Supplementary Material for
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
Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kouhoku-ku, Yokohama 223-8522, Japan and interdisciplinary Theoretical & Mathematical Sciences Program (iTHEMS) RIKEN 2-1, Hirosawa, Wako, Saitama 351-0198, Japan

Kohtaro Kato† and Fernando G. S. L. Brandão‡
Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA

CONTENTS

I. PROOF OF THEOREM 1

A. Preliminaries

We here recall the setup. We consider a quantum spin system with n spins, where each of the spin sits on a vertex of the graph $G = (V, E)$ with V the total spin set ($|V| = n$). For a partial set L of spins, we denote the cardinality, that is, the number of vertices contained in L, by $|L|$ (e.g. $L = \{i_1, i_2, \ldots, i_{|L|}\}$). We also denote the complementary subset of L by $L^{\mathrm{c}} := V \setminus L$. We denote the local Hilbert space by $\mathcal{H}^v$ ($v \in V$) with $\dim(\mathcal{H}^v) = d$ and the entire Hilbert space is given by $\mathcal{H} := \bigotimes_{v\in V} \mathcal{H}^v$ with $\dim(\mathcal{H}) = d^n$. We also define the local Hilbert space of the subset $L \subset V$ as $\mathcal{H}^L$ and denote the dimension by d_L, namely $d_L := d^{|L|}$. We define $\mathcal{B}(\mathcal{H})$ as the space of bounded linear operators on $\mathcal{H}$.

When we consider a reduced operator on a subsystem L, we denote it as

$$O^L = \mathrm{tr}_{L^{\mathrm{c}}}(O) \otimes \hat{1}_{L^{\mathrm{c}}} \in \mathcal{B}(\mathcal{H}) \tag{S.1}$$

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

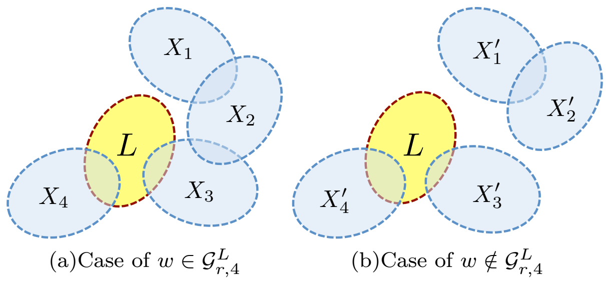

FIG. 1. Schematic pictures of clusters of $w \in \mathcal{G}^L_4$ and $w \notin \mathcal{G}^L_4$. Each of the elements $\{X_s | X_s \in E_r\}$ is a subset of the total set V (i.e., $X \subset V$). In (a), there there are no decompositions of $w = w_1 \sqcup w_2$ such that $(L \cup V_{w_1}) \cap V_{w_2} = \emptyset$ for $w = \{X_1, X_2, X_3, X_4\}$, whereas in (b) the decomposition $w' = w'_1 \sqcup w'_2$ with $w'_1 = \{X'_2, X'_3\}$ and $w'_2 = \{X'_1, X'_4\}$ satisfies $(L \cup V_{w_1}) \cap V_{w_2} = \emptyset$.

by using the superscript index, where $\hat{1}$ is the identity operator and tr_{L^c} is the partial trace operation with respect to the Hilbert space $\mathcal{H}^{L^c}$.

We also define the following set:

$$E^{(x)} := \{X \subset V | \mathrm{diam}(X) = x, \quad |X| \le k\} \tag{S.2}$$

with

$$\mathrm{diam}(X) := \max_{v_1, v_2 \in X} d_{v_1, v_2}, \tag{S.3}$$

where we defined $d_{A,B}$ as the shortest path length via E which connects A and B ($A \subset V$, $B \subset V$).

In the setup of Theorem 1, we consider the Hamiltonian as

$$H = \sum_{X \in E_r} h_X, \quad \text{with} \quad \sum_{X | X \ni v} \|h_X\| \le 1 \quad \text{for} \quad \forall v \in V \tag{S.4}$$

with

$$E_r := E^{(1)} \sqcup E^{(2)} \sqcup \cdots \sqcup E^{(r)} \quad (r \in \mathbb{N}). \tag{S.5}$$

Here, the Hamiltonian (S.4) describes an arbitrary k-body interacting systems with finite interaction length r.

Throughout the manuscript, we denote the natural logarithm by $\log(\cdot)$ for the simplicity, namely $\log(\cdot) = \log_e(\cdot)$.

1. *Cluster notation*

We then define several basic terminologies. On the graph (V, E), we call a multiset of subsystems $w = \{X_1, X_2, \ldots, X_{|w|}\}$ ($X_j \in E_r$ for $j = 1, 2, \ldots, |w|$) as "cluster", where $|w|$ is the cardinality of w. Note that each of the elements $\{X_j\}_{j=1}^{|w|}$ satisfies $\mathrm{diam}(X_j) \le r$ from the definition (S.5). We denote $\mathcal{C}_{r,m}$ by the set of w with $|w| = m$ and let $V_w \subseteq V$ and $E_w \subseteq E_r$ be the set of different vertices (or spins) and subsystems which are contained in w, respectively. Also, we define connected clusters as follows:

Definition 1. (Connected cluster) *For a cluster $w \in \mathcal{C}_{r,|w|}$, we say that w is a connected cluster if there are no decompositions of $w = w_1 \sqcup w_2$ such that $V_{w_1} \cap V_{w_2} = \emptyset$. We denote by $\mathcal{G}_{r,m}$ the set of the connected clusters with $|w| = m$.*

Definition 2. (Connected cluster to a region, Fig. 1) *Similarly, we say that $w \in \mathcal{C}_{r,|w|}$ is a connected cluster to a subsystem L if there are no decompositions of $w = w_1 \sqcup w_2$ such that $(L \cup V_{w_1}) \cap V_{w_2} = \emptyset$. We denote by $\mathcal{G}^L_{r,m}$ the set of the connected clusters to L with $|w| = m$.*

Definition 3. (Connected cluster with a link between two regions, Fig. 2) *Finally, for a connected cluster $w \in \mathcal{G}_{r,|w|}$, we say that w has links between A and B if there exist a path from A to B in E_w. We denote by $\mathcal{G}^{A,B}_{r,m}$ the set of the connected clusters with $|w| = m$ which have a link A and B.*

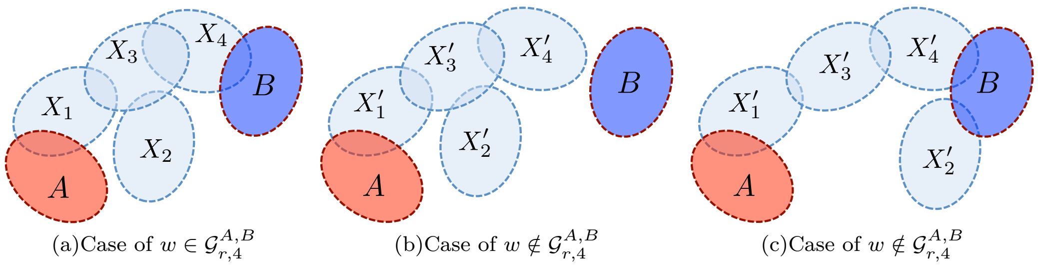

FIG. 2. Schematic pictures of clusters of $w \in \mathcal{G}_{r,4}^{A,B}$ and $w \notin \mathcal{G}_{r,4}^{A,B}$. In (a), subsystems A and B are connected with each other by the cluster w. On the other hand, in (b), the cluster w does not have a link between A and B, and in (c), the cluster has the link but is not connected.

2. *Basic lemmas for logarithmic operators*

Before going to the proof, we prove the following basic lemmas:

Lemma 1. *Let $O \in \mathcal{B}(\mathcal{H})$ be an arbitrary non-negative operator written as*

$$O = \Gamma_{L_1} \otimes \Gamma_{L_2} \otimes \cdots \otimes \Gamma_{L_m}, \tag{S.6}$$

where $\{\Gamma_{L_j}\}_{j=1}^{L} \in \mathcal{B}(\mathcal{H})$ are supported on the subsystems $\{L_j\}_{j=1}^{m}$, respectively and we assume $L_1 \sqcup L_2 \sqcup \cdots \sqcup L_m = V$. Then, for arbitrary subsystems $A, B, C \subset V$, we have

$$\log O^{AB} + \log O^{BC} - \log O^{ABC} - \log O^{B} = \sum_{j=1}^{m} (\log \Gamma_{L_j}^{AB} + \log \Gamma_{L_j}^{BC} - \log \Gamma_{L_j}^{ABC} - \log \Gamma_{L_j}^{B}). \tag{S.7}$$

Note that $\{O^{AB}, O^{BC}, O^{ABC}, O^{B}\}$ are reduced operators as defined in Eq. (S.1).

Proof of Lemma 1. We define $A_j := A \cap L_j$, $B_j := B \cap L_j$, and $C_j := C \cap L_j$ for $j = 1, 2, \dots m$. We notice that $\bigsqcup_{j=1}^{m} A_j = A$, $\bigsqcup_{j=1}^{m} B_j = B$ and $\bigsqcup_{j=1}^{m} C_j = C$ because of $\bigsqcup_{j=1}^{m} L_j = V$. Then, from the definition (S.1), the reduced operator of O with respect to the subsystem B is given by

$$\begin{aligned} O^{B} &= \mathrm{tr}_{L_1 \setminus B_1}(\Gamma_{L_1}) \otimes \mathrm{tr}_{L_2 \setminus B_2}(\Gamma_{L_2}) \otimes \cdots \otimes \mathrm{tr}_{L_m \setminus B_m}(\Gamma_{L_m}) \otimes \hat{1}_{B^c} \\ &= \Gamma_{B_1} \otimes \Gamma_{B_2} \otimes \cdots \otimes \Gamma_{B_m} \otimes \hat{1}_{B^c}, \end{aligned} \tag{S.8}$$

where $\Gamma_{B_j} := \mathrm{tr}_{L_j \setminus B_j}(\Gamma_{L_j}) \otimes \hat{1}_{B_j^c}$ for $j = 1, 2, \dots, m$. We define Γ_j^{BC}, Γ_j^{ABC} and Γ_j^{B} in the same way. We thus obtain

$$\log O^{B} = \sum_{j=1}^{m} \log \Gamma_{B_j}. \tag{S.9}$$

On the other hand, we have from the definition (S.1)

$$\Gamma_{L_j}^{B} = \mathrm{tr}_{B^c}(\Gamma_{L_j}) \otimes \hat{1}_{B^c} = d^{|B^c| - |L_j \setminus B_j|} \Gamma_{B_j} = d^{n - |B| - |L_j| + |B_j|} \Gamma_{B_j}, \tag{S.10}$$

which reduces Eq. (S.9) to

$$\log O^{B} = -(m-1)(n - |B|) \log(d) + \sum_{j=1}^{m} \log \Gamma_{L_j}^{B}. \tag{S.11}$$

We obtain the similar form to Eq. (S.11) for O^{AB}, O^{BC} and O^{ABC}. After a straightforward calculation, we prove the equation (S.7). □

Second, we prove the following lemma:

Lemma 2. *For an arbitrary non-negative operator $O \in \mathcal{B}(\mathcal{H})$ which is given by the form of*

$$O = O_L \otimes \hat{1}_{L^c} \tag{S.12}$$

with $L \cap C = \emptyset$, we have

$$\log O^{AB} + \log O^{BC} - \log O^{ABC} - \log O^{B} = 0. \tag{S.13}$$

Proof of Lemma 2. From the definition, we obtain

$$O^{ABC} = O^{AB} \otimes \hat{1}_C. \tag{S.14}$$

Thus, we obtain $\log O^{BC} = \log(O^B \otimes \hat{1}_C)$ and $\log O^{ABC} = \log O^{AB}$, and hence we immediately obtain Eq. (S.13). This completes the proof. □

B. Generalized cluster Expansion

We first parametrize H by using a parameter set $\vec{a} := \{a_X\}_{X\in E_r}$ as

$$H_{\vec{a}} = \sum_{X\in E_r} a_X h_X, \tag{S.15}$$

where $H = H_{\vec{1}}$ with $\vec{1} = \{1, 1, \dots, 1\}$. Note that there are $|E_r|$ parameters in total. By using Eq. (S.15), we define a parametrized Gibbs state $\rho_{\vec{a}}$ as

$$\rho_{\vec{a}} := \frac{e^{-\beta H_{\vec{a}}}}{Z_{\vec{a}}}, \tag{S.16}$$

where $Z_{\vec{a}} := \mathrm{tr}(e^{-\beta H_{\vec{a}}})$.

In the standard cluster expansion, we consider the Taylor expansion of $e^{-\beta H_{\vec{a}}}$ with respect to the parameters $\vec{a}$. It works well in analyzing a correlation function or tensor network representation, while it is not appropriate to analyze the entropy or effective Hamiltonian of a reduced density matrix. To overcome it, we generalize the standard cluster expansion. We 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 not only as a scholar function but also as a operator function. Here, we choose the conditional mutual information as the function $f_{\vec{a}}$. By using $\rho_{\vec{a}}$, we parameterize the conditional mutual information by $\mathcal{I}_{\vec{a}}(A : C|B)$ in the following form:

$$\begin{aligned}\mathcal{I}_{\vec{a}}(A : C|B) &= -\mathrm{tr}\left[\rho\left(\log\rho_{\vec{a}}^{AB} + \log\rho_{\vec{a}}^{BC} - \log\rho_{\vec{a}}^{ABC} - \log\rho_{\vec{a}}^{B}\right)\right] \\ &= -\mathrm{tr}\left[\rho\left(\log\tilde{\rho}_{\vec{a}}^{AB} + \log\tilde{\rho}_{\vec{a}}^{BC} - \log\tilde{\rho}_{\vec{a}}^{ABC} - \log\tilde{\rho}_{\vec{a}}^{B}\right)\right],\end{aligned} \tag{S.17}$$

where $\rho = \rho_{\vec{1}}$ and we define $\tilde{\rho}_{\vec{a}}$ as

$$\tilde{\rho}_{\vec{a}} := e^{-\beta H_{\vec{a}}} \tag{S.18}$$

with

$$\tilde{\rho}_{\vec{a}}^{L} = \left(e^{-\beta H_{\vec{a}}}\right)^{L} = \mathrm{tr}_{L^c}\left(e^{-\beta H_{\vec{a}}}\right) \otimes \hat{1}_{L^c}. \tag{S.19}$$

Note that we use the definition (S.1) for $\tilde{\rho}_{\vec{a}}^{L}$ $(L \subset V)$

In the following, we define

$$\tilde{H}_{\vec{a}}(A : C|B) := \log\tilde{\rho}_{\vec{a}}^{AB} + \log\tilde{\rho}_{\vec{a}}^{BC} - \log\tilde{\rho}_{\vec{a}}^{ABC} - \log\tilde{\rho}_{\vec{a}}^{B}, \tag{S.20}$$

which gives

$$\mathcal{I}_{\vec{a}}(A : C|B) = \mathrm{tr}\left[\rho\tilde{H}_{\vec{a}}(A : C|B)\right] \le \|\tilde{H}_{\vec{a}}(A : C|B)\|. \tag{S.21}$$

Then, the Taylor expansion with respect to $\vec{a}$ to the operator $\tilde{H}_{\vec{a}}(A : C|B)$ reads

$$\tilde{H}_{\vec{1}}(A : C|B) = \sum_{m=0}^{\infty} \frac{1}{m!}\left[\left(\sum_{X\in E_r}\frac{\partial}{\partial a_X}\right)^m \tilde{H}_{\vec{a}}(A : C|B)\right]_{\vec{a}=\vec{0}}, \tag{S.22}$$

where $\vec{0} = \{0, 0, \dots, 0\}$. By using the cluster notation, we obtain

$$\sum_{X_1, X_2, \dots, X_m \in E_r} = \sum_{w\in\mathcal{C}_{r,m}} n_w, \tag{S.23}$$

which yields

$$\tilde{H}_{\vec{1}}(A : C|B) = \sum_{m=1}^{\infty}\frac{1}{m!}\sum_{X_1, X_2, \dots, X_m\in E_r}\prod_{j=1}^{m}\frac{\partial}{\partial a_{X_j}}\tilde{H}_{\vec{a}}(A : C|B)\Big|_{\vec{a}=\vec{0}} = \sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{C}_{r,m}} n_w\mathcal{D}_w\tilde{H}_{\vec{a}}(A : C|B)\Big|_{\vec{a}=\vec{0}}, \tag{S.24}$$

where $w = \{X_1, X_2 \dots, X_m\}$ and n_w is the multiplicity that w appears in the summation, and we defined

$$\mathcal{D}_w := \prod_{j=1}^{m} \frac{\partial}{\partial a_{X_j}} \quad \text{with} \quad w = \{X_1, X_2 \dots, X_m\}. \tag{S.25}$$

We notice that the partial derivatives $\frac{\partial}{\partial a_X}$ and $\frac{\partial}{\partial a_{X'}}$ commute with each other because $\log(\tilde{\rho}^L_{\vec{a}})$ is a C^∞-smooth function with respect to $\vec{a}$ as long as the system size n is finite. The C^∞-smoothness of $\log(\tilde{\rho}^L_{\vec{a}})$ is proved as follows: For a finite system size n, the C^∞-smoothness of $e^{-\beta H_{\vec{a}}}$ is ensured, and hence $\tilde{\rho}^L_{\vec{a}}$ is also C^∞-smooth from the definition (S.19). Also, we can set

$$\|e^{-\tau \hat{1}} \tilde{\rho}^L_{\vec{a}}\| \le 1. \tag{S.26}$$

by choosing a finite energy $\tau < \infty$ appropriately. Notice that $e^{-\tau \hat{1}} \tilde{\rho}^L_{\vec{a}}$ is Hermitian and $e^{-\tau \hat{1}} \tilde{\rho}^L_{\vec{a}} \succeq 0$. This implies the absolute convergence of the following expansion:

$$\log(\tilde{\rho}^L_{\vec{a}}) = \tau \hat{1} + \log(e^{-\tau \hat{1}} \tilde{\rho}^L_{\vec{a}}) = \tau \hat{1} + \log(\hat{1} + e^{-\tau \hat{1}} \tilde{\rho}^L_{\vec{a}} - \hat{1}) = \tau \hat{1} + \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} (e^{-\tau \hat{1}} \tilde{\rho}^L_{\vec{a}} - \hat{1})^m. \tag{S.27}$$

Thus, the C^∞-smoothness of $\tilde{\rho}^L_{\vec{a}}$ implies of C^∞-smoothness of $\log(\tilde{\rho}^L_{\vec{a}})$.

Note that the case of $m = 0$ (i.e., $|w| = 0$) does not contribute to the expansion because of $\tilde{H}_{\vec{0}}(A : C|B) = 0$. In order to calculate the summation of $\sum_{w \in \mathcal{C}_{r,m}}$, we utilize the following proposition:

Proposition 3. *The cluster expansion (S.24) reduces to the summation of connected clusters which have links between A and C:*

$$\tilde{H}_{\vec{1}}(A : C|B) = \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{w \in \mathcal{G}^{A,C}_{r,m}} n_w \mathcal{D}_w \tilde{H}_{\vec{a}}(A : C|B) \Big|_{\vec{a}=\vec{0}}, \tag{S.28}$$

where the definition of $\mathcal{G}^{A,C}_{r,m}$ has been given in Def. 3.

From this proposition, we only need to estimate the contribution of clusters in $\mathcal{G}^{A,C}_{r,m}$ to upper-bound the conditional mutual information $\mathcal{I}_{\vec{1}}(A : C|B) = \mathrm{tr}[\rho \tilde{H}_{\vec{1}}(A : C|B)]$.

1. *Proof of Proposition 3*

We first introduce the notation $\vec{a}_w$ as a parameter vector such that the elements $\{a_X\}_{X \notin w}$ are vanishing, that is,

$$(\vec{a}_w)_X = 0 \quad \text{for} \quad X \notin w, \tag{S.29}$$

where we denote an element of a_X in $\vec{a}$ by $(\vec{a})_X$. We then obtain

$$\mathcal{D}_w \tilde{H}_{\vec{a}}(A : C|B) \Big|_{\vec{a}=\vec{0}} = \mathcal{D}_w \tilde{H}_{\vec{a}_w}(A : C|B) \Big|_{\vec{a}_w=\vec{0}}. \tag{S.30}$$

In the following, we aim to prove

$$\mathcal{D}_w \tilde{H}_{\vec{a}_w}(A : C|B) \Big|_{\vec{a}_w=\vec{0}} = 0 \quad \text{for} \quad w \notin \mathcal{G}^{A,C}_{r,|w|}. \tag{S.31}$$

We notice that if $w \notin \mathcal{G}^{A,C}_{r,|w|}$ the cluster w satisfies either one of the following two properties (see Figs. 2 (b) and (c)):

$$L_w \cap A = \emptyset \quad \text{or} \quad L_w \cap C = \emptyset \tag{S.32}$$

and

$$w \notin \mathcal{G}_{r,|w|}. \tag{S.33}$$

In the first case (S.32), we can immediately obtain $\tilde{H}_{\vec{a}_w}(A : C|B) = 0$ by choosing $O = e^{-\beta H_{\vec{a}_w}}$ in the lemma 2. In the second case (S.33), there exists a decomposition of $w = w_1 \sqcup w_2$ ($|w_1|, |w_2| > 0$) such that $V_{w_1} \cap V_{w_2} = \emptyset$. Hence, we have $e^{-\beta H_{\vec{a}_w}} = e^{-\beta H_{\vec{a}_{w_1}}} \otimes e^{-\beta H_{\vec{a}_{w_2}}}$, and from Lemma 1 we obtain

$$\tilde{H}_{\vec{a}_w}(A : C|B) = \tilde{H}_{\vec{a}_{w_1}}(A : C|B) + \tilde{H}_{\vec{a}_{w_2}}(A : C|B). \tag{S.34}$$

Because of $\mathcal{D}_{w_2} \tilde{H}_{\vec{a}_{w_1}}(A : C|B) = \mathcal{D}_{w_1} \tilde{H}_{\vec{a}_{w_2}}(A : C|B) = 0$, we have $\mathcal{D}_w \tilde{H}_{\vec{a}_w}(A : C|B) = 0$. This completes the proof of Proposition 3. □

[End of Proof of Proposition 3]

C. Estimation of the expanded terms

In order to estimate the summation (S.28) with respect to $\sum_{w\in\mathcal{G}^{A,C}_{r,m}}$, we consider a derivative of

$$\mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}}\Big|_{\vec{a}=\vec{0}} = \mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}_w}\Big|_{\vec{a}_w=\vec{0}} \tag{S.35}$$

for an arbitrary subsystem $L \subset V$. We choose the subsets AB, BC, ABC and B as L afterward. We here give an explicit form of the derivative $\mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}}$ in the following proposition 4.

Proposition 4. *Let us take $m-1$ copies of the partial Hilbert space $\mathcal{H}^{L^c}$ and distinguish them by $\{\mathcal{H}^{L^c}_j\}_{j=1}^m$. Then, we define the extended Hilbert space as $\mathcal{H}^L \otimes \mathcal{H}^{L^c}_{1:m}$ with*

$$\mathcal{H}^{L^c}_{1:m} := \mathcal{H}^{L^c}_1 \otimes \mathcal{H}^{L^c}_2 \otimes \cdots \otimes \mathcal{H}^{L^c}_m. \tag{S.36}$$

Then, for an arbitrary operator $O \in \mathcal{H}$, we extend the domain of definition and denote $O_{\tilde{\mathcal{H}}_s} \in \mathcal{B}(\mathcal{H}^L \otimes \mathcal{H}^{L^c}_{1:m})$ by the operator which acts only on the space $\mathcal{H}^L \otimes \mathcal{H}^{L^c}_s$. Now, for an arbitrary cluster $w = \{X_1, X_2, \ldots, X_m\}$, we have

$$\mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}}\big|_{\vec{a}=\vec{0}} = \frac{(-\beta)^m}{m! d^m_{L^c}} \mathcal{P}_m \mathrm{tr}_{L^c_{1:m}} \left(\tilde{h}^{(0)}_{X_1} \tilde{h}^{(1)}_{X_2} \cdots \tilde{h}^{(m-1)}_{X_m}\right), \tag{S.37}$$

where $\mathrm{tr}_{L^c_{1:m}}$ denotes the partial trace with respect to the Hilbert space $\mathcal{H}^{L^c}_{1:m}$ and we define

$$\tilde{O}^{(0)} := O_{\tilde{\mathcal{H}}_1}, \quad \tilde{O}^{(s)} := O_{\tilde{\mathcal{H}}_1} + O_{\tilde{\mathcal{H}}_2} + \cdots + O_{\tilde{\mathcal{H}}_s} - sO_{\tilde{\mathcal{H}}_{s+1}} \tag{S.38}$$

for $s = 1, 2, \ldots, m$. Note that $\mathcal{P}_m$ is the symmetrization operator as

$$\mathcal{P}_m \tilde{h}^{(0)}_{X_1} \tilde{h}^{(1)}_{X_2} \cdots \tilde{h}^{(m-1)}_{X_m} = \sum_{\sigma} \tilde{h}^{(0)}_{X_{\sigma_1}} \tilde{h}^{(1)}_{X_{\sigma_2}} \cdots \tilde{h}^{(m-1)}_{X_{\sigma_m}}, \tag{S.39}$$

where $\sum_\sigma$ denotes the summation of $m!$ terms which come from all the permutations.

1. *Proof of Proposition 4*

For the proof, we consider the Taylor expansion with respect to β:

$$\log \tilde{\rho}^L_{\vec{a}} = \sum_{m=0}^{\infty} \frac{\beta^m}{m!} \frac{\partial^m}{\partial \beta^m} \log \tilde{\rho}^L_{\vec{a}}\Big|_{\beta=0}. \tag{S.40}$$

Next, because of

$$\frac{\partial^m}{\partial \beta^m} \log(d_{L^c}) = 0 \quad \text{for} \quad m \geq 1, \tag{S.41}$$

we have

$$\frac{\partial^m}{\partial \beta^m} \log \tilde{\rho}^L_{\vec{a}}\Big|_{\beta=0} = \frac{\partial^m}{\partial \beta^m} \log \left[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}}/d_{L^c})\right]\Big|_{\beta=0} \tag{S.42}$$

for $m \geq 1$.

We aim to prove the following lemma which gives the explicit form of the derivatives with respect to β:

Lemma 5. *The derivatives of $\log \tilde{\rho}^L_{\vec{a}}$ with respect to β can be written as*

$$\frac{\partial^m}{\partial \beta^m} \log \left[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}}/d_{L^c})\right]\Big|_{\beta=0} = \frac{(-1)^m}{d^m_{L^c}} \mathrm{tr}_{L^c_{1:m}} \left(\tilde{H}^{(0)}_{\vec{a}} \tilde{H}^{(1)}_{\vec{a}} \cdots \tilde{H}^{(m-1)}_{\vec{a}}\right), \tag{S.43}$$

where the definitions of $\tilde{H}^{(s)}_{\vec{a}}$ $(s = 0, 1, 2, \ldots, m-1)$ and $\mathcal{H}^{L^c}_{1:m}$ have been given in Eqs. (S.38) and Eq. (S.36), respectively. We give the proof of the lemma afterward.

By assuming the above lemma, we can prove Eq. (S.37) as follows. In considering $\mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}}\big|_{\vec{a}=\vec{0}}$ with $|w|=m$, only the mth order terms of β in the expansion (S.40) contribute to the derivative. Hence, we have

$$\mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}}\big|_{\vec{a}=0} = \frac{\beta^m}{m!}\mathcal{D}_w \left(\frac{\partial^m}{\partial \beta^m} \log\left[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}}/d_{L^c})\right]\right)\bigg|_{\beta=0}. \tag{S.44}$$

By combining Eqs. (S.35), (S.43) and (S.44), we have

$$\begin{aligned}\mathcal{D}_w \log \tilde{\rho}^L_{\vec{a}}\big|_{\vec{a}=0} &= \frac{(-\beta)^m}{m!}\frac{1}{d^m_{L^c}}\mathcal{D}_w \mathrm{tr}_{L^c_{1:m}}\left(\tilde{H}^{(0)}_{\vec{a}}\tilde{H}^{(1)}_{\vec{a}}\cdots\tilde{H}^{(m-1)}_{\vec{a}}\right)\\ &= \frac{(-\beta)^m}{m!}\frac{1}{d^m_{L^c}}\mathcal{P}_m \mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}^{(0)}_{X_1}\tilde{h}^{(1)}_{X_2}\cdots\tilde{h}^{(m-1)}_{X_m}\right).\end{aligned} \tag{S.45}$$

We therefore obtain Eq. (S.37) in Proposition 4. This completes the proof. □

[Proof of Lemma 5] In order to prove Eq. (S.43), we first expand $\log\left[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}}/d_{L^c})\right]$ as follows:

$$\log\left[\frac{\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}})}{d_{L^c}}\right] = \log\left[\hat{1} + \sum_{m=1}^\infty \frac{(-\beta)^m}{m!}\frac{\mathrm{tr}_{L^c}(H^m_{\vec{a}})}{d_{L^c}}\right] = \sum_{q=1}^\infty \frac{(-1)^{q-1}}{q}\left(\sum_{m=1}^\infty \frac{(-\beta)^m}{m!}\frac{\mathrm{tr}_{L^c}(H^m_{\vec{a}})}{d_{L^c}}\right)^q, \tag{S.46}$$

where in the first equation we use the fact that 0th term of the expansion gives $\mathrm{tr}_{L^c}(\hat{1}/d_{L^c}) = \hat{1}$. We then pick up the terms of β^m. Because of

$$\begin{aligned}&\left(\sum_{m=1}^\infty \frac{(-\beta)^m}{m!}\frac{\mathrm{tr}_{L^c}(H^m_{\vec{a}})}{d_{L^c}}\right)^q\\ &= \sum_{m=q}^\infty \sum_{\substack{m_1+m_2+\cdots+m_q=m\\ m_1\ge 1, m_2\ge 1,\ldots,m_q\ge 1}} \frac{(-\beta)^{m_1+m_2+\cdots+m_q}}{m_1!m_2!\cdots m_q!}\frac{\mathrm{tr}_{L^c}(H^{m_1}_{\vec{a}})\mathrm{tr}_{L^c}(H^{m_2}_{\vec{a}})\cdots\mathrm{tr}_{L^c}(H^{m_q}_{\vec{a}})}{d^q_{L^c}},\end{aligned} \tag{S.47}$$

the mth-order term in Eq. (S.46) is given by

$$\beta^m \sum_{q=1}^m \frac{(-1)^{q-1}}{q}\sum_{\substack{m_1+m_2+\cdots+m_q=m\\ m_1\ge 1, m_2\ge 1,\ldots,m_q\ge 1}} \frac{(-1)^m}{m_1!m_2!\cdots m_q!}\frac{\mathrm{tr}_{L^c}(H^{m_1}_{\vec{a}})\mathrm{tr}_{L^c}(H^{m_2}_{\vec{a}})\cdots\mathrm{tr}_{L^c}(H^{m_q}_{\vec{a}})}{d^q_{L^c}}. \tag{S.48}$$

We thus obtain

$$\begin{aligned}&\frac{\partial^m}{\partial\beta^m}\log\left[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}}/d_{L^c})\right]\bigg|_{\beta=0}\\ &= \sum_{q=1}^m \frac{(-1)^{q-1}}{q}\sum_{\substack{m_1+m_2+\cdots+m_q=m\\ m_1\ge 1, m_2\ge 1,\ldots,m_q\ge 1}} \frac{m!(-1)^m}{m_1!m_2!\cdots m_q!}\frac{\mathcal{P}_q\mathrm{tr}_{L^c}(H^{m_1}_{\vec{a}})\mathrm{tr}_{L^c}(H^{m_2}_{\vec{a}})\cdots\mathrm{tr}_{L^c}(H^{m_q}_{\vec{a}})}{q!d^q_{L^c}},\end{aligned} \tag{S.49}$$

where $\mathcal{P}_q$ is the symmetrization operator with respect to $\{m_1, m_2, \ldots, m_q\}$. In the same manner, we can formally expand

$$\begin{aligned}&\frac{(-1)^m}{d^m_{L^c}}\mathrm{tr}_{L^c_{1:m}}\left(\tilde{H}^{(0)}_{\vec{a}}\tilde{H}^{(1)}_{\vec{a}}\cdots\tilde{H}^{(m-1)}_{\vec{a}}\right)\\ &= \sum_{q=1}^m \sum_{\substack{m_1+m_2+\cdots+m_q=m\\ m_1\ge 1, m_2\ge 1,\ldots,m_q\ge 1}} \mathcal{C}^{(q)}_{m_1,m_2,\ldots,m_q}\mathcal{P}_q\mathrm{tr}_{L^c}(H^{m_1}_{\vec{a}})\mathrm{tr}_{L^c}(H^{m_2}_{\vec{a}})\cdots\mathrm{tr}_{L^c}(H^{m_q}_{\vec{a}}).\end{aligned} \tag{S.50}$$

For the proof of Lemma 5, we need to check whether each of the coefficients of $\mathcal{P}_q\mathrm{tr}_{L^c}(H^{m_1}_{\vec{a}})\mathrm{tr}_{L^c}(H^{m_2}_{\vec{a}})\cdots\mathrm{tr}_{L^c}(H^{m_q}_{\vec{a}})$ for all the pairs of $\{m_1, m_2, \ldots, m_q\}$ is equal between Eqs. (S.49) and (S.50). Instead of directly writing down the explicit form of $\mathcal{C}^{(q)}_{m_1,m_2,\ldots,m_q}$, we will take the following step. First, we prove

$$\frac{\partial^m}{\partial\beta^m}\log\left[\mathrm{tr}_{L^c}(e^{-\beta H_{\vec{a}}}/d_{L^c})\right]\bigg|_{\beta=0} = \frac{(-1)^m}{d^m_{L^c}}\mathrm{tr}_{L^c_{1:m}}\left(\tilde{H}^{(0)}_{\vec{a}}\tilde{H}^{(1)}_{\vec{a}}\cdots\tilde{H}^{(m-1)}_{\vec{a}}\right) \tag{S.51}$$

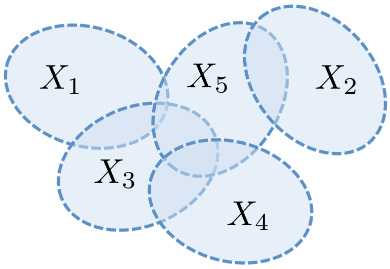

FIG. 3. $N_{X_s|w}$ is defined by a number of subsystems in w that have overlap with X_s. When $w = \{X_1, X_2, X_3, X_4, X_5\}$ is given as above, we have $N_{X_1|w} = 2$, $N_{X_2|w} = 1$, $N_{X_3|w} = 2$, $N_{X_4|w} = 2$ and $N_{X_5|w} = 4$.

in the case of $L^c = V$. The proof of Eq. (S.51) implies that the coefficients of $\mathcal{P}_q \mathrm{tr}_{L^c}(H_{\vec{a}}^{m_1})\mathrm{tr}_{L^c}(H_{\vec{a}}^{m_2})\cdots\mathrm{tr}_{L^c}(H_{\vec{a}}^{m_q})$ are equal between Eqs. (S.49) and (S.50) for $L^c = V$. Then, because the coefficients $\mathcal{C}^{(q)}_{m_1,m_2,\ldots,m_q}$ do not depend on the form of L^c, the proof in the case of $L^c = V$ also results in the proof in the other cases (i.e., $L^c \neq V$). Therefore, in the following, we aim to give the proof of Eq. (S.51) for $L^c = V$.

For $L^c = V$, we have

$$\frac{\partial}{\partial \beta} \log\left[\frac{\mathrm{tr}_V(e^{-\beta H_{\vec{a}}})}{d_V}\right] = -\mathrm{tr}(H_{\vec{a}}\rho_{\vec{a}}), \tag{S.52}$$

and hence our task is to calculate

$$\frac{\partial^m}{\partial \beta^m} \log\left[\frac{\mathrm{tr}_V(e^{-\beta H_{\vec{a}}})}{d_V}\right] = -\mathrm{tr}_V\left(H_{\vec{a}}\frac{\partial^{m-1}}{\partial\beta^{m-1}}\rho_{\vec{a}}\right). \tag{S.53}$$

By using Lemma 2 in Ref. [1], we have

$$\frac{\partial^{m-1}}{\partial\beta^{m-1}}\mathrm{tr}\left(H_{\vec{a}}\rho_{\vec{a}}\right)\Big|_{\beta=0} = \frac{(-1)^{m-1}}{d_V^m}\mathrm{tr}_{V^c_{1:m}}\left(\tilde{H}^{(0)}_{\vec{a}}\tilde{H}^{(1)}_{\vec{a}}\cdots\tilde{H}^{(m-1)}_{\vec{a}}\right), \tag{S.54}$$

where in the inequality (B.3) in [1], we choose as $m_1 = 0$, $m_2 = m-1$ and $\omega_X = H_{\vec{a}}$. We thus obtain the equation (S.51). This completes the proof of Lemma 5. $\square$

---[**End of Proof of Proposition 4**]

We then aim to obtain an upper bound of $\left\|\mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}^{(0)}_{X_1}\tilde{h}^{(1)}_{X_2}\cdots\tilde{h}^{(m-1)}_{X_m}\right)\right\|$. For the purpose, we utilize the following proposition.

Proposition 6. *Let $\{O_s\}_{s=0}^m$ be operators supported on a subset $w := \{X_s\}_{s=0}^m$, respectively. When they satisfy $\mathrm{tr}_{L^c}(O_s) = 0$ for $s = 0, 1, 2\ldots, m$, we obtain*

$$\frac{1}{d_{L^c}^m}\left\|\mathrm{tr}_{L^c_{1:m}}\left(\tilde{O}^{(0)}_0\tilde{O}^{(1)}_1\tilde{O}^{(2)}_2\cdots\tilde{O}^{(m-1)}_{m-1}\right)\right\| \le \|O_0\|\prod_{s=1}^m 2N_{X_s|w_L}\|O_s\|, \tag{S.55}$$

where we define $\tilde{O}^{(s)}_s$ as in Eq. (S.38). $N_{X_s|w}$ is a number of subsets in w that have overlap with X_s (Fig. 3):

$$N_{X_s|w} = \#\{X \in w | X \neq X_s, X \cap X_s \neq \emptyset\}. \tag{S.56}$$

The proof is the same as that of Proposition 3 in Ref. [1], which proves Ineq. (S.55) for $L^c = V$.

In order to apply Proposition (6) to $\mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}^{(0)}_{X_1}\tilde{h}^{(1)}_{X_2}\cdots\tilde{h}^{(m-1)}_{X_m}\right)$, the condition $\mathrm{tr}_{L^c}(h_X) = 0$ is necessary, whereas it is not generally satisfied. Thus, instead of considering h_X, we consider $\mathfrak{h}_X$ which is defined as follows:

$$\mathfrak{h}_X := h_X - \frac{h_X^L}{d_{L^c}} \quad \text{for} \quad X \in E_r, \tag{S.57}$$

where $\mathfrak{h}_X$ satisfies $\mathrm{tr}_{L^c}(\mathfrak{h}_X) = \mathrm{tr}_{L^c}(h_X) - h_X^L \mathrm{tr}_{L^c}(\hat{1})/d_{L^c} = h_X^L - h_X^L = 0$ from the definition (S.1). By using the notation of $\mathfrak{h}_X$, we obtain

$$\mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}^{(0)}_{X_1}\tilde{h}^{(1)}_{X_2}\cdots\tilde{h}^{(m-1)}_{X_m}\right) = \mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}^{(0)}_{X_1}\tilde{\mathfrak{h}}^{(1)}_{X_2}\cdots\tilde{\mathfrak{h}}^{(m-1)}_{X_m}\right) + \frac{h^L_{X_1}}{d_{L^c}}\otimes\mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}^{(1)}_{X_2}\tilde{\mathfrak{h}}^{(2)}_{X_3}\cdots\tilde{\mathfrak{h}}^{(m-1)}_{X_m}\right), \tag{S.58}$$

where we use $\tilde{\mathfrak{h}}_X^{(s)} = \tilde{h}_X^{(s)}$ for $s \geq 1$ which comes from the definition (S.38), and apply Eq. (S.57) to $\tilde{h}_{X_1}^{(0)}$. We then prove $\mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}_{X_2}^{(1)}\tilde{\mathfrak{h}}_{X_3}^{(2)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right) = 0$. By using the definition (S.38) for $\tilde{\mathfrak{h}}_{X_2}^{(1)}$, we have

$$\mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}_{X_2}^{(1)}\tilde{\mathfrak{h}}_{X_3}^{(2)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right) = \mathrm{tr}_{L^c_{1:m}}\left[\left(\tilde{\mathfrak{h}}_{X_2,\tilde{\mathcal{H}}_1} - \tilde{\mathfrak{h}}_{X_2,\tilde{\mathcal{H}}_2}\right)\tilde{\mathfrak{h}}_{X_3}^{(2)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right]. \tag{S.59}$$

Because the operator $\tilde{\mathfrak{h}}_X^{(s)}$ $(s \geq 2)$ is invariant under the swapping between the Hilbert spaces $H_1^{L^c}$ and $H_2^{L^c}$ (i.e., $\tilde{\mathcal{H}}_1 \leftrightarrow \tilde{\mathcal{H}}_2$), we have

$$\mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}_{X_2,\tilde{\mathcal{H}}_1}\tilde{\mathfrak{h}}_{X_3}^{(2)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right) = \mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}_{X_2,\tilde{\mathcal{H}}_2}\tilde{\mathfrak{h}}_{X_3}^{(2)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right). \tag{S.60}$$

Therefore, the term (S.59) vanishes and Eq. (S.58) reduces to

$$\mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}_{X_1}^{(0)}\tilde{h}_{X_2}^{(1)}\cdots\tilde{h}_{X_m}^{(m-1)}\right) = \mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}_{X_1}^{(0)}\tilde{\mathfrak{h}}_{X_2}^{(1)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right). \tag{S.61}$$

By using Proposition 6, we obtain an upper bound of $\mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}_{X_1}^{(0)}\tilde{h}_{X_2}^{(1)}\cdots\tilde{h}_{X_m}^{(m-1)}\right)$ as follows:

$$\begin{aligned}\frac{1}{d_{L^c}^m}\left\|\mathrm{tr}_{L^c_{1:m}}\left(\tilde{h}_{X_1}^{(0)}\tilde{h}_{X_2}^{(1)}\cdots\tilde{h}_{X_m}^{(m-1)}\right)\right\| =& \frac{1}{d_{L^c}^m}\left\|\mathrm{tr}_{L^c_{1:m}}\left(\tilde{\mathfrak{h}}_{X_1}^{(0)}\tilde{\mathfrak{h}}_{X_2}^{(1)}\cdots\tilde{\mathfrak{h}}_{X_m}^{(m-1)}\right)\right\| \\ \leq& \|\mathfrak{h}_{X_1}\|\prod_{s=2}^m 2N_{X_s|w}\|\mathfrak{h}_{X_s}\| \leq \frac{1}{2}\prod_{s=1}^m 4N_{X_s|w}\|h_{X_s}\|,\end{aligned} \tag{S.62}$$

where we use $\|\mathfrak{h}_X\| \leq 2\|h_X\|$ which comes from the definition (S.57). By combining the inequality (S.62) with Eq. (S.37), we obtain an upper bound of

$$\left\|\mathcal{D}_w \log\tilde{\rho}_{\vec{a}}^L\big|_{\vec{a}=\vec{0}}\right\| \leq \frac{1}{2}\prod_{s=1}^m 4\beta N_{X_s|w}\|h_{X_s}\|. \tag{S.63}$$

By applying the inequality (S.63) to the cases $L = AB$, $L = BC$, $L = ABC$ and $L = B$, we obtain the following inequality:

$$\left\|\mathcal{D}_w\tilde{H}_{\vec{a}}(A:C|B)\big|_{\vec{a}=\vec{0}}\right\| \leq 2(4\beta)^m\prod_{s=1}^m N_{X_s|w}\|h_{X_s}\|, \tag{S.64}$$

where $\tilde{H}_{\vec{a}}(A:C|B)$ has been defined in Eq. (S.20). Then, the final task is to upper-bound the summation with respect to $\sum_{w\in\mathcal{G}_{r,m}^{A,C}}$ in Eq. (S.28):

$$\begin{aligned}\left\|\tilde{H}_{\vec{1}}(A:C|B)\right\| \leq& \sum_{m=1}^\infty\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m}^{A,C}} n_w\left\|\mathcal{D}_w\tilde{H}_{\vec{a}}(A:C|B)\Big|_{\vec{a}=\vec{0}}\right\| \\ \leq& \sum_{m=1}^\infty\frac{2(4\beta)^m}{m!}\sum_{w\in\mathcal{G}_{r,m}^{A,C}} n_w\prod_{s=1}^m N_{X_s|w}\|h_{X_s}\|,\end{aligned} \tag{S.65}$$

where we use the proposition 3 in the first inequality.

For the estimation of the summation, we first focus on the fact that any cluster in $w \in \mathcal{G}_{r,m}^{A,C}$ must have overlaps with the surface regions of A and C, say ∂A_r and ∂C_r $(r \in \mathbb{N})$:

$$\partial A_r := \{v \in A | d_{v,A^c} \leq r\}, \quad \partial C_r := \{v \in C | d_{v,C^c} \leq r\}. \tag{S.66}$$

Second, because $d_{A,C}$ is the minimum path length on the graph (V, E) to connect the subsystems A and C, the condition $w \in \mathcal{G}_{r,m}^{A,C}$ implies $|w| \geq d_{A,C}/r$ as the necessary condition. From these two fact, we will replace the summation $\sum_{w\in\mathcal{G}_{r,m}^{A,C}}$ with $\sum_{v\in\partial A_r}\sum_{m\geq d_{A,C}/r}\sum_{w\in\mathcal{G}_{r,m}^v}$ by taking all the clusters with the sizes $|w| \geq d_{A,C}/r$ which have overlap with A into account:

$$\sum_{m=1}^\infty\frac{2(4\beta)^m}{m!}\sum_{w\in\mathcal{G}_{r,m}^{A,C}} n_w\prod_{s=1}^m N_{X_s|w}\|h_{X_s}\| \leq \sum_{v\in\partial A_r}\sum_{m\geq d_{A,C}/r} 2(4\beta)^m\sum_{w\in\mathcal{G}_{r,m}^v}\frac{n_w}{m!}\prod_{s=1}^m N_{X_s|w}\|h_{X_s}\|, \tag{S.67}$$

where the same inequality holds for the replacement of $\sum_{v\in\partial A_r}$ by $\sum_{v\in\partial C_r}$.

In order to estimate the summation of $\sum_{w\in\mathcal{G}_{r,m}^v}$, we utilize the following proposition which has been given in Ref. [1]:

Proposition 7 (Proposition 4 in Ref. [1]). *Let $\{o_X\}_{X\in E_\infty}$ be arbitrary operators such that*

$$\sum_{X|X\ni v} \|o_X\| \le g \quad \textit{for} \quad \forall v\in V, \tag{S.68}$$

where E_∞ is defined by Eq. (S.5) and it gives the set of all the subsystems $X\subset V$ with $|X|\le k$. Then, for an arbitrary subset L, we obtain

$$\sum_{w\in\mathcal{G}_m^L} \frac{n_w}{m!}\prod_{s=1}^m N_{X_s|w_L}\|o_{X_s}\| \le \frac{1}{2}e^{|L|/k}(2e^3gk)^m, \tag{S.69}$$

where w_L is defined as $w_L := \{L, X_1, X_2, \dots, X_{|w|}\}$ for $w=\{X_1,X_2,\dots,X_{|w|}\}$.

By applying Proposition 7 to the inequality (S.67), we have

$$\sum_{w\in\mathcal{G}_{r,m}^v} \frac{n_w}{m!}\prod_{s=1}^m N_{X_s|w}\|h_{X_s}\| \le \frac{1}{2}e^{1/k}(2e^3k)^m, \tag{S.70}$$

where we use $N_{X_s|w_L}\le N_{X_s|w}$ in (S.69) and the condition (S.4) gives $g=1$. Therefore, the inequality (S.67) reduces to

$$\begin{aligned}\sum_{m=0}^\infty \frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m}^{A,C}} n_w \left\|\mathcal{D}_w\tilde{H}_{\vec{a}}(A:C|B)\right\|_{\vec{a}=\vec{0}}\big| &\le \sum_{v\in\partial A_r}\sum_{m\ge d_{A,C}/r} e^{1/k}(8e^3k\beta)^m \\ &\le e|\partial A_r|\frac{(8e^3k\beta)^{d_{A,C}/r}}{1-8e^3k\beta},\end{aligned} \tag{S.71}$$

where we use $k\ge 1$. We notice that the same inequality holds for the replacement of $|\partial A_r|$ by $|\partial C_r|$. By combining the inequalities (S.21), (S.65) and (S.71), we prove Theorem 1. □

II. QUASI-LOCALITY OF EFFECTIVE HAMILTONIAN ON A SUBSYSTEM: PROOF OF THEOREM 3

We here consider the effective Hamiltonian on a subsystem L, which we define as

$$\tilde{H}_L := -\beta^{-1}\log\tilde{\rho}^L, \tag{S.72}$$

where $\tilde{\rho}^L$ is defined in Eq. (S.19). We prove the following theorem which refines the Theorem 3:

Theorem 8. *The effective Hamiltonian $\tilde{H}_L$ is given by a quasi-local operator*

$$\tilde{H}_L = H_L + \sum_{m=1}^\infty \sum_{w\in\mathcal{G}_{r,m}^{L,L^{\rm c}}} n_w h_{L_w} - \frac{\hat{1}}{\beta}\log Z_{L^{\rm c}} \tag{S.73}$$

with

$$H_L := \sum_{X\subset L} h_X, \quad Z_{L^{\rm c}} := \frac{1}{d_L}\mathrm{tr}(e^{-\beta H_{L^{\rm c}}}\otimes \hat{1}_L) \tag{S.74}$$

for $L\subset V$, where each of $\{h_{L_w}\}_{w\in\mathcal{G}_{r,m}^{L,L^{\rm c}}}$ is supported on the subsystem $L_w := L\cap V_w$ (see Def. (S.89)) and $\mathcal{G}_{r,m}^{L,L^{\rm c}}$ is defined as a cluster subset defined in Def. 3. The effective interaction terms $\{h_{L_w}\}_{w\in\mathcal{G}_{r,m}^{L,L^{\rm c}}}$ is exponentially localized around the boundary:

$$\sum_{m>m_0}^\infty \sum_{w\in\mathcal{G}_{r,m}^{L,L^{\rm c}}} n_w\|h_{L_w}\| \le \frac{e}{4\beta}\frac{(\beta/\beta_c)^{m_0+1}}{1-\beta/\beta_c}|\partial L_r| \tag{S.75}$$

for an arbitrary m_0.

From Eq. (S.73), the effective interaction term Φ_L is given by

$$\Phi_L = \sum_{m=1}^{\infty} \sum_{w\in \mathcal{G}_{r,m}^{L,L^c}} n_w h_{L_w} - \frac{\hat{1}}{\beta} \log Z_{L^c}. \tag{S.76}$$

Because of $\mathrm{diam}(V_w) \le mr$, the subsystem $L \cap V_w$ ($w \in \mathcal{G}_{r,m}^{L,L^c}$) is separated from the boundary ∂L at most by a distance mr, namely $L \cap V_w \subseteq \partial L_{mr}$, where ∂L_l was defined as follows:

$$\partial L_l := \{v \in L | d_{v,L^c} \le l\}. \tag{S.77}$$

Hence, by defining $\Phi_{\partial L_l}$ as

$$\Phi_{\partial L_l} = \sum_{m\le \lfloor l/r \rfloor} \sum_{w\in \mathcal{G}_{r,m}^{L,L^c}} n_w h_{L_w} - \frac{\hat{1}}{\beta} \log Z_{L^c}, \tag{S.78}$$

we have

$$\|\Phi_L - \Phi_{\partial L_l}\| \le \frac{e}{4\beta} \frac{(\beta/\beta_c)^{l/r}}{1-\beta/\beta_c} |\partial L_r|. \tag{S.79}$$

This gives the proof of Theorem 3.

A. Proof of Theorem 8

In order to apply the generalized cluster expansion, we first parametrize $\tilde{H}_L$ as

$$\tilde{H}_{L,\vec{a}} := -\beta^{-1} \log \tilde{\rho}_{\vec{a}}^L. \tag{S.80}$$

As in Eq. (S.24), the generalized cluster expansion for $\tilde{H}_{L,\vec{a}}$ reads

$$\tilde{H}_{L,\vec{1}} = -\frac{1}{\beta} \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{w\in \mathcal{C}_{r,m}} n_w \mathcal{D}_w \tilde{H}_{L,\vec{a}} \Big|_{\vec{a}=\vec{0}}. \tag{S.81}$$

We can now prove the following proposition:

Proposition 9. *The summation with respect to the clusters $\sum_{w\in\mathcal{C}_{r,m}}$ reduces to the following form:*

$$\tilde{H}_{L,\vec{1}} = H_L - \frac{1}{\beta} \log Z_{L^c} + \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{w\in \mathcal{G}_{r,m}^{L,L^c}} n_w \mathcal{D}_w \tilde{H}_{L,\vec{a}} \Big|_{\vec{a}=\vec{0}}, \tag{S.82}$$

where $H_L := \sum_{X\subset L} h_X$ and $Z_{L_c} := d_L^{-1} \mathrm{tr}(e^{-\beta H_{L^c}})$.

1. Proof of Proposition 9

For the proof, we first prove

$$\mathcal{D}_w \log(\tilde{\rho}_{\vec{a}_w}^L) = 0 \quad \text{for} \quad w \notin \mathcal{G}_{r,|w|}. \tag{S.83}$$

The proof is given as follows. Due to the existence of decomposition $w = w_1 \sqcup w_2$ such that $V_{w_1} \cap V_{w_2} = \emptyset$, we have $e^{-\beta H_{\vec{a}_w}} = e^{-\beta H_{\vec{a}_{w_1}}} \otimes e^{-\beta H_{\vec{a}_{w_2}}}$ and hence,

$$\log(\tilde{\rho}_{\vec{a}_w}^L) = \log(\tilde{\rho}_{\vec{a}_{w_1}}^L) + \log(\tilde{\rho}_{\vec{a}_{w_2}}^L) - \log d_{L^c}. \tag{S.84}$$

Because $\mathcal{D}_{w_2} \log(\tilde{\rho}_{\vec{a}_{w_1}}^L) = \mathcal{D}_{w_1} \log(\tilde{\rho}_{\vec{a}_{w_2}}^L) = 0$, we obtain Eq. (S.83).

We then consider the cases of $V_w \subseteq L$ and $V_w \subseteq L^c$ in Eq. (S.81). In the case of $V_w \subseteq L$, the definition (S.19) gives

$$\log(\tilde{\rho}_{L,\vec{a}_w}) = -\beta H_{\vec{a}_w} + \log d_{L^c}. \tag{S.85}$$

Therefore, we have $\mathcal{D}_w \log(\tilde{\rho}^L_{\vec{a}_w})$ vanishes for $m \geq 2$, and

$$-\frac{1}{\beta}\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}} = \sum_{X\subset L} h_X = H_L. \tag{S.86}$$

On the other hand, in the case of $V_w \subseteq L^{\rm c}$, $\log(\tilde{\rho}^L_{\vec{a}_w})$ becomes a constant operator (i.e., $\log(\tilde{\rho}^L_{\vec{a}_w}) \propto \hat{1}$). Hence, we obtain

$$-\frac{1}{\beta}\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L^{\rm c}} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}} = -\frac{1}{\beta}\log[{\rm tr}_{L^{\rm c}}(e^{-\beta H_{L^{\rm c}}})] = -\frac{\log Z_{L_{\rm c}}}{\beta}\hat{1}. \tag{S.87}$$

Thus, the summation (S.81) reduces to

$$\begin{aligned}\tilde{H}_{L,\vec{1}} =& -\frac{1}{\beta}\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}} - \frac{1}{\beta}\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L^{\rm c}} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}} \\ &-\frac{1}{\beta}\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}^{L,L^{\rm c}}_{r,m}} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}} \\ =& H_L - \frac{1}{\beta}\log Z_{L^{\rm c}} - \frac{1}{\beta}\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}^{L,L^{\rm c}}_{r,m}} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}.\end{aligned} \tag{S.88}$$

This completes the proof. □

———————————————————————————————— **[End of Proof of Proposition 9]**

We now define h_{L_w} as

$$h_{L_w} := \frac{-\beta^{-1}}{m!}\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}, \tag{S.89}$$

where $w \in \mathcal{G}^{L,L^{\rm c}}_{r,m}$. Note that the operator h_{L_w} is supported on the subsystem $L_w = L\cap V_w$. Then, the effective Hamiltonian $\tilde{H}_{L,\vec{1}}$ is formally written by

$$\tilde{H}_{L,\vec{1}} = H_L - \frac{1}{\beta}\log Z_{L^{\rm c}} + \sum_{m=1}^{\infty}\sum_{w\in\mathcal{G}^{L,L^{\rm c}}_{r,m}} n_w h_{L_w}. \tag{S.90}$$

By using the proposition 4 with the inequalities (S.62) and (S.70), we have

$$\begin{aligned}\sum_{w\in\mathcal{G}^{v}_{r,m}} n_w\|h_{L_w}\| &\leq \frac{\beta^{-1}}{m!}\sum_{w\in\mathcal{G}^{v}_{r,m}}\frac{n_w}{2}\prod_{s=1}^{m}4\beta N_{X_s|w}\|h_{X_s}\| \\ &\leq (4\beta)^{m-1}e^{1/k}(2e^3k)^m \leq \frac{e}{4\beta}(\beta/\beta_c)^m,\end{aligned} \tag{S.91}$$

where we use $e^{1/k} \leq e$ due to $k \geq 1$. By using the above inequality, the contribution of mth order terms in the expansion (S.82) is bounded from above by

$$\sum_{w\in\mathcal{G}^{L,L^{\rm c}}_{r,m}} n_w\|h_{L_w}\| \leq \sum_{v\in\partial L_r}\sum_{w\in\mathcal{G}^{v}_{r,m}} n_w\|h_{L_w}\| \leq \frac{e}{4\beta}(\beta/\beta_c)^m|\partial L_r|, \tag{S.92}$$

where ∂L_r has been defined in Eq. (S.77).

$$\sum_{m>m_0}^{\infty}\sum_{w\in\mathcal{G}^{L,L^{\rm c}}_{r,m}} n_w\|h_{L_w}\| \leq \frac{e|\partial L_r|}{4\beta}\sum_{m=m_0+1}^{\infty}(\beta/\beta_c)^m = \frac{e|\partial L_r|}{4\beta}\frac{(\beta/\beta_c)^{m_0+1}}{1-\beta/\beta_c}. \tag{S.93}$$

This completes the proof of Theorem 8. □

B. Computational cost of cluster summation

We here show the computational cost to estimate the effective Hamiltonian $\tilde{H}_L$. For this aim, we start from a slightly weaker expression than Eq. (S.82) as follows

$$\Phi_L = \tilde{H}_{L,\vec{1}} - H_L = -\frac{1}{\beta}\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L^c} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}} - \frac{1}{\beta}\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m}^{L,L^c}} n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}, \tag{S.94}$$

where we use the second and third terms in the first equation of (S.88). Our task is to estimate the computational cost of $n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\big|_{\vec{a}=\vec{0}}$ and the number of clusters in $\{w\in\mathcal{G}_{r,m}, V_w\subseteq L^c\}$ and $w\in\mathcal{G}_{r,m}^{L,L^c}$.

First, we consider $n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\big|_{\vec{a}=\vec{0}}$. As defined in Eq. (S.24), n_w is immediately calculated, and hence we need to estimate the computational cost to calculate the multiderivative

$$\mathcal{D}_w\tilde{H}_{L,\vec{a}_w}\Big|_{\vec{a}_w=\vec{0}} = \prod_{j=1}^{m}\frac{\partial}{\partial a_{X_j}}\tilde{H}_{L,\vec{a}_w}\Big|_{\vec{a}_w=\vec{0}} \tag{S.95}$$

with $w=\{X_s\}_{s=1}^m$ by using numerical differentiation. The operator $\tilde{H}_{L,\vec{a}_w}$ is given by

$$\tilde{H}_{L,\vec{a}_w} = -\beta^{-1}\log\tilde{\rho}_{\vec{a}_w}^L = -\beta^{-1}\mathrm{tr}_{L^c}\left(e^{-\beta H_{\vec{a}_w}}\right)\otimes\hat{1}_{L^c}, \tag{S.96}$$

where we use the definition (S.19). Note that $H_{\vec{a}_w}$ is supported on $V_w\subset V$. Hence, the computational cost to calculate $\tilde{H}_{L,\vec{a}_w}$ is at most of $d^{\mathcal{O}(|V_w|)}$. In order to perform the differentiation, we need to calculate $2^{|w|}$ values of $\tilde{H}_{L,\vec{a}_w}$ for $a_{X_s}=\pm\Delta$ ($\Delta\to+0$) for $s=1,2,\dots,|w|$. Thus, for the numerical differentiation we need the computational cost of $2^{|w|}\cdot d^{\mathcal{O}(|V_w|)} = d^{\mathcal{O}(mk)}$ with $|w|=m$, where we use $|V_w|\le|w|k$.

We then need to sum up the contributions from all the clusters in $\{w\in\mathcal{G}_{r,m}, V_w\subseteq L^c\}$ and $w\in\mathcal{G}_{r,m}^{L,L^c}$. For the purpose, we first prove the following theorem on the number of clusters:

Proposition 10. *The total number of different clusters in $\mathcal{G}_{r,m}^{L^c}$ is bounded as follows:*

$$\#\left\{w\in\mathcal{C}_{r,m}\middle| w\in\mathcal{G}_{r,m}, V_w\subseteq L^c \quad \text{or} \quad w\in\mathcal{G}_{r,m}^{L,L^c}\right\} \le |L^c|\left(3\cdot 2^k d_G^{rk}\right)^m. \tag{S.97}$$

This roughly gives the total number by $|L^c|d_G^{\mathcal{O}(rkm)}$,

In total, the computation of the m-th order in the expansion (S.94) is performed with the runtime bounded from above by

$$d^{\mathcal{O}(mk)}\cdot|L^c|d_G^{\mathcal{O}(rkm)} \le n(d\cdot d_G^r)^{mk}. \tag{S.98}$$

Also, the convergence of the expansion (S.94) is estimated as in (S.92) and (S.93)

$$\begin{aligned}&\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L^c}\left\|n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}\right\| - \sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m}^{L,L^c}}\left\|n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}\right\| \\ &\le \sum_{v\in L^c}\sum_{w\in\mathcal{G}_{r,m}^v} n_w\|h_{L_w}\| \le \frac{e}{4\beta}(\beta/\beta_c)^m|L^c| \le \frac{e}{4\beta}(\beta/\beta_c)^m n,\end{aligned} \tag{S.99}$$

which yields

$$\sum_{m>m_0}^{\infty}\sum_{w\in\mathcal{G}_{r,m},V_w\subseteq L^c}\left\|n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}\right\| - \sum_{m>m_0}^{\infty}\sum_{m=1}^{\infty}\frac{1}{m!}\sum_{w\in\mathcal{G}_{r,m}^{L,L^c}}\left\|n_w\mathcal{D}_w\tilde{H}_{L,\vec{a}}\Big|_{\vec{a}=\vec{0}}\right\| \le \frac{en}{4\beta}\frac{(\beta/\beta_c)^{m_0+1}}{1-\beta/\beta_c}. \tag{S.100}$$

Therefore, we need to choose $m=\mathcal{O}(\log(1/\epsilon))$ to calculate Φ_L up to an error $n\epsilon$ as long as $\beta<\beta_c$. Hence, the computational cost is estimated as

$$n(d\cdot d_G^r)^{k\mathcal{O}(\log(1/\epsilon))} = n(1/\epsilon)^{\mathcal{O}(k\log(dd_G^r))}. \tag{S.101}$$

This completes the derivation of the computational cost in Theorem 3 for calculating Φ_L. □

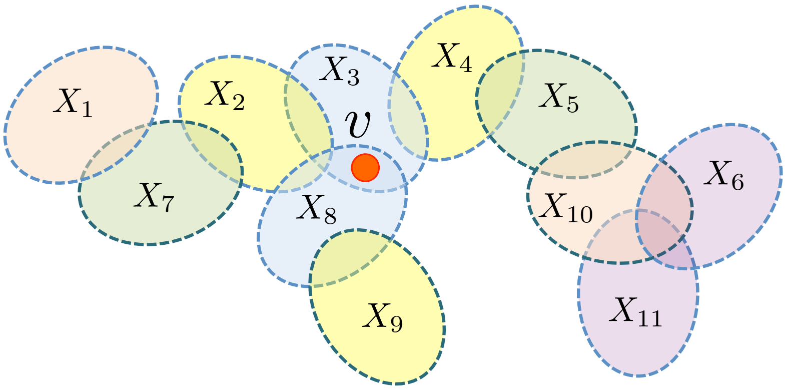

FIG. 4. Decomposition of w in $\mathcal{G}^v_{r,m}$ as in Eq. (S.107). In the picture, we have $w_0 = \{X_3, X_8\}$, $w_1 = \{X_2, X_4, X_9\}$, $w_2 = \{X_5, X_7\}$, $w_3 = \{X_1, X_{10}\}$, $w_4 = \{X_6, X_{11}\}$.

1. *Proof of Proposition 10*

We here prove Proposition 10 which gives an upper bound of the number of cluster connecting to a subset L^{c}. For the purpose, we estimate the number of clusters in $\mathcal{G}^v_{r,m}$, which gives an upper bound of

$$\#\left\{w \in \mathcal{C}_{r,m} \middle| w \in \mathcal{G}_{r,m}, V_w \subseteq L^{\mathrm{c}} \quad \text{or} \quad w \in \mathcal{G}^{L,L^{\mathrm{c}}}_{r,m}\right\} \le \sum_{v \in L^{\mathrm{c}}} \#\left\{w \middle| w \in \mathcal{G}^v_{r,m}\right\}. \tag{S.102}$$

First, we count the number of clusters $w = \{X_s\}_{s=1}^q$ which satisfy $X_s \cap Y \neq \emptyset$ for $\forall X_s$ $(s = 1, 2, \dots, q)$, where Y is an arbitrary subset in V. The number is bounded from above by

$$\#\left\{w \in \mathcal{C}_{r,q} | X_s \cap Y \neq \emptyset,\ s = 1, 2, \dots, q\right\} \le \sum_{\{v_1, v_2, \dots, v_q\} \subseteq Y} \prod_{s=1}^q \deg(v_s), \tag{S.103}$$

where we define $\deg(v)$ as $\deg(v) := \#\left\{X \in E_r | X \ni v\right\}$. By using the graph degree d_G, we can upper-bound $\deg(v)$ by

$$\deg(v) = \#\left\{X \in E_r | X \ni v\right\} \le \binom{d_G^r}{k} \le d_G^{rk}, \tag{S.104}$$

where d_G^r is the upper bound of the number of vertices $\{v'\}_{v' \in V}$ such that $d_{v,v'} \le r$. Also, note that $X \in E_r$ implies $|X| \le k$ from the definitions (S.4) and (S.5). The summation with respect to $\{v_1, v_2, \dots, v_q\}$ is equal to the q_1-multicombination from a set of $|L|$ vertices, which is equal to

$$\sum_{\{v_1, v_2, \dots, v_q\} \subseteq Y} = \left(\!\binom{|Y|}{q}\!\right) = \binom{q + |Y| - 1}{q} \le 2^{q + |Y| - 1}. \tag{S.105}$$

By combining the inequalities (S.104) and (S.105) with (S.103), we obtain

$$\#\left\{w \in \mathcal{C}_{r,q} | X_s \cap Y \neq \emptyset,\ s = 1, 2, \dots, q\right\} \le 2^{|Y|-1}(2d_G^{rk})^q. \tag{S.106}$$

We then consider the following decomposition of $w \in \mathcal{G}^v_{r,m}$ (see Fig. 4):

$$w = w_0 \sqcup w_1 \sqcup w_2 \sqcup \cdots \sqcup w_l, \quad 0 \le l \le m - 1, \tag{S.107}$$

where $w_j \subset w_L$ satisfy $d(w_j, v) = j$ for $j = 0, 1, 2, \dots, l$. Here, we define $d(w_j, w_0)$ as the shortest path length in the cluster $w_0 \sqcup w_1 \sqcup \cdots \sqcup w_{j-1}$ which connects from w_j to v. We also define $q_j := |w_j|$ with $q_j \ge 1$. We notice that all the clusters $w \in \mathcal{G}^v_{r,m}$ can be decomposed into the from of (S.107).

For fixed $\{q_0, q_1, \dots, q_l\}$, the number of clusters $\{w_1, w_2, \dots, w_l\}$ defined as in Eq. (S.107) is bounded by

$$\begin{aligned}
&\#\left\{w \in \mathcal{C}_{r,q_0} | X_{0,s} \cap v \neq \emptyset,\ s = 1, 2, \dots, q_0\right\} \prod_{j=1}^l \max_{w_{j-1} \in \mathcal{C}_{r,q_{j-1}}} \left(\#\left\{w \in \mathcal{C}_{r,q_j} | X_{s_j} \cap V_{w_{j-1}} \neq \emptyset,\ s_j = 1, 2, \dots, q_j\right\}\right) \\
\le& (2d_G^{rk})^{q_0} \prod_{j=1}^l \left[2^{kq_{j-1}-1}(2d_G^{rk})^{q_j}\right] \le 2^{-l}\left(2^{k+1} d_G^{rk}\right)^m,
\end{aligned} \tag{S.108}$$

where we denote $w_j = \{X_{s_j}\}_{s_j=1}^{q_j}$; note that $\sum_{j=0}^{l} q_j = m$. Then, by taking the summation with respect to $\{q_0, q_1, \dots, q_l\}$ and l, we finally obtain the upper bound of $\#\{w | w \in \mathcal{G}_{r,m}^v\}$ as follows:

$$\begin{aligned} \#\{w | w \in \mathcal{G}_{r,m}^v\} &\le \sum_{l=0}^{m-1} \sum_{\substack{q_0+q_1+\cdots+q_l=m \\ q_0 \ge 1, q_1 \ge 1, \dots, q_l \ge 1}} 2^{-l} \left(2^{k+1} d_G^{rk}\right)^m \\ &= \sum_{l=0}^{m-1} \left(\!\!\binom{l+1}{m-l-1}\!\!\right) 2^{-l} \left(2^{k+1} d_G^{rk}\right)^m \\ &= \sum_{l=0}^{m-1} \binom{m-1}{l} 2^{-l} \left(2^{k+1} d_G^{rk}\right)^m \le \left(3 \cdot 2^k d_G^{rk}\right)^m, \end{aligned} \tag{S.109}$$

where the summation with respect to $\{q_0, q_1, \dots, q_l\}$ ($q_0 \ge 1, q_1 \ge 1, \dots, q_l \ge 1$) is equal to the $(m-l-1)$-multicombination from a set of $l+1$ elements:

$$\sum_{\substack{q_0+q_1+\cdots+q_l=m \\ q_0 \ge 1, q_1 \ge 1, \dots, q_l \ge 1}} = \left(\!\!\binom{l+1}{m-l-1}\!\!\right) = \binom{m-1}{l}. \tag{S.110}$$

By applying the above upper bound to the inequality (S.102), we obtain the main inequality (S.97). This completes the proof. □

———————————————————— **[End of Proof of Proposition 10]**

III. PROOF OF THEOREM 5

We here show the proof of Theorem 5 which upper bounds the conditional mutual information in long-range interacting systems. We rewrite the Hamiltonian with the power-law decay interaction by using the notations (S.2) and (S.5):

$$H = \sum_{X \in E_\infty} h_X = \sum_{l=1}^{\infty} \sum_{X \in E^{(l)}} h_X. \tag{S.111}$$

We here define $\tilde{g}_l$ as

$$\tilde{g}_l := \max_{v \in V} \sum_{X \in E^{(l)} | X \ni v} \|h_X\|. \tag{S.112}$$

Then, the assumption

$$f(R) = R^{-\alpha} \quad (\alpha > 0) \tag{S.113}$$

in the main manuscript implies

$$\sum_{l \ge R}^{\infty} \sum_{X \in E^{(l)} | X \ni v} \|h_X\| \le \sum_{l \ge R} \tilde{g}_l \le R^{-\alpha}. \tag{S.114}$$

We again show the statement that we would like to prove:

Theorem 11. *Let A, B and C be arbitrary subsystems in V $(A, B, C \subset V)$. Then, under the assumption that the inverse temperature satisfies*

$$\beta < \beta_c/11 = \frac{1}{88 e^3 k}, \tag{S.115}$$

the Gibbs state ρ satisfies the approximate Markov property as follows:

$$\mathcal{I}_\rho(A:C|B) \le \beta \min(|A|, |C|) \frac{11 e^{1/k}/\beta_c}{1 - 11\beta/\beta_c} d_{A,C}^{-\alpha}, \tag{S.116}$$

where we assume that $d_{A,C} \ge 2\alpha$.

A. Details of the proof

We start from Eq. (S.24). By parametrizing the Hamiltonian as

$$H_{\vec{a}} = \sum_{X \in E_\infty} a_X h_X = \sum_{l=1}^{\infty} \sum_{X \in E^{(l)}} a_X h_X, \tag{S.117}$$

we have

$$\begin{aligned}
\tilde{H}_{\vec{1}}(A:C|B) &= \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{X_1, X_2, \ldots, X_m \in E_\infty} \prod_{j=1}^{m} \frac{\partial}{\partial a_{X_j}} \log(\tilde{\rho}^L_{\vec{a}}) \Big|_{\vec{a}=\vec{0}} \\
&= \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{l_1, l_2, \ldots, l_m = 1}^{\infty} \sum_{X_1 \in E^{(l_1)}, X_2 \in E^{(l_2)}, \ldots, X_m \in E^{(l_m)}} \prod_{j=1}^{m} \frac{\partial}{\partial a_{X_j}} \log(\tilde{\rho}^L_{\vec{a}}) \Big|_{\vec{a}=\vec{0}} \\
&= \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{l_0=m}^{\infty} \sum_{w \in \mathcal{C}_m(l_0)} n_w \mathcal{D}_w \tilde{H}_{\vec{a}}(A:C|B) \Big|_{\vec{a}=\vec{0}},
\end{aligned} \tag{S.118}$$

where we define $\mathcal{C}_m(l_0) \subset \mathcal{C}_{\infty,m}$ as

$$\mathcal{C}_m(l_0) = \left\{ w = \{X_1, X_2, \ldots, X_m\} \in \mathcal{C}_{\infty,m} \middle| X_j \in E^{(l_j)},\ j = 1, 2, \ldots, m \quad \text{s.t.} \quad \sum_{j=1}^{m} l_j = l_0 \right\}. \tag{S.119}$$

See Eq. (S.2) and Sec. I A 1 for the definitions of $\mathcal{C}_{\infty,m}$ and $E^{(l)}$.

Next, from Eq. (S.118), we can derive a similar statement to the proposition 3:

$$\begin{aligned}
\tilde{H}_{\vec{1}}(A:C|B) &= \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{l_0=m}^{\infty} \sum_{w \in \mathcal{C}_m(l_0)} n_w \mathcal{D}_w \tilde{H}_{\vec{a}}(A:C|B) \Big|_{\vec{a}=\vec{0}} \\
&= \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{l_0 \ge d_{A,C}} \sum_{w \in \mathcal{G}_m^{A,C}(l_0)} n_w \mathcal{D}_w \tilde{H}_{\vec{a}}(A:C|B) \Big|_{\vec{a}=\vec{0}},
\end{aligned} \tag{S.120}$$

where we define $\mathcal{G}_m^{A,C}(l_0) \subset \mathcal{G}_{\infty,m}^{A,C}$ as

$$\mathcal{G}_m^{A,C}(l_0) = \left\{ w = \{X_1, X_2, \ldots, X_m\} \in \mathcal{G}_{\infty,m}^{A,C} \middle| X_j \in E^{(l_j)},\ j = 1, 2, \ldots, m \quad \text{s.t.} \quad \sum_{j=1}^{m} l_j = l_0 \right\}. \tag{S.121}$$

Notice that we have $w \notin \mathcal{G}_m^{A,C}(l_0)$ if $l_0 < d_{A,C}$ from the above definition.

By following the same discussions in the derivation of Ineq. (S.67), we obtain

$$\|\tilde{H}_{\vec{1}}(A:C|B)\| \le \sum_{v \in A} \sum_{m=1}^{\infty} \sum_{l_0 \ge d_{A,C}} 2(4\beta)^m \sum_{w \in \mathcal{G}^v_{r,m}(l_0)} \frac{n_w}{m!} \prod_{s=1}^{m} N_{X_s|w} \|h_{X_s}\|, \tag{S.122}$$

where in this case, the summation of $v \in \partial A_r$ is replaced by $v \in A$ due to $\partial A_\infty = A$ (see Eq. (S.66)). Then, by using the inequality (S.70), obtain

$$\sum_{w \in \mathcal{G}^v_{r,m}(l_0)} \frac{n_w}{m!} \prod_{s=1}^{m} N_{X_s|w} \|h_{X_s}\| \le \frac{e^{1/k}(2e^3 k)^m}{2} \sum_{l_1 + l_2 + \ldots + l_m = l_0} \prod_{j=1}^{m} \tilde{g}_{l_j}, \tag{S.123}$$

where we defined $\tilde{g}_l$ in Eq. (S.112). By combining the inequalities (S.122) and (S.123), we obtain

$$\|\tilde{H}_{\vec{1}}(A:C|B)\| \le \sum_{v \in A} \sum_{m=1}^{\infty} \sum_{l_1 + l_2 + \cdots + l_m \ge d_{A,C}} e^{1/k} (8e^3 k\beta)^m \prod_{j=1}^{m} \tilde{g}_{l_j}. \tag{S.124}$$

We can prove the following inequality (see Sec. III A 1 for the proof):

$$\sum_{l_1 + l_2 + \cdots + l_m \ge l_0} \prod_{j=1}^{m} \tilde{g}_{l_j} \le 11^m l_0^{-\alpha} \tag{S.125}$$

for arbitrary $l_0 \geq 2\alpha$. By using the above inequality, we obtain

$$\sum_{m=1}^{\infty} \sum_{l_1+l_2+\cdots+l_m \geq d_{A,C}} e^{1/k}(8e^3 k\beta)^m \prod_{j=1}^{m} \tilde{g}_{l_j} \leq d_{A,C}^{-\alpha} \sum_{m=1}^{\infty} e^{1/k}(11\beta/\beta_c)^m \leq \frac{11e^{1/k}\beta/\beta_c}{1-11\beta/\beta_c} d_{A,C}^{-\alpha}. \tag{S.126}$$

By combining the inequalities (S.124) and (S.126), we finally obtain

$$\|\tilde{H}_{\vec{1}}(A:C|B)\| \leq \beta|A| \frac{11e^{1/k}/\beta_c}{1-11\beta/\beta_c} d_{A,C}^{-\alpha}. \tag{S.127}$$

In the same way, we can derive the inequality such that $|A|$ is replaced by $|C|$ in (S.127). By combining the above inequality with (S.21), we prove Theorem 5. □

1. *Proof of the inequality* (S.125)

For the proof, we start from the following form:

$$\sum_{l_1+l_2+\cdots+l_m \geq l_0} \prod_{j=1}^{m} \tilde{g}_{l_j} \leq \eta_m l_0^{-\alpha}. \tag{S.128}$$

We, in the following, construct a recurrence relation to determine η_m. First, Eq. (S.114) immediately implies

$$\sum_{l_1+l_2+\cdots+l_m \geq l_0} \prod_{j=1}^{m} \tilde{g}_{l_j} \leq \prod_{j=1}^{m} \sum_{l_j=1}^{\infty} \tilde{g}_{l_j} \leq 1. \tag{S.129}$$

Based on the inequalities (S.128) and (S.129), we consider the case of $m+1$ as

$$\begin{aligned}
\sum_{l_1+l_2+\cdots+l_{m+1} \geq l_0} \prod_{j=1}^{m+1} \tilde{g}_{l_j} &\leq \sum_{l_{m+1}=1}^{\infty} \tilde{g}_{l_{m+1}} \sum_{l_1+l_2+\cdots+l_m \geq l_0 - l_{m+1}} \prod_{j=1}^{m} \tilde{g}_{l_j} \\
&\leq \eta_m \sum_{l_{m+1}=1}^{\infty} \tilde{g}_{l_{m+1}} \max\left[(l_0 - l_{m+1})^{-\alpha}, 1\right] \\
&\leq \eta_m \sum_{l=1}^{l_0-1} \tilde{g}_l (l_0 - l)^{-\alpha} + \eta_m \sum_{l \geq l_0} \tilde{g}_l \leq \eta_m \sum_{l=1}^{l_0-1} \tilde{g}_l (l_0 - l)^{-\alpha} + \eta_m l_0^{-\alpha},
\end{aligned} \tag{S.130}$$

where the last inequality comes from the inequality (S.114) with $R = l_0$. In order to upper-bound the first term, we decompose the summation as follows:

$$\sum_{l=1}^{l_0-1} \tilde{g}_l (l_0 - l)^{-\alpha} = \left(\sum_{l \in [1,l_1)} + \sum_{l \in [l_1,l_2)} + \sum_{l \in [l_2,l_3)} + \sum_{l \in [l_3,l_0)} \right) \tilde{g}_l (l_0 - l)^{-\alpha}, \tag{S.131}$$

for $\alpha > 2$, where $l_1 = \lceil l_0/\alpha \rceil$, $l_2 = \lceil l_0/2 \rceil$, $l_3 = \lceil l_0 - l_0/\alpha \rceil$. For $\alpha \leq 2$, we decompose as

$$\sum_{l=1}^{l_0-1} \tilde{g}_l (l_0 - l)^{-\alpha} = \left(\sum_{l \in [1,l_2)} + \sum_{l \in [l_2,l_0)} \right) \tilde{g}_l (l_0 - l)^{-\alpha}. \tag{S.132}$$

Next, for arbitrary choice of $[x, y)$ $(1 \leq x \leq y \leq l_0 - 1)$, we have

$$\sum_{l \in [x,y)} \tilde{g}_l (l_0 - l)^{-\alpha} \leq (l_0 - y + 1)^{-\alpha} \sum_{l \in [x,y)} \tilde{g}_l \leq (l_0 - y + 1)^{-\alpha} \sum_{l \geq x} \tilde{g}_l \leq (l_0 - y + 1)^{-\alpha} x^{-\alpha}, \tag{S.133}$$

which reduces the inequality (S.131) to

$$\begin{aligned}
\sum_{l=1}^{l_0-1} \tilde{g}_l (l_0 - l)^{-\alpha} \leq& (l_0 - \lceil l_0/\alpha \rceil + 1)^{-\alpha} + (l_0 - \lceil l_0/2 \rceil + 1)^{-\alpha} \lceil l_0/\alpha \rceil^{-\alpha} \\
&+ (l_0 - \lceil l_0 - l_0/\alpha \rceil + 1)^{-\alpha} \lceil l_0/2 \rceil^{-\alpha} + \lceil l_0 - l_0/\alpha \rceil^{-\alpha} \\
\leq& 2(l_0 - l_0/\alpha)^{-\alpha} + 2(l_0/2)^{-\alpha}(l_0/\alpha)^{-\alpha} \\
=& 2l_0^{-\alpha} \left[\frac{1}{(1-1/\alpha)^{\alpha}} + \left(\frac{2\alpha}{l_0} \right)^{\alpha} \right] \leq 10 l_0^{-\alpha}
\end{aligned} \tag{S.134}$$

for $\alpha > 2$, where we use $1/(1-1/x)^x \le 4$ for $x \ge 2$ and $l_0 \ge 2\alpha$ from the condition of the theorem. For $\alpha \le 2$, we also obtain

$$\sum_{l=1}^{l_0-1} \tilde{g}_l (l_0 - l)^{-\alpha} \le 2(l_0/2)^{-\alpha} \le 8 l_0^{\alpha} \tag{S.135}$$

from the decomposition (S.132), where we use $2^\alpha \le 4$ for $\alpha \le 2$.

By applying the inequalities (S.134) and (S.135) to the inequality (S.130), we obtain

$$\sum_{l_1+l_2+\cdots+l_{m+1} \ge l_0} \prod_{j=1}^{m+1} \tilde{g}_{l_j} \le 11 \eta_m l_0^{-\alpha}, \tag{S.136}$$

which gives rise to

$$\eta_{m+1} \le 11 \eta_m. \tag{S.137}$$

This yields the inequality (S.125). This completes the proof. □

---

[1] Tomotaka Kuwahara and Keiji Saito, "*Gaussian concentration bound and Ensemble equivalence in generic quantum many-body systems including long-range interaction*," arXiv preprint arXiv:1906.10872 (2019), arXiv:1906.10872.