Nonadditive generalization of the quantum Kullback-Leibler divergence for measuring the degree of purification

Sumiyoshi Abe
Institute of Physics, University of Tsukuba, Ibaraki 305-8571, Japan
(Received 13 January 2003; published 10 September 2003)

The Kullback-Leibler divergence offers an information-theoretic basis for measuring the difference between two given distributions. Its quantum analog, however, fails to play a corresponding role for comparing two density matrices, if the reference states are pure states. Here it is shown that nonadditive quantum information theory inspired by nonextensive statistical mechanics is free from such a difficulty and the associated quantity, termed the quantum q-divergence, can in fact be a good information-theoretic measure of the degree of state purification. The correspondence relation between the ordinary divergence and the q-divergence is violated for the pure reference states, in general.

DOI: 10.1103/PhysRevA.68.032302 PACS number(s): 03.67.−a, 03.65.Ud, 05.20.−y, 05.30.−d

Purification is of fundamental relevance to quantum error correction, which is important for quantum computation and quantum communication. Specifically, a task is to purify a state of a subsystem of a composite system decayed into a mixed state (see [1–4], for example).

In such a situation, it is essential to quantify the degree of purification, that is, to compare a mixed-state density matrix with a reference pure-state density matrix. This problem is often treated by the use of the concept of fidelity [5,6]. For two density matrices, ρ and σ, it is given by

$$F[\sigma,\rho]=[\mathrm{Tr}(\sqrt{\sigma}\rho\sqrt{\sigma})^{1/2}]^2, \tag{1}$$

which is also related to the Bures metric between ρ and σ as $d_B^2=2-2\sqrt{F[\sigma,\rho]}$. For a pure state, $\sigma=|\psi\rangle\langle\psi|$, the fidelity becomes $F[|\psi\rangle\langle\psi|,\rho]=\langle\psi|\rho|\psi\rangle$.

On the other hand, in classical information theory, a comparison of two distributions is customarily discussed by employing the Kullback-Leibler divergence. Its quantum-mechanical counterpart is the quantum divergence of a density matrix ρ with respect to a reference density matrix σ, which is given by [7]

$$K[\rho\|\sigma]=\mathrm{Tr}[\rho(\ln\rho-\ln\sigma)]\geqslant 0, \tag{2}$$

where the equality holds if and only if $\rho=\sigma$. However, *this quantity turns out to be inadequate for measuring degree of purification, since* $\ln\sigma$ *is a singular quantity if the reference state* σ *is a pure state*. (More generically, $K[\rho\|\sigma]$ can be well defined only when the support of σ is equal to or larger than that of ρ [7].)

In this paper we study a generalized information-theoretic approach to quantifying the degree of purification based on nonadditive quantum information theory, which has been initiated in [8] and applied to the problems of quantum entanglement [9–11]. In particular, we discuss the nonadditive generalization of the quantum divergence, termed the quantum q-divergence, and explicitly show how it is superior to the one in Eq. (2).

Let us start our discussion with noting that the ordinary quantum divergence in Eq. (2) can be rewritten as follows:

$$K[\rho\|\sigma]=\frac{d}{dx}\mathrm{Tr}(\rho^x\sigma^{1-x})\bigg|_{x\to 1-0}. \tag{3}$$

The quantum q-divergence is obtained by replacing the derivative in Eq. (3) with the Jackson q-derivative:

$$K_q[\rho\|\sigma]=D_q\mathrm{Tr}(\rho^x\sigma^{1-x})|_{x\to 1-0}, \tag{4}$$

where D_q denotes the Jackson differential operator defined by

$$D_q f(x)=\frac{f(qx)-f(x)}{x(q-1)}, \tag{5}$$

which satisfies the following q-deformed Leibniz rule:

$$D_q[f(x)g(x)]=[D_qf(x)]g(x)+f(x)[D_qg(x)]+x(q-1)\times[D_qf(x)][D_qg(x)]. \tag{6}$$

In the limit $q\to 1$, D_qf tends to the ordinary derivative, df/dx. Equation (4) is found to be

$$K_q[\rho\|\sigma]=\frac{1}{1-q}\mathrm{Tr}[\rho^q(\rho^{1-q}-\sigma^{1-q})]=\mathrm{Tr}[\rho^q(\ln_q\rho-\ln_q\sigma)]. \tag{7}$$

In this equation, q is a positive parameter termed the entropic index, and $\ln_q x$ stands for the q-logarithmic function defined by

$$\ln_q x=\frac{1}{1-q}(x^{1-q}-1), \tag{8}$$

which converges to the ordinary logarithmic function, $\ln x$, in the limit $q\to 1$. Therefore, K_q might also be expected to converge to K in such a limit. (However, we shall see that this is not the case, in general.) Since K_q should not be too sensitive to small eigenvalues of ρ and σ, the range of the entropic index must be taken to be

$$0<q<1. \tag{9}$$

Several comments are in order. Firstly, the classical counterpart of Eq. (7) has been proposed independently and almost simultaneously in [12–14]. Secondly, the construction in Eq. (4) reminds us of that of the Tsallis entropy [15] developed in [16]. This is due to the fact that K_q in Eq. (6) is the relative entropy associated with the Tsallis entropy, $S_q[\rho]=(\mathrm{Tr}\rho^q-1)/(1-q)$, analogously to the relationship between K in Eq. (2) and the von Neumann entropy, $S[\rho]=-\mathrm{Tr}(\rho\ln\rho)$. Thirdly, it should be noted that the Tsallis entropy gives the basis for nonextensive generalization of traditional Boltzmann-Gibbs statistical mechanics. This theory, termed nonextensive statistical mechanics [17], is considered to statistically describe complex systems in their nonequilibrium stationary states and, in fact, is currently accumulating its successful applications (see the URL, http://tsallis.cat.cbpf.br/TEMUCO.pdf, for the comprehensive list of references). At first glance, one might feel that one-parameter generalization of the von Neumann entropy of this kind is completely *ad hoc* and there are huge ambiguities behind it. Quite remarkably, however, it is not the case at all. It has been rigorously shown that the (classical) Tsallis entropy is a unique quantity, which is consistent with the principles of thermodynamics [18–20], satisfies the stability condition [21], and is characterized by the generalized Shannon-Khinchin axioms and the uniqueness theorem [8,22]. In addition, it has analytically been shown by the renormalization group technique [23] that the Kolmogorov-Sinai entropy has to be replaced by the Tsallis entropy when nonlinear dynamical systems are prepared at the edge of chaos. Taking into account these developments and recalling the philosophy of Jaynes [24] for building a bridge between statistical mechanics and physical information theory, clearly it is of importance to explore if nonadditive information theory based on the Tsallis entropy and its associated q-divergence K_q have any points superior to the additive von Neumann theory.

Let us summarize the basic properties of the quantum q-divergence. $K_q[\rho\|\sigma]$ is jointly convex:

$$K_q\left[\sum_i \lambda_i\rho^{(i)}\middle\|\sum_j \lambda_j\sigma^{(j)}\right]\leqslant\sum_i \lambda_i K_q[\rho^{(i)}\|\sigma^{(i)}],$$

where $\lambda_i>0$ and $\Sigma_i\,\lambda_i=1$. This follows from Lieb's theorem [25], which states that $\mathrm{Tr}(L^{1-x}M^x)$ with $0<x<1$ is jointly concave in any positive operators, L and M.

Quite recently, it has been proved [26] that K_q monotonically decreases by projective measurements. K_q is nonadditive in the sense that for the factorized joint density matrices of a composite system (A, B), $\rho(A,B)=\rho_1(A)\otimes\rho_2(B)$ and $\sigma(A,B)=\sigma_1(A)\otimes\sigma_2(B)$, it yields

$$K_q[\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2]=K_q[\rho_1\|\sigma_1]+K_q[\rho_2\|\sigma_2]+(q-1)K_q[\rho_1\|\sigma_1]K_q[\rho_2\|\sigma_2], \tag{10}$$

which essentially has its origin in the q-deformed Leibniz rule in Eq. (6). Thus, the value of $1-q$ indicates the degree of nonadditivity.

Let us see that $K_q[\rho\|\sigma]$ is non-negative for any two density matrices, ρ and σ. For this purpose, consider the diagonal decompositions of ρ and σ [27]:

$$\rho=\sum_a r(a)|a\rangle\langle a|,\quad \sigma=\sum_b s(b)|b\rangle\langle b|, \tag{11}$$

where $\{|a\rangle\}$ and $\{|b\rangle\}$ are the orthonormal complete bases, $0\leqslant r(a),\ s(b)\leqslant 1$, and $\Sigma_a\, r(a)=\Sigma_b\, s(b)=1$. A straightforward calculation shows that

$$K_q[\rho\|\sigma]=\frac{1}{1-q}\sum_{a,b}|\langle a|b\rangle|^2 r(a)\left[1-\left(\frac{s(b)}{r(a)}\right)^{1-q}\right]. \tag{12}$$

Making use of the inequality, $(1-x^p)/p\geqslant 1-x$ $(x\geqslant 0,\ 0<p<1)$ with the equality for $x=1$ (Theorem 42 in [28]), we arrive at the conclusion

$$K_q[\rho\|\sigma]\geqslant\sum_{a,b}|\langle a|b\rangle|^2 r(a)\left[1-\frac{s(b)}{r(a)}\right]=0. \tag{13}$$

From the above discussion, it is clear that non-negativity of $K_q[\rho\|\sigma]$ itself holds for any positive values of q and the restriction in Eq. (9) is not necessary.

$K_q[\rho\|\sigma]$ is not symmetric in ρ and σ unless $q=1/2$. Its symmetrization, $d_q^2=K_q[\rho\|\sigma]+K_q[\sigma\|\rho]$ looks like the squared "distance" between ρ and σ. However, to our knowledge, the triangle inequality is not established yet for d_q with $q\neq 1$.

Now, a point of crucial difference between $K[\rho\|\sigma]$ and $K_q[\rho\|\sigma]$ is that, in marked contrast with $\ln\sigma$, $\ln_q\sigma$ is a well-defined quantity for a pure reference state $\sigma=|\psi\rangle\langle\psi|$. In fact, in this case, $\sigma^{1-q}=\sigma(0<q<1)$, whereas $\ln\sigma=(\sigma-I)\zeta(1)$, which is divergent, where I and $\zeta(s)$ are the unit matrix and the Riemann ζ function, respectively. Accordingly, Eq. (7) is seen to be

$$K_q[\rho\||\psi\rangle\langle\psi|]=\frac{1}{1-q}(1-\langle\psi|\rho^q|\psi\rangle). \tag{14}$$

It is important to note that the additive limit $q\to 1$ cannot be taken in this equation anymore.

Here, we wish to consider the particular case when ρ is also a pure state, $\rho=|\phi\rangle\langle\phi|$. Then, Eq. (14) further becomes

$$K_q[|\phi\rangle\langle\phi|\||\psi\rangle\langle\psi|]=\frac{d_{FS}^2}{1-q}, \tag{15}$$

where

$$d_{FS}^2=1-|\langle\phi|\psi\rangle|^2 \tag{16}$$

is the Fubini-Study metric in the projective Hilbert space, which may give the geometric interpretations to quantum uncertainty and correlation [29]. In addition, the transition probability on the right-hand side of Eq. (16) coincides with the value of the fidelity in this case.

Finally, let us examine the quantum q-divergence for measuring the degree of purification of the Werner state [2]. The Werner state is a state of a bipartite spin-1/2 system, i.e., two qubits, given as follows [30]:

$$\rho_W = F|\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{3}(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle \times\langle\Phi^-|), \tag{17}$$

where $|\Psi^\pm\rangle$ and $|\Phi^\pm\rangle$ are the Bell states: $|\Psi^\pm\rangle = 2^{-1/2}(|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle)$, $|\Phi^\pm\rangle = 2^{-1/2}(|\uparrow\uparrow\rangle \pm |\downarrow\downarrow\rangle)$. F is the fidelity with respect to the reference state $\sigma = |\Psi^-\rangle\langle\Psi^-|$. Its allowed range is $1/4 \leqslant F \leqslant 1$, and ρ_W is known to be separable if and only if $F \leqslant 1/2$. In a recent paper [31], it has been discussed how to experimentally prepare such a state.

Now, the quantum q-divergence of ρ_W with respect to the reference state $\sigma = |\Psi^-\rangle\langle\Psi^-|$ is immediately calculated to be

$$K_q[\rho_W \| |\Psi^-\rangle\langle\Psi^-|] = \frac{1}{1-q}(1-F^q) \geqslant 0, \tag{18}$$

where the zero value is realized when $F=1$ or $q \to +0$. However, as already stressed, the limit $q \to 1-0$ is singular and does not commute with the limit $F \to 1-0$.

In conclusion, we have discussed a possible information-theoretic measure of the degree of state purification based on nonadditive quantum information theory. We have analyzed the properties of the quantum q-divergence and have found that, for pure reference states, it is superior to the ordinary quantum divergence. In particular, we have seen that the additive limit cannot be taken in such a situation, and the correspondence relation between the ordinary divergence and the q-divergence is violated.

The author thanks Dr. A. K. Rajagopal for discussions.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[2] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[3] C. H. Bennett, D. P. DiVicenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[4] V. Korepin and J. Terilla, e-print quant-ph/0202054.
[5] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
[6] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[7] H. Umegaki, Kodai Math. Sem. Rep. **14**, 59 (1962).
[8] S. Abe, Phys. Lett. A **271**, 74 (2000). See also, S. Abe and A. K. Rajagopal, Physica A **289**, 157 (2001); Chaos, Solitons Fractals **13**, 431 (2002).
[9] S. Abe, Phys. Rev. A **65**, 052323 (2002); Physica A **306**, 316 (2002).
[10] C. Tsallis, S. Lloyd, and M. Baranger, Phys. Rev. A **63**, 042104 (2001); F. C. Alcaraz and C. Tsallis, Phys. Lett. A **301**, 105 (2002); C. Tsallis, D. Prato, and C. Anteneodo, Eur. Phys. J. B **29**, 605 (2002).
[11] N. Canosa and R. Rossignoli, Phys. Rev. Lett. **88**, 170401 (2002); R. Rossignoli and N. Canosa, Phys. Rev. A **66**, 042306 (2002).
[12] C. Tsallis, Phys. Rev. E **58**, 1442 (1998); L. Borland, A. R. Plastino, and C. Tsallis, J. Math. Phys. **39**, 6490 (1998); **40**, 2196(E) (1999).
[13] S. Abe, in *Proceedings of the 5th International Wigner Symposium*, Vienna, 1997, edited by P. Kasperkovitz and D. Grau (World Scientific, Singapore, 1998), p. 66; see also A. K. Rajagopal and S. Abe, Phys. Rev. Lett. **83**, 1711 (1999).
[14] M. Shiino, J. Phys. Soc. Jpn. **67**, 3658 (1998).
[15] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[16] S. Abe, Phys. Lett. A **224**, 326 (1997); **244**, 229 (1998); see also R. S. Johal, Phys. Rev. E **58**, 4147 (1998); Phys. Lett. A **253**, 47 (1999).
[17] *Nonextensive Statistical Mechanics and Its Applications*, edited by S. Abe and Y. Okamoto (Springer-Verlag, Heidelberg, 2001).
[18] S. Abe, Phys. Rev. E **63**, 061105 (2001).
[19] S. Abe, Continuum Mech. Thermodyn. (to be published), e-print cond-mat/0305087.
[20] S. Abe and A. K. Rajagopal, Phys. Rev. Lett. (to be published), e-print cond-mat/0304066.
[21] S. Abe, Phys. Rev. E **66**, 046134 (2002).
[22] R. J. V. dos Santos, J. Math. Phys. **38**, 4104 (1997).
[23] F. Baldovin and A. Robledo, Phys. Rev. E **66**, 045104(R) (2002); Europhys. Lett. **60**, 518 (2002).
[24] *E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, edited by R. D. Rosenkrantz (Kluwer, Dordrecht, 1989).
[25] E. H. Lieb, Adv. Math. **11**, 267 (1973).
[26] S. Abe, Phys. Lett. A **312**, 336 (2003).
[27] A. K. Rajagopal, in [18]. See also, A. K. Rajagopal and R. W. Rendell, Phys. Rev. A **66**, 022104 (2002).
[28] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed. (Cambridge University Press, Cambridge, 1952).
[29] S. Abe, Phys. Rev. A **46**, 1667 (1992); **48**, 4102 (1993).
[30] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[31] Y.-S. Zhang, Y.-F. Huang, C.-F. Li, and G.-C. Guo, Phys. Rev. A **66**, 062315 (2002).