Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Sharp Continuity Bounds for Entropy and Conditional Entropy

by

Zhi-Hao Ma, Zhi-Hua Chen, Ismail Nikoufar, and Shao-Ming Fei

Preprint no.: 15 2017

Sharp Continuity Bounds for Entropy and Conditional Entropy

Zhihua Chen[1,2], Zhihao Ma[3,4], Ismail Nikoufar[5] and Shao-Ming Fei[6,7]

[1]Department of Science, Zhijiang College, Zhejiang University of Technology, Hangzhou 310024, China

[2]Centre for Quantum Technologies, National University of Singapore, 117543 Singapore

[3]Department of Mathematics, Shanghai Jiaotong University, Shanghai 200240, China

[4]Department of Physics and Astronomy, University College London, WC1E 6BT London, UK

[5]Department of Mathematics, Payame Noor University, 19395-3697 Tehran, Iran

[6]School of Mathematical Sciences, Capital Normal University, Beijing 100048, China

[7]Max-Planck-Institute for Mathematics in the Sciences, Leipzig 04103, Germany

Abstract

We study the continuity estimation of the Tsallis entropy. An inequality relating the Tsallis entropy difference of two quantum states to their trace norm distance is derived. This inequality is shown to be tight in the sense that equality can be attained for every prescribed value of the trace norm distance. It includes the sharp Fannes inequality for von Neumann entropy as a special case.

PACS: 03.65.Ud, 03.67.Mn

The Tsallis entropy plays an essential role in nonextensive statistics with many important findings [1]. Besides many interesting properties [2, 3], the Tsallis entropy is tightly related to many physical phenomena such as the distribution characterizing the motion of cold atoms in dissipative optical lattices [4, 5], the fluctuations of the magnetic field in the solar wind [6], the velocity distributions in driven dissipative dusty plasma [7] and spin glass relaxation [8]. The well known von Neumann entropy and linear entropy are just two special cases of the Tsallis entropy. We study the continuity estimation of the Tsallis entropy and present a tight inequality connecting the Tsallis entropy difference of two quantum states and their trace norm distance, which includes the sharp Fannes inequality for von Neumann entropy as a special case.

The von Neumann entropy of a quantum state ρ is defined by

$$S(\rho) := -\operatorname{Tr}[\rho \log_2 \rho]. \tag{1}$$

For classical probability distributions, the von Neumann entropy reduces to the Shannon entropy,

$$H(p) := \sum_{i}^{d} H(p_i) = -\sum_{i}^{d} p_i \log_2 p_i, \tag{2}$$

where $p := (p_i) = (p_1, p_2, ..., p_d)$ is a d-dimensional probability vector, $p_i \geq 0$, $\sum_i^d p_i = 1$ and $H(p_i) := -p_i \log_2 p_i$.

In [9, 10] Fannes proved the famous Fannes inequality for the continuity of the von Neumann entropy,

$$|S(\rho) - S(\sigma)| \leq 2T \log_2(d) - 2T \log_2(2T), \tag{3}$$

where $T = ||\rho - \sigma||_1/2$, $||\rho - \sigma||_1 = \mathrm{Tr}[\sqrt{(\rho - \sigma)^\dagger(\rho - \sigma)}]$ is the trace norm distance between the states ρ and σ, $T \in [0, 1]$. The inequality (3) is valid for $0 \leq T \leq 1/2e$. The inequality (3) is further improved to be a sharp one by Audenaert [11]:

$$|S(\rho) - S(\sigma)| \leq T \log_2(d - 1) + H((T, 1 - T)). \tag{4}$$

The Tsallis entropy is a more general form of the von Neumann entropy,

$$\mathcal{T}_\alpha(\rho) := \frac{1}{1 - \alpha}[\mathrm{Tr}(\rho^\alpha) - 1], \qquad \alpha > 0, \tag{5}$$

When α goes to one, the Tsallis entropy becomes the von Neumann entropy. When $\alpha = 2$, the Tsallis entropy is just the linear entropy (up to a factor 2), $\mathcal{T}_2(\rho) = 1 - \mathrm{Tr}(\rho^2)$, which is a measure of the mixedness of quantum states. In the following, we show that for the Tsallis entropy, an improved sharp Fannes type inequality exists.

[**Theorem**]. For all d-dimensional states ρ and σ,

$$|\mathcal{T}_\alpha(\rho) - \mathcal{T}_\alpha(\sigma)| \leq \frac{1}{|1 - \alpha|}[1 - (1 - T)^\alpha - (d - 1)^{1-\alpha} T^\alpha], \tag{6}$$

where T is half of the trace norm distance of ρ and σ.

In proving the Theorem, for simplicity, we denote $S_\alpha(\rho) := (1 - \alpha)T_\alpha(\rho) = [\mathrm{Tr}(\rho^\alpha) - 1]$. Hence in stead of (6), we prove the following inequality,

$$|S_\alpha(\rho) - S_\alpha(\sigma)| \leq 1 - (1 - T)^\alpha - (d - 1)^{1-\alpha} T^\alpha. \tag{7}$$

Let λ_i, $i = 1, 2, ..., d$, be the eigenvalues of ρ. One has $S_\alpha(\rho) = \sum\limits_i [(\lambda_i)^\alpha - \lambda_i] := \sum\limits_i S_\alpha(\lambda_i)$, where $S_\alpha(\lambda_i) := [(\lambda_i)^\alpha - \lambda_i]$. First we show that for any probability vectors $p = (p_i)$ and $q = (q_i)$, the following inequality holds:

$$|S_\alpha(p) - S_\alpha(q)| \le 1 - (1-T)^\alpha - (d-1)^{1-\alpha} T^\alpha, \tag{8}$$

where $T = \frac{1}{2} \sum\limits_{i=1}^{d} |p_i - q_i|$, $S_\alpha(p) = \sum\limits_i [(p_i)^\alpha - p_i] = \sum\limits_i S_\alpha(p_i)$, $S_\alpha(q) = \sum\limits_i [(q_i)^\alpha - q_i] = \sum\limits_i S_\alpha(q_i)$, $p_i \ge 0$, $q_i \ge 0$ and $\sum\limits_i p_i = \sum\limits_i q_i = 1$.

Let $q = p + \delta^+ - \delta^-$, where $\delta^+ = (\delta_i^+)$ and $\delta^- = (\delta_i^-)$ are two vectors such that $\delta_i^+ \ge 0$, $\delta_i^- \ge 0$, $i = 1, 2, ..., d$, $\delta^+ \cdot \delta^- = 0$, $\sum_i \delta_i^+ = T$.

(I) For $\alpha < 1$, $S_\alpha(p)$ is concave. Hence $S_\alpha(p + \delta^+ - \delta^-) - S_\alpha(p)$ is a concave function of δ^+. And it will get its minimum at a certain point, say, $\delta^+ = e_1 = (1, 0, \cdots, 0)^T$.

Set $p = (p_1, (1-p_1)r)$, $q = (p_1 + T, (1-p_1)r - Ts)$, where r and s are $d-1$ dimensional probability vectors such that $p_1 + T \le 1$, $(1-p_1)r - Ts \ge 0$ and $Ts = \delta^-$. We have

$$S_\alpha(q) - S_\alpha(p) = S_\alpha(p_1 + T) - S_\alpha(p_1) + S_\alpha((1-p_1)r - Ts) - S_\alpha((1-p_1)r).$$

Denote $(1-p_1)r - Ts = (1 - p_1 - T)\eta$. Then

$$S_\alpha((1-p_1)r - Ts) - S_\alpha((1-p_1)r) = S_\alpha((1-p_1-T)\eta) - S_\alpha((1-p_1-T)\eta + Ts). \tag{9}$$

Since $S_\alpha(x) - S_\alpha(x+y)$ is concave and a monotonously increasing function of x, the right hand side of (9) gets its minimum at certain point of η, say, $\eta = e_1 = (1, 0, \cdots, 0, 1)^T$ is a d dimensional vector whose first entry is 1 and other entries are 0.

Let $s = (s_1, (1-s_1)\phi)$ with ϕ a $d-2$ dimensional probability vector. We have

$$S_\alpha((1-p_1-T)\eta) - S_\alpha((1-p_1-T)\eta + Ts) = S_\alpha(1-p_1-T) - S_\alpha(1-p_1-T(1-s_1)) - S_\alpha(T(1-s_1)\phi) \triangleq \Delta.$$

As $S_\alpha(T(1-s_1)\phi)$ gets its maximum when $\phi = (1, 1, \cdots, 1)/(d-2)$ is the uniform distribution, Δ has the minimum,

$$\Delta = S_\alpha(1-p_1-T) - S_\alpha(1-p_1-T(1-s_1)) - (d-2)^{1-\alpha}(T(1-s_1))^\alpha + T(1-s_1).$$

From

$$\frac{\partial \Delta}{\partial s_1} = -T\alpha[(1-p_1-T(1-s_1))^{\alpha-1} - (d-2)^{1-\alpha}(T(1-s_1))^{1-\alpha}] = 0,$$

we get

$$T(1-s_1)=\frac{(1-p_1)(d-2)}{(d-1)}\equiv\omega. \tag{10}$$

From (10) we see that for $0<T<\omega$, there is no local minimum of Δ. For $T(1-s_1)=T$, i.e. $s_1=0$, Δ has a minimum, $-(d-2)^{(1-\alpha)}T^\alpha+T$. When $\omega\le T\le 1-p_1$, (10) can be satisfied and Δ gets it minimum $(1-p_1-T)^\alpha+T-(1-p_1)^\alpha/(d-1)^{\alpha-1}$.

Therefore when $0<T<\omega$, $S_\alpha(q)-S_\alpha(p)$ gets its minimum $S_\alpha(p_1+T)-S_\alpha(p_1)-(d-2)^{1-\alpha}T^\alpha+T$. Moreover, due to that $S_\alpha(p_1+T)-S_\alpha(p_1)$ is decreasing function of p_1, when $p_1=1-\frac{(d-1)T}{d-2}$, $S_\alpha(q)-S_\alpha(p)$ gets its minimum $(1-\frac{T}{d-2})^\alpha-(1-\frac{(d-1)T}{d-2})^\alpha-(d-2)^{1-\alpha}T^\alpha$.

When $\omega\le T<1-p_1$, $S_\alpha(q)-S_\alpha(p)$ gets its minimum

$$-S_\alpha(p_1)+S_\alpha(p_1+T)+(1-p_1-T)^\alpha-\frac{(1-p_1)^\alpha}{(d-1)^{\alpha-1}}+T.$$

The derivative of the above formula with respect to p_1 is less than zero. Hence when $p_1=1-T$, $S_\alpha(q)-S_\alpha(p)$ gets its minimum $1-(1-T)^\alpha-(d-1)^{1-\alpha}T^\alpha$.

Since $S_\alpha(x)-S_\alpha(x-T)$ is a decreasing function of x, and $1>1-\frac{T}{d-2}$, $1-(1-T)^\alpha\le(1-\frac{T}{d-2})^\alpha-(1-\frac{(d-1)T}{d-2})^\alpha$. Therefore $1-(1-T)^\alpha-(d-1)^{1-\alpha}T^\alpha\le(1-\frac{T}{d-2})^\alpha-(1-\frac{(d-1)T}{d-2})^\alpha-(d-2)^{1-\alpha}T^\alpha$. $1-(1-T)^\alpha-(d-1)^{1-\alpha}T^\alpha$ is minimum of $S_\alpha(q)-S_\alpha(p)$.

(II) When $1\le\alpha<2$, $S_\alpha(p)-S_\alpha(q)=S_\alpha(p)-S_\alpha(p+\delta^+-\delta^-)$ is concave with respect to δ^+. Take $\delta^+=e_1=(1,0,\cdots,0,1)^T$. We have

$$S_\alpha(p)-S_\alpha(q)=S_\alpha(p_1)-S_\alpha(p_1+T)+S_\alpha((1-p_1)r)-S_\alpha((1-p_1)r-Ts),$$

in which

$$S_\alpha((1-p_1)r)-S_\alpha((1-p_1)r-Ts)=S_\alpha((1-p_1-T)\eta+Ts)-S_\alpha((1-p_1-T)\eta).$$

Since $S_\alpha(x+y)-S_\alpha(x)$ is concave for $\alpha<2$, $S_\alpha((1-p_1-T)\eta+Ts)-S_\alpha((1-p_1-T)\eta)$ gets its minimum at $\eta=e_1=(1,0,\cdots,0,0)^T$.

Let $s=(s_1,(1-s_1)\phi)$ with ϕ a $d-2$ dimensional probability vector. We get

$$S_\alpha((1-p_1-T)\eta+Ts)-S_\alpha((1-p_1-T)\eta)=S_\alpha(1-p_1-T(1-s_1))+S_\alpha(T(1-s_1)\phi)-S_\alpha(1-p_1-T)\triangleq\Delta.$$

When $\phi=(1,1,\cdots,1)/(d-2)$, $S_\alpha(T(1-s_1)\phi)$ gets its minimum, and Δ gets its minimum, $\Delta=S_\alpha(1-p_1-T(1-s_1))-S_\alpha(1-p_1-T)+(d-2)^{1-\alpha}(T(1-s_1))^\alpha-T(1-s_1)$. From $\frac{\partial\Delta}{\partial s_1}=0$, we have the formula (10) again.

Δ has no local minimum for $0 < T < \omega$. For $T(1-s_1) = T$, Δ has a minimum $(d-2)^{(1-\alpha)}T^\alpha - T$. For $\omega \leq T \leq 1-p_1$, at $T(1-s_1) = \omega$, Δ gets it minimum $\frac{(1-p_1)^\alpha}{(d-1)^{\alpha-1}} - (1-p_1-T)^\alpha - T$. Correspondingly, when $0 < T < \omega$, $S_\alpha(p) - S_\alpha(q)$ gets its minimum, $S_\alpha(p_1) - S_\alpha(p_1+T) + (d-2)^{1-\alpha}T^\alpha - T$, which takes the minimum value $(1-\frac{(d-1)T}{d-2})^\alpha - (1-\frac{T}{d-2})^\alpha + (d-2)^{1-\alpha}T^\alpha$ at $p_1 = 1 - \frac{(d-1)T}{d-2}$. When $\omega \leq T < 1-p_1$, $S_\alpha(q) - S_\alpha(p)$ gets its minimum $S_\alpha(p_1) - S_\alpha(p_1+T) - (1-p_1-T)^\alpha - T + \frac{(1-p_1)^\alpha}{(d-1)^{\alpha-1}} = (1-T)^\alpha + (d-1)^{1-\alpha}T^\alpha - 1$ at $p_1 = 1-T$. Therefore

$$(1-T)^\alpha + (d-1)^{1-\alpha}T^\alpha - 1 < (1-\frac{(d-1)T}{d-2})^\alpha - (1-\frac{T}{d-2})^\alpha + (d-2)^{1-\alpha}T^\alpha$$

and (6) is valid.

(III) When $\alpha \geq 2$, $S_\alpha((1-p_1)r) - S_\alpha((1-p_1)r - Ts)$ is concave with respect to s. Hence $S_\alpha((1-p_1)r) - S_\alpha((1-p_1)r - Ts)$ gets its minimum in one of the extreme points of s, say, $s = e_1 = (1, 0, \cdots, 0)$.

Let $r = (r_1, (1-r_1)\phi)$ with ϕ a $d-2$ dimensional probability vector. Then

$$S_\alpha((1-p_1)r) - S_\alpha((1-p_1)r - Ts) = S_\alpha((1-p_1)r_1) - S_\alpha((1-p_1)r_1 - T) \triangleq \nabla.$$

Since

$$\frac{\partial \nabla}{\partial r_1} = \alpha(1-p_1)^\alpha r_1^{\alpha-1} - \alpha(1-p_1)((1-p_1)r_1 - T)^{\alpha-1} > 0,$$

when $(1-p_1)r_1 = T$, ∇ has the minimum T^α. Therefore $S_\alpha(p) - S_\alpha(q)$ takes its minimum $p_1^\alpha - (p_1+T)^\alpha + T^\alpha$. Because $p_1^\alpha - (p_1+T)^\alpha + T^\alpha$ decreases as p_1 decreases, $S_\alpha(p) - S_\alpha(q)$ takes its minimum $T^\alpha + (1-T)^\alpha - 1$ at $p_1 = 1-T$.

We have proved the inequality (8), namely the inequality (7) for the case that both states ρ and σ are diagonal ones. For general ρ and σ, the inequality can be directly proved accounting to the fact that the Tsallis entropy is unitary invariant [11].

We have investigated the continuity estimation of the Tsallis entropy, by presenting an inequality which relates the Tsallis entropy difference of two quantum states to their trace norm distance. In our inequality, equality can be attained for every prescribed value of the trace norm distance. It is direct to verify that for $\alpha \to 1$, our inequality (6) gives rise to the the sharp Fannes inequality for von Neumann entropy. Our inequality also solves the problem of the continuity estimation of linear entropy ($\alpha = 2$).

Acknowledgments This work is supported by the NSFC under number 11275131, 10901103 and 11201427.

References

[1] C. Tsallis et al., Nonextensive Statistical Mechanics and Its Applications, S. Abe and Y. Okamoto, Eds. Heidelberg, Germany: Springer-Verlag, 2001.

[2] S. Furuichi, IEEE Transactions on information theory 51 (2005) 3638.

[3] G. Ciuperca, V. Girardin, and L. Lhote, IEEE Transactions on information theory 57 (2011) 4026.

[4] E. Lutz, Phys. Rev. A 67 (2003) 51402 (R).

[5] P. Douglas, S. Bergamini, F. Renzoni, Phys. Rev. Lett. 96 (2006) 110601.

[6] L. F. Burlaga, A. F. Vinas, Physica A 356 (2005) 375.

[7] B. Liu, J. Goree, Phys. Rev. Lett. 100 (2008) 055003.

[8] R. Pickup, R. Cywinski, C. Pappas, B. Farago, P. Fouquet, Phys. Rev. Lett. 102 (2009) 097202.

[9] M. Fannes, Commun. Math. Phys. 31 (1973) 291.

[10] R. Alicki, M. Fannes, J. Phys. A: Math.Gen.37 (2004) L55.

[11] K. M R Audenaert, J. Phys. A: Math. Theor. 40 (2007) 8127.