where $\psi(z)$ is the digamma function defined in terms of the Gamma function $\Gamma(z)$ by

$$\psi(z) = \frac{d}{dz}\ln\Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}. \tag{11}$$

It is easy to verify that the quantity η_1 is exactly identical with one of the MLEs for the pdf in Eq. (4) in statistics.

On the other hand, the parameter a is not the natural one. Taking the derivative of $\Psi(a,b)$ with respect to a, one gets the second dual coordinate

$$\eta_2 = -\frac{\partial}{\partial a}\Psi(a,b) = \left\langle \frac{2ab}{x^2+a^2}\right\rangle = \frac{2b-1}{a}. \tag{12}$$

One may define the second dual coordinate $\eta_2^{(0)}$ in stead of η_2 by factoring out $2ab$ as

$$\eta_2^{(0)} = \left\langle \frac{1}{x^2+a^2}\right\rangle = \frac{1}{a^2}\left(1-\frac{1}{2b}\right). \tag{13}$$

It is easy to see that the two parameters a and b in the pdf in Eq. (4) are estimated by the two MLEs in Eqs. (10) and (13) [or Eq. (12)] in the parameter space $(a>0, b>1/2)$, since $b-1/2=\alpha/2D_p>0$. They are free from divergence. Therefore, we can regard them as the generalized FDT for the GCP in Eq. (1).

E. Higher order statistics and Fisher information matrix

To discuss the accuracy of estimated parameters a and b by the MLEs in Eqs. (10) and (13) [or Eq. (12)], let us derive the higher order statistics,

$$I_m(a,b) \equiv \langle [\ln(x^2+a^2)]^m \rangle, \tag{14}$$

and

$$J_m(a,b) \equiv \langle [1/(x^2+a^2)]^m \rangle \tag{15}$$

$(m=2,3,\ldots)$, and the Fisher information matrix,

$$g(a,b) \equiv \langle \partial_i \ell(x;a,b) \partial_j \ell(x;a,b) \rangle = -\langle \partial_i \partial_j \ell(x;a,b) \rangle, \tag{16}$$

where $\ell(x;a,b)=\ln P_s(x)$, $\partial_i=\partial/\partial\theta_i$, $(i=1,2;\theta_1=b$ and $\theta_2=a)$.

The higher order moments $I_m(a,b)$ $(m=2,3)$ are given by (see Appendix A for derivation) $I_2(a,b)=\langle \ln(x^2+a^2)\rangle^2-[\psi'(b)-\psi'(b-1/2)]$ and $I_3(a,b)=-2\langle \ln(x^2+a^2)\rangle^3+3\langle [\ln(x^2+a^2)]^2\rangle\langle \ln(x^2+a^2)\rangle+[\psi''(b)-\psi''(b-1/2)]$. One can get compact expressions for the moments $I_m(a,b)$ $(m=1,2,3)$ in terms of the cumulants as

$$\langle \ln(x^2+a^2)\rangle_c = \psi(b)-\psi(b-1/2)+2\ln a, \tag{17}$$

$$\langle [\ln(x^2+a^2)]^2\rangle_c = -[\psi'(b)-\psi'(b-1/2)], \tag{18}$$

and

$$\langle [\ln(x^2+a^2)]^3\rangle_c = \psi''(b)-\psi''(b-1/2), \tag{19}$$

where $\psi(z)$ is the digamma function, $\psi'(z)$ is the trigamma function, and $\psi''(z)$ is the tetragamma function. Note here that (i) the first cumulant involves the parameters a and b and (ii) the second and third cumulants involve only the parameter b.

On the other hand, the higher order mth moment of $J_m(a,b)$ $(m=2,3)$ is given in terms of the beta function by $J_m(a,b)=(1/a^{2m})(B(b+m-1/2,1/2)/B(b-1/2,1/2))$. Namely, $J_2(a,b)=(1-1/2b)(1-1/2(b+1))/a^4$ and $J_3(a,b)=(1-1/2b)(1-1/2(b+1))(1-1/2(b+2))/a^6$. In terms of the cumulants with the constant factor $2ab$ [cf. Eq. (12)], they are expressed as

$$\left\langle \left(\frac{2ab}{x^2+a^2} \right)^2 \right\rangle_c = \frac{2b-1}{a^2(b+1)}$$

and

$$\left\langle \left(\frac{2ab}{x^2+a^2} \right)^3 \right\rangle_c = -4\frac{(2b-1)(b-1)}{a^3(b+1)(b+2)}. \tag{20}$$

The Fisher information matrix $g(a,b)$ should be evaluated according to the definition in Eq. (16) since the parameter a is not the natural one. The result is given by

TABLE I. Typical values of polygamma function.

z	$\psi(z)$	$\psi'(z)$	$\psi''(z)$
1/4	$-\gamma-3\ln 2-\pi/2$	17.197	−129.328
1/2	$-\gamma+2\ln 2$	$\pi^2/2$	$-14\zeta(3)$
1	$-\gamma$	$\pi^2/6$	$-2\zeta(3)$
3/4	$-\gamma-3\ln 2+\pi/2$	2.542	−5.303
3/2	$2-2\ln 2-\gamma$	$-4+\pi^2/2$	$16-14\zeta(3)$

$$g(a,b) \equiv \begin{pmatrix} \langle[\ln(x^2+a^2)]^2\rangle_c, & * \\ \langle \ln(x^2+a^2)(2ab/(x^2+a^2))\rangle_c, & \langle(2ab/(x^2+a^2))^2\rangle_c \end{pmatrix} = \begin{pmatrix} -[\psi'(b)-\psi'(b-1/2)], & * \\ -1/ab, & (2b-1)/a^2(b+1) \end{pmatrix}. \tag{21}$$

The first diagonal element becomes the same value as $\partial^2\Psi(a,b)/\partial b^2=-[\psi'(b)-\psi'(b-1/2)]$. On the other hand, the second diagonal element in Eq. (21) differs from $\partial^2\Psi(a,b)/\partial a^2=(2b-1)/a^2$. The cross cumulant in the off-diagonal element $\langle 2ab/(x^2+a^2)\cdot\ln(x^2+a^2)\rangle_c \equiv \langle 2ab/(x^2+a^2)\cdot\ln(x^2+a^2)\rangle - \langle 2ab/(x^2+a^2)\rangle\langle\ln(x^2+a^2)\rangle = -1/ab$ which is different from $\partial^2\Psi(a,b)/\partial a\partial b=-2/a$. This $g_{ij}(\theta)$ represents the geodesic distance between the two distributions $p=p(x,\theta)$ and $p'=p(x,\theta+d\theta)$, $ds^2=|\vec{pp}'|^2=g_{ij}(\theta)d\theta^i d\theta^j$ in the Riemanian space $\theta\in\Theta$.[34–37] The Riemanian metric $g_{ij}(\theta)$ gives also the Cramer-Rao lowest bound[34–37] which expresses the degree of accuracy for estimated parameters. It is easily verified that the minimum value is attained $Cov[\overline{F_i(x)}-\eta_i,\overline{F_j(x)}-\eta_j]=g_{ij}(\theta)/N$ for $\overline{F_1(x)}=\Sigma_i\ln(x_i^2+a^2)/N$ and $\overline{F_2(x)}=\Sigma_i 2ab/(x_i^2+a^2)/N$ by the MLEs in the case of statistical inferrence on the fat-tailed distribution $P_s(x,\theta)$ with N independent observations $\{x_i\}$ even though the non-natural parameter a is involved.

F. Numerical examples

Some typical values of polygamma function for $b=3/4$, $b=1$, and $b=3/2$ are summarized in Table I. Also, typical values of $\langle\ln(x^2+a^2)\rangle_c$, $\langle[\ln(x^2+a^2)]^2\rangle_c$, and $\langle[\ln(x^2+a^2)]^3\rangle_c$ are given in Table II. In this table, the numeric values of $\langle[\ln(x^2+a^2)]^2\rangle_c$ and $\langle[\ln(x^2+a^2)]^3\rangle_c$ are given for $b=3/4$, since their corresponding analytic expression cannot be obtained. The function $\zeta(z)$ is the Riemann zeta function defined by $\zeta(s)\equiv\Sigma_{k=1}^{\infty}k^{-s}$.

The values of $U_1\equiv\{\langle\ln(x^2+a^2)\rangle_c-2\ln a\}$, $U_2\equiv\langle[\ln(x^2+a^2)]^2\rangle_c$, and $U_3\equiv\langle[\ln(x^2+a^2)]^3\rangle_c$, which depend on only b, are depicted in Fig. 1 (U_1, solid line; U_2, dashed line; U_3, dotted line). In the parameter range of $1/2<b<3/2$, $U_1<U_2<U_3$. This fact implies that large sampling error is expected as the value of $U_i(i=1,2,3)$ increases in this region. On the other hand, values of U_i tend to converge to the order of $O(0.1)$ in the region $b>3/2$. This means the associated sampling errors also converge to values of the same order. When a long-time record of data is available, the effect of sampling errors may be reduced. In this case, the third order cumulant U_3 will be used as information to infer value of the parameters.

TABLE II. Typical values of the first, the second, and the third order cumulants for $b=3/4$, 1, and 3/2. $\zeta(z)$ is the Riemann zeta function.

b	$\langle\ln(x^2+a^2)\rangle_c$	$\langle[\ln(x^2+a^2)]^2\rangle_c$	$\langle[\ln(x^2+a^2)]^3\rangle_c$
3/4	$\pi+2\ln a$	14.655	124.025
1	$2\ln 2+2\ln a$	$\pi^2/3$	$12\zeta(3)$
3/2	$2-2\ln 2+2\ln a$	$4-\pi^2/3$	$16-12\zeta(3)$

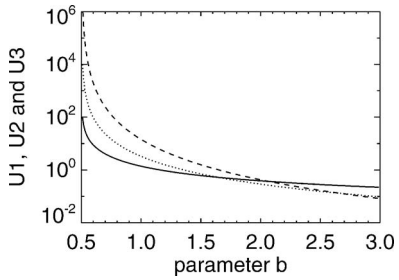

FIG. 1. The universal linear-log curves of $U_1(=\langle \ln(x^2+a^2)\rangle_c - 2\ln a)$ (solid line), $U_2(=\langle [\ln(x^2+a^2)]^2\rangle_c)$ (dotted line), and $U_3(=\langle [\ln(x^2+a^2)]^3\rangle_c)$ (dased line) as a function of parameter b are depicted. For $1/2<b<3/2$, there is a monotonic order: $U_1<U_2<U_3$.

III. GENERALIZED CAUCHY PROCESS 2

A. Model with two correlated noises

Next, let us consider the GCP under the influence of the noises $F_p(t)$ and $F_a(t)$ in Eq. (1) which are mutually correlated,

$$\langle F_p(t)F_a(t')\rangle = 2D_{ap}\delta(t-t'). \tag{22}$$

The parameters and the other constraints imposed on the fluctuating forces are preserved.

B. Probability density

The FP equation for the correlated noises becomes

$$\frac{\partial}{\partial t}P(x) = \frac{\partial}{\partial x}[K(x)P(x,t)] + \frac{\partial^2}{\partial x^2}[(D(x)P(x,t)\ ], \tag{23}$$

where $K(x)=(\alpha-D_p)$ and $D(x)=D_px^2+2D_{ap}x+D_a$. The pdf at the steady state is given by

$$P_s(x) = P_0(D_px^2+2D_{ap}x+D_a)^{-\alpha/2D_p-1/2}\exp\left\{\frac{(\alpha-D_p)D_{ap}}{D_p\sqrt{D_pD_a-D_{ap}^2}}\arctan\frac{D_px+D_{ap}}{\sqrt{D_aD_p-D_{ap}^2}}\right\}, \tag{24}$$

where P_0 is the normalization constant. To derive Eq. (24), the natural assumption among the strengths of noises, $D_pD_a>D_{ap}^2$, is adopted. To simplify the expression (24), let us put $a^2=D_a/D_p$, $b=1/2+\alpha/2D_p$ in the same way in Eq. (5), and define new parameters as

$$A=\sqrt{a^2-c^2}, \quad c=\frac{D_{ap}}{D_p},$$

and

$$f=\frac{2(b-1)c}{A}. \tag{25}$$

After transforming into the new variable ξ via

$$\xi = x+c, \tag{26}$$

the pdf takes in the compact form,

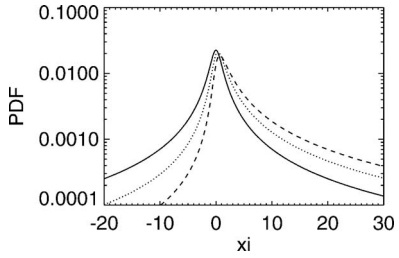

FIG. 2. The pdf profile $P_s(\xi)$ as a function of parameter f [$f=0$ (solid line), $f=0.5$ (dotted line), $f=1$ (dashed line)] with $a=1$ and $b=3/4$.

$$P_s(\xi) = \frac{(2A)^{2b-1}|\Gamma(b+if/2)|^2}{2\pi\Gamma(2b-1)} \cdot \frac{\exp[f\arctan(\xi/A)]}{(\xi^2+A^2)^b}, \tag{27}$$

where A, b, and f are the three parameters defined in Eq. (25) (cf. the original number of parameters are four; α, D_a, D_p, and D_{ap}). Due to the existence of the correlation $D_{ap} \neq 0$, the symmetry breaking of the pdf profile takes place. To obtain the analytical expression of the normalization factor, we have used the variable transformation,

$$y = \arctan(\xi/A), \tag{28}$$

and the integral formula,

$$\int_0^{\pi/2} \cos^\beta y \cosh(fy)dy = \frac{\pi\Gamma(\beta+1)}{2^{\beta+1}|\Gamma(\{\beta+2+if\}/2)|^2}. \tag{29}$$

In the limit $f \to 0$, the pdf in Eq. (27) is reduced to the one in Eq. (4); the normalization factor is reduced to $a^{2b-1}B(b-1/2,1/2)$. The duplication formula for the Gamma function, $\Gamma(2(b-1/2)) = (2\pi)^{-1/2}2^{2(b-1/2)-1/2}\Gamma(b-1/2)\Gamma(b)$, is available to prove this formula in Eq. (29).

The most probable value (or the mode) of the pdf in Eq. (27) is placed at

$$\xi_p = \frac{fA}{2b} = \left(1-\frac{1}{b}\right)c. \tag{30}$$

In the original x coordinate in Eq. (24), $x_p = -(c/b)$, forms Eq. (26). The feature of deformation in the pdf (27) as a function of f [$f=0$ (solid line); $f=0.5$ (dotted line); $f=1$ (dashed line)] for $A=1$ and $b=1$ is displayed in Fig. 2. This result means that a careful detection of the position of the mode x_p gives us available information on (c/b).

C. Moments

Taking the same procedure as described in Sec. II, one obtains the expressions of the second and the fourth moment in the form

$$\langle \xi^2 \rangle = \frac{1}{2}\frac{A^2}{b-3/2} + \frac{A^2}{(b-1)(b-3/2)}\left(\frac{f}{2}\right)^2 \tag{31}$$

and

$$\langle \xi^4 \rangle = \frac{3}{4}\frac{A^4}{(b-3/2)(b-5/3)} + \frac{A^4}{(b-1)(b-2)(b-3/2)(b-5/2)}\left(\frac{f}{2}\right)^2\left\{\left(\frac{f}{2}\right)^2 + 3b - 5\right\}. \tag{32}$$

It is clear from the results that there are constraints $2b-3>0$ in Eq. (31) and $2b-5>0$ in Eq. (32). These constraints are same as the case of independent noises in Eqs. (7) and (8) described in Sec. II.

The features of deformation from the case of $f=0$ in the pdf might influence upon the values of odd order moments ($\langle \xi^{2m+1} \rangle, (m=0,1,2,\cdots)$). It is not easy to get useful analytical expressions of them.

D. Maximum likelihood estimator

The IGP for the pdf in Eq. (27) can be defined in the form

$$\Psi(A,b,f) = -(2b-1)\ln(2A) + \ln(2\pi) + \ln\Gamma(2b-1) - \ln\Gamma(b+if/2) - \ln\Gamma(b-if/2). \tag{33}$$

The two dual coordinates η_1 and η_2[34,35] for the natural (canonical) parameters (b and f) are given by

$$\eta_1 = -\frac{\partial}{\partial b}\Psi(A,b,f) = \langle \ln(\xi^2 + A^2) \rangle \tag{34}$$

$$= 2\ln(2A) + \psi(b+if/2) + \psi(b-if/2) - 2\psi(2b-1), \tag{35}$$

and

$$\eta_2 = \frac{\partial}{\partial f}\Psi(A,b,f) = \langle \arctan(\xi/A) \rangle \tag{36}$$

$$= \frac{i}{2}\{\psi(b-if/2) - \psi(b+if/2)\}. \tag{37}$$

The first dual coordinate (34) depends on the parameters A, b, and f. On the other hand, the second one (36) depends only on two parameters b and f. The appearance of irrelevance of f to A is a formal nature since $f=2(b-1)/A$, as defined in Eq. (25).

Since A is the non-natural parameter, the third dual coordinate η_3 is expressed in terms of subcoordinates both $\eta_3^{(1)} \equiv \langle 1/(\xi^2+A^2) \rangle$ and $\eta_3^{(2)} \equiv \langle \xi/(\xi^2+A^2) \rangle$ as

$$\eta_3 = 2bA\,\eta_3^{(1)} + f\,\eta_3^{(2)}. \tag{38}$$

This relation is derived by noticing the following two equations: (i) Differentiating both sides of $\int_{-\infty}^{\infty} P_s(\xi)d\xi = 1$ [cf. Eq. (40)] with respect to A, one obtains $2bA\,\eta_3^{(1)} + f\,\eta_3^{(2)} = (2b-1)/A$; (ii) $\eta_3 = -\partial/\partial A \Psi(A,b,f) = (2b-1)/A$.

The subcoordinate $\eta_3^{(1)}$ is easily evaluated as

$$\eta_3^{(1)} = \left\langle \frac{1}{\xi^2 + A^2} \right\rangle = \frac{N_0(A,b+1,f)}{N_0(A,b,f)} = \frac{b(b-1/2)}{A^2(b^2+f^2/4)}, \tag{39}$$

where the function $N_0(A,b,f)$ is defined by

$$N_0(A,b,f) = \int_{-\infty}^{\infty} \frac{\exp(f\arctan(\xi/A))}{(\xi^2+A^2)^b} d\xi = \frac{2\pi\Gamma(2b-1)}{(2A)^{2b-1}|\Gamma(b+if/2)|^2}. \tag{40}$$

By noticing the constraint $2bA\,\eta_3^{(1)} + f\,\eta_3^{(2)} = (2b-1)/A$, we obtain

$$\eta_3^{(2)} = \left\langle \frac{\xi}{\xi^2 + A^2} \right\rangle = \frac{(b-1/2)f}{2A(b^2+f^2/4)}. \tag{41}$$

These analytical expressions depend on A, b, and f are available to infer the parameters. In this case, the ratio of these becomes [cf. eq. (30)]

$$\frac{\eta_3^{(2)}}{\eta_3^{(1)}} = \frac{Af}{2b} = \left(1 - \frac{1}{b}\right)c. \tag{42}$$

In the limit $f \to 0$, the first dual coordinate in Eq. (35) $\eta_1 \to \psi(b) - \psi(b-1/2) + 2\ln a$. One can prove this with the aid of the duplication formula, $\psi^{(m)}(2b-1) = (\psi^{(m)}(b-1/2) + \psi^{(m)}(b))/2^{m+1} + \delta_{m,0}\ln 2$ for the polygamma function $[\psi^{(0)}(z) \equiv \psi(z),\ \psi^{(1)}(z) \equiv \psi'(z)$ and $\psi^{(2)}(z) \equiv \psi''(z)]$. The second dual coordinate in Eq. (37) $\eta_2 \to 0$ for $f \to 0$. The third dual coordinate η_3 is expressed in terms of both $\eta_3^{(1)} = \langle 1/(\xi^2 + A^2)\rangle$ and $\eta_3^{(2)} = \langle \xi/(\xi^2 + A^2)\rangle$ in Eq. (38). In the limit $f \to 0$, $\eta_3^{(1)} \to (1 - 1/2b)/a^2$ [cf. Eq. (13)] and $\eta_3^{(2)} \to 0$. We can regard the MLEs as the for the GCP in Eq. (1) with the correlated noises $F_p(t)$ and $F_a(t)$.

E. Higher order statistics and Fisher information matrix

The expressions of the higher order cumulants $(m = 2, 3, \ldots)$ for the two natural parameters b and f can be evaluated by taking derivatives of the IGP as described in Sec. II. Since the integral formula $N_0(A, b, f)$ is given in Eq. (40), the integral of moment,

$$K_m(A,b,f) = N_0^{-1}(A,b,f)\int_{-\infty}^{\infty} [\ln(\xi^2 + A^2)]^m \frac{\exp(f \arctan(\xi/A))}{(\xi^2 + A^2)^b} d\xi, \tag{43}$$

can be evaluated by using the recursion relation $N_m(A,b,f) = -\partial N_{m-1}(A,b,f)/\partial b$ $(m = 1, 2, \ldots)$. The results of cumulants $C_m(A,b,f) = \langle [\ln(\xi^2 + A^2)]^m \rangle_c$ $(m = 2, 3)$ are given by

$$C_m = (-1)^{m-1}\{\psi^{(m-1)}(b + if/2) + \psi^{(m-1)}(b - if/2) - 2^m \psi^{(m-1)}(2b-1)\}. \tag{44}$$

Then, the cross-cumulant associated with the two natural parameters b and f is given by

$$\eta_4 = -\frac{\partial^2}{\partial f \partial b}\Psi(A,b,f) = \langle \ln(\xi^2 + A^2)\arctan(\xi/A)\rangle_c \tag{45}$$

$$= \frac{i}{2}\{\psi'(b + if/2) - \psi'(b - if/2)\}. \tag{46}$$

This statistics seems to be sensitive to the degree of broken symmetry in the pdf. The universal statistics described here in the present GCP take real values though they are expressed in terms of the polygamma functions with imaginary argument.

The analytical expressions of all elements of the Fisher information matrix are obtained (see Appendix B for derivation) according to the definition $g(A,b,f) = -\langle \partial_i \partial_j \ell(\xi; A, b, f)\rangle$, where $\ell(\xi; A, b, f) = \ln P_s(\xi)$, $\partial_i = \partial/\partial\theta_i (i = 1, 2, 3; \theta_1 = b, \theta_2 = f, \theta_3 = A)$,

$$g(A,b,f) \equiv \begin{pmatrix} \langle [\ln(\xi^2 + A^2)]^2\rangle_c, & *, & * \\ \langle \arctan(\xi/A)\ln(\xi^2 + A^2)\rangle_c, & \langle [\arctan(\xi/A)]^2\rangle_c, & * \\ \left\langle \left(\frac{2Ab}{\xi^2 + A^2} + \frac{f\xi}{\xi^2 + A^2}\right)\ln(\xi^2 + A^2)\right\rangle_c, & \left\langle \left(\frac{2Ab}{\xi^2 + A^2} + \frac{f\xi}{\xi^2 + A^2}\right)\arctan(\xi/A)\right\rangle_c, & \left\langle \left[\frac{2Ab}{\xi^2 + A^2} + \frac{f\xi}{\xi^2 + A^2}\right]^2\right\rangle_c \end{pmatrix}$$

$$= \begin{pmatrix} C_2, & *, & * \\ -\eta_4, & (1/4)[\psi'(b - if/2) + \psi'(b + if/2)], & * \\ -\frac{b + f^2}{A(b^2 + f^2/4)}, & \frac{(2b-1)f}{4A(b^2 + f^2/4)}, & \frac{(2b-1)(b + 1 + f^2/4)}{A^2[(b+1)^2 + f^2/4]} \end{pmatrix}. \tag{47}$$

The matrix elements $\{(3, j), j = 1, 2, 3\}$ are different from $\partial^2\Psi(A,b,f)/\partial b \partial A = -2/A$, $\partial^2\Psi(A,b,f)/\partial f \partial A = 0$, and $\partial^2\Psi(A,b,f)/\partial A^2 = (2b-1)/A^2$, respectively. The evaluation of cross cumulants in the off-diagonal matrix elements in $g_{ij}(\theta)$ is reduced to calculating $\langle 1/(\xi^2 + A^2)\rangle$, $\langle 1/(\xi^2 + A^2)^2\rangle$, $\langle \xi/(\xi^2 + A^2)\rangle$, and $\langle \xi/(\xi^2 + A^2)^2\rangle$ (see Appendix B).

With the use of $g(A,b,f)$, one can verify also that the minimum value of the Cramer-Rao's inequality[35–37] is attained $Cov[\overline{F_i(\xi)} - \eta_i, \overline{F_j(\xi)} - \eta_j] = g_{ij}(\theta)/N$ for $\overline{F_1(\xi)} = \Sigma_i \ln(\xi_i^2 + A^2)/N$, $\overline{F_2(\xi)}$

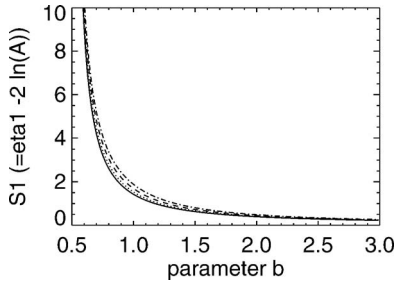

FIG. 3. The universal curve $S_1(\equiv\{\eta_1-2\ln(A)\})$ as a function of parameter b for $f=0.25$ (solid line), $f=0.5$ (dotted line), $f=0.75$ (dashed line), and $f=1.0$ (dash-dot line).

$=\Sigma_i \arctan(\xi/A)/N$, and $\overline{F_3(\xi)}=\Sigma_i[2Ab/(\xi_i^2+A^2)+f\xi/(\xi^2+A^2)]/N$ by the MLEs in the case of statistical inferrence on the fat-tailed distribution $P_s(\xi,\theta)$ with N independent observations $\{\xi_i\}$ even though the pdf has a broken symmetry, and the non-natural parameter A is involved.

F. Numerical examples

The curves of $S_1\equiv\{\eta_1-2\ln(A)\}$, η_2, and η_4 for $f=0.25$, 0.5, 0.75, and 1.0 as a function of parameter b are displayed in Figs. 3–5.

There are marked difference among the pdf profiles for $f=0.5$ and 1.0, as shown in Fig. 2. However, as far as the universal curve of S_1 in Fig. 3 is concerned, there are no marked difference among the values of S_1 as a function of f ($f=0.25,0.5,0.75,1.0$). Namely, the curves of $S_1=\eta_1-2\ln(A)$ as a function of f and b are located in the narrow band. One can say that S_1 is not sensitive statistic to variation of the parameter f. The effect of sampling errors in S_1 takes the same order of magnitude for the variations of f (=0.25, 0.5, 0.75, and 1).

In Fig. 4, η_2 shows monotonic decrease as a function of b. Also, η_2 is sensitive to the parameter f. Namely, the separability of each curve for different values of f as a function of b is clear. For η_2, large variations as a function of f can be seen in Fig. 4. The increase of relative magnitude of sampling errors is expected as the value of f becomes large. But, it is not worry about the absolute values of them.

In Fig. 5, the curves of η_4 as a function of f and b are depicted. The monotonic dependence of η_4 can be observed in the whole region of b in the case $f=0.25$. On the other hand, the curves of $f=0.50$ intersect those of $f=0.75$ and $f=1.0$ in the region of $0.5<b<0.7$. It is important to notice that $\eta_4(=\langle\ln(\xi^2+A^2)\arctan(\xi/A)\rangle_c)$ takes the form of the gradient of η_2 with respect to b. Although η_4 is a higher order statistics for the MLEs η_1, η_2, and η_3, the cross cumulant η_4 may be helpful to infer value of the parameters by virtue of medium sensitivity to variation of f.

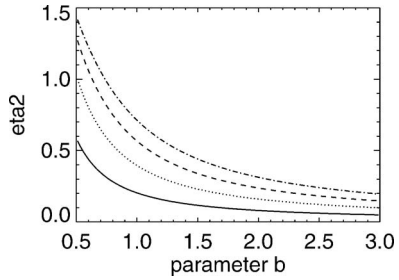

FIG. 4. The universal curve $\eta_2(\equiv\langle\arctan(\xi/A)\rangle)$ as a function of parameter b for $f=0.25$ (solid line), $f=0.5$ (dotted line), $f=0.75$ (dashed line), and $f=1$ (dash-dot line).

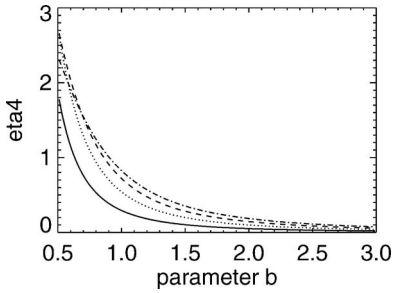

FIG. 5. The universal curve $\eta_4(\equiv\langle \ln(\xi^2+A^2)\arctan(\xi/A)\rangle_c)$ as a function of parameter b for $f=0.25$ (solid line), $f=0.5$ (dotted line), $f=0.75$ (dashed line), and $f=1$ (dash-dot line).

IV. DISCUSSIONS

A. Physical significance of dual coordinate

Let us discuss the physical significance of the dual coordinate in the GCP before considering the difference between the natural and nonnatural parameters in terms of the FDT in the GCP. To get a clear physical picture, we introduce a Langevin equation with a multiplicative noise $F_m(t)$, which is equivalent to the case of correlated noises ($D_{ap}\neq 0$) [cf. FP equation in Eq. (23)],

$$\frac{d}{dt}x = -\alpha x + \sqrt{2(D_p x^2 + 2D_{ap}x + D_a)}F_m(t). \tag{48}$$

For brevity of discussion, the Wong-Zakai (WZ) noise correction (the Ito correction) is not accounted here. Taking the transformation of variable with

$$X = \frac{x+c}{A}, \tag{49}$$

we have

$$\frac{d}{dt}X = -\alpha X + \beta + \sqrt{2D_p(X^2+1)}F_m(t), \tag{50}$$

where

$$\langle F_m(t)\rangle = 0, \quad \langle F_m(t)F_m(t')\rangle = \delta(t-t'), \tag{51}$$

and β is the parameter in terms of α, D_p, D_{ap}, and D_a [A and c are defined in Eq. (25)] defined by

$$\beta = \frac{\alpha c}{A}. \tag{52}$$

Then taking the nonlinar transformation with

$$X = \sinh(Y), \tag{53}$$

we have the Langevin equation with the additive noise $F_m(t)$,

$$\frac{d}{dt}Y = -\alpha \tanh(Y) + \beta\,\mathrm{sech}(Y) + \sqrt{2D_p}F_m(t). \tag{54}$$

When the WZ noise correction (the Ito correction) is taken, α in Eq. (54) is replaced by $\hat{\alpha}\equiv\alpha - D_p$. Without bypassing the FP equation, one can write down immediately the pdf for Y,

$$P_s(Y) = P_0 \exp\{-V(Y)/D_p\}, \tag{55}$$

where P_0 is the normalization constant, and the potential is defined by

$$V(Y) = \alpha \ln(\cosh(Y)) - \beta \arctan(\sinh(Y)). \tag{56}$$

By virtue of the above transformations of variables, the number of parameters in the pdf for Y reduces to two. They are defined by

$$\theta_1 = \frac{\alpha}{D_p}$$

and

$$\theta_2 = \frac{\beta}{D_p} (= \theta_1 s), \tag{57}$$

where s is the nondimensional scale factor $s = 1/\sqrt{r-1}$ with $r = D_p D_a / D_{ap}^2 > 1$. It is important to notice that these $\{\theta_1, \theta_2\}$ are both natural parameters for the pdf in the coordinate Y. We can say that the above Langevin equation for Y is the canonical equation in the GCP. The dual coordinates $\langle \ln(\cosh(Y)) \rangle$ [i.e., $\langle \ln(X^2+1) \rangle$ in the X coordinate] and $\langle \arctan(\sinh(Y)) \rangle$ [i.e., $\langle \arctan(X) \rangle$ in the X coordinate] are expressed in terms of the two natural parameters $\{\theta_1, \theta_2\}$ that they are not embedded in nonlinear function. It is important to see how to eliminate non-natural parameter that is embedded in linear/nonlinear function, in addition to get the physical significance.

Without taking the variable transformation, it is not avoidable to have complicated expressions in the MLEs with natural and non-natural parameters and to give supplementary coordinates such as $\langle 1/(x^2+a^2) \rangle$ and $x/(x^2+a^2)$ that are shown in Sec. III.

B. Physical relation to fluctuation-dissipation theorem

By observing the Langevin equation for Y in Eq. (54) and the corresponding pdf in Eq. (55), one can see that the natures of statistical quantities in the Y coordinate might be combined with that of the potential $V(Y)$. Since the potential function is concave up for positive $\alpha > 0$. It is easily expected that the nondivergent property of the statistical quantities in the Y coordinate is ascribed to the global stability of the highly nonlinear potential $V(Y)$. The MLE relations in the Y coordinate reduce to

$$\langle \ln(\cosh(Y)) \rangle = \ln 2 + \tfrac{1}{2}\{\psi(\theta_1 + i\theta_2/2) + \psi(\theta_1 - i\theta_2/2))\} - \psi(2\theta_1 - 1) \tag{58}$$

and

$$\langle \arctan(\sinh(Y)) \rangle = \frac{i}{2}\{\psi(\theta_1 - i\theta_2/2) - \psi(\theta_1 + i\theta_2/2)\}. \tag{59}$$

These are obtained from Eqs. (35) and (37) with substitutions $\xi \to X$, $A \to 1$, $b \to \theta_1$, $f \to \theta_2$, and with noticing the relations $\ln(\cosh(Y)) = (1/2)\ln(1+X^2)$ and $\arctan(\sinh(Y)) = \arctan(X)$.

The MLE relations obtained in Eqs. (58) and (59) can be regarded as the exact expression of "a generalized fluctuation-dissipation theorem (GFDT)" with the two natural parameters $\theta_1 = \alpha/D_p$ and $\theta_2 = \beta/D_p$ [i.e., the ratios of the strengths of fluctuation (D_p, D_a) and the coefficient of gross nonlinear dissipation α $[\mathrm{s}^{-1}]$ and that of gross nonlinear excitation (negative dissipation) $\beta = \alpha c/A$ $[\mathrm{s}^{-1}]$) for the "canonical" equation in Eq. (50) in the Y coordinates. The canonical equation has only the two natural parameters α and β. Namely, in the Langevin equation in Eq. (54), two different processes of a nonlinear dissipation $\alpha \tanh(Y)$ ("kink-type" nonlinear normal mode) and a nonlinear excitation (negative dissipation) $\beta \,\mathrm{sech}(Y)$ ("soliton-type" nonlinear normal mode) are competing under the influence of the additive noise $F_m(t)$. Therefore, the GFDTs in Eqs. (58) and (59) are those for the gross nonlinear dissipation of "kink" and for the gross nonlinear excitation (negative dissipation) of "soliton" in terms of the two natural parameters for

the nonlinear normal modes in the Y coordinate. For the special case with only the kink-type nonlinear mode is relevant, $\theta_1=1$ and $\theta_2=0$ [i.e., $P_s(Y)=\exp(-\ln(\cosh(Y)))/\pi=\mathrm{sech}(Y)/\pi$], the GFDT reduces to $\langle \ln(\cosh(Y))\rangle=\ln 2$ and $\langle \arctan(\sinh(Y))\rangle=0$.

In the limit of small Y, the expansion of $\ln(\cosh(Y))$ with respect to Y ($\ln(\cosh(Y))\approx Y^2+O(Y^4)$) leads to the Einstein relation $\langle Y^2\rangle=D_p/\alpha+\beta^2/\alpha^2$, i.e., the FDT in the Y coordinate for the linear system with the constant bias β. In the same limit, the expansion of $\langle \arctan(\sinh(Y))\rangle$ ($\arctan(\sinh(Y))\approx Y+O(Y^3)$) gives the first moment $\langle Y\rangle=\beta/\alpha$. Namely, the MLE relations reduce to the traditional FDT in the Y coordinate in this limit.

The physical and mathematical arguments described in the Y coordinate are used to interpret the MLE relations as the GFDT for the nonlinear normal modes in the X coordinate for the Langevin equation in Eq. (50) on account of the nonlinear transformation $X=\sinh(Y)$ from the Y coordinate to the X coordinate. Namely, the nonlinear normal modes in the X coordinates reduce to $X/(1+X^2)$ and $1/\sqrt{1+X^2}$. The associated dual coordinate for the two modes are $(1/2)\langle \ln(1+X^2)\rangle$ and $\langle \arctan(X)\rangle$ in the GDFT of the X coordinate on account of the integral formulas $\int dX[X/(1+X^2)]=(1/2)\ln(1+X^2)$ and $\int dX[1/\sqrt{1+X^2}]=\arctan(X)$.

Finally, let us argue with the MLE relations for Eq. (1) in the original x coordinate (or the ξ coordinate) in connection with the GFDT in terms of natural and unnatural parameters. It is clear physically that the multiplicative noise $xF_p(t)$ plays the role in giving rise to (i) intermittent burst in temporal profile and (ii) fat tail in pdf. The interplay between the two noises $F_p(t)$ and $F_a(t)$ is the origin of the natural D_p and the unnatural parameter D_a in the diffusion coefficient $D(x)=D_px^2+2D_{ap}x+D_a$ in the FP equation. One should note that D_{ap} is not counted as unnatural parameter since it is removed by introducing the shift $\xi=x+c$ as shown in the pdf of the ξ coordinate in Eq. (27). In observing the Langevin equation of (48) in the x coordinate [this is equivalent to Eq. (1)], the X coordinate is connected with the original x coordinate by the affine transformation $\xi=x+c=AX$ in Eq. (49). Following the argument described in the Y and X coordinates, the nonlinear "collective" mode in the original x coordinate associated with the Langevin equation of Eq. (48) is defined by $Coll(x)\equiv \alpha x/(D_px^2+2D_{ap}x+D_a)$ [the ratio of the dissipation αx to the strength of fluctuations $D(x)$]. Consequently, the statistical measures associated with the collective mode become $\langle \ln(D_px^2+2D_{ap}x+D_a)\rangle$ and $\langle \arctan([(D_px+D_{ap})/\sqrt{D_pD_a-D_{ap}^2}])\rangle$ [cf. the pdf in Eq. (24)]. With noticing the integral formula $\int dx x/(D_px^2+2D_{ap}x+D_a)=(1/2D_p)\ln(D_px^2+2D_{ap}x+D_a)-(D_{ap}/2D_p\sqrt{D_pD_{ap}-D_{ap}^2})\arctan\{(D_px+D_{ap})/\sqrt{D_pD_{ap}-D_{ap}^2}\}$, the function $Coll(x)$ expresses a nonlinear mixed mode. The interplay among the noises and the state variable x makes things complex. Thus the MLE relations in the x coordinate are regarded as a GFDT for the nonlinear collective (mixed) modes $Coll(x)$ for the GCP in Eq. (1).

Okabe[38] proposed a generalized FDT for one-dimensional nonlinear diffusion process. But the FDT cannot be utilized in the situation that the ordinary second moment diverges.

C. Procedures to estimate parameters

To give the way of thinking to estimate parameters in the GCP, let us exhibit first the parameter estimation for the amplitude distribution for a stochastic Ginzburg-Landau equation: $P_s(R)=[2\beta^{(\gamma+1)/2}/\Gamma((\gamma+1)/2)]R^{(\gamma-1)}\exp(-\beta R^2)$.[39,40] The parameters β and γ are natural ones. In this case, the MLE relations become $\langle R^2\rangle=(\gamma+1)/2\beta$ and $\langle \ln R\rangle=\psi((\gamma+1)/2)-(1/2)\ln\beta$. The parameter estimation can be done successfully with the two MLEs. The cross cumulant between them (the off-diagonal element of Fisher information) is given by $\langle R^2\ln R\rangle_c=1/2\beta$. In this case, one can also estimate successfully values of these parameters by combining $\langle R^2\rangle$ and $\langle R^2\ln R\rangle_c$ with simple algebraic calculation and without bypassing the digamma function $\psi(z)$.[39,40]

The method of parameter estimation from time series data in the GCP should be recursive even in the case of independent noise since there is the non-natural parameter a in the pdf for the GCP. One can take the MLEs $\langle \ln(x^2+a^2)\rangle=\psi(b)-\psi(b-1/2)+2\ln a$ in Eq. (10) and $\langle 1/(x^2+a^2)\rangle=(1-1/2b)/a^2$ in Eq. (13). Also, one can take the combination of $\langle 1/(x^2+a^2)\rangle$ and $\langle \ln(x^2+a^2)/(x^2+a^2)\rangle_c=-1/2a^2b^2$ in Eq. (21) (the off-diagonal element of Fisher information). The latter combination gives an estimation procedure without the di-gamma function $\psi(z)$. But, due to the

existence of the non-natural parameter a, a recursive procedure with the Newton method starting from a set of initial values (a,b) is required to get convergence in any case. In estimating value of $\psi(z)$, it is convenient to take the series expansion form, $\psi(z)=-\gamma-\Sigma_{n=0}^{\infty}(1/(z+n)-1/(n+1))$ [for the polygamma function, $\psi^{(m)}(z)=\Sigma_{n=0}^{\infty}((-1)^{m+1}m!/(z+n)^{m+1})$ is available]. Numerical test teaches us that the other combinations with higher order moments (HOMs) do not show good performance due to large sampling error in HOMs.

The procedure of parameter estimation in the case of dependent noises in the GCP seems to be more confusing. It is desirable if one can start an estimation from the canonical equation in the Y (or X) coordinate with only the natural parameters. Unfortunately, this is not possible since the values of four parameters α, D_p, D_{ap}, and D_a are not given *a priori*. Then one can start from the pdf in the coordinate ξ in Eq. (27) with three parameters A, b, and f since the shifting parameter c can be easily determined at a glance provided that the pdf in the x coordinate is given. Then combining the MLEs, η_1, η_2, and $\eta_3^{(j)}$ $(j=1,2)$ in Eqs. (35), (37), (39), and (41), one can determine values of the three parameters. To determine them, those four equations seem to be redundant. Noticing the relation of the ratio of Eqs. (39) and (41) given in Eq. (42), i.e., $\eta_3^{(2)}/\eta_3^{(1)}=(1-1/b)c$, the parameter b is estimated. Then combing Eqs. (35) and (37), for example, the remaining parameters A and f are estimated. An another procedure with the combination of η_1, η_2, and η_4 in Eq. (46) is possible. The other procedure without use of the digamma function, for example, is also possible by taking $\langle 1/(\xi^2+a^2)\rangle$, $\langle \xi/(\xi^2+a^2)\rangle$, and $\langle 1/(\xi^2+a^2)^2\rangle$.

So there may arise a question: Which is the best choice among many possibilities. In general, there are large sampling errors in the higher order statistical quantities than those in the lower order ones. Then recursive procedure is not avoidable in the GCP in any case. In the iterative procedure to get convergence, the digamma function with complex argument in the form of the series expansion works well. Numerical test also teaches us that the use of the MLEs η_1, η_2, and $\eta_3^{(j)}(j=1,2)$ in Eqs. (35), (37), (39), and (41) seems to be the best choice. To confirm reliability of estimated values of the parameters (A,b,f), Fisher information given in Eq. (47) is available.

D. Relation to Tsallis statistics

Tsallis *et al.*[14] introduced a generalized entropy,

$$S_q[p]=k\frac{1-\int dx[p(x)]^q}{q-1}. \tag{60}$$

Under the constraint [$\phi(x)$ being arbitrary function of x],

$$\int dx\phi(x)p(x)^q=\langle\phi(x)\rangle, \tag{61}$$

one can derive the relevant pdf in the form

$$p_q(x)=[1-\beta(1-q)\phi(x)]^{1/(1-q)}/Z_q, \tag{62}$$

where $Z_q=\int dx[1-\beta(1-q)\phi(x)]^{1/(1-q)}$. The traditional statistics are obtained in the limit of $q\to 1$. There are three different types of the definition of expectation value.[41] The case $\phi(x)=x^2$ with the definition of expectation value in Eq. (61) is the original one due to Tsallis *et al.*[14]

$$\langle x^2\rangle_q\equiv\int d(x/\sigma)x^2[\sigma p_q(x)]^q. \tag{63}$$

Under the definition of this q-deformed statistics, Tsallis and Bukman proposed a nonlinear FP equation $(q\neq 1)$ in the following form:[28]

$$\frac{\partial}{\partial t}[P(x,t)] = -\frac{\partial}{\partial x}\{F(x)[P(x,t)]\} + D\frac{\partial^2}{\partial x^2}[P\ (x,t))]^q, \tag{64}$$

where $\phi(x) = -\int F(x)dx$. Physical and mathematical significance of the nonlinear FP equation has not been shown in conjunction with the incorporation of the multiplicative noise up to now.

With the aid of the MLEs described in Sec. II, one can infer the parameters q and β in Eq. (62). However, there is no clear certification that the temporal profile (the stochastic process) coincides with the observed one derived from the nonlinear FP equation (the corresponding SDE does not exist in the rigorous sense). Since the nature of intermittency and of anomalous diffusion[28] for Eq. (64) is usually quite different from the one with the multiplicative noise though the pdf takes the same form [cf. Eqs. (4) and (62)]. Also, the value at $q=5/3$ may give an additional problem in practical applications.

E. Connection to analysis on intermittent chaos

Jensen *et al.*[26] and Halsay *et al.*[27] proposed a thermodynamic formalism and applied it to characterize various chaotic time series data. The theory is based on the partition function,

$$\chi(q,\ell) \equiv \int d\alpha \ell^{(q-1)\alpha} p(\alpha,\ell), \tag{65}$$

where $p(\alpha,\ell) = \ell^{\alpha - f(\alpha)} p(\overline{\alpha},\ell)$, and the multifractal dimension (MFD) in terms of $\chi(q,\ell)$ defined by

$$D(q) = \frac{1}{q-1}\frac{\ln[\chi(q,\ell)]}{\ln \ell}. \tag{66}$$

The theory of MFD has been successfully applied to characterize fractal natures of chaotic attractors, self-similar structures of time series in the phase space (cf. Table II in Fujisaka and Inoue[24]) for systems of low-dimensional chaos.[25] Various simple models which generate intermittent chaos (types I, II, and III) by chaotic modulation have been studied.[23] However, the related inverse problems for inferring parameters of models with various intermittent nature only from time series data are quite laborious, in general. The intermittent natures in more sophisticated models are not well described both physically and mathematically. Actually, it is known that the statistical mechanical description based on the coarse-grained stochastic processes of fluctuations (e.g., soliton number, soliton amplitude, interdistance between solitons) is still in the qualitative level though it is shown that the pairing of solitons is universal nature. Specifically, in the case of "soliton turbulence," wherein annihilation/creation of solitonlike (or holelike) objects takes place in PDEs, e.g., a driven nonlinear Schrödinger equation,[42,43] complex Ginzburg-Landau equation,[44] Kuramoto-Sivashinsky equation,[45] Benny equation,[46] it is shown that the pdfs of amplitude and phase fluctuations for their PDEs can be successfully described by a class of simple SDEs (Refs. 46 and 42 with multiplicative noise.

F. Availability in a stochastic complex Ginzburg-Landau model

Consider the stochastic Ginzburg-Landau model under the influence of multiplicative noise $F(t)(=F_x(t)+iF_y(t))$ in the complex number,[39,40]

$$\frac{d}{dt}A = (a+i\omega_0)A - (b+ic)|A|^2 A + AF(t). \tag{67}$$

The random force of complex number $F(t)$ represents the existence of both amplitude and frequency modulations: $\langle F_R(t)F_R(t')\rangle = 2D\delta(t-t')$ and $\langle F_\Phi(t)F_\Phi(t')\rangle = 2D\delta(t-t')$. The amplitude R and the phase Φ equation in Eq. (67) after the Wong-Zakai[31] noise correction become

$$\frac{d}{dt}R = (a+D)R - bR^3 + RF_R(t) \tag{68}$$

and

$$\frac{d}{dt}\Phi = \omega_0 - cR^2 + F_\Phi(t). \tag{69}$$

An expedient procedure of parameter estimation for the parameters a, b, and D from the amplitude-pdf $P_s(R)$ is demonstrated in our previous papers.[7,39,40] However, the detailed description of parameter estimation for the pdf of the phase-velocity has not been given in detail.

Since the amplitude equation in Eq. (68) can be rewritten in the closed form with only R^2. Therefore, R^2 can be eliminated with the use of R^2 in Eq. (69). Thus, the equation of the phase-velocity $v_\phi = d\phi/dt$ ($\Phi \equiv \phi + \omega_0$) reduces exactly to

$$\dot{v}_\phi = 2\{a + F_p(t)\}v_\phi + \frac{2b}{c}v_\phi^2 + F_a(t), \tag{70}$$

where

$$F_p(t) = F_R(t) - \frac{2b}{c}F_\Phi(t) \tag{71}$$

and

$$F_a(t) = -\dot{F}_\Phi(t) + 2aF_\Phi(t) - \frac{2b}{c}(F_\Phi(t))^2 + 2F_\Phi(t)F_R(t). \tag{72}$$

It is clear that Eq. (70) with the parametric (71) and the additive noise (72) may be relevant to the Cauchy process with the correlated noises. At least, under the assumptions that $a>0$ and $b>0$ and $c<0$, $av_\phi + (2b/c)v_\phi^2$ can be replaced by the effective linear damping $-\gamma v_\phi$ with $\gamma>0$, $D_p \approx 4D(1+4b^2/c^2)$, $D_{ap} \approx 8abD/c$, and $D_p \approx 4a^2D$ after neglecting the higher order correlations of the noise sources $F_R(t)$, $F_\Phi(t)$, and $\dot{F}_\Phi(t)$.

The detailed analysis of the model in conjunction with real electroencephalography signals will be published elsewhere.

V. SUMMARY AND REMARKS

We have studied the universal statistics of a GCP in Eq. (1) with two independent/dependent random forces with the aid of the method of information geometry in statistics.[34–37] The geometrical concept becomes more useful in the detailed analysis of statistical inferences based on the Fisher information matrix obtained in this paper.

In the case of independent random forces, we have derived the three universal statistics, the first, the second, and the third order cumulant of $\ln(x^2+a^2)$ for the natural parameter b. The universal statistics for the non-natural parameter a are the moments of $1/(x^2+a^2)$. Fisher information matrix is also obtained. It is shown that (i) they are free from the divergence; (ii) the second and the third order cumulant depend only on the parameter b; (iii) the availability for estimating the parameters with combining the MLEs $\eta_1 = \langle \ln(x^2+a^2) \rangle$, $\eta_2 = \langle 1/(x^2+a^2) \rangle$, and the cross cumulant $\langle \ln(x^2+a^2)/(x^2+a^2) \rangle$ is also discussed in Sec. IV C.

Then, in the case of dependent random forces, we have derived also all the universal statistic quantities (the dual coordinate[34,35]); (η_1:) the first moments of $\ln(\xi^2+A^2)$, (η_2:)$\arctan(\xi/A)$, (η_3:)$1/(\xi^2+A^2)$, and $\xi/(\xi^2+A^2)$. It is also obtained all the matrix elements of Fisher information including (η_4:) the mutual correlation between $\ln(\xi^2+A^2)$ and $\arctan(\xi/A)$. It is further shown that (i) nondivergent natures of them and (ii) feasible procedures of estimating parameters with the MLEs and Fisher information are discussed in Sec. IV C.

Discussions are also given for physical significance of dual coordinate, physical relation to FDT, relation to Tsallis statistics, connection to analysis on intermittent chaos, and relevance to the phase-velocity distribution in a stochastic complex Ginzburg-Landau system in conjunction with the GCP.[47–50] The universal statistics and Fisher information derived in this paper might be useful to analyze economic time series data when $F_a(t)$ and $F_p(t)$ can be modeled by Gaussian-white noises.

The MLEs under the influence of colored noises for complex nonlinear dynamical systems with a cubic nonlinear term such as nuclear power reactors and electroencephalography/magnetoencephalography in human brains are under investigation based on the GCP in the canonical form in Eq. (50) [or Eq. (54)]. Their results will be published elsewhere.

ACKNOWLEDGMENTS

This work was partially supported by JSPS for Scientific Research (c), No. 16500169. We are greatly benefited from the comments from referees to improve the manuscript.

APPENDIX A: INTEGRAL FORMULA USED TO DERIVE EQ. (14)

The mth order moments are defined by $I_m(a,b)=P_0K_m(a,b), K_m(a,b)\equiv\int_{-\infty}^{\infty}[\ln(x^2+a^2)]^m(x^2+a^2)^{-b}dx (m=1,2,\ldots)$ in Eq. (14), where P_0 is the normalization constant. One should note that $K_m(a,b)$ of $m=0$ is just given in terms of beta function $B(\mu,\nu)$ in the form

$$K_0(a,b)=\int_{-\infty}^{\infty}(x^2+a^2)^{-b}dx=a^{1-2b}B(b-1/2,1/2)\equiv P_0^{-1} \tag{A1}$$

for real constants of $a>0$ and $b>1/2$. Thus P_0^{-1} is expressed as $P_0^{-1}=\exp(\ln B(b-1/2,1/2)-(2b-1)\ln a)=\exp(\ln\Gamma(1/2)+\ln\Gamma(b-1/2)-\ln\Gamma(b)-(2b-1)\ln a)$. By noticing the recursion relation $K_m(a,b)=-\partial K_{m-1}(a,b)/\partial b$, one obtains for $m=1$ as

$$K_1(a,b)=\int_{-\infty}^{\infty}\ln(x^2+a^2)(x^2+a^2)^{-b}dx=-\frac{\partial}{\partial b}P_0^{-1}=(\psi(b)-\psi(b-1/2)+2\ln a)P_0^{-1}.$$

Hence we have

$$I_1(a,b)=P_0K_1(a,b)=[\psi(b)-\psi(b-1/2)+2\ln a]. \tag{A2}$$

The recursion relation for $K_m(a,b)$ is applied for $m=2$ and 3. The results are given by

$$K_2(a,b)=\int_{-\infty}^{\infty}[\ln(x^2+a^2)]^2(x^2+a^2)^{-b}dx=-\frac{\partial}{\partial b}K_1(a,b),$$

$$I_2(a,b)=P_0K_2(a,b)=\{[\psi(b)-\psi(b-1/2)+2\ln a]^2-[\psi'(b)-\psi'(b-1/2)]\}, \tag{A3}$$

and

$$K_3(a,b)=\int_{-\infty}^{\infty}[\ln(x^2+a^2)]^3(x^2+a^2)^{-b}dx=-\frac{\partial}{\partial b}K_2(a,b),$$

$$\begin{aligned}I_3(a,b)=P_0K_3(a,b)=\{&[\psi(b)-\psi(b-1/2)+2\ln a]^3\\&-3(\psi'(b)-\psi'(b-1/2))[\psi(b)-\psi(b-1/2)+2\ln a]+\psi''(b)-\psi''(b-1/2)\}.\end{aligned} \tag{A4}$$

APPENDIX B: DERIVATION OF FISHER INFORMATION MATRIX IN EQ. (47)

The matrix elements of Fisher information $g(A,b,f)$ for the two natural parameters b and f are obtained by taking the derivative of the IGP, i.e., $g_{ij}=\partial_i\partial_j\Psi(A,b,f)$ $(i,j=1,2;\ \partial_1=\partial/\partial b$ and $\partial_2=\partial/\partial f)$.

For the non-natural parameters, one needs to evaluate them directly from $g(A,b,f)=-\langle\partial_i\partial_j\ell(A,b,f)\rangle$, $\ell=\ln P_s(\xi;A,b,f)$. Specifically,

$$\left\langle\frac{\partial^2\ell}{\partial A^2}\right\rangle=-\frac{2b-1}{A^2}-2b\left\langle\frac{1}{\xi^2+A^2}\right\rangle+4A^2b\left\langle\frac{1}{(\xi^2+A^2)^2}\right\rangle+2Af\left\langle\frac{\xi}{(\xi^2+A^2)^2}\right\rangle, \tag{B1}$$

$$\left\langle\frac{\partial^2\ell}{\partial A\partial b}\right\rangle=\frac{2}{A}-2A\left\langle\frac{1}{\xi^2+A^2}\right\rangle, \tag{B2}$$

and

$$\left\langle\frac{\partial^2\ell}{\partial A\partial f}\right\rangle=-\left\langle\frac{\xi}{(\xi^2+A^2)^2}\right\rangle. \tag{B3}$$

The statistic $\langle 1/(\xi^2+A^2)\rangle$ is given in Eq. (39). The statistic $\langle 1/(\xi^2+A^2)^2\rangle$ is evaluated by the integral formula $N_0(A,b,f)$ defined in Eq. (40) as

$$\left\langle\frac{1}{(\xi^2+A^2)^2}\right\rangle=\frac{N_0(A,b+2,f)}{N_0(A,b,f)}=\frac{(b+1)(b+1/2)b(b-1/2)}{A^4(b^2+f^2/4)[(b+1)^2+f^2/4]}. \tag{B4}$$

Therefore, when $\langle\xi/(\xi^2+A^2)^2\rangle$ is given, all the matrix elements are determined. Differentiate Eq. (40) twice with respect to A, we have

$$(f^2-2b)\left\langle\frac{1}{\xi^2+A^2}\right\rangle+(4b+4b^2-f^2)A^2\left\langle\frac{1}{(\xi^2+A^2)^2}\right\rangle+(2Af+4Afb)\left\langle\frac{\xi}{(\xi^2+A^2)^2}\right\rangle=\frac{2b(2b-1)}{A^2} \tag{B5}$$

Inserting Eqs. (39) and (B4) into Ee. (B5), we have

$$\left\langle\frac{\xi}{(\xi^2+A^2)^2}\right\rangle=\frac{(b-1/2)b(b+1/2)f}{2A^3(b^2+f^2/4)[(b+1)^2+f^2/4]}. \tag{B6}$$

Thus, inserting Eqs. (39), (41), (B4), and (B6) into Eqs. (B1)–(B3), all the matrix elements are obtained, as described in Eq. (47).

[1] B. Alder and T. E. Wainwright, Phys. Rev. Lett. **18**, 988 (1967).
[2] A. Schenzle and H. Brand, Phys. Rev. A **20**, 1628 (1979).
[3] Y. Hamada, Prog. Theor. Phys. **64**, 850 (1981).
[4] M. Suzuki, K. Kaneko, and S. Takesue, Prog. Theor. Phys. **67**, 1756 (1982).
[5] E. Wong, *Stochastic Processes in Information and Dynamical Systems* (McGraw-Hill, New York, 1971).
[6] T. Ozaki, *Handbook of Statistics* (North-Holland, Amsterdam, 1985), Vol. 5, pp. 25–83.
[7] H. Konno and P. S. Lomdahl, J. Phys. Soc. Jpn. **73**, 573 (2004).
[8] J. M. Deutsch, Physica A **208**, 433 (1994).
[9] S. C. Venkataramani, T. M. Antonsen, Jr., E. Ott, and J. C. Sommerer, Physica D **96**, 66 (1996).
[10] H. Takayasu, A.-H. Sato, and M. Takayasu, Phys. Rev. Lett. **79**, 966 (1997).
[11] M. Levy and S. Solomon, Int. J. Mod. Phys. C **7**, 595 (1996).
[12] H. Nakao, Phys. Rev. E **58**, 1591 (1998).
[13] H. Konno, S. Kanamoto, and Y. Takeuch, Prog. Nucl. Energy **43**, 201 (2003).
[14] C. Tsallis, S. V. F. Levy, A. M. C. Souza, and R. Maynard, Phys. Rev. Lett. **75**, 3589 (1995).
[15] M. F. Shlesinger and B. D. Hughes, Physica A **109**, 597 (1981); E. W. Montroll and M. F. Shlesinger, J. Stat. Phys. **32**, 209 (1983); *Nonequilibrium Phenomena II: From Stochastic to Hydrodynamics*, edited by J. L. Lebowitz and E. W. Montroll (North-Holland, Amsterdam, 1984).
[16] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
[17] T. H. Solomon, E. R. Weeks, and H. L. Swinney, Phys. Rev. Lett. **71**, 3975 (1993).

[18] C. K. Peng, J. Mietus, J. M. Housdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger, Phys. Rev. Lett. **70**, 1343 (1993).
[19] T. J. P. Penna, Phys. Rev. E **51**, R (1995).
[20] Y. H. Taguchi and H. Takayasu, Europhys. Lett. **30**, 499 (1995).
[21] R. Frierich and J. Pinke, Physica D **102**, 147 (1997).
[22] K. Kiyono, Z. R. Struzik, N. Aoyagi, S. Sakata, J. Hayano, and Y. Yamamoto, Phys. Rev. Lett. **93**, 178103 (2004).
[23] H. Fujisaka and T. Yamada, Prog. Theor. Phys. **77**, 1045 (1987).
[24] H. Fujisaka and M. Inoue, Prog. Theor. Phys. **77**, 1334 (1987).
[25] H. Mori and Y. Kuramoto, *Dissipative Structures and Chaos* (Springer-Verlag, Berlin, 2001).
[26] M. H. Jensen, L. P. Kadanoff, and A. Libchaber, Phys. Rev. Lett. **55**, 2798 (1985).
[27] T. C. Halsay, L. P. Kadanoff, I. Procaccia and B. I. Shraiman, Phys. Rev. A **33**, 1141 (1986).
[28] C. Tsallis, D. J. Bukman, Phys. Rev. E **54**, R2197 (1996).
[29] H. Mori, Prog. Theor. Phys. **33**, 283 (1965).
[30] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).
[31] E. Wong and M. Zakai, Ann. Math. Stat. **36**, 1560 (1965).
[32] G. Kitagawa, J. Am. Stat. Assoc. **93**, 1023 (1998).
[33] T. Higuchi and G. Kitagawa, IEICE Trans. Inf. Syst. **E83-D**, 36 (2000).
[34] S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics Vol. 28 (Springer-Verlag, Berlin, 1985).
[35] S. Amari and H. Nagaoka, *Methods of Information Geometry* (AMS, New York, 2000).
[36] M. K. Murrey and J. M. Rice, *Differential Geometry and Statistics* (Chapman, New York, 1993).
[37] R. E. Kass and P. Vos, *Geometrical Foundations of Asymptotic Inference* (Wiley, New York, 1997).
[38] Y. Okabe, Commun. Math. Phys. **98**, 449 (1985).
[39] F. Watanabe and H. Konno, Ann. Nucl. Energy **30**, 1267 (2003).
[40] F. Watanabe and H. Konno, Ann. Nucl. Energy **31**, 375 (2004).
[41] M. Shiino, J. Math. Phys. **42**, 2540 (2001).
[42] H. Konno and P. S. Lomdahl, Phys. Lett. A **193**, 185 (1994).
[43] K. Nozaki and N. Bekki, Physica D **21**, 381 (1986).
[44] H. Shate, Nonlinearity **7**, 185 (1994).
[45] S. Toh, J. Phys. Soc. Jpn. **56**, 949 (1987).
[46] H. Konno and P. S. Lomdahl, J. Phys. Soc. Jpn. **69**, 1629 (2000).
[47] L. Arnold, *Stochastic Differential Equations: Theory and Applications* (Wiley, New York, 1974).
[48] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
[49] R. A. Ibrahim, *Parametric Random Vibration* (Research Studies, Letchworth, 1985).
[50] W. Horsthemke and R. Lefever, *Noise-Induced Transitions* (Springer, Berlin, 1984).