Generalized Wigner–von Neumann entropy and its typicality

Zhigang Hu (胡志刚),[1] Zhenduo Wang (王朕铎),[1] and Biao Wu (吴飙)[1,2,3,*]

[1]International Center for Quantum Materials, School of Physics, Peking University, Beijing 100871, China
[2]Collaborative Innovation Center of Quantum Matter, Beijing 100871, China
[3]Wilczek Quantum Center, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

(Received 4 March 2019; published 14 May 2019)

We propose a generalization of the quantum entropy introduced by Wigner and von Neumann [Z. Phys. **57**, 30 (1929)]. Our generalization is applicable to both quantum pure states and mixed states. When the dimension N of the Hilbert space is large, this generalized Wigner–von Neumann (GWvN) entropy becomes independent of the choices of basis and is asymptotically equal to $\ln N$ in the sense of typicality. The dynamic evolution of our entropy is also typical, and is reminiscent of quantum H theorem proved by von Neumann. For a composite system, the GWvN entropy is typically additive; for the microcanonical ensemble, it is equivalent to the Boltzmann entropy; and for a system entangled with environment, it is consistent with the familiar von Neumann entropy, which is zero for pure states. In addition, the GWvN entropy can be used to derive the Gibbs ensemble.

DOI: 10.1103/PhysRevE.99.052117

I. INTRODUCTION

As an empirical understanding of macroscopic irreversible processes, the second law of thermodynamics has stood the test of time. However, its microscopic origin is still perplexing physicists. Boltzmann gave us a clear understanding of how the second law could arise from the time-reversible Newton's laws of motion. Unfortunately, we now know that macroscopic objects are made of microscopic particles that obey the laws of quantum mechanics. It is therefore desirable and imperative to understand the second law from the perspective of quantum dynamics [1–15].

In 1929, von Neumann made the first attempt to understand the second law quantum mechanically by proving the quantum ergodic theorem and the quantum H theorem [1,2,16]. In proving the H theorem, von Neumann introduced an entropy for quantum pure states but acknowledged that this definition came from Wigner's unpublished work [1,2]. Therefore, we call such quantum entropy Wigner–von Neumann (WvN) entropy. The well-known von Neumann entropy was not used because it is always zero for quantum pure states and it can not describe the relaxation and fluctuations in isolated macroscopic quantum systems. After introducing WvN entropy, von Neumann showed with the help of typicality arguments that for the overwhelming majority of bases, the H-theorem holds without exception for all states. The typicality argument [17–21] is mathematically known as a measure concentration [22] and Levy's lemma [5].

However, WvN entropy involves vaguely defined coarse graining, and it does not apply to systems with spins. It was shown in Ref. [11] that the quantum H theorem still holds for a generalized WvN entropy that does not involve coarse graining. In this work, we generalize WvN entropy further so that it applies to any quantum systems including spin systems. We choose the eigenstates of a given observable as a complete basis set. The generalized WvN (GWvN) entropy for a quantum state is defined with the probability distribution of this state over the chosen basis. We show analytically that the GWvN entropy of almost all quantum states in a Hilbert space of dimension N lies around $\ln N$ with a variance of order $1/N$. This means that when N is large, there is typicality for the GWvN entropy. The GWvN entropy can also be readily defined for a mixed state, for which it has a similar typical behavior. When it applies to a system entangled with the environment, the GWvN entropy is consistent with the familiar von Neumann entropy. When the quantum state can sample adequately the Hilbert space during its dynamical evolution, the GWvN entropy will typically change quickly from its initial value and saturate around $\ln N$, reminiscent of the quantum H theorem [1,2,11]. In the sense of typicality, the GWvN entropy is additive. Therefore, when it is applied to a microcanonical ensemble of N quantum states in an energy shell, the GWvN entropy is not only identical to the Boltzmann entropy but it also shares its property of being additive. In the end, we show that the Gibbs ensemble can be derived from the GWvN entropy with the maximal entropy principle [23,24].

II. GENERALIZED WIGNER–VON NEUMANN ENTROPY AND ITS TYPICALITY

In 1929, von Neumann introduced an entropy for pure quantum states, but he acknowledged that the idea was from Wigner's unpublished work [1,2]. To define the entropy, Wigner and von Neumann chose a pair of macroscopic position and macroscopic momentum operators that commute. Their common eigenstates are wave functions localized on individual Planck cells in phase space, and they form a complete basis set. A wave function is then mapped unitarily

*wubiao@pku.edu.cn

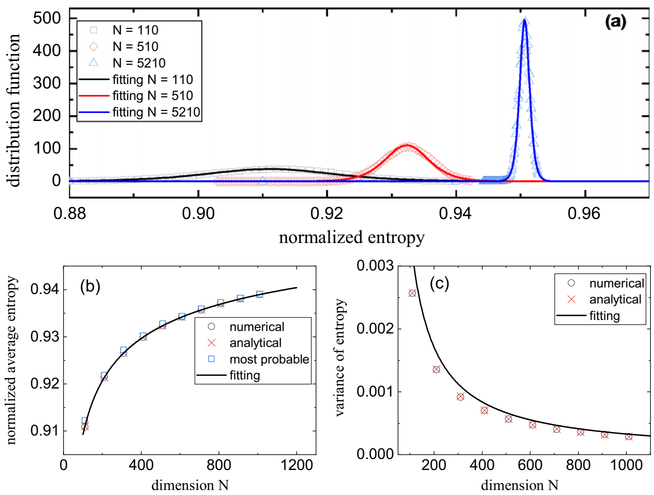

FIG. 1. Statistical properties of the GWvN entropy S over a Hilbert space of dimension N. (a) The distributions of the normalized entropy $s = S/\ln N$. (b) The average entropy obtained by numerical calculations, analytical results, and fittings. The blue rectangle represents the most probable entropy. (c) The variance of the entropy obtained by numerical calculations, analytical results, and fittings.

to the phase space with this basis. The resulting probability distribution in the phase space is coarse-grained and used to define the WvN entropy. This WvN entropy was generalized in Ref. [11] with two improvements: (i) there is no more coarse graining, and (ii) an efficient way is found to compute the basis as a set of Wannier functions [25].

Here we choose an operator A whose eigenstates $\{|\phi_i\rangle\}$ form a complete basis. For a quantum state $|\varphi\rangle$, we use its probability distribution over the basis $\{|\phi_i\rangle\}$ to define an entropy

$$S(\varphi) \equiv -\sum_{j=1}^{N} |\langle\varphi|\phi_j\rangle|^2 \ln |\langle\varphi|\phi_j\rangle|^2. \tag{1}$$

Entropy of this form has appeared in many different contexts and had different names [11,26–31]. In particular, without the minus sign, it was defined as the information of operator A in Ref. [32]. In our judgment, it is fair to regard Eq. (1) as the generalization of the quantum entropy proposed by Wigner and von Neumann in 1929 [1,2]. We will show in this work that this GWvN entropy is independent of the choice of A in the sense of typicality.

Due to the normalization rule, all quantum states in a Hilbert space of N dimensions lie on the $(2N-1)$-dimensional hypersphere $\mathbb{S}^{2N-1}$,

$$\sum_{j=1}^{N} |z_j|^2 = 1, \quad z_j = \langle\phi_j|\varphi\rangle\,. \tag{2}$$

We are interested in the statistical average, variance, and distribution of the GWvN entropy when the quantum states are sampled on this hypersphere uniformly at random. Let us look at the numerical results first. Shown in Fig. 1(a) are the distributions of the entropies for $N = 110, 510, 5210$. The figure shows that the distribution $P(S)$ becomes narrower and the average gets closer to $\ln N$ quickly as N increases. These two trends are further demonstrated in Figs. 1(b) and 1(c). These results indicate that when N is large, which is the usual case for a quantum many-body system, the GWvN entropy is the same for the overwhelming majority of quantum states with a value that is very close to $\ln N$. It is clear that these results are independent of the choice of operator A.

We find that the distribution $P(S)$ of the GWvN entropies obtained numerically in Fig. 1 can be fit very well with a function $\mathcal{P}(s) = d\,\mathrm{Pr}(s)/ds$, where $\mathrm{Pr}(s)$ is a Fermi-Dirac-like function

$$\mathrm{Pr}(s) = \frac{1}{1+\exp\left[-c_N\left(s-\frac{\mu_N}{\ln N}\right)\right]} \tag{3}$$

with $s = S/\ln N$ being the normalized GWvN entropy and the parameters

$$\mu_N = \frac{\Gamma'(N+1)}{\Gamma(N+1)} - \frac{\Gamma'(2)}{\Gamma(2)}, \quad c_N = \mu_N\sqrt{\frac{\pi^2 N}{\pi^2-9}}. \tag{4}$$

Here $\Gamma'(z)$ is the derivative of the gamma function with respect to z. As shown in Fig. 1, all three essential features of the entropy distribution—shape, average value, and variance—are captured very well by $\mathcal{P}(s)$. The empirical distribution in Eq. (3) is discussed and compared with rigorous results of the concentration of measure and Levy's lemma in Appendix A. We note that a distribution similar to Fig. 1(a) was computed in a very different context in Ref. [33].

We turn to analytical results, which can be obtained for the average value and the variance of the GWvN entropy over all quantum states on $\mathbb{S}^{2N-1}$. This is achieved with the introduction of an auxiliary function $H_\lambda = -\sum_j |z_j|^{2\lambda}$ [34], which is related to the GWvN entropy as

$$S(\varphi) = \left.\frac{dH_\lambda}{d\lambda}\right|_{\lambda=1}. \tag{5}$$

We average $|z_j|^{2\lambda}$ over all points on $\mathbb{S}^{2N-1}$ and obtain (see Appendix B for details)

$$\langle |z_j|^{2\lambda}\rangle = \frac{\Gamma(N)\Gamma(1+\lambda)}{\Gamma(N+\lambda)}. \tag{6}$$

Therefore, the average GWvN entropy reads

$$\langle S\rangle = -N\left.\frac{d\langle |z_i|^{2\lambda}\rangle}{d\lambda}\right|_{\lambda=1} = \mu_N. \tag{7}$$

Asymptotically, we have $\mu_N = \ln N + \gamma - 1 + O(1/N)$, where $\gamma \approx 0.577$ is the Euler-Mascheroni constant. This confirms our numerical and fitting results [see Fig. 1(b)]. In fact, when N is of order 10^{19} (about 63 spins in a spin-1/2 model, still a small system), $\langle S\rangle$ is already $0.99 \ln N$, less than 1% off the maximal value.

Similarly, with the auxiliary function H_λ, we find the variance of the GWvN entropy according to (7), (B7), and (B8),

$$\sigma_S^2 = \left.\frac{\partial^2\langle H_\lambda H_\zeta\rangle}{\partial\lambda\partial\zeta}\right|_{\lambda=\zeta=1} - \langle S\rangle^2$$
$$= N(N-1)\left.\frac{\partial^2\langle |z_i|^{2\lambda}|z_j|^{2\zeta}\rangle}{\partial\lambda\partial\zeta}\right|_{\lambda=\zeta=1,i\neq j}$$

$$+N\frac{\partial^2\langle|z_i|^{2(\lambda+\zeta)}\rangle}{\partial\lambda\partial\zeta}\bigg|_{\lambda=\zeta=1}-\langle S\rangle^2$$
$$=\frac{1}{N+1}\left(\frac{\pi^2}{3}-2\right)-\Psi_1(N+1), \tag{8}$$

where $\Psi_1(z)=d^2\ln\Gamma(z)/dz^2$ is a trigamma function. Asymptotically, $\Psi_1(N+1)=1/(N+1)+O(1/N^2)$. So, the leading term of the variance of GWvN entropy is $\sim 0.3/N$ [26]. It agrees with the numerical result and the fitting result [see Fig. 1(c)].

All of the above results, analytical or numerical, demonstrate clearly that there is typicality for the GWvN entropy when the dimension of the occupied Hilbert space N is large enough. In other words, for a closed many-body quantum system, the GWvN entropy of almost any of its quantum state is very close to $\ln N$ and the exception is very rare. These results are independent of the choice of operator A as long as the eigenstates of A form a complete basis. When N can be regarded as the number of microstates (e.g., energy eigenstates) of a quantum system, GWvN entropy is consistent with the well-known Boltzmann entropy. We will discuss further the physical implication of these results.

III. GWvN ENTROPY OF SUBSYSTEMS

The GWvN entropy can be extended to quantum subsystems that are described by mixed states. We consider a closed quantum system that consists of two subsystems. We focus on one of the subsystems and call the other subsystem the environment. The whole system is in a pure quantum state $|\varphi\rangle=\sum_{i,\alpha}c_{i\alpha}|\phi_i,\psi_\alpha\rangle$ with normalization $\sum_{i,\alpha}|c_{i\alpha}|^2=1$. Here the Roman indices represent the subsystem and the Greek indices the environment. By convention, $\{\phi_i, i=1,2,\dots,n\}$ is a complete basis for the subsystem and $\{\psi_\alpha,\alpha=1,2,\dots,m\}$ is a complete basis for the environment. Tracing out the environment in the density matrix of the system, $\rho=\sum_{ij,\alpha\beta}c_{i\alpha}c^*_{j\beta}|\phi_i,\psi_\alpha\rangle\langle\phi_j,\psi_\beta|$, we obtain the reduced density matrix for the subsystem,

$$\rho_s=\mathrm{Tr}_e\rho=\sum_{ij}p_{ij}|\phi_i\rangle\langle\phi_j|, \tag{9}$$

where $p_{ij}=\sum_{\alpha=1}^m c_{i\alpha}c^*_{j\alpha}$. The definition of GWvN entropy for the subsystem is

$$S(\rho_s)=-\sum_{i=1}^n\mathrm{Tr}(\rho_s\,|\phi_i\rangle\,\langle\phi_i|)\ln\mathrm{Tr}(\rho_s\,|\phi_i\rangle\,\langle\phi_i|)$$
$$=-\sum_{i=1}^n p_{ii}\ln p_{ii}. \tag{10}$$

We are interested in the average value and the variance of $S(\rho_s)$ when the quantum state $|\varphi\rangle$ of the whole system is sampled uniformly over the hypersphere $\mathbb{S}^{2N-1}$. Here $N=nm$.

We introduce another auxiliary function $K_\lambda=-\sum_{i=1}^n p_{ii}^\lambda$, which is related to the entropy as

$$S(\rho_s)=\frac{dK_\lambda}{d\lambda}\bigg|_{\lambda=1}. \tag{11}$$

Direct computation (see Appendix B for details) shows that

$$\langle p_{ii}^\lambda\rangle=\frac{\Gamma(mn)\Gamma(m+\lambda)}{\Gamma(mn+\lambda)\Gamma(m)}. \tag{12}$$

When $\lambda=1$, we have $\langle p_{ii}\rangle=1/n$ as expected. This leads to

$$-\frac{d\langle p_{ii}^\lambda\rangle}{d\lambda}\bigg|_{\lambda=1}=\frac{1}{n}\left[\frac{\Gamma'(mn+1)}{\Gamma(mn+1)}-\frac{\Gamma'(m+1)}{\Gamma(m+1)}\right]. \tag{13}$$

With Eq. (11) we have the average of the GWvN entropy,

$$\langle S(\rho_s)\rangle=\frac{\Gamma'(mn+1)}{\Gamma(mn+1)}-\frac{\Gamma'(m+1)}{\Gamma(m+1)}\approx\ln n, \tag{14}$$

where the approximation is asymptotic and it holds when both n and m are very large. The variance of $S(\rho_s)$ can also be derived analytically. With Eqs. (14), (B7), and (B8), we have

$$\sigma^2_{S(\rho_s)}=\frac{\partial^2\langle K_\lambda K_\zeta\rangle}{\partial\lambda\partial\zeta}\bigg|_{\lambda=\zeta=1}-\langle S(\rho_s)\rangle^2$$
$$=n(n-1)\frac{\partial^2\langle p_{ii}^\lambda p_{jj}^\zeta\rangle}{\partial\lambda\partial\zeta}\bigg|_{\lambda=\zeta=1,i\neq j}$$
$$+n\frac{\partial^2\langle p_{ii}^{\lambda+\zeta}\rangle}{\partial\lambda\partial\zeta}\bigg|_{\lambda=\zeta=1}-\langle S(\rho_s)\rangle^2$$
$$=\frac{m+1}{N+1}\Psi_1(m+1)-\Psi_1(N+1). \tag{15}$$

When $N>m\gg 1$, we have

$$\sigma^2_{S(\rho_s)}\approx\frac{1}{2mN}+O\left(\frac{1}{N^2}\right). \tag{16}$$

Here we have used the condition $m>n$, which is usually the case. This shows that the variance of a subsystem's entropy is effectively controlled by the environment and the whole system. The reason is that we are averaging over the Hilbert space of the whole system, where the overwhelming majority of quantum states are almost maximally entangled [35]. If the subsystem and the environment are not entangled, then the subsystem can be regarded as an isolated system, which was already discussed in the preceding section. It is clear from these results that the GWvN entropy of a subsystem has typicality. Similarly, the GWvN entropy of the environment is typical, with the average being $\langle S(\rho_e)\rangle\approx\ln m$ and the variance of order $1/(nN)$.

It is interesting to compare these results on mixed states with the familiar von Neumann entropy. For this purpose, we assume that $m\gg n$ since the environment should usually be much larger than the subsystem. Page computed the average value of the von Neumann entropy $S_v(\rho_s)=-\mathrm{Tr}\rho_s\ln\rho_s$ when the quantum state $|\varphi\rangle$ is sampled randomly [35]. He found that the average value is approximately $\ln n-n/(2m)$, which is consistent with the GWvN entropy asymptotically. Page did not compute the variance of $S_v(\rho_s)$. This shows that the GWvN entropy for a system with a large environment is consistent with the von Neumann entropy, which was already noticed in Ref. [11]. However, for the environment, its von Neumann entropy $S_v(\rho_e)=S_v(\rho_s)\approx\ln n$, very different from its GWvN entropy $\ln m$.

Strictly speaking, the GWvN entropy is not additive. However, in the sense of typicality it is additive. This is evident in the following result:

$$\begin{aligned}\langle S(\rho_s)+S(\rho_e)-S(\rho)\rangle\\ &=\frac{\Gamma'(mn+1)}{\Gamma(mn+1)}-\frac{\Gamma'(m+1)}{\Gamma(m+1)}-\frac{\Gamma'(n+1)}{\Gamma(n+1)}+\frac{\Gamma'(2)}{\Gamma(2)}\\ &=1-\gamma+O\left(\frac{m+n}{mn}\right).\end{aligned} \tag{17}$$

When N, m, and n are large, we can safely ignore the constant $1-\gamma$ and therefore have $\langle S(\rho_s)\rangle+\langle S(\rho_e)\rangle=\langle S(\rho)\rangle$. The result can be easily generalized to multipartite systems where each subsystem has a Hilbert space of large dimension.

The dimension N of the Hilbert space in the above discussion should be regarded as the dimension of a subspace that is physically relevant. As an example, and also as an important application, let us consider a quantum microcanonical ensemble, which is characterized by an energy shell $[E, E+\Delta E]$ with a large but finite number of energy eigenstates [36,37]. Suppose that N is the number of energy eigenstates in the shell. According to the above results, the GWvN entropy of any quantum state (pure or mixed) in the microcanonical ensemble is typically $\ln N$ with a very small variance. This is consistent with the well known Boltzmann entropy.

In the above discussion, we have fixed the basis $\{|\phi_i\rangle\}$ (or operator A) while sampling the quantum states in the Hilbert space. It is clear that we can equivalently fix the quantum state while sampling all possible bases. In fact, this is exactly what von Neumann did in his 1929 paper [1,2].

IV. DYNAMIC EVOLUTION OF GWvN ENTROPY

Our results for the GWvN entropy so far are kinematic and have nothing to do with the Hamiltonian of a quantum system. In this section, we investigate how the GWvN entropy evolves dynamically, and we find that the GWvN entropy has dynamical typicality, similar to the dynamical behavior of observables found in Refs. [38,39]. For an isolated quantum system with a set of energy eigenstates $|E_j\rangle$, its dynamical evolution is given by

$$|\varphi(t)\rangle=\sum_j a_j e^{-iE_jt/\hbar}\,|E_j\rangle\,, \tag{18}$$

where $a_j=\langle E_j|\Psi(0)\rangle$ is determined by the initial quantum state $|\varphi(0)\rangle$. In general, because a_j quickly approaches zero as $j\to\infty$, there are only a finite number of energy eigenstates occupied. During the dynamical evolution, this number of occupied states does not change as a_j is independent of time. This means that when a quantum system evolves dynamically, its dynamical path in Hilbert space will lie entirely in this sub-Hilbert space of occupied states.

Let us consider a quantum system with a macroscopic number of particles. In this case, the dimension N of the sub-Hilbert space of occupied states is in general enormously large. It is natural to expect that almost all the quantum states on the dynamical path are typical and their GWvN entropies are very close to $\ln N$ with small fluctuations. This is exactly what is implied in the quantum H theorems proved in Refs. [1,2,11]. However, according to these proofs, there are exceptions that occur when the system's Hamiltonian has a great deal of degeneracy in its eigenenergies and eigenenergy differences. These degeneracies are shown to be closely connected to the integrability of the Hamiltonians [40]. In other words, when the system is integrable, its quantum dynamics will be restricted by various good quantum numbers and cannot adequately sample the sub-Hilbert space. When the system is nonintegrable, its quantum dynamics can adequately sample the sub-Hilbert space so that the quantum states involved in the dynamics are typical.

We now show how the GWvN entropy relaxes dynamically in a quantum chaotic system. We choose a complete orthonormal basis, $\{|\psi_0\rangle,|\xi_i\rangle\}(i=1,\dots,N-1)$, where $|\psi_0\rangle$ is the initial state. In this case, the GWvN entropy is zero initially. As the system evolves under a unitary operator $U(t)=\exp(-iHt/\hbar)$, its GWvN entropy changes with time as

$$\overline{S(t)}=-\frac{d}{d\beta}\left[p^\beta(t)+\sum_{i=1}^{N-1}\overline{|\langle\psi(t)|\xi_i\rangle|^{2\beta}}\right]_{\beta=1}, \tag{19}$$

where $p(t)=|\langle\psi|\psi(t)\rangle|^2$ is the surviving probability (also called fidelity) [41,42] and the overline denotes averaging over the other $(N-1)$ basis $\{|\xi_i\rangle\}$. The averaging is justified by the fact that the system is quantum chaotic, and its dynamics can sample adequately in the subspace spanned by $\{|\xi_i\rangle\}$.

Note that $|\langle\psi(t)|\xi_i\rangle|^2=[1-p(t)]|\langle\psi'(t)|\xi_i\rangle|^2$, where $|\psi'(t)\rangle$ is the normalized state projected by $|\psi(t)\rangle$ onto the subspace spanned by $\{|\xi_i\rangle\}$. Similar to Eq. (6), we have

$$\overline{|\langle\psi(t)|\xi_i\rangle|^{2\beta}}=[1-p(t)]^\beta\frac{\Gamma(N-1)\Gamma(1+\beta)}{\Gamma(N-1+\beta)}. \tag{20}$$

Eventually we find

$$\begin{aligned}\overline{S(t)}&=f(p(t))+[1-p(t)]\left[\frac{\Gamma'(N)}{\Gamma(N)}-\frac{\Gamma'(2)}{\Gamma(2)}\right]\\ &=f(p(t))+[1-p(t)]\mu_{N-1},\end{aligned} \tag{21}$$

where $f(p)=-p\ln p-(1-p)\ln(1-p)$ and asymptotically $\mu_{N-1}=\ln(N-1)+\gamma-1+O(1/N)$. The deviation is negligible with the order $O(1/N)$ by typicality arguments. More precisely, it is a direct result from (8). A very interesting fact is that this expression of $S(t)$ explicitly verifies the validity of the conjectured form proposed by Flambaum and Izrailev [43].

V. CONNECTION TO THERMODYNAMICS

The GWvN entropy can also be used to derive the Gibbs ensemble with the maximal entropy principle [23,24]. We choose the operator $A=H$, and its eigenstates $|E_i\rangle$ form a complete basis. For a typical quantum state ρ, its GWvN entropy is $S=-\sum_i \mathrm{Tr}(\rho P_{E_i})\ln\mathrm{Tr}(\rho P_{E_i})$, where $P_{E_i}=|E_i\rangle\langle E_i|$. We want to maximize it under the condition that $\langle H\rangle=\mathrm{Tr}(\rho H)$ is constant. Mathematically, this can be done as

$$\frac{d}{d\varepsilon}\sum_i\mathrm{Tr}\big[(\rho+\varepsilon\Omega)P_{E_i}\big]\ln\big\{\mathrm{Tr}\big[(\rho+\varepsilon\Omega)P_{E_i}\big]\big\}\Big|_{\varepsilon=0}=0, \tag{22}$$

with the restriction $\mathrm{Tr}(\Omega H)=0$ and $\mathrm{Tr}(\Omega)=0$. This leads to

$$\mathrm{Tr}\left(\Omega\sum_i \ln\left[\mathrm{Tr}\left(\rho P_{E_i}\right)\right]P_{E_i}\right)=0. \tag{23}$$

It can be inferred that $\sum_i \ln[\mathrm{Tr}(\rho P_{E_i})]P_{E_i}$ is a linear function of H. Therefore, the density matrix has the general form

$$\rho=\alpha\exp(-\beta H), \tag{24}$$

where α is the normalization factor and β is the inverse temperature. Explicitly, we have

$$\alpha=\frac{1}{\mathrm{Tr}[\exp(-\beta H)]},\quad \beta=\frac{\delta S}{\delta\langle H\rangle}. \tag{25}$$

From a standard textbook in statistical physics [36,37], Eq. (24) represents the well-known Gibbs ensemble. And from (25) we have

$$\Delta\langle H\rangle=\frac{\delta\langle H\rangle}{\delta S}\Delta S=T\Delta S, \tag{26}$$

where the temperature $T=1/\beta$. This is exactly the first law of thermodynamics. Similar derivations can be done for the general operator A and its eigenstates $|\phi_i\rangle$ with the restriction of fixed $\langle A\rangle$. We would then have the generalized Gibbs ensemble $\rho=\alpha\exp(-\lambda_A A)$, where α is a normalization factor and λ_A is the generalized inverse temperature. With the Gibbs ensemble, other thermodynamical relations can also be easily derived. Our result here is consistent with previous results in Refs. [18,29,44].

VI. CONCLUSION

We have generalized the quantum entropy proposed by Wigner and von Neumann in 1929. Although the definition uses a specific complete basis, the generalized Wigner–von Neumann entropy becomes typical when the dimension N of the Hilbert space is large. We have shown analytically that when we sample a quantum state uniformly at random in the Hilbert space, the average of its GWvN entropy is asymptotic to $\ln N$ and its variance is of order $1/N$. As a result, when the GWvN entropy is applied to a microcanonical ensemble, it is equivalent to the Boltzmann entropy. When it is extended to a subsystem with a large environment, it is consistent with the von Neumann entropy. In the end, with the maximal entropy principle we have shown that the GWvN entropy can be used to obtain the Gibbs ensemble from which all thermodynamic relations can be derived.

ACKNOWLEDGMENTS

We acknowledge stimulating discussion with Chushun Tian. This work was supported by The National Key R&D Program of China (Grants No. 2017YFA0303302 and No. 2018YFA0305602).

Z.H. and Z.W. contributed equally to this work.

APPENDIX A: CONCENTRATION OF MEASURE AND LEVY'S LEMMA

The concentration of measure states as a function that depends in a Lipschitz way on many independent variables is almost constant. As a special form of concentration of measure, Levy's lemma is commonly used in the typicality related literature [14,33,45]. Here we compare the distribution of the GWvN entropy S [Eq. (3)] to Levy's lemma. According to Levy's lemma, for entropy $S(\varphi)$, where φ is a random point drawn uniformly from hypersphere $\mathbb{S}^{2N-1}$, the upper bound of the deviations from an expected value $\langle S\rangle$ is given by

$$\Pr[|S(\varphi)-\langle S\rangle|\geqslant\delta]\leqslant 2\exp\left(\frac{-2N\delta^2}{9\pi^3\eta^2}\right), \tag{A1}$$

where $\eta=\sup|\nabla_\varphi S|$ is a Lipschitz constant. Straightforward computation shows that

$$\begin{aligned}\eta^2&=4\sum_{j=1}^N|z_j|^2(\ln|z_j|^2+1)^2\\&=4\left(1+2\sum_{j=1}^N|z_j|^2\ln|z_j|^2+\sum_{j=1}^N|z_j|^2\ln^2|z_j|^2\right)\\&\leqslant 4(1-\ln N)^2.\end{aligned} \tag{A2}$$

The last inequality is true for dimension $N>e^2$, and the equality holds only when those $|z_j|^2$ are all the same. This means that $\Pr(|S-\langle S\rangle|\geqslant\delta)$ has a sub-Gaussian tail, which presents stronger convergence than a Poisson-like tail in the distribution of Eq. (3). However, the GWvN entropy is far away from zero only in a small region given by $|S-\mu_N|\sim\ln N/c_N\sim N^{-1/2}$. In this region, our distribution is very accurate, as indicated in Fig. 1(a).

Moreover, the variance of our distribution [Eq. (3)] is consistent with Levy's lemma. Among all the distributions that satisfy Levy's lemma, the one that has the maximal variance should be given by

$$\rho(x=|S-\langle S\rangle|)=\begin{cases}0, & x<\sqrt{\ln 2/f_N},\\ 4f_Nxe^{-f_Nx^2}, & x\geqslant\sqrt{\ln 2/f_N},\end{cases} \tag{A3}$$

where $f_N=2N/9\pi^3\eta^2$ is the factor in the exponent of Eq. (A1). It leads to the maximal variance that Levy's lemma allows,

$$\mathrm{var}_{\max}=\int_{\sqrt{\ln 2/f_N}}^{+\infty}4f_Nx^3\exp(-f_Nx^2)dx \tag{A4}$$

$$=\frac{9\pi^3(1+\ln 2)\eta^2}{2N}. \tag{A5}$$

The variance (of order $1/N$) of our distribution gives a tighter bound than the above variance bound (of order η^2/N, namely $\ln^2 N/N$).

APPENDIX B: AVERAGING IN HILBERT SPACE

In this Appendix, we give out the derivation details that are needed in the main text, particularly those related to Eqs. (6), (8), (12), and (15).

Any quantum state $|\varphi\rangle$ in an N-dimensional Hilbert space can be expanded in an orthonormal basis with coefficients $z_1,\dots,z_N$. z_i is usually a complex number, and we denote the real part and the imaginary part as x_{2i-1} and x_{2i}. Due to the normalization, the quantum state $|\varphi\rangle$

corresponds to point $(x_1, x_2, \dots, x_{2N})$ on the $(2N-1)$-dimensional hypersphere $\mathbb{S}^{2N-1}$.

The average of a function in N-dimensional Hilbert space is to integrate it on the hypersphere $\mathbb{S}^{2N-1}$. We can connect the integration on the sphere with an integration in the ball [46,47]. More precisely, let $B^n = \{x \in R^n : |x| \leqslant 1\}$. If $f : B^n \to R$ is continuous, then

$$\int_{B^n} f(x) dx_1 dx_2, \dots, dx_n = \int_0^1 r^{n-1} dr \int_{S^{n-1}} f(rs) d\sigma_{n-1}(s), \tag{B1}$$

where $d\sigma_{n-1}(s)$ denotes an element on hypersphere S^{n-1}. If $f(rx) = r^\alpha f(x)$ is a homogeneous function of degree α, we have

$$\int_{B^n} f(x) d(x_1, \dots, x_n) = \frac{1}{\alpha + n} \int_{S^{n-1}} f(s) d\sigma_{n-1}(s). \tag{B2}$$

Consider a function

$$g(x_1, \dots, x_{2N}) = \prod_{k=1}^{n} T_k^{\alpha_k}, \tag{B3}$$

where $N = nm$ and

$$T_k = \sum_{i=1}^{m} \left(x^2_{2(k-1)m+2i-1} + x^2_{2(k-1)m+2i}\right). \tag{B4}$$

Then the average $\langle g \rangle$ of $g(x_1, \dots, x_{2N})$ on sphere $\mathbb{S}^{2N-1}$ reads

$$\begin{aligned}
\langle g \rangle &= \frac{1}{\Omega_{2N-1}} \int_{\mathbb{S}^{2N-1}} \prod_{k=1}^{n} T_k^{\alpha_k} d\sigma_{2N-1}(s) = \frac{\Gamma(N)}{2\pi^N} (2\alpha_1 + 2\alpha_2 + \cdots + 2\alpha_n + 2N) \int_{B^{2N}} \prod_{k=1}^{n} T_k^{\alpha_k} dx_1 dx_2 \cdots dx_{2N} \\
&= \frac{\Gamma(N)(\alpha_1 + \cdots + \alpha_n + N)}{\pi^N} \int_{B^{2N-2m}} \prod_{k=1}^{n-1} T_k^{\alpha_k} \int_0^{\sqrt{1-\sum_{k=1}^{n-1} T_k}} \rho^{2\alpha_n} \rho^{2m-1} d\rho\, d\sigma_{2m-1}\, dx_1 dx_2 \cdots dx_{2N-2m} \\
&= \frac{\pi^m \Gamma(N)(\alpha_1 + \cdots + \alpha_n + N)}{\pi^N \Gamma(m)(m + \alpha_n)} \int_{B^{2N-2m}} \prod_{k=1}^{n-1} T_k^{\alpha_k} \left(1 - \sum_{k=1}^{n-1} T_k\right)^{m+\alpha_n} dx_1 dx_2 \cdots dx_{2N-2m} \\
&= \frac{\Gamma(N)(\alpha_1 + \cdots + \alpha_n + N)}{\pi^{N-m} \Gamma(m)(m + \alpha_n)} \int_0^1 r^{2(\alpha_1 + \cdots + \alpha_{n-1} + N - m - \frac{1}{2})} (1 - r^2)^{m+\alpha_n} dr \int_{\mathbb{S}^{2N-2m-1}} \prod_{k=1}^{n-1} T_k^{\alpha_k} d\sigma_{2N-2m-1} \\
&= \frac{\Gamma(N)(\alpha_1 + \cdots + \alpha_n + N)}{\pi^{N-m} \Gamma(m)(m + \alpha_n)} \frac{\Gamma(\alpha_1 + \cdots + \alpha_{n-1} + N - m)\Gamma(\alpha_n + m + 1)}{2\Gamma(\alpha_1 + \cdots + \alpha_n + N + 1)} \frac{2\pi^{N-m}}{\Gamma(N - m)} \left\langle \prod_{k=1}^{n-1} T_k^{\alpha_k} \right\rangle_{2N-2m} \\
&= \frac{\Gamma(N)}{\Gamma(m)\Gamma(N - m)} \frac{\Gamma(\alpha_1 + \cdots + \alpha_{n-1} + N - m)\Gamma(\alpha_n + m)}{\Gamma(\alpha_1 + \cdots + \alpha_n + N)} \left\langle \prod_{k=1}^{n-1} T_k^{\alpha_k} \right\rangle_{2N-2m} \\
&= \frac{\Gamma(N)}{\Gamma^n(m)} \frac{\Gamma(\alpha_n + m)\Gamma(\alpha_{n-1} + m) \cdots \Gamma(\alpha_1 + m)}{\Gamma(\alpha_1 + \cdots + \alpha_n + N)} \langle T_1^{\alpha_1} \rangle_{2m} \\
&= \frac{\Gamma(N)}{\Gamma^n(m)} \frac{\Gamma(\alpha_n + m)\Gamma(\alpha_{n-1} + m) \cdots \Gamma(\alpha_1 + m)}{\Gamma(\alpha_1 + \cdots + \alpha_n + N)},
\end{aligned} \tag{B5}$$

where $\Omega_{2N-1} = \frac{2\pi^N}{\Gamma(N)}$ is the surface area of $\mathbb{S}^{2N-1}$.

It is useful to calculate some derivatives of special forms of $g(x_1, \dots, x_{2N}) = \prod_{k=1}^{n} T_k^{\alpha_k}$,

$$\frac{d}{d\alpha_1} \langle T_1^{\alpha_1} \rangle = \frac{\Gamma(N)\Gamma(\alpha_1 + m)}{\Gamma(m)\Gamma(\alpha_1 + N)} [\Psi(\alpha_1 + m) - \Psi(\alpha_1 + N)], \tag{B6}$$

$$\frac{\partial^2}{\partial\alpha_1 \partial\alpha_2} \langle T_1^{\alpha_1+\alpha_2} \rangle = \frac{\Gamma(N)\Gamma(\alpha_1 + \alpha_2 + m)}{\Gamma(m)\Gamma(\alpha_1 + \alpha_2 + N)} ([\Psi(\alpha_1 + \alpha_2 + m) - \Psi(\alpha_1 + \alpha_2 + N)]^2 + [\Psi_1(\alpha_1 + \alpha_2 + m) - \Psi_1(\alpha_1 + \alpha_2 + N)]), \tag{B7}$$

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha_1 \partial\alpha_2} \langle T_1^{\alpha_1} T_2^{\alpha_2} \rangle = \frac{\Gamma(N)\Gamma(\alpha_1 + m)\Gamma(\alpha_2 + m)}{\Gamma^2(m)\Gamma(\alpha_1 + \alpha_2 + N)} &([\Psi(\alpha_1 + m) - \Psi(\alpha_1 + \alpha_2 + N)] \\
&\times [\Psi(\alpha_2 + m) - \Psi(\alpha_1 + \alpha_2 + N)] - \Psi_1(\alpha_1 + \alpha_2 + N)).
\end{aligned} \tag{B8}$$

Here $\Psi(z) = \Gamma'(z)/\Gamma(z)$ and $\Psi_1(z) = d^2 \ln \Gamma(z)/dz^2$ are digamma and trigamma functions, respectively.

[1] J. v. Neumann, Z. Phys. **57**, 30 (1929).
[2] J. von Neumann, Eur. Phys. J. H **35**, 201 (2010).
[3] M. Srednicki, Phys. Rev. E **50**, 888 (1994).
[4] J. Gemmer and G. Mahler, Eur. Phys. J. B **31**, 249 (2003).
[5] S. Popescu, A. J. Short, and A. Winter, Nat. Phys. **2**, 754 (2006).
[6] M. Rigol, V. Dunjko, and M. Olshanii, Nature (London) **452**, 854 (2008).
[7] P. Reimann, Phys. Rev. Lett. **101**, 190403 (2008).
[8] C. Gogolin, arXiv:1003.5058.
[9] S. Goldstein, T. Hara, and H. Tasaki, Phys. Rev. Lett. **111**, 140401 (2013).
[10] S. Goldstein, T. Hara, and H. Tasaki, New J. Phys. **17**, 045002 (2015).
[11] X. Han and B. Wu, Phys. Rev. E **91**, 062106 (2015).
[12] C. Tian, K. Yang, and J. Wang, arXiv:1606.08371.
[13] C. Gogolin and J. Eisert, Rep. Prog. Phys. **79**, 056001 (2016).
[14] P. Fang, L. Zhao, and C. Tian, Phys. Rev. Lett. **121**, 140603 (2018).
[15] C. Tian, K. Yang, P. Fang, H.-J. Zhou, and J. Wang, Phys. Rev. E **98**, 060103(R) (2018).
[16] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Eur. Phys. J. H **35**, 173 (2010).
[17] S. Lloyd, Ph.D. thesis, Rockefeller University, 1988, arXiv:1307.0378.
[18] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, Phys. Rev. Lett. **96**, 050403 (2006).
[19] P. Reimann, Phys. Rev. Lett. **99**, 160404 (2007).
[20] P. Reimann, New J. Phys. **12**, 055027 (2010).
[21] H. Tasaki, J. Stat. Phys. **163**, 937 (2016).
[22] V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces* (Springer, Berlin, 1986).
[23] E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[24] J. von Neumann, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1927**, 273 (1927).
[25] Y. Fang, F. Wu, and B. Wu, J. Stat. Mech.: Theor. Exp. (2018) 023113.
[26] B. Mirbach and H. Korsch, Ann. Phys. **265**, 80 (1998).
[27] A. Polkovnikov, Ann. Phys. **326**, 486 (2011).
[28] L. F. Santos, A. Polkovnikov, and M. Rigol, Phys. Rev. Lett. **107**, 040601 (2011).
[29] F. Anzà and V. Vedral, Sci. Rep. **7**, 44066 (2017).
[30] D. Safranek, J. M. Deutsch, and A. Aguirre, Phys. Rev. A **99**, 012103 (2019).
[31] S. Alterman, J. Choi, R. Durst, S. M. Fleming, and W. K. Wootters, J. Phys. A **51**, 345301 (2018).
[32] H. Everett III, in *The Many-Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and N. Graham (Princeton University Press, Princeton, NJ, 1973).
[33] U. Singh, L. Zhang, and A. K. Pati, Phys. Rev. A **93**, 032125 (2016).
[34] K. R. W. Jones, J. Phys. A **23**, L1247 (1990).
[35] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[36] L. D. Landau and E. Lifshitz, *Statistical Physics*, 3rd ed., Pt. 1, Course of Theoretical Physics Volume 5 (Butterworth-Heinemann, Oxford, 1980).
[37] K. Huang, *Statistical Mechanics* (Wiley, New York, 1987).
[38] C. Bartsch and J. Gemmer, Phys. Rev. Lett. **102**, 110403 (2009).
[39] P. Reimann, Phys. Rev. Lett. **120**, 230601 (2018).
[40] D. Zhang, H. T. Quan, and B. Wu, Phys. Rev. E **94**, 022150 (2016).
[41] V. V. Flambaum and F. M. Izrailev, Phys. Rev. E **64**, 026124 (2001).
[42] E. J. Torres-Herrera, M. Vyas, and L. F. Santos, New J. Phys. **16**, 063010 (2014).
[43] V. V. Flambaum and F. M. Izrailev, Phys. Rev. E **64**, 036220 (2001).
[44] F. Anza, C. Gogolin, and M. Huber, Phys. Rev. Lett. **120**, 150603 (2018).
[45] P. Hayden, D. W. Leung, and A. Winter, Commun. Math. Phys. **265**, 95 (2006).
[46] J. A. Baker, Am. Math. Mon. **104**, 36 (1997).
[47] G. B. Folland, Am. Math. Mon. **108**, 446 (2001).