Clustering of Conditional Mutual Information for Quantum Gibbs States above a Threshold Temperature

Tomotaka Kuwahara*
Mathematical Science Team, RIKEN Center for Advanced Intelligence Project (AIP), 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan and Interdisciplinary Theoretical & Mathematical Sciences Program (iTHEMS) RIKEN 2-1, Hirosawa, Wako, Saitama 351-0198, Japan

Kohtaro Kato†
Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA

Fernando G. S. L. Brandão‡
Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA and Amazon Web Services, USA

(Received 4 November 2019; revised manuscript received 31 January 2020; accepted 8 May 2020; published 1 June 2020)

We prove that the quantum Gibbs states of spin systems above a certain threshold temperature are approximate quantum Markov networks, meaning that the conditional mutual information decays rapidly with distance. We demonstrate the exponential decay for short-ranged interacting systems and power-law decay for long-ranged interacting systems. Consequently, we establish the efficiency of quantum Gibbs sampling algorithms, a strong version of the area law, the quasilocality of effective Hamiltonians on subsystems, a clustering theorem for mutual information, and a polynomial-time algorithm for classical Gibbs state simulations.

DOI: 10.1103/PhysRevLett.124.220601

Introduction.—Quantum Gibbs states describe the thermal equilibrium properties of quantum systems. The advent of quantum information science opened up new investigation avenues in the study of Gibbs states, such as the stability of topological quantum memory [1–4], thermalization in isolated quantum systems [5–10], and Hamiltonian complexity [11–14]. Efficient methods to prepare quantum Gibbs states in quantum computers have also proved useful in giving quantum speed-ups for problems such as semidefinite programming [15–17] and quantum machine learning [18–22].

Quantum Gibbs states also inherit the locality of their parent Hamiltonians, which allows for an efficient classical description in many cases. One of the simple characterizations is the exponential decay of bipartite correlation functions, which is true in general one-dimensional quantum spin lattices [23] and in higher dimensions above a threshold temperature [24–28]. Another characterization is that at arbitrary finite temperatures, the mutual information between a region and its complement obeys the area law [29]. Quantum Gibbs states also have efficient representations in terms of tensor networks [30,31].

In classical systems, there are even stronger structural results for Gibbs states. For instance, the Hammersley-Clifford theorem [32] states that classical Gibbs states are equivalent to a class of probability distributions called Markov networks. They satisfy the Markov property; that is, a site is independent from all others conditioned on its neighbors. Therefore, for classical Gibbs states, all the correlations between two separated vertices are induced by the intermediate vertices connecting them.

Although the notion of conditional probability distribution is missing in quantum systems, we can still generalize Markov networks to quantum systems using the (quantum) conditional mutual information:

$$\mathcal{I}_\rho(A:C|B) \coloneqq S(\rho^{AB}) + S(\rho^{BC}) - S(\rho^{ABC}) - S(\rho^{B}), \quad (1)$$

where ρ^{AB} is the reduced density matrix in the subsystem $(AB = A \cup B)$, and $S(\rho^{AB})$ is the von Neumann entropy, namely, $S(\rho^{AB}) \coloneqq -\mathrm{tr}(\rho^{AB}\log\rho^{AB})$. In classical systems, the conditional mutual information becomes zero if and only if the state is conditionally independent. In quantum cases, the conditional mutual information is related to the approximate recoverability [33]; hence, it is widely used as the measure of conditional independence in quantum systems.

The quantum version of the Hammersley-Clifford theorem has been established for the case where the Hamiltonian is short range and commuting [34,35]: Any quantum Gibbs state of such a Hamiltonian on a triangle-free graph is a Markov network and vise versa.

More recently, it has been shown that the Hammersley-Clifford theorem *approximately* holds in a one-dimensional lattice [36], in the sense that the conditional mutual information of any Gibbs state decays subexponentially with respect to distance.

In the present Letter, we will establish the approximate Markov property for quantum Gibbs states in spin systems interacting on generic graphs. In our Letter, we consider not only short-range interactions but also long-range (i.e., power-law decaying) interactions on graphs. We prove that above a certain threshold temperature, the conditional mutual information decays exponentially (polynomially) for short-ranged (long-ranged) models. Our result will strengthen the 1D result obtained in Ref. [36], the area law for mutual information [29], and the standard clustering theorem [27,28]. Moreover, our result immediately implies a quasi-polynomial-time quantum Gibbs sampling algorithm by following the discussion in Ref. [37]. Finally, for computing thermodynamic quantities (e.g., the partition function), we develop a polynomial-time classical algorithm for the first time.

Setup.—We consider a quantum system with n spins, where each spin has a d-dimensional Hilbert space. We assume that the spins sit on the vertices of a graph $G = (V, E)$ where V is the total spin set ($|V| = n$). For arbitrary subsystems $A, B \subset V$, we define $d_{A,B}$ as the shortest path length on the graph that connects A and B. If $A \cap B \neq \emptyset$, $d_{A,B} = 0$. We define the surface region of an arbitrary subsystem $L \subseteq V$ as $\partial L_l \subseteq V$ ($l \in \mathbb{N}$):

$$\partial L_l := \{v \in L | d_{v,L^c} \leq l\}, \tag{2}$$

where L^c is the complementary set of L (i.e., $L \cup L^c = V$).

We define the system Hamiltonian H as

$$H = \sum_{|X| \leq k} h_X, \tag{3}$$

where each interaction term $\{h_X\}$ acts on the spins in $X \subset V$. The Hamiltonian [Eq. (3)] describes generic k-body interacting systems. We characterize the locality of the interactions as follows:

$$\sum_{\substack{X | X \ni v \\ \mathrm{diam}(X) \geq R}} \|h_X\| \leq f(R) \quad \text{with} \quad f(1) \leq g \tag{4}$$

for $\forall\, v \in V$, where $\|\cdot\|$ is the operator norm and $\mathrm{diam}(X)$ is the diameter of X, namely, $\mathrm{diam}(X) := \max_{\{v_1, v_2\} \in X} d_{v_1, v_2}$ for $X \subset V$. The parameter g corresponds to one-spin energy since $\sum_{X | X \ni v} \|h_X\| \leq f(1) \leq g$. By taking the energy unit appropriately, we set $g = 1$. For example, if $k = 2$ and $f(R) = 0$ for $R \geq 2$, the Hamiltonian is described by bipartite nearest-neighbor interactions as $H = \sum_{\{i,j\} \in E} h_{i,j}$. We consider the Gibbs state for the Hamiltonian H at an inverse temperature β as follows:

$$\rho := \frac{1}{Z} e^{-\beta H}, \qquad Z := \mathrm{tr}(e^{-\beta H}).$$

Our purpose is to discuss the Markov property of Gibbs states. Let $V_0 \subseteq V$ be an arbitrary subsystem. Consider a tripartite partitioning of V_0 as $V_0 = ABC$, where we denote $A \cup B$ by AB for simplicity. We notice that the subsystems $\{A, B, C\}$ are *not* necessarily concatenated on the graph (see Fig. 1). If any two nonadjacent subsystems A and C are conditionally independent of the other subsystem B ($= V_0 \backslash AC$), we say that ρ^{V_0} is the quantum Markov network on V_0. Mathematically, this implies $\mathcal{I}_\rho(A:C|B) = 0$ for $d_{A,C} > 0$ [34,38], where $\mathcal{I}_\rho(A:C|B)$ is defined in Eq. (1). It is noteworthy that the Markov property of ρ^{V_0} strongly depends on the selection of the subsystem $V_0 \subseteq V$. To prove this point, let us consider a one-dimensional graph. Then, the GHZ state is a Markov network for $\forall\, V_0 \subset V$, but not globally, namely, $\mathcal{I}_\rho(A:C|B) = 1$ for $ABC = V$. In contrast, the cluster state [39] is globally a Markov network, but not for particular selections of V_0 [40,41] (see also Ref. [42]). Based on the example of the cluster state, which has a finite correlation length and is described by the matrix product state with bond dimension two [43], we cannot ensure the Markov property only using the clustering theorem and matrix product (or tensor network) representation of the quantum Gibbs state.

The Markov property has a clear operational meaning in terms of a recovery map as follows: If ρ^{V_0} is a Markov network, we can always find a quantum channel $\tau_{B \to BC}$ referred to as the Petz recovery map [38,46], which recovers ρ^{ABC} from ρ^{AB} ($V_0 = ABC$):

$$\tau_{B \to BC}(\rho^{AB}) = \rho^{ABC}.$$

The above local reconstruction is not possible for generic quantum states. Note that any quantum Markov network on a tree graph can be constructed from a sequence of n local quantum channels.

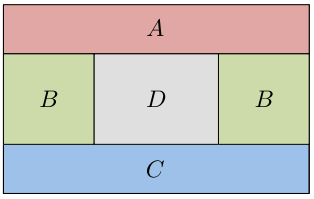

FIG. 1. Decomposition of the total system into A, B, C, and D. It is possible that in a quantum state, there is no correlation between A and C when considering only the subsystems A and C; however, there is a strong correlation when looking at them via the subsystem B. This kind of correlation between A and B related to C is measured by conditional mutual information (1). Physically, conditional mutual information characterizes tripartite correlations between A, B, and C. A representative example is the topological entanglement entropy [44,45], which is a special form of the conditional mutual information.

In realistic situations, we frequently encounter cases where the density matrix is not given by the exact Markov network but by an approximate Markov network; that is, the conditional mutual information $\mathcal{I}_\rho(A:C|B)$ approaches zero as the distance $d_{A,C}$ increases. In the case where $\mathcal{I}_\rho(A:C|B)=\epsilon$, the celebrated Fawzi-Renner theorem [33] (see also Refs. [47–54]) ensures the existence of the recovery map such that

$$\|\tau_{B\to BC}(\rho^{AB})-\rho^{ABC}\|_1^2 \le \epsilon \log 2, \tag{5}$$

where $\|\cdot\|_1$ is the trace norm [i.e., $\|O\|_1 \coloneqq \mathrm{tr}(\sqrt{O^\dagger O})$ for an operator O]. Here, the form of $\tau_{B\to BC}$ is given by the rotated Petz map [54]. Based on this theorem, we can still relate the approximate Markov property to the local reconstruction of the state.

The main purpose of this study was to characterize the decay rate of the conditional mutual information $\mathcal{I}_\rho(A:C|B)$ with respect to the distance $d_{A,C}$ on the graph. To explain the physics of the theorems, we have provided the proofs of our main theorems in the Supplemental Material [55].

Main result.—We proved the exponential decay of the conditional mutual information above a temperature threshold:

Theorem 1: Let us consider finite-range interaction up to a finite length r; that is, we consider a function $f(R)$ in inequality (4) such that $f(R)=0$ for $R>r\in\mathbb{N}$. Then, the condition

$$\beta<\beta_c \coloneqq \frac{1}{8e^3k} \tag{6}$$

implies that the Gibbs state ρ is an approximate Markov network on an arbitrary subset $V_0\subseteq V$ in the sense that

$$\mathcal{I}_\rho(A:C|B)\le e\min(|\partial A_r|,|\partial C_r|)\frac{(\beta/\beta_c)^{d_{A,C}/r}}{1-\beta/\beta_c}, \tag{7}$$

where $V_0=ABC$ and the subset ∂A_r (∂C_r) is defined by Eq. (2) with $l=r$ and $L=A$ ($L=C$).

We notice that if we select B as an empty set (i.e., $B=\varnothing$), the conditional mutual information reduces to bipartite mutual information:

$$\mathcal{I}_\rho(A:C|\varnothing)=\mathcal{I}_\rho(A:C),$$

where $\mathcal{I}_\rho(A:C)\coloneqq S(\rho^A)+S(\rho^C)-S(\rho^{AC})$. Therefore, inequality (7) also implies the exponential decay of the mutual information between two separated subsystems. It is an improved version of the standard clustering theorem for the bipartite operator correlation $\mathrm{Cor}_\rho(O_A,O_B)\coloneqq \mathrm{tr}(\rho O_A O_B)-\mathrm{tr}(\rho O_A)\mathrm{tr}(\rho O_B)$, where O_A and O_B are arbitrary operators with unit norm (i.e., $\|O_A\|=\|O_B\|=1$) supported on subsystems A and B, respectively. From the relation $[\mathrm{Cor}_\rho(O_A,O_B)]^2\le 2\mathcal{I}_\rho(A:B)$ [29], the clustering theorem can be derived from the exponential decay of the mutual information. Moreover, it is well known [56,57] in the context of data hiding that even if the operator correlation is arbitrarily small in a quantum state, the state may still be highly correlated in terms of the mutual information [29].

An important implication of this theorem is related to the quantum sampling of Gibbs states. Based on the Fawzi-Renner theorem (5), an approximate Markov network can be efficiently reconstructed from its reduced density matrix using a quantum computer. According to Ref. [37], the clustering and Markov properties ensure an efficient preparation of quantum Gibbs states on finite-dimensional lattices. By combining our Theorem 1 with Theorem 5 in Ref. [37], we obtain the following statement.

Let us consider the case where the graph G is given by a D-dimensional lattice, where D is the spatial dimension. Then, under the assumption of Theorem 1, there exists a completely positive trace-preserving (CPTP) map of $\mathbb{F}=\mathbb{F}_{D+1}\cdots\mathbb{F}_2\mathbb{F}_1$ such that

$$\|\mathbb{F}(\psi)-\rho\|_1=1/\mathrm{poly}(n),$$

where ψ is an arbitrary quantum state, and each of $\{\mathbb{F}_s\}_{s=1}^{D+1}$ is given by a direct product of quasilocal CPTP maps that act on $\mathcal{O}(\log^D n)$ spins.

The number of the elementary gates for the each quasilocal channel $\{\mathbb{F}_s\}_{s=1}^{D+1}$ is on the order $\exp[\mathcal{O}(\log^D n)]=n^{\mathcal{O}(\log^{D-1} n)}$ [58,59]. This also provides the computational time of Gibbs sampling by the quantum computer. This algorithm requires only quasipolynomial computational time, and it is considerably better than a few existing algorithms [60,61], which require at least subexponential computational time. Our algorithm still performs slightly worse than the algorithms proposed in Refs. [62,63], which require polynomial computational time. However, our method has advantages in the following senses: The method in Ref. [62] is applicable only to commuting Hamiltonians, and the method in Ref. [63] requires twice the number of qubits (i.e., $2n$ qubits) for implementation.

The second implication of the theorem is the strengthening of the area law. The area law for mutual information was derived at arbitrary temperatures in Ref. [29] in the following form:

$$\mathcal{I}_\rho(A:B)\le c\beta|\partial A|, \tag{8}$$

where $AB=V$ and c is an $\mathcal{O}(1)$ constant. The area law implies that $\mathcal{I}_\rho(A:B')$ saturates as $B'\subset B$ grows to B; however, Eq. (8) does not provide the saturation rate. Our result implies it saturates exponentially fast, and the mutual information between two subsystems is exponentially localized around the boundary between A and B. For further understanding, let us decompose B into l_0 slices,

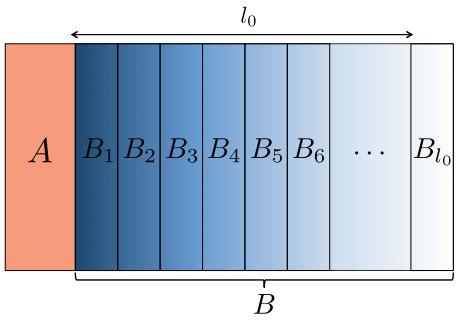

FIG. 2. Strengthening of the area law resulting from the Markov property. In the figure, we consider a 2D system and decompose it into A and $B = B_1 B_2, \ldots, B_{l_0}$ with $d_{A,B_l} = l$ $(1 \le l \le l_0)$.

$B_1 B_2, \ldots, B_{l_0}$, with $d_{A,B_l} = l$ for $l = 1, 2, \ldots, l_0$ (see Fig. 2). Then, the question is how rapidly the mutual information $\mathcal{I}_\rho(A\!:\!B_1 B_2 \cdots B_l)$ saturates to $\mathcal{I}_\rho(A\!:\!B)$. From the relation $\mathcal{I}(A\!:\!C|B) = \mathcal{I}(A\!:\!BC) - \mathcal{I}(A\!:\!B)$ and inequality (7), we have

$$\begin{aligned}&\mathcal{I}_\rho(A\!:\!B_1 \cdots B_l) - \mathcal{I}_\rho(A\!:\!B_1 \cdots B_{l-1}) \\ &\quad = \mathcal{I}_\rho(A\!:\!B_l|B_1 \cdots B_{l-1}) \sim (\beta/\beta_c)^{l/r}, \end{aligned} \tag{9}$$

which shows exponential decay with respect to l.

Effective Hamiltonian on subsystem and classical simulation of Gibbs state.—Theorem 1 is related to the locality of the effective Hamiltonian. We define the effective Hamiltonian of the local reduced density matrix as

$$\tilde{H}_L := -\beta^{-1} \log \mathrm{tr}_{L^c}(e^{-\beta H}). \tag{10}$$

We formally describe $\tilde{H}_L$ as

$$\tilde{H}_L = H_L + \Phi_L, \tag{11}$$

where H_L is composed of the original interacting terms in H on subsystem L, namely, $H_L = \sum_{X \subset L} h_X$, and Φ_L is the effective interaction term. We are interested in the locality of Φ_L. Typically, it is computationally difficult to determine the effective term, even in classical Gibbs states [64]. Our present question is whether the (quasi)locality of Φ_L can be ensured (Fig. 3). In classical Gibbs states or systems with commuting Hamiltonians, Φ_L is exactly localized around the surface region of L (not necessarily localized along the boundary). This point is crucial for the Gibbs states to be the exact Markov networks [34,37]. Additionally, for systems with noncommuting Hamiltonians, the quasilocality of Φ_L is numerically verified in Ref. [66]. By following the same analysis as the proof of Theorem 1, we can rigorously prove the quasilocality of Φ_L not only in the direction orthogonal to the boundary but also along the boundary.

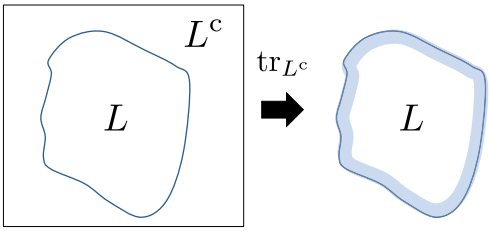

FIG. 3. Effective Hamiltonian $\tilde{H}_L$ for the reduced density matrix ρ^L. We decompose $\tilde{H}_L$ as $\tilde{H}_L = H_L + \Phi_L$, where H_L is the original Hamiltonian in L and Φ_L is the effective term that originates outside L. Theorem 2 implies that Φ_L is exponentially localized around the surface region of L.

Theorem 2: Using the setup and assumption of Theorem 1, Φ_L is approximated using a localized operator $\Phi_{\partial L_l}$ as follows:

$$\|\Phi_L - \Phi_{\partial L_l}\| \le \frac{e}{4\beta} \frac{(\beta/\beta_c)^{l/r}}{1 - \beta/\beta_c} |\partial L_r|,$$

where $\Phi_{\partial L_l}$ is supported on the region ∂L_l that has been defined in Eq. (2). In addition, $\Phi_{\partial L_l}$ is composed of local operators that act on at most $(k\lfloor l/r \rfloor)$ spins (see Supplemental Material [55] for an explicit form of $\Phi_{\partial L_l}$). Moreover, computation of Φ_L up to a norm error of $n\epsilon$ is performed with the runtime bounded from above by

$$n(1/\epsilon)^{\mathcal{O}(k \log(d d_G^r))},$$

where d_G is the degree of the graph G.

This theorem immediately implies that the classical simulation of the Gibbs states is possible in polynomial time within an error of $1/\mathrm{poly}(n)$. We note that the definition (10) implies $\Phi_L = -\beta^{-1} \log(Z)$ for $L = \varnothing$; i.e., we can calculate the partition function using the same algorithm. We can also calculate the expectation values of local observables or the local entropy by explicitly obtaining the expression $\rho^L = e^{-\beta \tilde{H}_L}$. This is summarized in the following corollary.

Corollary 1: Thermodynamic properties such as local observables (e.g., energy and magnetization), the partition function $\log(Z)$, and local entropy $-\mathrm{tr}(\rho^L \log \rho^L)$ are classically simulated in polynomial time $\mathrm{poly}(n)$ as long as an error of $1/\mathrm{poly}(n)$ is allowed.

From Ref. [31], we can prepare tensor network representations for arbitrary Gibbs states in the polynomial time of $n^{\mathcal{O}(\beta)}$. However, the classical simulation of the tensor network is $\#P$ complete problem [67,68] except in 1D cases. To the best of our knowledge, our result, for the first time, provides the fully-polynomial-time approximation scheme (FPTAS [69]) for the classical simulation of quantum Gibbs states, which is a quantum generalization of the FPTAS for classical Gibbs states [70–72].

Long-range interacting systems.—Finally, we extend Theorem 1 from short-range interacting systems to long-range interacting systems. We define the Hamiltonian with

the power-law decay interaction assuming that $f(R)$ in Eq. (4) is given by

$$f(R) = R^{-\alpha},$$

where $\alpha > 0$. To consider a more general form as $f(R) = gR^{-\alpha}$, we must only scale the inverse temperature from β to β/g. For example, we can consider the following Hamiltonian on a graph with a D-dimensional structure:

$$H = \sum_{i,j\in V} \frac{J}{R_{i,j}^{\alpha+D}} h_{i,j} \quad \text{with} \quad \|h_{i,j}\| = 1,$$

where $R_{i,j}$ is the distance between spins i and j defined by the graph structure (V, E), and J is determined so that inequality (4) is satisfied. This type of Hamiltonian is now controllable in realistic experiments and attracts considerable attention both in experimental [73–77] and theoretical aspects [78–82].

Similar to the case of short-range interacting systems, we prove the decay of the conditional mutual information for long-range interacting systems for $\alpha > 0$.

Theorem 3: Let A, B, and C be arbitrary subsystems in V $(A, B, C \subset V)$. Then, under the assumptions of $\beta < \beta_c/11$ and $d_{A,C} \geq 2\alpha$, the Gibbs state ρ satisfies the approximate Markov property as follows:

$$\mathcal{I}_\rho(A:C|B) \leq \beta \min(|A|, |C|) \frac{C_\beta}{d_{A,C}^\alpha}, \tag{12}$$

where $C_\beta := [(11e^{1/k}/\beta_c)/(1 - 11\beta/\beta_c)]$ and β_c is as defined in Eq. (6).

By selecting $B = \emptyset$, we can also derive the power-law decay of the mutual information between two separated subsystems. To the best of our knowledge, the clustering theorem for the Gibbs state with long-range interaction is limited for classical cases [83–88] and special quantum cases [89,90]. Our result provides the first general proof of the clustering theorem at finite temperatures in long-range interacting quantum systems.

Proof ideas of the main theorems.—We finally show the proof ideas to obtain the decay of the conditional mutual information. The proof utilizes a high-temperature expansion. The difficulty lies in the fact that the standard cluster expansion technique cannot be applied to the logarithm of the reduced density matrix (e.g., ρ^L with $L \subset \Lambda$). We introduce a new technique of the generalized cluster expansion, which allows us to systematically treat logarithmic operators (see Sec. I B in Ref. [55]). Here, we parametrize the Hamiltonian (3) as $H_{\vec{a}} = \sum_{X\in\Lambda} a_X h_X$ with $\vec{a} = \{a_X\}_{X\in\Lambda}$. We then parametrize a target function of interest by $f_{\vec{a}}$ and directly expand it with respect to $\vec{a}$, where $f_{\vec{a}}$ can be chosen as a scholar function and also as an operator function. Here, we choose the conditional mutual information as the function $f_{\vec{a}}$. The challenge in the generalized cluster expansion is to estimate the convergence radius of the expansion, where we need to consider a multiple derivative of the operators like $\log[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}})]$ with $L \subset \Lambda$. Our technical contributions are the systematical expression of the multiple derivative in the generalized cluster expansion (e.g., Propositions 3 and 4 in Ref. [55]) and the estimation of the convergence radius (see Ref. [55] for the details).

Future perspective.—Here we mention an open problem. The most important problem is the Markov property in low-temperature regimes, where our present analytical technique (i.e., the generalized cluster expansion [55]) breaks down. It is no longer desirable that the Markov property holds for the arbitrary selections of the subregions A, B, and C because the topological order can exist at finite temperatures in four-dimensional systems [1]. Further, we hope to apply the current analyses to other essential problems, such as the contraction problem of the projected entangled pair states [91–94] and efficiency guarantee of the heuristic classical algorithms for the quantum Gibbs states [95–99].

We thank Keiji Saito for valuable discussions on this work. The work of T. K. was supported by the RIKEN Center for AIP and JSPS KAKENHI Grant No. 18K13475. K. K. acknowledges funding provided by the Institute for Quantum Information and Matter, a NSF Physics Frontiers Center (NSF Grant No. PHY-1733907). F. B. is supported by the NSF.

Note added.—Regarding the classical simulations of quantum Gibbs states, we identified a related result obtained using a similar approach [100] at the same time of our submission.

*tomotaka.kuwahara@riken.jp
†kokato@caltech.edu
‡fbrandao@caltech.edu

[1] M. B. Hastings, Topological Order at Nonzero Temperature, Phys. Rev. Lett. **107**, 210501 (2011).
[2] O. Landon-Cardinal and D. Poulin, Local Topological Order Inhibits Thermal Stability in 2D, Phys. Rev. Lett. **110**, 090502 (2013).
[3] S. Bravyi and J. Haah, Quantum Self-Correction in the 3D Cubic Code Model, Phys. Rev. Lett. **111**, 200501 (2013).
[4] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, Quantum memories at finite temperature, Rev. Mod. Phys. **88**, 045005 (2016).
[5] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, Nat. Phys. **2**, 754 (2006).
[6] M. P. Müller, E. Adlam, L. Masanes, and N. Wiebe, Thermalization and canonical typicality in translation-invariant quantum lattice systems, Commun. Math. Phys. **340**, 499 (2015).

[7] F. G. S. L. Brandao and M. Cramer, Equivalence of statistical mechanical ensembles for non-critical quantum systems, arXiv:1502.03263.

[8] H. Tasaki, On the local equivalence between the canonical and the microcanonical ensembles for quantum spin systems, J. Stat. Phys. **172**, 905 (2018).

[9] T. Kuwahara and K. Saito, Ensemble Equivalence and Eigenstate Thermalization from Clustering of Correlation, Phys. Rev. Lett. **124**, 200604 (2020).

[10] T. Kuwahara and K. Saito, Gaussian concentration bound and ensemble equivalence in generic quantum many-body systems including long-range interaction, arXiv:1906.10872.

[11] D. Poulin and M. B. Hastings, Markov Entropy Decomposition: A Variational Dual for Quantum Belief Propagation, Phys. Rev. Lett. **106**, 080403 (2011).

[12] T. J. Osborne, Hamiltonian complexity, Rep. Prog. Phys. **75**, 022001 (2012).

[13] D. Aharonov, I. Arad, and T. Vidick, Guest column: The quantum PCP conjecture, ACM SIGACT News **44**, 47 (2013).

[14] S. Gharibian, Y. Huang, Z. Landau, and S. W. Shin, Quantum Hamiltonian complexity, Found. Trends® Theor. Comput. Sci. **10**, 159 (2015).

[15] F. G. S. L. Brandão and K. M. Svore, Quantum speed-ups for solving semidefinite programs, in *Proceedings of the 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2017), pp. 415–426.

[16] J. Van Apeldoorn, A. Gilyén, S. Gribling, and R. de Wolf, Quantum SDP-solvers: Better upper and lower bounds, in *Proceedings of the 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2017), pp. 403–414.

[17] F. G. S. L. Brandao, A. Kalev, T. Li, C. Y. Y. Lin, K. M. Svore, and Xiaodi Wu, Quantum SDP solvers: Large speed-ups, optimality, and applications to quantum learning, arXiv:1710.02581.

[18] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, Nature (London) **549**, 195 (2017).

[19] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, Quantum Boltzmann Machine, Phys. Rev. X **8**, 021050 (2018).

[20] M. Kieferová and N. Wiebe, Tomography and generative training with quantum Boltzmann machines, Phys. Rev. A **96**, 062327 (2017).

[21] H.-J. Song, T. Song, Q.-K. He, Y. Liu, and D. L. Zhou, Geometry and symmetry in the quantum Boltzmann machine, Phys. Rev. A **99**, 042307 (2019).

[22] D. Crawford, A. Levit, N. Ghadermarzy, J. S. Oberoi, and P. Ronagh, Reinforcement learning using quantum Boltzmann machines, Quantum Inf. Comput. **18**, 0051 (2018).

[23] H. Araki, Gibbs states of a one dimensional quantum lattice, Commun. Math. Phys. **14**, 120 (1969).

[24] L. Gross, Decay of correlations in classical lattice models at high temperature, Commun. Math. Phys. **68**, 9 (1979).

[25] Y. M. Park and H. J. Yoo, Uniqueness and clustering properties of Gibbs states for classical and quantum unbounded spin systems, J. Stat. Phys. **80**, 223 (1995).

[26] D. Ueltschi, Cluster expansions and correlation functions, Moscow Math. J. **4**, 511 (2004).

[27] M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, and J. Eisert, Locality of Temperature, Phys. Rev. X **4**, 031019 (2014).

[28] J. Fröhlich and D. Ueltschi, Some properties of correlations of quantum lattice systems in thermal equilibrium, J. Math. Phys. (N.Y.) **56**, 053302 (2015).

[29] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Area Laws in Quantum Systems: Mutual Information and Correlations, Phys. Rev. Lett. **100**, 070502 (2008).

[30] M. B. Hastings, Solving gapped Hamiltonians locally, Phys. Rev. B **73**, 085115 (2006).

[31] A. Molnar, N. Schuch, F. Verstraete, and J. I. Cirac, Approximating Gibbs states of local Hamiltonians efficiently with projected entangled pair states, Phys. Rev. B **91**, 045138 (2015).

[32] J. M. Hammersley and P. Clifford, Markov fields on finite graphs and lattices (to be published).

[33] O. Fawzi and R. Renner, Quantum conditional mutual information and approximate Markov chains, Commun. Math. Phys. **340**, 575 (2015).

[34] W. Brown and D. Poulin, Quantum Markov networks and commuting Hamiltonians, arXiv:1206.0755.

[35] F. G. Jouneghani, M. Babazadeh, R. Bayramzadeh, and H. Movla, Investigation of commuting Hamiltonian in quantum Markov network, Int. J. Theor. Phys. **53**, 2521 (2014).

[36] K. Kato and F. G. S. L. Brandão, Quantum approximate Markov chains are thermal, Commun. Math. Phys. **370**, 117 (2019).

[37] F. G. S. L. Brandão and M. J. Kastoryano, Finite correlation length implies efficient preparation of quantum thermal states, Commun. Math. Phys. **365**, 1 (2019).

[38] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Structure of states which satisfy strong subadditivity of quantum entropy with equality, Commun. Math. Phys. **246**, 359 (2004).

[39] R. Raussendorf and H. J. Briegel, A One-Way Quantum Computer, Phys. Rev. Lett. **86**, 5188 (2001).

[40] D. J. Williamson, A. Dua, and M. Cheng, Spurious Topological Entanglement Entropy from Subsystem Symmetries, Phys. Rev. Lett. **122**, 140506 (2019).

[41] L. Zou and J. Haah, Spurious long-range entanglement and replica correlation length, Phys. Rev. B **94**, 075151 (2016).

[42] For the cluster state on a ring with n as an even integer, let ρ^{V_0} be the reduced density matrix with respect to the odd spins, $V_0 = \{1, 3, 5, \ldots, 2s-1\}$ ($s \coloneqq n/2$). Then, we obtain the von Neumann entropy for ρ^{V_0} as $S(\rho^{V_0}) = (|V_0| - 1)\log 2 = (s-1)\log 2$. On the contrary, for arbitrary subset $V_0' \subset V_0$, the reduced density matrix $\rho^{V_0'}$ is given by the identity matrix, and hence, $S(\rho_{V_0'}) = |V_0'|\log 2$. Therefore, for an arbitrary decomposition $V_0 = ABC$, we have $\mathcal{I}(A:C|B) = \log 2$ and the Markov property breaks down.

[43] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations, Quantum Inf. Comput. **7**, 401 (2007).

[44] A. Kitaev and J. Preskill, Topological Entanglement Entropy, Phys. Rev. Lett. **96**, 110404 (2006).

[45] M. Levin and X.-G. Wen, Detecting Topological Order in a Ground State Wave Function, Phys. Rev. Lett. **96**, 110405 (2006).

[46] D. Petz, Sufficient subalgebras and the relative entropy of states of a von Neumann algebra, Commun. Math. Phys. **105**, 123 (1986).

[47] F. G. S. L. Brandão, A. W. Harrow, J. Oppenheim, and S. Strelchuk, Quantum Conditional Mutual Information, Reconstructed States, and State Redistribution, Phys. Rev. Lett. **115**, 050501 (2015).

[48] M. M. Wilde, Recoverability in quantum information theory, Proc. R. Soc. A **471**, 20150338 (2015).

[49] M. Berta, M. Lemm, and M. M. Wilde, Monotonicity of quantum relative entropy and recoverability, Quantum Inf. Comput. **15**, 1333 (2015).

[50] M. Berta, K. P. Seshadreesan, and M. M. Wilde, Rényi generalizations of the conditional quantum mutual information, J. Math. Phys. (N.Y.) **56**, 022205 (2015).

[51] D.Sutter, O. Fawzi, and R. Renner, Universal recovery map for approximate Markov chains, Proc. R. Soc. A **472**, 20150623 (2016).

[52] M. Berta and M. Tomamichel, The fidelity of recovery is multiplicative, IEEE Trans. Inf. Theory **62**, 1758 (2016).

[53] D. Sutter and R. Renner, Necessary criterion for approximate recoverability, Ann. Inst. Henri Poincaré **19**, 3007 (2018).

[54] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, Universal recovery maps and approximate sufficiency of quantum relative entropy, Ann. Inst. Henri Poincaré **19**, 2955 (2018).

[55] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.220601 for the details of all the rigorous proofs.

[56] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Randomizing quantum states: Constructions and applications, Commun. Math. Phys. **250**, 371 (2004).

[57] M. B. Hastings, Entropy and entanglement in quantum ground states, Phys. Rev. B **76**, 035114 (2007).

[58] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Elementary gates for quantum computation, Phys. Rev. A **52**, 3457 (1995).

[59] J. J. Vartiainen, M. Möttönen, and M. M. Salomaa, Efficient Decomposition of Quantum Gates, Phys. Rev. Lett. **92**, 177902 (2004).

[60] D. Poulin and P. Wocjan, Sampling from the Thermal Quantum Gibbs State and Evaluating Partition Functions with a Quantum Computer, Phys. Rev. Lett. **103**, 220502 (2009).

[61] E. Bilgin and S. Boixo, Preparing Thermal States of Quantum Systems by Dimension Reduction, Phys. Rev. Lett. **105**, 170405 (2010).

[62] M. J. Kastoryano and F. G. S. L. Brandão, Quantum Gibbs samplers: The commuting case, Commun. Math. Phys. **344**, 915 (2016).

[63] Y. Ge, A. Molnár, and J. I. Cirac, Rapid Adiabatic Preparation of Injective Projected Entangled Pair States and Gibbs States, Phys. Rev. Lett. **116**, 080503 (2016).

[64] When Φ_L is given for all $L \subset V$ with $|L| \leq k$, we can compute the average energy $\mathrm{tr}(\rho H)$ in polynomial time. However, the computational complexity of calculating the average energy is known to be non-deterministic polynomial time hard [65].

[65] F. Barahona, On the computational complexity of Ising spin glass models, J. Phys. A **15**, 3241 (1982).

[66] E. Bilgin and D. Poulin, Coarse-grained belief propagation for simulation of interacting quantum systems at all temperatures, Phys. Rev. B **81**, 054106 (2010).

[67] N.t Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Computational Complexity of Projected Entangled Pair States, Phys. Rev. Lett. **98**, 140506 (2007).

[68] J. Haferkamp, D. Hangleiter, J. Eisert, and M. Gluza, Contracting projected entangled pair states is average-case hard, Phys. Rev. Res. **2**, 013010 (2020).

[69] FPTAS is a type of deterministic algorithm which gives an approximation to a given problem up to an error ϵ with the runtime bounded from above by $\mathrm{poly}(n) \times \mathrm{poly}(1/\epsilon)$ with n the input size.

[70] D. Weitz, Counting independent sets up to the tree threshold, in *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '06 (Association for Computing Machinery, New York, 2006), pp. 140–149.

[71] A. Sinclair, P. Srivastava, and M. Thurley, Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs, in *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12 (Society for Industrial and Applied Mathematics, 2012), pp. 941–953.

[72] J. Liu, A. Sinclair, and P. Srivastava, The Ising partition function: Zeros and deterministic approximation, in *Proceedings of the 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2017), pp. 986–997.

[73] B. Yan, S. A. Moses, B. Gadway, J. P. Covey, K. R. A. Hazzard, A. M. Rey, D. S. Jin, and J. Ye, Observation of dipolar spin-exchange interactions with lattice-confined polar molecules, Nature (London) **501**, 521 (2013).

[74] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, and C. Monroe, Non-local propagation of correlations in quantum systems with long-range interactions, Nature (London) **511**, 198 (2014).

[75] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, Quasiparticle engineering and entanglement propagation in a quantum many-body system, Nature (London) **511**, 202 (2014).

[76] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks, and C. Monroe, Emergence and frustration of magnetism with variable-range interactions in a quantum simulator, Science **340**, 583 (2013).

[77] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature (London) **551**, 601 (2017).

[78] T. Koffel, M. Lewenstein, and L. Tagliacozzo, Entanglement Entropy for the Long-Range Ising Chain in a Transverse Field, Phys. Rev. Lett. **109**, 267203 (2012).

[79] D. Vodola, L. Lepori, E. Ercolessi, A. V. Gorshkov, and G. Pupillo, Kitaev Chains with Long-Range Pairing, Phys. Rev. Lett. **113**, 156402 (2014).
[80] T. Kuwahara, Exponential bound on information spreading induced by quantum many-body dynamics with long-range interactions, New J. Phys. **18**, 053034 (2016).
[81] T. Kuwahara, I. Arad, L. Amico, and V. Vedral, Local reversibility and entanglement structure of many-body ground states, Quantum Sci. Technol. **2**, 015005 (2017).
[82] T. Kuwahara and K. Saito, Area law of non-critical ground states in 1D long-range interacting systems, arXiv:1908.11547.
[83] R. B. Israel and C. R. Nappi, Exponential clustering for long-range integer-spin systems, Commun. Math. Phys. **68**, 29 (1979).
[84] C. Cammarota, Decay of correlations for infinite range interactions in unbounded spin systems, Commun. Math. Phys. **85**, 517 (1982).
[85] H. Spohn and W. Zwerger, Decay of the two-point function in one-dimensional $O(N)$ spin models with long-range interactions, J. Stat. Phys. **94**, 1037 (1999).
[86] A. Procacci and B. Scoppola, On decay of correlations for unbounded spin systems with arbitrary boundary conditions, J. Stat. Phys. **105**, 453 (2001).
[87] A. Kargol, Decay of correlations in multicomponent ferromagnets with long-range interaction, Rep. Math. Phys. **56**, 379 (2005).
[88] G. Menz and R. Nittka, Decay of correlations in 1D lattice systems of continuous spins and long-range interaction, J. Stat. Phys. **156**, 239 (2014).
[89] A. Kargol, Decay of correlations in N-vector ferromagnetic quantum models with long-range interactions, J. Stat. Mech. (2014) P10006.
[90] S. Hernández-Santana, C. Gogolin, J. I. Cirac, and A. Acín, Correlation Decay in Fermionic Lattice Systems with Power-Law Interactions at Nonzero Temperature, Phys. Rev. Lett. **119**, 110601 (2017).
[91] M. Lubasch and J. I. Cirac and M.-C. Bañuls, Unifying projected entangled pair state contractions, New J. Phys. **16**, 033014 (2014).
[92] A. Anshu, I. Arad, and A. Jain, How local is the information in tensor networks of matrix product states or projected entangled pairs states, Phys. Rev. B **94**, 195143 (2016).
[93] M. Schwarz, O. Buerschaper, and J. Eisert, Approximating local observables on projected entangled pair states, Phys. Rev. A **95**, 060102(R) (2017).
[94] M. J. Kastoryano, A. Lucia, and D. Perez-Garcia, Locality at the boundary implies gap in the bulk for 2D PEPS, Commun. Math. Phys. **366**, 895 (2019).
[95] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems, Phys. Rev. Lett. **93**, 207204 (2004).
[96] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, Rev. Mod. Phys. **73**, 33 (2001).
[97] M. Rigol, T. Bryant, and R. R. P. Singh, Numerical Linked-Cluster Approach to Quantum Lattice Models, Phys. Rev. Lett. **97**, 187202 (2006).
[98] B. Tang, E. Khatami, and M. Rigol, A short introduction to numerical linked-cluster expansions, Comput. Phys. Commun. **184**, 557 (2013).
[99] S. R. White, Minimally Entangled Typical Quantum States at Finite Temperature, Phys. Rev. Lett. **102**, 190601 (2009).
[100] A.Harrow, S. Mehraban, and M. Soleimanifar, Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems, arXiv:1910.09071.