Information geometry and synchronization phase transition in the Kuramoto model

Artem Alexandrov [*]
Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia
and Laboratory of Complex Networks, Brain and Consciousness Research Center, Moscow, Russia

Alexander Gorsky
Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia;
Institute for Information Transmission Problems, Moscow 127994, Russia;
and Laboratory of Complex Networks, Brain and Consciousness Research Center, Moscow, Russia

(Received 24 December 2022; accepted 10 April 2023; published 24 April 2023)

We discuss how the synchronization in the Kuramoto model can be treated in terms of information geometry. We argue that the Fisher information is sensitive to synchronization transition; specifically, components of the Fisher metric diverge at the critical point. Our approach is based on the recently proposed relation between the Kuramoto model and geodesics in hyperbolic space.

DOI: 10.1103/PhysRevE.107.044211

I. INTRODUCTION

The Kuramoto model is the paradigmatic model for synchronization phenomena in nonlinear systems. Despite its deceptively simple form, there are still enough white spots, concerning different aspects of the model. Some of these aspects turn around the synchronization in systems with complicated topology, which were comprehensively described in Ref. [1]. Other aspects are mostly focused on the rigorous description of the Kuramoto model with all-to-all couplings (complete graph case). It was shown that in the case of identical oscillators the Kuramoto model exhibits low-dimensional dynamics. The emergent dimensional reduction occurs due to existence of integrals of motion, initially discovered by trial and error in Refs. [2,3]. The investigation of such integrals of motion was done comprehensively in Ref. [4], where it was demonstrated that they emerge due to the invariance of dynamics under the Möbius transformation. The Kuramoto model dynamics can be represented as the motion on the Möbius group orbit. Moreover, recently it was noticed that the corresponding low-dimensional dynamics can be described in terms of gradient flows on a two-dimensional hyperbolic manifold [5].

However, the origin of the hyperbolic manifold was not established. At first glance, this manifold simply inherits symmetries of the Möbius group, and there are no special features. We shall interpret the hyperbolic manifold as being induced by probability measures, which is usually referred to as a statistical manifold (see Ref. [6] for a review). In particular, the AdS_2 metric emerges immediately for the shifted Gaussian distribution. In a generic situation the hyperbolic metric emerges as the Fisher metric on the parameter space. The Fisher metric is the probabilistic counterpart of the quantum metric which is familiar in quantum mechanics, where one introduces the so-called complex quantum tensor which has the quantum metric as the real part and the Berry curvature as the imaginary part [7,8]. It is built from the quantum wave function depending on some parameter space; however, its classical analog exists as well.

In this paper we focus on the hyperbolic geometry description of the Kuramoto model and its connection to information geometry. In particular, we explain the hyperbolic geometry involved as the geometry of the statistical manifold. We shall argue that the equation of motion for the Kuramoto model has the form of gradient flow for the Kullback-Leibler divergence. The properties of the induced metric or, in other words, the response of the system to the perturbation of parameters serve as an indicator of the phase transition [9–12]. In particular, the diagonal elements of the quantum metric provide a good tool in the search for critical curves, which has been demonstrated in several examples [13–17]. We shall demonstrate in a clear-cut manner that the components of the Fisher information metric identify in a similar manner the critical curve for the phase transition in the Kuramoto-Sakaguchi model. Hence the Fisher information metric yields the proper order parameter for the synchronization phase transition.

The paper is organized as follows. In Sec. II we recall the Kuramoto model, focusing on its description in terms of the Möbius group. The dimensional reduction provides the description of the model in terms of motion of a point at a hyperbolic disk. In Sec. III we shall present the key aspects of the information geometry. In Sec. IV we apply the information geometry to the Kuramoto and Kuramoto-Sakaguchi models. We shall treat the coordinates of a point on the disk as the two parameters which the distribution involved depends on. It turns out that the Cauchy distribution governs the Kuramoto model while the von Mises distribution is relevant for the Kuramoto-Sakaguchi model. In both cases we shall demonstrate that the singularity of the Fisher metric coincides

[*]aleksandrov.aa@phystech.edu

with the synchronization transition, adding an example of the relation between the singularities of the Fisher metric and the phase transitions. In Sec. V we formulate the results and the open questions. General comments concerning the relations between the distributions and the possible role of the q-Gaussian distributions in the Kuramoto model can be found in Appendixes A and B.

II. BRIEF REVIEW OF THE KURAMOTO MODEL

The Kuramoto model on a complete graph with identical oscillators has the following equations of motion:

$$\dot{\theta}_i = \omega + \frac{\lambda}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \tag{1}$$

where N is the number of oscillators, $\theta_i \in [0, 2\pi)$, $\lambda > 0$ is the coupling constant, and ω is the oscillator eigenfrequency. The order parameter (synchronization measure) is defined by

$$r(t) = \frac{1}{N}\sum_{j=1}^{N} z_j(t), \tag{2}$$

where we introduce the complex variable $z_i(t) = \exp(i\theta_i(t))$. The model exhibits so-called *synchronization transition* from a given initial state to a synchronized state with the amplitude of order parameter $|r| \to 1$, which does not depend on time. In terms of particles on the unit circle, this means that in the synchronized state all the particles are concentrated at one point that moves with a certain average frequency.

The equations of motion become

$$\dot{z}_i = i\omega z_i + \frac{\lambda}{2}\left(r - \overline{r}z_i^2\right). \tag{3}$$

The evolution of $z_i = z_i(t)$ can be represented as the action of the Möbius transformation,

$$z_i(t) = \mathcal{M}\left(z_i^0\right), \quad z_i^0 \equiv \exp(i\theta_i(t=0)), \tag{4}$$

where the parameters of the Möbius transformation depend on time. Following Ref. [5], this transformation can be parametrized as follows:

$$\mathcal{M}(z) = \zeta\frac{z - w}{1 - \overline{w}z}, \tag{5}$$

where $\zeta \in \mathbb{C}$, $|\zeta| = 1$, $w \in \mathbb{C}$, and $|w| < 1$. The time derivative of the phase is

$$\dot{z}_i = -\frac{\dot{w}\zeta}{1 - |w|^2} + \left(\dot{\zeta}\overline{\zeta} + \frac{\dot{\overline{w}}w - \dot{w}\overline{w}}{1 - |w|^2}\right)z_i + \frac{\dot{\overline{w}}\zeta}{1 - |w|^2}z_i^2. \tag{6}$$

Matching the right-hand side of Eq. (6) and the right-hand side of Eq. (3), we obtain the system of equations

$$\dot{w} = -\frac{\lambda}{2}(1 - |w|^2)\overline{\zeta}r, \quad \dot{\zeta} = i\omega\zeta - \frac{\lambda}{2}(\overline{w}r - w\overline{r}\zeta^2). \tag{7}$$

The key feature of these equations is that they are decoupled,

$$\overline{\zeta}r = \overline{\zeta}\frac{1}{N}\sum_{j=1}^{N}\zeta\frac{z_j^0 - w}{1 - \overline{w}z_j^0} = |\zeta|^2\frac{1}{N}\sum_{j=1}^{N}\frac{z_j^0 - w}{1 - \overline{w}z_j^0} = \frac{1}{N}\sum_{j=1}^{N}\frac{z_j^0 - w}{1 - \overline{w}z_j^0}; \tag{8}$$

hence this means that the equation for w does not feel the variable ζ. This equation is our object of interest, and it was shown previously [5] that it can be rewritten as the gradient flow on the hyperbolic disk $\mathbb{D} = \{z : |z| < 1\}$,

$$\dot{w} = -\frac{\lambda}{2}(1 - |w|^2)^2\frac{\partial}{\partial\overline{w}}\left\{\frac{1}{N}\sum_{j=1}^{N}\ln P\left(w, z_j^0\right)\right\}, \tag{9}$$

where $P = P(w, z_0)$ is the Poisson kernel,

$$P(w, z_0) = \frac{1 - |w|^2}{|w - z_0|^2}, \quad z_0 \in \mathbb{S}^1, \quad w \in \mathbb{D}, \quad \mathbb{S}^1 = \{z : |z| = 1\}. \tag{10}$$

To determine how the Poisson kernel appears, one should substitute expression (8) into Eq. (7) and perform integration over $\overline{w}$. The authors of Ref. [5] have already noticed that for $|w| \neq 1$ the fixed point of Eq. (9) coincides with the conformal barycenter of N points on $\mathbb{S}^1$, whose existence and uniqueness are guaranteed [18]. Also, one should note that the $|w| \to 1$ limit corresponds to the synchronized state.

Using this representation of the Kuramoto model dynamics, we would like to discuss several features of the model that are devoted to the interconnections between information geometry and synchronization.

III. INFORMATION GEOMETRY AND THE KURAMOTO MODEL

A. Basic concepts of information geometry

In this section we discuss the basic concepts of information geometry and develop an information geometry view of the Kuramoto model. We mainly focus on the continuum $N \to \infty$ limit of the Kuramoto model.

To begin with, let us provide some essentials from information geometry. We follow Ref. [6], where more details can be found. Let us consider the family of probability density functions $p = p(\xi; x)$, where ξ is the vector of parameters and x is the vector of variables. For all possible values of ξ, this family forms the manifold $\mathbb{M} = \{p(\xi; x)\}$, which is called the statistical manifold. To determine how two distributions $p(\xi_1; x)$ and $p(\xi_2; x)$ differ from each other, one can consider the divergence function $D[\xi_1; \xi_2]$. It is possible to define different divergence functions, but these functions make sense only if they are invariant and decomposable [19].

The large class of such divergence functions is called the standard f-divergences, which can be represented as

$$D_f[p; q] = \int dx\, p(x) f\left(\frac{q(x)}{p(x)}\right), \quad f(0) = 1, \quad f(1) = 1, \quad f''(1) = 1, \tag{11}$$

where f is the convex function. For $f(v) = -\ln v$ it is easy to see that the corresponding f-divergence is nothing more than Kullback-Leibler divergence. Another important type of divergence is the so-called α-divergence, which is defined by (see Refs. [6,19] for details)

$$f_\alpha(v) = \frac{4}{1-\alpha^2}[1 - v^{(1+\alpha)/2}], \quad \alpha \neq 1. \tag{12}$$

The important property of the divergence function is that in the case of two close enough points, $\xi_1 = \xi + d\xi$ and $\xi_2 = \xi$, it has the following Taylor expansion at point ξ:

$$D[\xi_1; \xi_2] = \tfrac{1}{2} g_{ij}(\xi) d\xi_i d\xi_j + O(|d\xi|^3), \tag{13}$$

where g_{ij} is the metric of manifold $\mathbb{M}$. Any standard f-divergence gives the same metric g_{ij}, which coincides with the Fisher metric g^F_{ij},

$$g^{\mathrm{F}}_{ij} = \int dx\, p(x;\xi) \frac{\partial \ln p(x;\xi)}{\partial \xi_i} \frac{\partial \ln p(x;\xi)}{\partial \xi_j}. \tag{14}$$

Strictly speaking, to completely describe the statistical manifold $\mathbb{M}$, we need one more object, which is called the skewness tensor T^{F},

$$T^{\mathrm{F}}_{ijk} = \int dx\, p(x;\xi) \frac{\partial \ln p(x;\xi)}{\partial \xi_i} \frac{\partial \ln p(x;\xi)}{\partial \xi_j} \frac{\partial \ln p(x;\xi)}{\partial \xi_k}. \tag{15}$$

Finally, the statistical manifold is the triplet $(\mathbb{M}, g, T)$. For a given standard f-divergence, the metric tensor g (obtained from the Taylor expansion) coincides with g^{F}, whereas the corresponding skewness tensor is given by $T = \alpha T^{\mathrm{F}}$ with $\alpha = 2f'''(1) + 3$.

The pair of g^{F} and T^{F} allows us to define so-called α-connections on the statistical manifold, which are given by

$$\Gamma^{\pm\alpha}_{ijk} = \Gamma^0_{ijk} - \frac{\alpha}{2} T^{\mathrm{F}}_{ijk}. \tag{16}$$

The statistical manifold is called α-flat if the corresponding α-Riemannian curvature tensor vanishes everywhere with condition $R^{\alpha}_{ijkl} = R^{-\alpha}_{ijkl}$. Note that in contrast with the usual Riemannian manifold, the statistical manifold can have nonzero torsion. It was shown that such α-flat manifolds have dual affine structure, which gives a dual (via Legendre transformation) coordinate system on the statistical manifold. Such a manifold is also called a dual flat manifold. For a dual flat manifold, the metric and skewness tensors are given by

$$g^{\mathrm{F}}_{ij} = \partial_i \partial_j \psi(\xi), \quad T^{\mathrm{F}}_{ijk} = \partial_i \partial_j \partial_k \psi(\xi), \quad \partial_i \equiv \frac{\partial}{\partial \xi_i}, \tag{17}$$

where $\psi(\xi)$ is a convex function. For a dual flat manifold, the canonical divergence is the so-called the Bregman divergence.

B. Gradient flow on a statistical manifold

For a dual flat manifold, the existence of $\psi(\xi)$ allows us to consider the gradient flow with respect to the Fisher metric [20,21],

$$\frac{d\xi}{dt} = -(g^F)^{-1} \partial_\xi \psi(\xi). \tag{18}$$

Some examples of gradient flows on statistical manifolds are discussed in Ref. [20].

The simplest illustrative example of the mentioned gradient flow corresponds to the Gaussian distribution statistical manifold. It is straightforward to check that the statistical manifold formed by Gaussian distributions, $\mathbb{M}_{\mathrm{G}} = \{p(\mu, \sigma; x) | \mu \in \mathbb{R},\ \sigma > 0\}$, is α-flat for $\alpha = \pm 1$. One can also generalize this statement to the exponential distribution family,

$$p(\xi; x) = \exp\left[\xi_i g_i(x) + h(x) - \psi(\xi)\right], \tag{19}$$

which forms a dual flat manifold with canonical coordinate system ξ and the function $\psi = \psi(\xi)$. In the case of a Gaussian distribution, the coordinate system is (ξ_1, ξ_2) with $\xi_1 = \mu/\sigma^2$, $\xi_2 = -1/(2\sigma^2)$ and $\psi = \mu^2/(2\sigma^2) + \ln(\sqrt{2\pi}\sigma)$. In the case of the exponential distributions family the Bregman divergence coincides with the Kullback-Leibler divergence. The gradient flow for the Gaussian statistical manifold looks like

$$\dot{\mu} = -\mu; \quad \dot{\sigma} = -\frac{\sigma^2 + \mu^2}{2\sigma}. \tag{20}$$

Note that this system has the conserved integral of motion $H = \sigma^2/\mu - \mu$, which corresponds to the arc radius.

C. Cauchy distribution

One can show that the family of univariate elliptic distributions forms α-flat statistical manifolds [22]. However, there is one orphan distribution in this family: the Cauchy distribution,

$$p_{\mathrm{C}}(\gamma, \beta; x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \beta)^2}, \quad \gamma > 0,\ \beta \in \mathbb{R}, \tag{21}$$

which can be written in the more-convenient-for-our-goals McCullagh's representation [23],

$$p_{\mathrm{C}}(w; x) = \frac{\operatorname{Im} w}{\pi |x - w|^2}, \quad w = \beta + i\gamma. \tag{22}$$

The corresponding Fisher metric looks like

$$g^{\mathrm{F}}_{ij} = \frac{1}{2\gamma^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{23}$$

It is straightforward to verify that for the Cauchy distribution the skewness tensor T^{F} vanishes everywhere; so we cannot find any value of α to obtain a dual flat structure at first glance.

However, we can interpret the Cauchy distribution as a member of the q-Gaussian family, whose probability density function is defined as

$$\mathcal{N}_q(\beta, x) = \frac{\sqrt{\beta}}{C_q} \exp_q\{-\beta x^2\}, \quad \beta > 0, \tag{24}$$

where C_q is the normalization constant. The Cauchy distribution can be considered as a q-Gaussian with $q = 2$, which was done in Ref. [24] (we discuss some properties of q-Gaussians in Appendix B),

$$p_{\mathrm{C}}(\gamma, \beta; x) = \mathcal{N}_q(\gamma^{-2}, (x - \beta)^2)|_{q=2}. \tag{25}$$

Using the definition of the q-exponent, it is straightforward to verify that the distribution function $p_{\mathrm{C}}(\gamma, \beta; x)$ can be also represented as

$$p_{\mathrm{C}}(\gamma, \beta; x) = \exp_q\left(\xi_i g_i(x) - \psi(\xi)\right)|_{q=2}, \tag{26}$$

where $\xi_1 = 2\pi\beta/\gamma$, $\xi_2 = -\pi/\gamma$, $g_1(x) = x$, and $g_2(x) = x^2$ and the potential function $\psi(\xi)$ is

$$\psi(\xi) = -\frac{\pi^2}{\xi_2} - \frac{\xi_1^2}{4\xi_2} - 1. \tag{27}$$

Next, it was shown that the divergence

$$D_{\text{B-T}}[\xi_1;\xi_2] = \left(\int dx\, p(\xi_2;x)^2\right)^{-1}\left[\int dx\,\frac{p(\xi_2;x)^2}{p(\xi_1;x)} - 1\right] \tag{28}$$

is the Bregman divergence, i.e., describes the dual flat manifold. Such a divergence is called a Bregman-Tsallis divergence. The metric corresponding to the Bregman-Tsallis divergence does not coincide with the Fisher metric and can be computed as

$$g^q_{ij} = \frac{\partial^2\psi(\xi)}{\partial\xi_i\partial\xi_j}, \tag{29}$$

where ξ corresponds to the canonical coordinate system on the dual flat manifold. In the case of Cauchy distributions, we have

$$g^q_{ij} = \frac{2\pi}{\gamma}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \tag{30}$$

and it is clear that g^q and $g^{\rm F}_{\rm C}$ are related via conformal transformation. The dual coordinates η_i are defined by

$$\eta_i = \frac{\partial\psi(\xi)}{\partial\xi_i} = \left(-\frac{\xi_1}{2\xi_2}, \frac{4\pi^2+\xi_1^2}{4\xi_2^2}\right) = (\beta,\ \gamma^2+\beta^2) \tag{31}$$

with corresponding dual potential

$$\phi(\eta) = 1 - 2\pi\sqrt{\eta_2 - \eta_1^2}. \tag{32}$$

The gradient flow (18) for the Cauchy distribution looks like

$$\begin{cases}\dot\xi_1 = \frac{\xi_1}{8\pi^2}(4\pi^2+\xi_1^2)\\ \dot\xi_2 = \frac{\xi_2}{8\pi^2}(4\pi^2-\xi_1^2)\end{cases} \leftrightarrow \begin{cases}\dot\beta = -\beta\\ \dot\gamma = \frac{\beta^2-\gamma^2}{2\gamma}.\end{cases} \tag{33}$$

Note that these equations are quite similar to the Gaussian case, Eq. (20). Having introduced all relevant concepts from information geometry, we can turn to the interpretation of the Kuramoto model in terms of the information geometry.

IV. FISHER METRIC AS THE ORDER PARAMETER FOR SYNCHRONIZATION TRANSITION

A. Fisher metric and phase transitions

The quantum tensor is now an effective tool for analysis of the topological and critical phenomena in complicated many-body systems. It is defined as

$$Q_{ij} = \langle\partial_i\Psi|\partial_j\Psi\rangle - \langle\partial_i\Psi|\Psi\rangle\langle\Psi|\partial_j\Psi\rangle, \tag{34}$$

where the i, j indices correspond to the coordinates in the parameter space. Its real part is the quantum metric, while its imaginary part is the Berry curvature [7,8]

$$Q_{ij} = g_{ij} + iF_{ij}. \tag{35}$$

The quantum metric can be thought of as a two-point correlator or fidelity whose behavior is expected to quantify the properties of the system [9–11]. This generic argument can be made more precise if we focus on the geometry of the metric, and it was argued [9,10] that the singularity of the metric or Ricci curvature corresponds to the position of the phase transition. The kind of classification of induced geometries attributed to the ground state can be found in Ref. [13]. The relation between the singularities of the metric and the phase transitions is not a completely rigorous statement, but there is a convincing list of examples, say, Refs. [16,17] for the Dicke and Hubbard models. A review of this subject can be found in Ref. [14]. The parameters of the Hamiltonian or the momenta can be used to evaluate the components of the quantum metric. More recently, it was recognized that the behavior of the metric near the singular point can distinguish the integrable or chaotic behavior [15]. The singularities of the different components of the metric provide information concerning the criticality in the different directions of the parameter space.

The Fisher information metric is parallel to the quantum metric in quantum mechanics when the probability substitutes the squared modulus of the wave function. The probability obeys the Fokker-Planck (FP) equation instead of the Schrödinger one; hence we shall focus below on solutions to FP equations. The notion of the phase transition in the stochastic problems also requires some care; however, the relation between the quantum and semiclassical criticalities is seen in the temperature component of the information metric [25]. It turns out that the synchronization phase transition provides another clear-cut example when the behavior of the Fisher metric yields the identification of the phase transition in the classical system. In this section we shall demonstrate that the representation of the Kuramoto model in terms of the distributions does the job. In the pure Kuramoto model we shall see a kind of classical version of the quantum phase transition at zero temperature, while in the Kuramoto-Sakaguchi model we have an effective temperature due to noise.

B. Fisher metric in the Kuramoto model

The authors of Ref. [5] have introduced the hyperbolic description of the Kuramoto model. We argue that the hyperbolic space arises as the statistical manifold of wrapped Cauchy distributions. The normalized version of the Poisson kernel is given by

$$p_{\rm wC}(w;z) = \frac{1}{2\pi}\frac{1-|w|^2}{|w-z|^2}, \quad |z| = 1,$$
$$\oint_{|z|=1} dz\,\frac{p_{\rm wC}(w,z)}{iz} = 1, \tag{36}$$

where wC denotes "wrapped Cauchy." It is convenient to use a polar representation, $w = re^{i\phi}$ and $z = e^{i\theta}$, which gives

$$p_{\rm wC}(r,\phi;\theta) = \frac{1}{2\pi}\frac{1-r^2}{1-2r\cos(\theta-\phi)+r^2},$$
$$0 \leqslant r < 1,\ \phi = \phi \bmod 2\pi. \tag{37}$$

This is a probability density function (PDF) with two parameters, r and ϕ. All such PDFs form a manifold $\mathbb{M}_{\rm wC}$,

$$\mathbb{M}_{\rm wC} = \{p_{\rm wC}(r,\phi;\theta)\,|\,r\in[0,1), \phi = \phi \bmod 2\pi\},$$

and the Fisher metric g_{ij}^{F} is given by (14), which can be represented in more compact form,

$$g_{ij}^{\mathrm{F}} = 4\int_{-\pi}^{+\pi} d\theta\, \partial_i \sqrt{p_{\mathrm{wC}}(r,\phi;\theta)}\partial_j\sqrt{p_{\mathrm{wC}}(r,\phi;\theta)}.$$

It is straightforward to compute this integral and obtain components of the metric,

$$g_{rr}^{\mathrm{F}} = \frac{2}{(1-r^2)^2}; \quad g_{\phi\phi}^{\mathrm{F}} = \frac{2r^2}{(1-r^2)^2}; \quad g_{r\phi}^{\mathrm{F}} = g_{\phi r}^{\mathrm{F}} = 0. \tag{38}$$

Hence, reintroducing $w = re^{i\phi}$, we can write

$$ds_{\mathrm{F}}^2 = \frac{2(dr^2 + r^2 d\phi^2)}{(1-r^2)^2} \equiv \frac{2\,dw\,d\overline{w}}{(1-|w|^2)^2}, \tag{39}$$

which coincides with the hyperbolic disk metrics. Also, the direct computation of (15) shows that $T^{\mathrm{F}} \equiv 0$. This is not surprising, because the wrapped Cauchy distribution is simply related to the usual one and the manifold formed by the usual Cauchy distributions is nothing more than upper half-plane model of AdS_2 space. The wrapped Cauchy distributions also form AdS_2 space but in terms of the Poincaré disk model. Two such models can be mapped to each other via conformal transformation.

We also can represent the wrapped Cauchy distribution as a q-exponent with $q = 2$, which can be checked in a straightforward way. Computing the canonical coordinates (ξ_1, ξ_2) and then deriving the potential function, we can find the metric g_{ij}^{q}. In $(w, \overline{w})$ coordinates the metric obtained from the corresponding potential function is given by

$$g_{ij}^{q} = \frac{2\pi}{1-|w|^2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow ds_q^2 = \frac{4\pi\, dw\, d\overline{w}}{1-|w|^2}; \tag{40}$$

so the q-metric again is conformally equivalent to the Fisher metric.

The wrapped Cauchy distribution plays a crucial role in the Kuramoto model dynamics since it is invariant under Möbius group action. Let z be a wrapped random variable with Cauchy distribution, $z \sim p_{\mathrm{wC}}(w; z)$. Let z' be the image of z produced by the Möbius transformation $\mathcal{M}$, $z' = \mathcal{M}(z)$. It was proven that if $z \sim p_{\mathrm{wC}}(w; z)$, then $z' \sim p_{\mathrm{wC}}(w; z')$. So, wrapped Cauchy distributions are closed with respect to the action of Möbius transformation. This fact drives us to conclude that the wrapped Cauchy distribution is a kind of universal distribution for the Kuramoto model dynamics. Such an implication resonates with the role of the Ott-Antonsen (O-A) ansatz [26].

We would like to emphasize that the Fisher information blows up in the synchronized phase, which corresponds to the limit $|w| \rightarrow 1$. In the Kuramoto-Sakaguchi model the relation between the singularity of the Fisher metric and the phase transition will be made more transparent.

C. Fisher metric in the Kuramoto-Sakaguchi model

Consider now the Kuramoto-Sakaguchi model [27], i.e., the Kuramoto model with noise,

$$\dot{\theta}_i = \omega_i + \frac{\lambda}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i) + \eta_i(t), \tag{41}$$

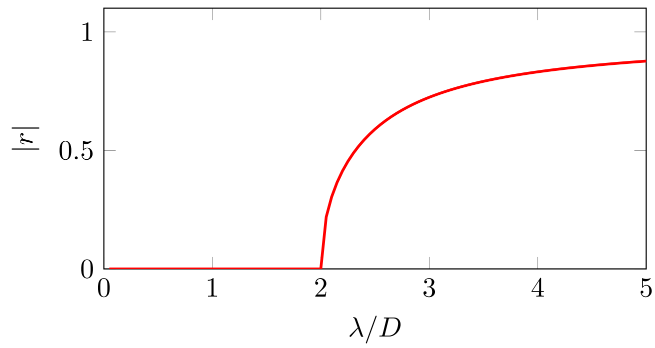

FIG. 1. Continuous phase transition in the Sakaguchi-Kuramoto model.

where $\eta_i = \eta_i(t)$ is a stochastic term with properties $\langle \eta_i(t)\rangle = 0$, $\langle \eta_i(t)\eta_j(t')\rangle = 2D\delta_{ij}\delta(t-t')$, where $D > 0$ is the noise amplitude. This model was initially considered by Sakaguchi, and it was shown that there is a continuous phase transition for the conventional order parameter r with respect to the coupling constant λ. The critical coupling could be obtained by the analysis of the self-consistency equation or by stability analysis of the incoherent state and is given by

$$\frac{1}{\lambda_c} = \frac{D}{2}\int_{-\infty}^{+\infty} \frac{d\omega\, g(\omega)}{\omega^2 + D^2}. \tag{42}$$

Both methods deal with the Fokker-Planck equation, which arises in the continuum limit of the Kuramoto-Sakaguchi model. In the case of identical frequencies, i.e., $g(\omega) = \delta(\omega)$, the critical coupling becomes $\lambda_c = 2D$. Moreover, in such a case one could easily find the stationary solution of the Fokker-Planck equation [28] in the rotating reference frame. The stationary solution ρ_0 reads as

$$\rho_0(\theta,\omega) = \frac{1}{2\pi I_0(\lambda|r|/D)} \exp\left\{\frac{\lambda|r|}{D}\cos(\arg r - \theta)\right\}. \tag{43}$$

The self-consistency equation for the Kuramoto-Sakaguchi model is very simple,

$$|r| = \frac{I_1(\lambda|r|/D)}{I_0(\lambda|r|/D)}, \tag{44}$$

where it is straightforward to notice that the nontrivial solution exists for $\lambda > \lambda_c = 2D$ (see Fig. 1). This equation is similar to the self-consistency equation that appeared in the XY model and to the self-consistency equation that appeared in the stationary Hamiltonian mean field model. This is not a surprise, because the noise strength D plays the role of temperature. In the case of nonidentical frequencies, the stationary distribution could be found explicitly [29], but the Fisher metric can be evaluated only numerically. The described stationary distribution is nothing more than the von Mises distribution belonging to the exponential family. In general form, it looks like

$$p_{\mathrm{vM}}(\kappa,\mu;\theta) = \frac{1}{2\pi I_0(\kappa)}\exp\{\kappa\cos(\theta-\mu)\},$$
$$\kappa \geqslant 0, \quad \mu = \mu \bmod 2\pi. \tag{45}$$

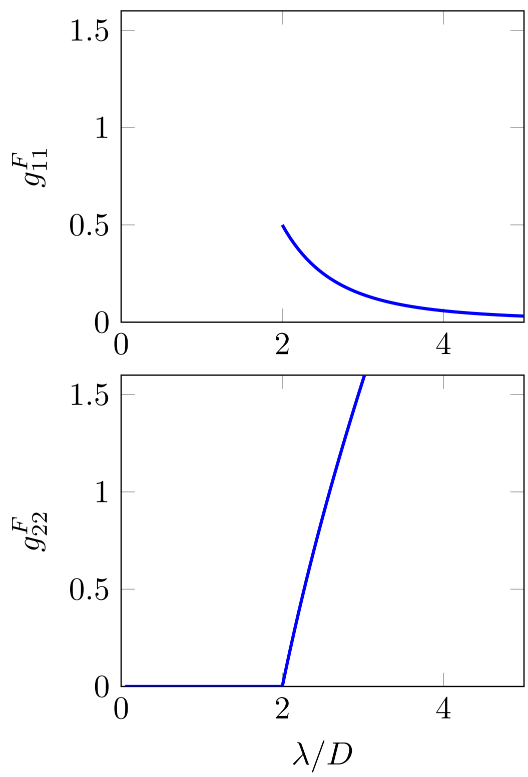

FIG. 2. Nonzero components of the Fisher metric for the Kuramoto-Sakaguchi model as the function of λ/D with $\lambda_c = 2D$.

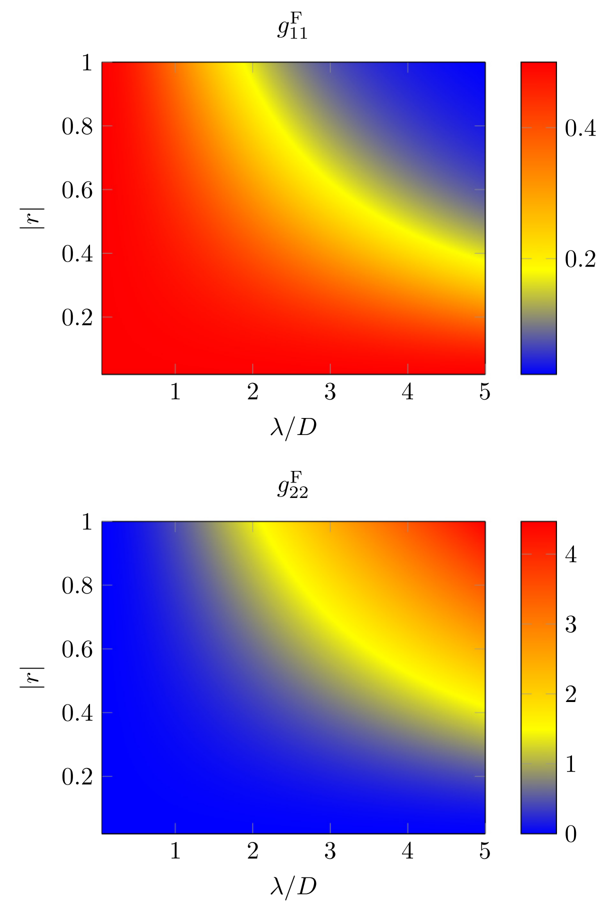

FIG. 3. Nonzero components of the Fisher metric for the Kuramoto-Sakaguchi model as a function of $|r|$ and λ/D with $\lambda_c = 2D$.

The computation of the Fisher metric for the von Mises distribution is straightforward,

$$g_{ij}^{\mathrm{F}} = \begin{pmatrix} 1 - \frac{I_1(\kappa)^2}{I_0(\kappa)^2} - \frac{I_1(\kappa)}{\kappa I_0(\kappa)} & 0 \\ 0 & \frac{\kappa I_1(\kappa)}{I_0(\kappa)} \end{pmatrix}, \tag{46}$$

where $(i, j) = (\kappa, \mu)$. Setting $\kappa = \lambda|r|/D$, we obtain the Fisher metric corresponding to the stationary solution of the Kuramoto-Sakaguchi model. We can interpret this in a twofold way: We can set $|r|$ to be a fixed value and then plot the components of the metric as functions of λ and D, or we can solve the self-consistency equation (44), find the order parameter $|r|$ as a function of λ and D, and then consider the components of the metric tensor as functions of λ and D only. The second interpretation gives us the following dependencies (Fig. 2): The component g_{11}^{F} does not exist for $\lambda < \lambda_c$, whereas the component g_{22}^{F} is identically zero for $\lambda < \lambda_c$. The first interpretation tells us to plot the components of the Fisher metric as functions of order parameter $|r|$ and λ/D; see Fig. 3. The transition point coincides with $\lambda_c = 2D$; hence the Fisher metric provides the order parameter for the synchronization transition.

D. Kullback-Leibler divergence in the Kuramoto model

After discussion of the Fisher metric, we would like to notice the following fact: The proposed hyperbolic space description of the Kuramoto model dynamics coincides with the gradient flow of the Kullback-Leibler divergence on the statistical manifold.

Indeed, in the continuum limit, $N \to \infty$, we deal with the initial distribution of phases $f_0 = f_0(z)$, $z = e^{i\theta}$. The gradient flow in the limit $N \to \infty$ becomes

$$\dot{w} = -\frac{\lambda}{2}(1 - |w|^2)^2 \frac{\partial}{\partial \overline{w}} \oint_{|z|=1} \frac{dz\, f_0(z) \ln p_{\mathrm{wC}}(w; z)}{iz}, \tag{47}$$

where we rewrite the integral over θ as the contour integral over the unit circle $|z| = 1$ in the complex plane. Our key observation is that the integral on the right-hand side of Eq. (47) is nothing more than the so-called cross entropy of two distributions with PDF $f_0(z)$ and $p_{\mathrm{wC}}(w; z)$,

$$S_{\mathrm{cross}} = \oint_{|z|=1} \frac{dz\, f_0(z) \ln p_{\mathrm{wC}}(w; z)}{iz}. \tag{48}$$

This quantity is tightly connected to the well-known Kullback-Leibler divergence,

$$S_{\text{K-L}}[f, g] = -\oint_{|z|=1} dz \frac{f(z)}{iz} \ln \frac{g(z)}{f(z)} = -S_0[f(z)] - S_{\mathrm{cross}}, \tag{49}$$

where $f = f(z)$ and $g = g(z)$ are two different PDFs of wrapped distributions and $S_0[f(z)]$ denotes the entropy of distribution $f(z)$. We can notice that the gradient with respect to $\overline{w}$ in Eq. (47) has the form

$$\frac{\partial S_{\text{K-L}}}{\partial \overline{w}} = -\frac{\partial}{\partial \overline{w}} \oint_{|z|=1} \frac{dz\, f(z) \ln g(z)}{iz}. \quad (50)$$

So we conclude that the gradient flow can be expressed in terms of the Kullback-Leibler divergence,

$$\dot{w} = +\frac{\lambda}{2}(1-|w|^2)^2 \frac{\partial S_{\text{K-L}}[f_0, p_{\text{wC}}]}{\partial \overline{w}}, \quad (51)$$

where $f_0 = f_0(z)$ is the initial distribution and $p_{\text{wC}} = p_{\text{wC}}(w; z)$ is the wrapped Cauchy distribution.

As we have mentioned earlier, the Kullback-Leibler divergence in general does not coincide with the Bregman divergence. Nevertheless, Eq. (51) clearly says that the Kuramoto model dynamics can be treated as the gradient flow of the Kullback-Leibler divergence on the statistical manifold formed by wrapped Cauchy distributions. Therefore, to understand such dynamics, we can compute the Kullback-Leibler divergence between a given initial distribution $f_0(z)$ and the wrapped Cauchy distribution $p_{\text{wC}}(w; z)$. Let us notice that the uniform distribution $p_{\text{U}}(z) = (2\pi)^{-1}$ is nothing more than the wrapped Cauchy distribution with $w = 0$ and the Dirac-delta distribution is also the wrapped Cauchy distribution with $w = 1$.

If the initial distribution $f_0(z)$ is uniform, so that $f_0(z) = (2\pi)^{-1}$, the Kullback-Leibler divergence it is easy to compute, which gives

$$S_{\text{K-L}}[f_0, p_{\text{wC}}] = -\ln(1-|w|^2). \quad (52)$$

Substituting this expression into the gradient flow, we find

$$\dot{w} = \frac{\lambda}{2}(1-|w|^2)w. \quad (53)$$

The obtained equation is interesting for two reasons. First, it is similar to the equation that appeared in the Ott-Antonsen ansatz. Second, it tells us that in the large-N limit the Kuramoto model with all-to-all couplings and identical frequencies is *dual* to the *single* Landau-Stuart oscillator. It is quite straightforward to obtain the solution $w = w(t)$, and then we can see how the Kullback-Leibler divergence evolves in time. Until the transition moment, the divergence is negligibly small (this is quite clear because an instant distribution on the unit circle is still similar to uniform f_0). At the transition moment, the Kullback-Leibler divergence starts to grow significantly and finally blows up: The final distribution on the unit circle is drastically different from the uniform distribution, which corresponds to the complete synchronization.

V. CONCLUSION

In this paper we revised the hyperbolic description of the Kuramoto model proposed in Ref. [5] relating it to the statistical manifolds. The key observation of our study is that the singularity of the Fisher information metric captures the synchronization transition. Namely, the metrics corresponding to the Kuramoto and Kuramoto-Sakaguchi models blow up in the synchronized state. Therefore we extend the list of examples when the phase transitions get identified via the singularities of the Fisher metric. We have focused on the information geometry describing the all-to-all Kuramoto-Sakaguchi model, but it would be interesting to recognize the information metric for more general graph architecture such as the star when the exact solution is available [29].

The presented approach can also be useful for a case with complicated topology, i.e., when an analytical treatment is extremely tough. Based on the fact that the Fisher metric has a singularity at the transition point, it seems possible to extract an explicit expression for the model parameters, which raises the synchronized state. Moreover, there is no restriction for generalizations of the Kuramoto model, which is emphasized by the described example of a model with noise. This fact allows us to use the information geometry framework for generalizations, which include phase lags, external forces, and so on.

We have noticed that the gradient flow in the Kuramoto model can be represented as a gradient flow on a statistical manifold. However, instead of a potential function, one deals with the Kullback-Leibler divergence. As was emphasized, the Kullback-Leibler divergence is not the Bregman divergence for the (wrapped) Cauchy distribution, and it is reasonable to represent such flow in terms of the Bregman-Tsallis divergence. However, the metric obtained via the potential function of the Bregman-Tsallis divergence does not coincide with the Fisher metric (it is conformally equivalent, but different). We leave a closer treatment of the connection between synchronization and the Bregman-Tsallis divergence for further research. Next, the discussion of the Cauchy distribution as a member of the q-exponential family naturally raises a question concerning wrapped q-exponential distributions. As far as we know, the properties of such distributions have not been examined. It seems that this research could be interesting in the information geometry context.

In this paper the coordinates on the $\mathrm{SL}(2, \mathbb{R})$ orbit play the role of the parameters due to the dimensional reduction of dynamics. Hence, to some extent, one could be confused as to why the effective degrees of freedom play the role of the effective parameters. However, we know of a very precise example of the same nature: the duality between the inhomogeneous spin chains and an integrable many-body system with long-range interaction of Calogero-Ruijsenaars type which has been formulated in probabilistic terms in Ref. [30]. In that case the parameters and inhomogeneities in the spin chains get identified with the coordinates of the particles in the Calogero model, which, on the other hand, are the coordinates on the group orbit. That is, the Fisher metric on the parameters in the spin chain gets mapped into the clear-cut object on the Calogero-Ruijsenaars side. We shall discuss this analogy and relation elsewhere.

The hyperbolic geometry and the large number of degrees of freedom at its boundary invite a holographic description. However, in this case there are some subtle points since the boundary theory is classical and the equations are of the first order in the time derivatives. Nevertheless, the description in terms of the conformal barycenter has a lot in common with the dynamics of a kind of baryonic vertex since the Poisson kernels are related to the geodesics connecting the points at the boundary and the bulk. We postpone these issues for a separate study. A general discussion concerning the relation of the probabilistic manifold with holography can be found in Ref. [31].

Our concluding remark concerns the application of information geometry to neuroscience. The idea is not new [6]; however, our findings provide alternative perspectives for this issue. The Kuramoto model is widely used as the model for the synchronization of the functional connectomes; hence one could expect that the properties of the Fisher metric for the Kuramoto model on more general graphs shall provide an alternative approach for brain rhythm generation. In particular, it would be very interesting to investigate the synchronization of the several functional connectomes via the properties of the Fisher metric and possible dependence of the synchronization of the several functional connectomes on the architecture of the structural connectome.

ACKNOWLEDGMENTS

We are grateful to A. Polkovnikov for useful discussions, to S. Kato for discussions of Kato-Jones distribution properties, and to F. Nielsen for comments concerning the Cauchy distribution statistical manifold. A.G. thanks IHES and Nordita, where part of the work was done, for the hospitality and support. A.A. is thankful for a grant from the Brain Program of the IDEAS Research Center and Grant No. 18-12-00434П from the Russian Science Foundation. A.G. is thankful for Grant No. 075-15-2020-801 from the Ministry of Science and Higher Education of the Russian Federation and a grant from the Basis Foundation.

APPENDIX A: ON MÖBIUS TRANSFORMATION AND WRAPPED DISTRIBUTIONS

In this Appendix, let us describe some facts concerning the wrapped distributions which can be obtained from non-wrapped ones via so-called "compactification." For a random variable $x \in (-\infty, +\infty)$ with probability density function $p(x)$, we introduce the complex variable $z = e^{ix}$ and consider its phase $\theta = \arg z$, $\theta \in (-\pi, +\pi]$, which has *wrapped* probability density function $p_{\mathrm{w}}(\theta)$. The connection between $p(x)$ and its wrapped version is the following:

$$p_{\mathrm{w}}(\theta) = \sum_{k=-\infty}^{+\infty} p(\theta + 2\pi k). \tag{A1}$$

All the common concepts of probability theory work for wrapped distributions as well.

There are (at least) four common wrapped distributions that are usually discussed in the context of the Kuramoto model: the uniform distribution, the wrapped Cauchy distribution, the von Mises distribution, and the recently appearing Kato-Jones distribution. The uniform distribution is trivial, and the wrapped Cauchy distribution is well known. The von Mises distribution arises in the Sakaguchi-Kuramoto model as the stationary solution of the corresponding Fokker-Planck equation. The Kato-Jones distribution was introduced in Ref. [32], and it is the Möbius-transformed von Mises distribution. This distribution is invariant under the action of the Möbius group. The abovementioned list of distributions is easy to extend. For the sake of completeness, we can also mention cardioid distribution and its Möbius image and wrapped normal distribution and its Möbius image. The invariance under Möbius group action is crucial in the Kuramoto model. So here we focus on the triplet of the wrapped Cauchy distribution, von Mises distribution, and Kato-Jones distribution. We would like to establish the interconnections between members of such a triplet in the information geometry context. First of all, let us write down the probability density functions for each member of the triplet,

$$p_{\mathrm{wC}}(r,\phi;\theta) = \frac{1}{2\pi}\frac{1-r^2}{1-2r\cos(\theta-\phi)+r^2}, \qquad 0 \leqslant r < 1, \quad \phi = \phi \bmod 2\pi, \tag{A2}$$

$$p_{\mathrm{vM}}(\kappa,\mu;\theta) = \frac{1}{2\pi I_0(\kappa)}\exp\{\kappa\cos(\theta-\mu)\}, \quad \kappa \geqslant 0, \quad \mu = \mu \bmod 2\pi, \tag{A3}$$

$$p_{\text{K-J}}(\kappa,r,\nu,\mu;\theta) = \frac{1}{2\pi I_0(\kappa)}\frac{1-r^2}{1-2r\cos(\theta-\mu-\nu)+r^2}\exp\left\{\frac{\kappa\cos(\mu+\theta)+\kappa r^2\cos(\mu+2\nu-\theta)-2r\kappa\cos\nu}{1-2r\cos(\theta-\mu-\nu)+r^2}\right\},$$
$$\kappa \geqslant 0,\ 0 \leqslant r < 1, \quad \mu = \mu \bmod 2\pi, \quad \nu = \nu \bmod 2\pi. \tag{A4}$$

From these expressions it is straightforward to see the following facts: The wrapped Cauchy and von Mises distributions form two-dimensional statistical manifolds, the von Mises distribution belongs to the exponential family, and the Kato-Jones distribution forms a four-dimensional statistical manifold. The wrapped Cauchy distribution can be derived from the Kato-Jones distribution by setting $\kappa \to 0$, whereas the von Mises distribution can be derived by setting $r \to 0$. Therefore we can treat the statistical manifolds corresponding to the wrapped Cauchy distribution and to the von Mises distribution as the submanifolds of the Kato-Jones statistical manifold. Wrapped distributions p_{wC} and $p_{\text{K-J}}$ belong to Möbius-invariant statistical manifold $\mathbb{M}_{\mathcal{M}}$, whereas p_{vM} does not. However, from $p_{\text{K-J}}$ one can obtain p_{vM} and p_{wC} taking an appropriate limit, $r \to 0$ and $\kappa \to 0$, respectively (see Fig. 4).

APPENDIX B: ON *q*-GAUSSIAN DISTRIBUTIONS IN THE KURAMOTO MODEL

The Kuramoto model shares some similarities with the famous Hamiltonian mean field (HMF) model [33,34], which describes N interacting particles on $\mathbb{S}^1$ via a cosine potential with all-to-all couplings,

$$H_{\mathrm{HMF}} = \sum_{i=1}^{N}\frac{p_i^2}{2m} + \frac{\lambda}{2N}\sum_{i<j}\cos(\theta_j-\theta_i), \tag{B1}$$

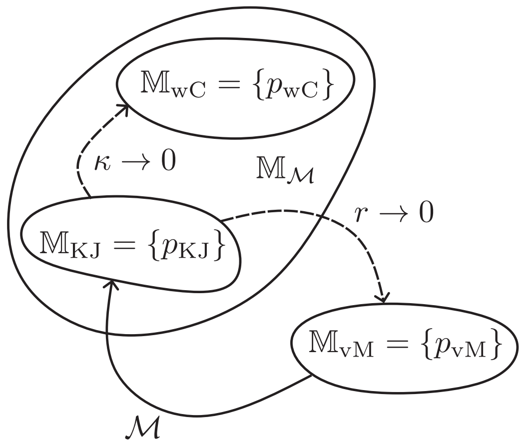

FIG. 4. Kuramoto-model-related statistical manifolds. One can obtain $p_{\text{K-J}}$ via Möbius transformation $\mathcal{M}$ and then reduce $M_{\text{K-J}}$ to the wrapped Cauchy submanifold $\mathbb{M}_{\rm wC}$ and the von Mises submanifold $\mathbb{M}_{\rm vM}$. Both $p_{\text{K-J}}$ and $p_{\rm wC}$ belong to the Möbius-invariant submanifold $\mathbb{M}_{\rm M}$, whereas $p_{\rm vM}$ does not.

where λ is the coupling constant. This model is prototypical for the study of systems with long-range interactions (LRIs). In the HMF model, LRIs are caused by all-to-all couplings. The effects of LRIs in the HMF model are quite well known: It was shown that LRIs cause a violent relaxation phenomenon and the existence of quasistationary states [35]. It is worth mentioning that the quantum version of the HMF model also exhibits such phenomena. One quite interesting consequence of LRIs is the appearance of so-called Tsallis q-statistics [36]. Due to the presence of LRIs, the entropy of a system becomes nonextensive. Tsallis and co-workers developed the appropriate machinery to describe this nonextensivity [37]. Roughly speaking, the parameter q measures the nonextensivity in a system. The q-statistics deals with so-called q-Gaussian distributions (see Ref. [37] for a detailed discussion). Such q-Gaussian distributions differ from the usual Gaussian distribution for $q \neq 1$: They have heavier tails (see Fig. 5), and in the limit $q \to 1$ one restores the usual Gaussian distribution. Next, the authors of Ref. [38] presented a method to capture correlations between degrees of freedom based on q-statistics. Despite the fact that there are no mathematically rigorous derivations of q-statistics for a system with LRIs, such a method allows one to verify (at least numerically) that q-statistics takes place. For instance, it was shown that q-statistics appears in the HMF model [39,40].

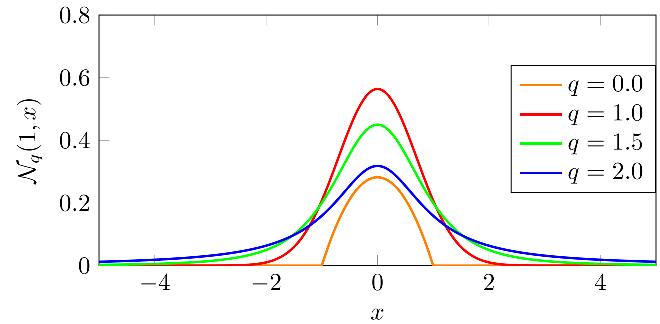

FIG. 5. q-Gaussians.

Based on the idea that the HMF model and Kuramoto model are related to each other, the authors of Ref. [41] examined fingerprints of q-statistics in the Kuramoto model. The key observation was the following. From the phases $\theta_i(t)$ one should construct the so-called central limit theorem (CLT) variable y_i (see Ref. [38] for details),

$$y_i = \frac{1}{\sqrt{M}} \sum_{k=1}^{M} \theta_i(k\delta), \tag{B2}$$

where $\delta > 0$ is the predefined time interval. In the Kuramoto model it was demonstrated that for $\lambda < \lambda_c$ the variables y_i obey a q-Gaussian distribution with $q \approx 1.7$, i.e., q-CLT appears. The authors considered the case of *nonidentical* frequencies and focused on the case of a uniform distribution and the case of a Gaussian distribution. In the case of $\lambda > \lambda_c$ the variables y_i were fitted by the usual Gaussian distribution.

Having briefly discussed the q-Gaussians, we would like to draw attention to the fact that q-Gaussian distributions have already appeared in the context of the Kuramoto model. Indeed, the continuum limit of the Kuramoto model can be described in the Ott-Antonsen framework [26]. In the case of identical oscillators, the distribution function for the continuum limit can be explicitly obtained. This distribution function is nothing more than the normalized Poisson kernel, i.e., the wrapped Cauchy distribution. Of course, parameters of the distribution depend on time and initial conditions. As we have shown, the usual Cauchy distribution belongs to the q-exponential family with $q = 2$, so the wrapped Cauchy distribution also inherits properties of the q-exponential family (wrapped and nonwrapped versions of the Cauchy distribution are related to each other via conformal transformation).

[1] F. A. Rodrigues, T. K. D. Peron, P. Ji, and J. Kurths, Phys. Rep. **610**, 1 (2016).
[2] S. Watanabe and S. H. Strogatz, Phys. Rev. Lett. **70**, 2391 (1993).
[3] S. Watanabe and S. H. Strogatz, Phys. D (Amsterdam) **74**, 197 (1994).
[4] S. A. Marvel, R. E. Mirollo, and S. H. Strogatz, Chaos **19**, 043104 (2009).
[5] B. Chen, J. R. Engelbrecht, and R. Mirollo, J. Phys. A: Math. Theor. **50**, 355101 (2017).
[6] S.-I. Amari, *Information Geometry and Its Applications*, Applied Mathematical Sciences Vol. 194 (Springer, New York, 2016).
[7] J. Provost and G. Vallee, Commun. Math. Phys. **76**, 289 (1980).
[8] M. V. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984).
[9] P. Zanardi and N. Paunković, Phys. Rev. E **74**, 031123 (2006).
[10] P. Zanardi, P. Giorda, and M. Cozzini, Phys. Rev. Lett. **99**, 100603 (2007).
[11] W.-L. You, Y.-W. Li, and S.-J. Gu, Phys. Rev. E **76**, 022101 (2007).

[12] S. Garnerone, D. Abasto, S. Haas, and P. Zanardi, Phys. Rev. A **79**, 032302 (2009).
[13] M. Kolodrubetz, V. Gritsev, and A. Polkovnikov, Phys. Rev. B **88**, 064304 (2013).
[14] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, Phys. Rep. **697**, 1 (2017).
[15] M. Pandey, P. W. Claeys, D. K. Campbell, A. Polkovnikov, and D. Sels, Phys. Rev. X **10**, 041017 (2020).
[16] P. Buonsante and A. Vezzani, Phys. Rev. Lett. **98**, 110601 (2007).
[17] A. Dey, S. Mahapatra, P. Roy, and T. Sarkar, Phys. Rev. E **86**, 031137 (2012).
[18] J. Cantarella and H. Schumacher, SIAM J. Appl. Algebra Geom. **6**, 503 (2022).
[19] S.-I. Amari, IEEE Trans. Inf. Theory **55**, 4925 (2009).
[20] Y. Nakamura, Jpn. J. Indust. Appl. Math. **10**, 179 (1993).
[21] A. Fujiwara and S.-I. Amari, Phys. D (Amsterdam) **80**, 317 (1995).
[22] A. F. Mitchell, Int. Stat. Rev. **56**, 1 (1988).
[23] P. McCullagh, Biometrika **79**, 247 (1992).
[24] F. Nielsen, Entropy **22**, 713 (2020).
[25] P. Zanardi, L. C. Venuti, and P. Giorda, Phys. Rev. A **76**, 062318 (2007).
[26] E. Ott and T. M. Antonsen, Chaos **18**, 037113 (2008).
[27] H. Sakaguchi, Prog. Theor. Phys. **79**, 39 (1988).
[28] D. S. Goldobin, I. V. Tyulkina, L. S. Klimenko, and A. Pikovsky, Chaos **28**, 101101 (2018).
[29] A. Alexandrov, Chaos, Solitons Fractals **167**, 113056 (2023).
[30] A. Gorsky, M. Vasilyev, and A. Zotov, J. High Energy Phys. 04 (2022) 159.
[31] J. Erdmenger, K. Grosvenor, and R. Jefferson, SciPost Phys. **8**, 073 (2020).
[32] S. Kato and M. Jones, J. Am. Stat. Assoc. **105**, 249 (2010).
[33] G. Miritello, A. Pluchino, and A. Rapisarda, Europhys. Lett. **85**, 10007 (2009).
[34] A. Pluchino and A. Rapisarda, Phys. A (Amsterdam) **365**, 184 (2006).
[35] Y. Y. Yamaguchi, J. Barré, F. Bouchet, T. Dauxois, and S. Ruffo, Phys. A (Amsterdam) **337**, 36 (2004).
[36] C. Tsallis and U. Tirnakli, in *Mathematical Aspects of Generalized Entropies and their Applications, Kyoto, Japan*, Journal of Physics: Conference Series Vol. 201 (Institute of Physics, London, 2010), p. 012001.
[37] S. Umarov, C. Tsallis, and S. Steinberg, Milan J. Math. **76**, 307 (2008).
[38] U. Tirnakli, C. Beck, and C. Tsallis, Phys. Rev. E **75**, 040106(R) (2007).
[39] A. Pluchino, A. Rapisarda, and C. Tsallis, Europhys. Lett. **85**, 60006 (2009).
[40] A. Pluchino, A. Rapisarda, and C. Tsallis, Phys. A (Amsterdam) **387**, 3121 (2008).
[41] G. Miritello, A. Pluchino, and A. Rapisarda, Phys. A (Amsterdam) **388**, 4818 (2009).