Gaussian concentration bound and Ensemble equivalence in generic quantum many-body systems including long-range interaction

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, Kohoku-ku, Yokohama 223-8522, Japan

interdisciplinary Theoretical & Mathematical Sciences Program (iTHEMS) RIKEN 2-1, Hirosawa, Wako, Saitama 351-0198, Japan

Keiji Saito
Department of Physics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Japan 223-8522

Abstract

This work explores fundamental statistical and thermodynamic properties in both of the short-range and long-range interacting systems. The purpose of this study is two folds. Firstly, we rigorously prove a Gaussian concentration bound (or Chernoff-Hoeffding inequality) of probability distribution for arbitrary few-body observables above a threshold temperature. This bound is derived for arbitrary Gibbs states of systems with long-range interactions. Second, we establish a quantitative relationship between the concentration bound of the Gibbs state and the equivalence of the canonical and the micro-canonical ensembles. For thermodynamic quantities, we evaluate difference of the averages between the canonical and the micro-canonical ensembles. Under the assumption of the Gaussian concentration bound on the canonical ensemble, the difference is upperbounded by $\left[ n^{-1} \log(n^{3/2} \Delta^{-1}) \right]^{1/2}$ with $n$ the system size and $\Delta$ the width of the energy shell of the micro-canonical ensembles. Our estimation gives a non-trivial upper bound even for an exponentially small energy width with respect to the system size. By combining these two results, we prove the ensemble equivalence as well as the weak eigenstate thermalization for arbitrary long-range interacting systems above a threshold temperature.

Keywords:
Long-range interacting systems, Concentration bound, Chernoff-Hoeffding inequality, Ensemble equivalence, Eigenstate thermalization hypothesis, Weak eigenstate thermalization

PACS: 05.30.Ch, 05.30.-d, 65.40.Gr, 02.50.-r,

Highlights:
* Foundations of thermodynamics and statistical mechanics in generic long-range interacting systems
* Proof of the Gaussian concentration bound for probability distribution of observables
* The ensemble equivalence between canonical and micro-canonical ensembles is proved in long-range interacting systems
* The weak eigenstate thermalization is proved in long-range interacting systems
* Width of the energy shell can be taken exponentially small with respect to the system size

1. Introduction

In recent years, systems with long-range interactions are ubiquitous in various experimental setups such as atomic, molecular, and optical systems [1, 2, 3, 4, 5, 6, 7]. Such systems often exhibit new physics which cannot appear in short-range interacting systems [8, 9, 10, 11, 12]. Both in experimental and theoretical aspects, long-range interacting systems play crucial roles in modern physics and attract more and more attentions. At the same time, we need non-trivial modifications to most of the existing analyses in short-range interacting systems. For example,
Figure 1: Schematic picture of the canonical and the micro-canonical ensembles. The canonical ensemble is characterized by the distribution $e^{-\beta E}/Z$ with $Z$ the partition function (green curve). The micro-canonical ensemble is defined by the uniform distribution in an energy shell (orange region). We show the precise definitions in Eqs. (24). The ensemble equivalence discusses whether these two ensembles have similar expectation values for thermodynamic quantities such as the magnetization. Our purpose is to quantitatively evaluate dependence of the ensemble equivalence on the system size $n$ and the width of the energy shell $\Delta$. We show in Theorem 2 that the ensemble equivalence for Gibbs states is deeply related to the concentration bound as (2). By proving the concentration bound with $\gamma = 2$ in generic long-range interacting systems above a threshold temperature (Corollary 1), we rigorously prove the ensemble equivalence in long-range interacting systems.

In long-range interacting systems, the ensemble equivalence can be violated [32, 33]. Then, we aim to identify the condition where the ensemble equivalence is reliably ensured. In analyzing the ensemble equivalence, we need to discuss properties of the canonical state (i.e., the Gibbs state or the thermal equilibrium state) at finite temperatures:

$$\rho = \frac{e^{-\beta H}}{Z}$$

with $Z := \text{tr}(e^{-\beta H})$, where $H$ and $\beta$ are the system Hamiltonian and the inverse temperature, respectively. At temperatures above a critical threshold (or in high-temperature phases), the clustering property has been proved both in classical [34] and quantum [35, 36, 37, 38, 39] spin systems with short-range interactions. However, the
long-range interacting systems do not usually have a finite correlation length at arbitrary temperatures, and hence we need to rely on an alternative one to the clustering property.

In the present paper, we prove the concentration bound and utilize it to discuss the ensemble equivalence. If the spins are independent with each other, the following Chernoff-Hoeffding concentration inequality [40, 41] is known to hold. Roughly speaking, it says that the probability distribution for a macroscopic observable is sharply concentrated around the average value. Let us consider a product state $\rho_0 = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$ on an $n$-spin system. Then, the Chernoff-Hoeffding inequality upperbounds the probability distribution of a one-body observable $A = \sum_i a_i$ with $\|a_i\| = 1$ in the Gaussian form:

$$\int_{x_0+(A)}^{\infty} \text{tr}[\rho_0 \delta(A-x)]dx \leq \exp\left[-C \left(\frac{x_0}{\sqrt{n}}\right)^\gamma\right] \quad (x_0 > 0),$$

with $\gamma = 2$, where $\delta(x)$ is the delta function, $\langle A \rangle := \text{tr}(\rho_0 A)$ and $C$ is a constant independent of the system size $n$. Our question is whether Ineq. (2) holds beyond the setup of product states and one-body observables. In weakly correlated spin systems, the generalization of the inequality (2) has been obtained in several ways. First, for product states or short-range entangled states (see [42] for the definition), the inequality (2) with $\gamma = 2$ has been proved for probability distributions of generic few-body observables [43, 44]. If we consider more general class of states, the concentration inequality has been derived in weaker ways (i.e., $\gamma < 2$): $\gamma = 1$ for gapped ground states [45, 46] and $\gamma = 1/(D + 1)$ ($D$: the system dimension) for states with the clustering [44]. In these works, the locality of interaction in the Hamiltonian plays central roles [47]. Moreover, if we restrict ourselves to classical spin systems with short-range interactions, the concentration inequalities have been extensively investigated [48, 49, 50, 51] both at high temperatures ($\gamma = 2$) and low temperatures ($\gamma < 2$).

In this paper, through the generalized cluster expansion, we derive the Gaussian concentration bound, for a generic many-body systems including long-range systems. Below, we list our findings in this paper:

1. The Gaussian concentration inequality ($\gamma = 2$) is rigorously proved for a generic systems including long range interaction systems above a threshold temperature (see Corollary 1).
2. Under the assumption of the concentration bound, we quantitatively prove the ensemble equivalence between the canonical and the micro-canonical ensembles (see Theorem 2).
3. By applying Theorem 2 to the high-temperature Gibbs states, the difference between the canonical and the micro-canonical ensembles is bounded from above by $\left[n^{-1} \log(n^{3/2} \Delta^{-1})\right]^{1/2}$. The ensemble equivalence holds for sufficiently large systems (or $n \gg 1$) as long as $\Delta = \exp(-n^{1-\epsilon})$ with $\epsilon > 0$.

The above three results solve the problems i) to iii) accurately in the viewpoint of the system-size dependence.. For the problem i), we can ensure the ensemble equivalence in long-range interacting systems above a threshold temperature (see Eq. (9)). For the problem ii), the quantitative difference of the averages between the canonical and the micro-canonical ensembles is bounded by $O(n^{-1/2})$ up to a logarithmic correction for $\Delta = 1/\text{poly}(n)$. Finally, for the problem iii), the ensemble equivalence approximately holds even for energy width of $\Delta = e^{-O(n)}$ (see Corollary 3). Because the density of states in energy spectrum is at most of $e^{O(n)}$, the energy gap smaller than $e^{-O(n)}$ implies that the individual eigenstates become visible and affect the ensemble equivalence. We note that the realization of ensemble equivalence for the infinitesimal limit of energy width leads to the ETH. However, the ETH cannot be proven without imposing specific properties such as the non-integrability of the system [30, 31]. Hence, it is plausible that one cannot reduce the energy gap smaller than $e^{-O(n)}$ in the present general framework. We thus conclude that our estimation for the limitation to the energy width is qualitatively tight.

This paper is organized in the following way. In sec.2, we explain the setup and main findings on the concentration bound using the cluster expansion. In sec.3, we apply our findings to the ensemble equivalence and weak version of the ETH. In sec. 4, we discuss the future perspective. In sec.5, we outline the mathematical structure to derive the results.

2. Setup and Main results

We consider a quantum spin system with $n$ spins, where each of the spins has $d$-dimensional Hilbert space. We let $V = \{1, 2, 3, \ldots, n\}$ be the total spin set, and denote the local Hilbert space by $\mathcal{H}^v$ ($v \in V$) with $\dim(\mathcal{H}^v) = d$. 

3
Now, the total Hilbert space is given by $\mathcal{H} := \bigotimes_{v \in V} \mathcal{H}^v$ with $\mathcal{D}_\mathcal{H} := \dim(\mathcal{H}) = d^n$. We define the space of linear operators on $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$. In order to characterize the interactions of spins, we write the system Hamiltonian $H \in \mathcal{B}(\mathcal{H})$ as

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

where each of $\{h_X\}_{|X| \leq k}$ denotes interaction between the spins in $X \subset V$. The Hamiltonian (3) describes generic $k$-body interacting systems. We define $\mathcal{E}$ as the set of all eigenstates and describe each of the energy eigenstates by $|E\rangle \in \mathcal{E}$ such that $H|E\rangle = E|E\rangle$.

We consider the Hamiltonian where the spectrum of the local Hamiltonian is finite. More precisely, we impose the condition

$$\sum_{X: X \ni v} \|h_X\| \leq g \text{ for } \forall v \in V,$$

where $\|\cdot\|$ is the operator norm and $\sum_{X: X \ni v}$ sums up all the interactions which contain the spin $v$. We can immediately obtain the following inequality for the total norm of the Hamiltonian:

$$\|H\| \leq \sum_{v \in V} \sum_{X: X \ni v} \|h_X\| \leq \sum_{v \in V} g = g|V| = gn$$

Thus, the inequality (4) upperbounds the one-spin energy by $g$.

The above class of the Hamiltonians includes the long-range interacting spin systems with a power-law decay on a lattice as well as short-range interacting case. For example, let us consider the following Hamiltonian on a $D$-dimensional lattice system which has interactions with a power-law decay of $1/r^\alpha$ ($r$: interaction length):

$$H = \frac{1}{N} \sum_{i,j} \frac{J}{r_{i,j}^{\alpha}} (\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z),$$

with $J = \mathcal{O}(1)$, where $r_{i,j}$ is the Manhattan distance between the spins $i$ and $j$ defined by the lattice geometry and $N$ is determined so that the finite norm (4) for the local Hamiltonian is satisfied. If the exponent $\alpha$ is larger than $D$, we have $N = \mathcal{O}(1)$. On the other hand, for $\alpha \leq D$, we need to take $N = \mathcal{O}(n^{D-\alpha})$ due to the condition (4). In this example, $k = 2$. This type of the interaction contains the Blume-Emery-Griffiths (BEG) model with infinite-range interactions (i.e., $\alpha = 0$) where the ensemble inequivalence has been investigated in Ref. [32]. Moreover, it is noteworthy that the Hamiltonian (3) can also treat the quantum systems on infinite-dimensional networks, where the breaking of the ensemble equivalence has been reported [52].

Throughout the paper, we consider the Gibbs state for the Hamiltonian $H$ with inverse temperature $\beta$ as follows:

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

Here, we aim to prove the following theorem below a certain threshold $\beta < \beta_c$, where $\beta_c$ does not depend on the system size $n$, but only on $k$ and $g$.

**Theorem 1.** Let $\Omega \in \mathcal{B}(\mathcal{H})$ be an arbitrary operator with the same condition as (4), namely

$$\Omega := \sum_{|X| \leq k} \omega_X \text{ with } \sum_{X: X \ni v} \|\omega_X\| \leq g \text{ for } v \in V.$$  

Then, under the assumption that the inverse temperature satisfies

$$\beta < \beta_c := \frac{1}{4e^3 g k},$$

the Gibbs state $\rho$ satisfies the following inequality:

$$\log \left[ \text{tr} (e^{-\tau \Omega}) \right] \leq -\tau \langle \Omega \rangle_\beta + \frac{\tau^2 \bar{\Omega}}{\beta_c - \beta - \tau},$$

where $\bar{\Omega}$ is the mean value of $\Omega$.
where

\[ \langle \Omega \rangle_\beta := \text{tr}(\Omega \rho) \quad (11) \]

and we assume \( \tau < \beta_c - \beta \) and \( \tilde{\Omega} \) are defined as \( \tilde{\Omega} := \sum_{|\omega| \leq k} \|\omega\| \). We notice that the same inequality holds for \( \text{tr}(e^{\tau \tilde{\Omega}} \rho) \).

For the sake of clear presentation, we provide the proof in the section 5 and here we concentrate on several physical applications of the theorem.

This theorem implies the following Chernoff-Hoeffding inequality:

**Corollary 1.** We assume the condition of Theorem 1 and let \( P_\rho(x) \) be

\[ P_\rho(x) := \text{tr}(\rho \delta(x - \Omega)) \quad (12) \]

with \( \delta(x) \) the delta function. We then obtain

\[ P_\rho(|x - \langle \Omega \rangle_\beta| \geq x_0) := \int_{|x - \langle \Omega \rangle_\beta| \geq x_0} P_\rho(x) dx \leq 2 \exp \left( -\frac{x_0^2}{c_\beta \Omega} \right) \quad (13) \]

where we define

\[ c_\beta := \frac{2}{\beta_c - \beta}. \quad (14) \]

We here compare the above concentration inequality with the previous works. Around the average value \( x_0 = O(n^{1/2}) \), the well-known the central limit theorem has been derived for several classes of quantum systems with translational invariance \([53, 54, 55, 56]\). It states that the distribution of a macroscopic observable is not only bounded by Gaussian function but also converges to a Gaussian normal distribution in the thermodynamic limit \((n \to \infty)\). As a refined statement, the Berry-Essen theorem \([57, 58]\) has been proved for arbitrary quantum states with the clustering property \([28]\). Both of the above theorems impose a stronger limitation than Ineq. (13) in that they prove the exact convergence to the Gaussian distribution in the limit of \( n \to \infty \). On the other hand, for finite \( n \), they cannot give a tight asymptotic behavior of the tail of probability distribution; indeed, the best convergence behavior is at most \( O(1/\sqrt{n}) \) as in Ref. \([28]\).

On the asymptotic behavior of finite systems for \( x_0 = O(n) \), there are various studies on the large deviation \([59, 60, 61, 62, 63]\). The large deviation theorem asserts that the probability becomes exponentially small as increasing the system size \( n \) as follows:

\[ P_\rho(|x - \langle \Omega \rangle_\beta| \geq x_0) = \exp \left[ -n I(x_0/n) + O(n^{1-\kappa}) \right] \quad (15) \]

with \( \kappa > 0 \), where the rate function \( I(\cdot) \) is a non-zero and smooth function. The large deviation theorem is stronger than the Chernoff-Hoeffding inequality (13) since it gives the correct asymptotic exponential decay of the probability for \( x_0 = O(n) \). However, the large deviation theory usually focuses on the large deviation function \( I(x_0/n) \) and does not discuss the decay rate around the average value due to the sub-leading term of \( O(n^{1-\kappa}) \) that is written in Eq. (15). This aspect can be crucial in discussing the finite-size effect on the ensemble equivalence between the canonical and the micro-canonical distributions (see Sec. 3).

**Proof of Corollary 1.** Without loss of generality, we here set \( \langle \Omega \rangle_\beta = \text{tr}(\Omega \rho) = 0 \) and \( \tau \leq (\beta_c - \beta)/2 \), and the inequality (10) reads

\[ \text{tr}(e^{-\tau\tilde{\Omega}} \rho) \leq e^{c_\beta \tau^2 \tilde{\Omega}}. \quad (16) \]

By using the above inequality, we obtain for \( \tau > 0 \) and \( x \geq 0 \)

\[ P_\rho(x \geq x_0) = \int_{x \geq x_0} \text{tr}[\rho \delta(x - \Omega)] dx = \int_{x \geq x_0} \text{tr}[\rho e^{\tau\tilde{\Omega}} e^{-\tau\tilde{\Omega}} \delta(x - \Omega)] dx \]

\[ \leq e^{-\tau x_0} \cdot \text{tr}(e^{\tau\tilde{\Omega}} \rho) \]

\[ \leq e^{-\tau x_0 + c_\beta \tau^2 \tilde{\Omega}} = \exp \left( c_\beta \tilde{\Omega}[\tau - \langle x_0/(c_\beta \tilde{\Omega}) \rangle]^2 - x_0^2/(c_\beta \tilde{\Omega}) \right). \quad (17) \]
By using $\bar{\Omega} \geq \|\Omega\| \geq x_0$, we have

$$\frac{x_0}{c_\beta \bar{\Omega}} \leq \frac{1}{c_\beta} = \frac{\beta_c - \beta}{2}.$$  \hspace{1cm} (18)

Thus, $\tau$ can be chosen as $\tau = x_0/(c_\beta \bar{\Omega}) \leq (\beta_c - \beta)/2$ in (17) and we obtain the inequality (13) for $x_0 \geq 0$. In the same way, we can prove the case of $x_0 \leq 0$. This completes the proof of Corollary 1. \hfill \Box

The Chernoff-Hoeffding inequality (13) also tells us information on the density of states:

**Corollary 2.** For an arbitrary few-body Hamiltonian as in Eq. (3) with (4), the total number of energy eigenstates in $E \in [-x_0, x_0]$ is bounded from below by

$$\# \{|E\} \in \mathcal{E} | E \in [-x_0, x_0]\} \geq \frac{1}{d^n} \geq 1 - 2 \exp \left(-\frac{x_0^2}{c_0 g n}\right),$$  \hspace{1cm} (19)

where $c_0 = 2/\beta_c$ and we set $\text{tr}(H) = 0$.

**Remark.** This corollary does not characterize the Gibbs states but the Hamiltonian itself. It rigorously proves that the density of the energy eigenstates follows the Gaussian concentration bound around the infinite-temperature value. Thus, the spectral distributions of *all the few-body Hamiltonians* are similar to those of one-body Hamiltonians. This result is a generalization of the result in Ref. [64] (see Theorem 2 in the reference) which proves the Gaussian concentration (19) for translation-invariant spin chains.

**Proof of Corollary 2.** By choosing the infinity temperature states (i.e., $\beta = 0$) in Corollary 1, we obtain for $\rho = \hat{1}/d^n$ and $\Omega = H$

$$\frac{1}{d^n} \int_{|x - \text{tr}(H)/d^n| \geq x_0} \text{tr}[^\delta(H - x)] dx \leq 2 \exp \left(-\frac{x_0^2}{c_0 \bar{H}}\right).$$  \hspace{1cm} (20)

Note that $\langle H \rangle_{\beta} = \text{tr}(H)/d^n$ for $\beta = 0$. By using the condition $\text{tr}(H) = 0$, we have

$$\# \{|E\} \in \mathcal{E} | E \in [-x_0, x_0]\} = \int_{|x| \leq x_0} \text{tr}[\delta(H - x)] dx = 1 - \int_{|x| > x_0} \text{tr}[\delta(H - x)] dx$$  \hspace{1cm} (21)

and the condition (4) gives

$$\bar{H} := \sum_{|X| \leq k} \|h_X\| \leq gn.$$  \hspace{1cm} (22)

Then, by applying Eq. (21) and Ineq. (22) to (20), we obtain the inequality (19) under the condition $\text{tr}(H) = 0$. This completes the proof. \hfill \Box

### 3. Concentration bound and Ensemble Equivalence between the canonical and the micro-canonical distributions

We here consider the ensemble equivalence between the canonical and the micro-canonical distributions. By following the setup of Refs. [28, 29], we first define the canonical and the micro-canonical average for an arbitrary operator $O \in \mathcal{B}(\mathcal{H})$

$$\langle O \rangle_{\beta} := \frac{1}{Z} \text{tr}(O e^{-\beta H}),$$  \hspace{1cm} (23)

$$\langle O \rangle_{U,\Delta} := \frac{1}{\mathcal{N}_{U,\Delta}} \sum_{E \in (U - \Delta, U]} \langle E | O | E \rangle,$$  \hspace{1cm} (24)

where (23) and (24) are averages of the observable $O$ over the canonical ensemble and microcanonical ensemble, respectively. The quantity $\mathcal{N}_{U,\Delta}$ is the total number of energy eigenstates in $E \in (U - \Delta, U]$, namely,

$$\mathcal{N}_{U,\Delta} := \text{tr} \left( \sum_{E \in (U - \Delta, U]} |E\rangle \langle E| \right).$$  \hspace{1cm} (25)
To characterize the micro-canonical ensemble, we choose \( U \) as
\[
U = \delta \nu^*, \quad \nu^* := \arg\max_{\nu \in \mathbb{Z}} \left( e^{-\beta \nu \Delta} N_{\nu, \delta, \delta} \right), \quad \delta := \min(\Delta, 1/\beta)
\] (26)

Note that if \( \Delta \leq 1/\beta \) the energy width of the energy shell \( \Delta \) is equal to \( \delta \).

We are here interested in the difference between the canonical average \( \langle \Omega \rangle_{\beta} \) and the micro-canonical average \( \langle \Omega \rangle_{U, \Delta} \). For the purpose, we aim to prove that almost all the eigenstates in the energy shell \( (U - \Delta, U) \) have the same expectation value as \( \langle \Omega \rangle_{\beta} \). We consider a probability distribution \( P_{U, \Delta}(x) \) such that
\[
P_{U, \Delta}(x) = \frac{1}{N_{U, \Delta}} \sum_{E \in (U - \Delta, U]} \delta(x - \langle E | \Omega | E \rangle).
\]
(27)

We now aim to derive an upper bound for the cumulative probability distribution as
\[
P_{U, \Delta}(|x| \geq x_0) := \int_{x_0}^{\infty} P_{U, \Delta}(x) dx + \int_{-\infty}^{x_0} P_{U, \Delta}(x) dx.
\]
(28)

Here, based on the concentration bound like (2), we can prove the following theorem:

**Theorem 2.** Let \( \Omega \in \mathcal{B}(\mathcal{H}) \) be a few-body operator as in Eq. (8). Under the assumption that a Gibbs state (7) satisfies the following concentration bound for \( \Omega \) such that
\[
P_p(|x - \langle \Omega \rangle_{\beta}| \geq x_0) = \int_{|x - \langle \Omega \rangle_{\beta}| \geq x_0} P_p(x) dx \leq \exp \left[ - \left( \frac{x_0}{\sqrt{\epsilon g n}} \right)^\gamma \right]
\]
(29)

with \( \gamma \) and \( \epsilon \) a positive constant of \( O(1) \), we have
\[
P_{U, \Delta}(|x - \langle \Omega \rangle_{\beta}| \geq x_0) \leq C_\delta \exp \left[ - \gamma \left( \frac{|x_0 - \langle \Omega \rangle_{\beta}|}{\sqrt{\epsilon g n}} \right)^\gamma \right],
\]
(30)

where
\[
C_\delta := \frac{16 e^2 \sqrt{g n}}{\gamma \sqrt{\epsilon}} \left( 1 + \frac{g n}{\delta} \right) = O(\delta^{-1} n^{3/2}).
\]
(31)

This theorem immediately leads to the following corollary:

**Corollary 3.** Under the assumption in Theorem 2, we have
\[
\frac{1}{n} |\langle \Omega \rangle_{U, \Delta} - \langle \Omega \rangle_{\beta}| \leq C_2 \frac{\log^{1/\gamma} (\delta^{-1} n^{3/2})}{\sqrt{n}},
\]
(32)

with \( C_2 \) a constant which depends only on the parameters \( g, \gamma \) and \( \epsilon \). For arbitrary \( \epsilon > O(n^{-1/2}) \), we have \( \frac{1}{n} |\langle \Omega \rangle_{U, \Delta} - \langle \Omega \rangle_{\beta}| \leq \epsilon \) as long as \( \delta \geq \exp \left[ -C(\epsilon^2 n)^{2/\gamma} \right] \) with \( C = O(1) \).

Before giving the proof, we introduce the following useful lemma which has been proved in Ref. [65].

**Lemma 1.** Let \( p(x) \) be an arbitrary probability distribution whose cumulative distribution is bounded from above:
\[
P(|x - a| \geq x_0) := \int_{|x - a| \geq x_0} p(x) dx \leq \min(1, e^{-x_0^2/\sigma + x_1}), \quad \gamma > 0, \sigma > 0, \ x_0 > 0.
\]
(33)

Subsequently, for an arbitrary \( k \in \mathbb{N} \), we obtain
\[
\int_{-\infty}^{\infty} |x - a|^k p(x) dx \leq (2\sigma x_1)^{k/\gamma} + \frac{k}{\gamma} (2\sigma)^{k/\gamma} \Gamma(k/\gamma)
\]
(34)

with \( \Gamma(\cdot) \) as the gamma function.
Proof of Corollary 3. By applying this lemma to the probability (30) with the parameter set as

$$\{a, \gamma, \sigma, x_1, k\} = \{(\Omega)_{\beta}, \gamma, (\tilde{c} ;eg n)^{\gamma/2}, \log(C_{\delta}), 1\},$$

we obtain the inequality (32). This completes the proof. □

There are several remarks on this theorem:

1. Theorem 2 does not necessarily assume the high-temperature condition.

2. The theorem is concerned with a specific choice of the operator $\Omega$ which satisfies the concentration inequality (29).

3. In the case where the Chernoff-Hoeffding inequality holds (i.e., $\gamma = 2$), the ensemble equivalence approximately holds for exponentially small energy width as $\delta = \exp(-C\epsilon^2 n)$.

4. If we consider a high-temperature Gibbs state with $\beta < \beta_c := 1/(4e^3 g k)$, from Corollary 1, the assumption (29) holds for arbitrary few-body operators with $\tilde{c} = c_\beta$ and $\gamma = 2$. Thus, from Corollaries 1 and 3, we can prove the ensemble equivalence in arbitrary long-range interacting systems above a temperature threshold $\beta_c$.

By applying this corollary to the high-temperature regime, we conclude that even for exponentially small energy width $\Delta$ the ensemble equivalence approximately holds. At first glance, it is contradictory to the counterexample of the ETH such as the many-body localization [31], where any single eigenstates do not have the thermal property. In our theorem, it is true that the energy width can be small as small as $e^{-C\epsilon^2 n}$. However, in this energy shell, there are still exponentially-large number of eigenstates as $\mathcal{N}_{U, \Delta} e^{-C\epsilon^2 n}$, where $\mathcal{N}_{U, \Delta}$ is typically as large as the total dimension of the Hilbert space $d^n$. In order to make the micro-canonical ensemble reduce to a single eigenstate, we have to choose $\Delta$ sufficiently small such that $\mathcal{N}_{U, \Delta} e^{-C\epsilon^2 n} = O(1)$, but such a choice of the energy width no longer gives a non-trivial bound on $\frac{1}{n!} \langle \Omega \rangle_{U, \Delta} - \langle \Omega \rangle_{\beta}$. This apparent contradiction can be resolved in this way.

On the other hand, in low-energy regions, the energy density $\mathcal{N}_{U, \Delta}$ can be much smaller than the total dimension of the Hilbert space $d^n$. If the concentration bound holds at low-temperatures, even a single eigenstate can resemble the canonical ensemble. It indeed occurs under the assumption of the clustering property at sufficiently small temperatures [65].

3.1. Weak eigenstate thermalization

Before going to the proof, we mention the weak eigenstate thermalization [66, 67]. In the eigenstate thermalization hypothesis, all the eigenstates in an energy shell have the same property as the canonical state, while the weak eigenstate thermalization argues that most of the eigenstates in an energy shell have the same property. As discussed in Ref. [67], we consider the the variance of $\langle E | \Omega | E \rangle$ in the energy shell:

$$\frac{1}{\mathcal{N}_{U, \Delta}} \sum_{E \in (U - \Delta, U]} \left( \frac{\langle E | \Omega | E \rangle}{n} - \frac{\langle \Omega \rangle_{U, \Delta}}{n} \right)^2$$

If this variance approaches to 0 in the limit of $n \to 0$, almost all the eigenstates have the same expectation value as the micro-canonical average $\langle \Omega/n \rangle_{U, \Delta}$. Our concern is the finite-size effect of the variance with respect to the system size $n$.

Here, we can prove the following corollary:

**Corollary 4.** Under the assumption in Theorem 2, we have

$$\frac{1}{\mathcal{N}_{U, \Delta}} \sum_{E \in (U - \Delta, U]} \left( \frac{\langle E | \Omega | E \rangle}{n} - \frac{\langle \Omega \rangle_{U, \Delta}}{n} \right)^2 \leq C'_2 \log^{2/\gamma} \left( \frac{\delta^{-1/2} n^{3/2}}{\epsilon} \right),$$

with $C'_2$ a constant which depends only on the parameters $g$, $\gamma$ and $\tilde{c}$.  

Therefore, provided that $\Delta = 1/\text{poly}(n)$ (i.e., $\delta = 1/\text{poly}(n)$ from Eq. (26)), this estimation provides the upper bound of the variance of $\Omega/n$ by $O(\log^{2/\gamma}(n)/n)$. Up to the logarithmic correction, this estimation is qualitatively tight since recent calculations by Alba [68] showed that $(1/2)$-spin isotropic Heisenberg chain expresses the variance of $O(1/n)$.

**Proof of Corollary 4.** We here set $\langle \Omega \rangle_\beta = 0$. By using the definition (30), we first obtain

$$
\frac{1}{N_{U,\Delta}} \sum_{E \in (U-\Delta, U]} \left( \frac{\langle E|\Omega|E \rangle}{n} - \frac{\langle \Omega \rangle_{U,\Delta}}{n} \right)^2 
= \frac{1}{n^2} \int_{-\infty}^\infty x^2 P_{U,\Delta}(x)dx
$$

$$
\leq \frac{1}{n^2} \int_{-\infty}^\infty x^2 P_{U,\Delta}(x)dx
\leq \frac{1}{n^2} \int_{-\infty}^\infty x^2 P_{U,\Delta}(x)dx
$$

Under the assumption, we obtain the inequality (29) for $P_{U,\Delta}(x)$ with $\langle \Omega \rangle_\beta = 0$. Hence, we can utilize Lemma 1 by choosing the parameters

$$
\{a, \gamma, \sigma, x_1, k\} = \{0, \gamma, (\tilde{c} g n)^{\gamma/2}, \log(C\delta), 2\}.
$$

Then, the inequality (34) gives (37). This completes the proof. □

### 3.2. Proof of Theorem 2

Throughout the proof, we set $\langle \Omega \rangle_\beta = 0$. We start from the $m$th moment function ($m$: even) as

$$
M_{U,\Delta}(m) := \int_{-\infty}^\infty x^n P_{U,\Delta}(x)dx.
$$

From the definition (27), we have

$$
M_{U,\Delta}(m) = \frac{1}{N_{U,\Delta}} \sum_{E \in (U-\Delta, U]} (\langle E|\Omega|E \rangle)^m
\leq \frac{1}{N_{U,\Delta}} \sum_{E \in (U-\Delta, U]} \langle E|\Omega^m|E \rangle = \langle \Omega^m \rangle_{U,\Delta},
$$

where we have used that for an arbitrary quantum state $|\psi\rangle$, we have

$$
(\langle \psi|\Omega|\psi \rangle)^m \leq \langle \psi|\Omega^m|\psi \rangle
$$

due to the convexity of $x^m$ ($m$: even).

Second, we consider the relation that is proved in the subsequent subsection

$$
\frac{\langle \tilde{O} \rangle_{U,\Delta}}{\langle \tilde{O} \rangle_\beta} \leq 2e \left(1 + \frac{gn}{\delta}\right).
$$

for arbitrary non-negative operators $\tilde{O} \in \mathcal{B}(\mathcal{H})$. By choosing $\tilde{O} = \Omega^m \geq 0$, we have

$$
M_{U,\Delta}(m) \leq \langle \Omega^m \rangle_{U,\Delta} \leq 2e \left(1 + \frac{gn}{\delta}\right) (\Omega^m)_\beta
\leq 8e \left(1 + \frac{gn}{\delta}\right) (\tilde{c} g n)^{m/2} \left(\frac{m+1}{\gamma}\right)^{(m+1)/\gamma},
$$

where we have used the fact that the assumption (29) implies

$$
(\Omega^m)_\beta \leq \frac{4(\tilde{c} g n)^{m/2}}{\gamma} \Gamma \left(\frac{m+1}{\gamma}\right) \leq \frac{4(\tilde{c} g n)^{m/2}}{\gamma} \left(\frac{m+1}{\gamma}\right)^{(m+1)/\gamma},
$$

with $\Gamma(\cdot)$ the gamma function.
Then, by using the inequality (44), we obtain
\[
P_{U,\Delta}(x \geq x_0) = \int_{x_0}^{\infty} P_{U,\Delta}(x) dx \leq \frac{1}{x_0^m} \int_{x_0}^{\infty} x^m P_{U,\Delta}(x) dx
\]
\[
\leq 8e \left( \frac{1+\gamma}{\delta} \right) \left( \frac{\gamma}{\delta} \right)^{m/2} \left( \frac{m+1}{\gamma} \right)^{(m+1)/\gamma}
\]
\[
= 8e \left( \frac{1+\gamma}{\delta} \right) \left( \frac{x_0^2}{\gamma \delta c e} \right)^{1/2} \left[ \left( \frac{m+1}{\gamma} \right)^{2/\gamma} \frac{\gamma}{\delta c e} \right]^{m+1/2}.
\]
We now choose
\[
m + 1 = \left\lfloor \gamma \left( \frac{x_0^2}{\gamma \delta c e} \right)^{\gamma/2} \right\rfloor \quad \text{or} \quad 1,
\]
and the inequality (46) reduces to
\[
P_{U,\Delta}(x \geq x_0) \leq 8e2^{\sqrt{\gamma n \delta}} \left( 1 + \frac{\gamma}{\delta} \right) \exp \left[ -\gamma \left( \frac{x_0}{\sqrt{\gamma c e}} \right)^{\gamma} \right],
\]
where we use \( x_0^2/\gamma \delta e \leq gn/\delta \) due to \( x_0 \leq gn \). We thus obtain the inequality (30) by combining the case of \( P_{U,\Delta}(x \leq x_0) \). \( \square \)

3.2.1. Proof of the inequality (43)
For the purpose, we consider
\[
\langle \hat{O} \rangle_{U,\Delta} = \frac{1}{\mathcal{N}_{U,\Delta}} \sum_{E \in \{U-\Delta, U\}} \langle E | \hat{O} | E \rangle
\]
\[
\leq e^{\beta U} \frac{\mathcal{N}_{U,\Delta}}{\mathcal{N}_{U,\delta}} \sum_{E \in \{U-\Delta, U\}} e^{-\beta E} \langle E | \hat{O} | E \rangle
\]
\[
\leq \frac{Ze^{\beta U}}{\mathcal{N}_{U,\delta}} \sum_{E \in (-\infty, \infty)} \frac{1}{Z} e^{-\beta E} \langle E | \hat{O} | E \rangle \leq \frac{Ze^{\beta U}}{\mathcal{N}_{U,\delta}} \langle \hat{O} \rangle_{\beta},
\]
where \( \mathcal{N}_{U,\Delta} \geq \mathcal{N}_{U,\delta} \) because of \( \delta = \min(\Delta, 1/\beta) \leq \Delta \).
To bound \( \frac{Ze^{\beta U}}{\mathcal{N}_{U,\delta}} \) from above, we consider
\[
Z = \sum_{E \in (-\|H\|,\|H\|)} e^{-\beta E} \leq \sum_{\nu \in \mathbb{Z}, \nu \delta \in (-\|H\|,\|H\|+\nu)} \mathcal{N}_{\nu,\delta,\nu} e^{-\beta (\nu-1)}
\]
\[
\leq e^{\beta \delta} \left( 2 + \frac{2\|H\|}{\delta} \right) \max_{\nu \in \mathbb{Z}} (\mathcal{N}_{\nu,\delta,\nu} e^{-\beta \nu})
\]
\[
\leq 2e \left( 1 + \frac{\gamma n}{\delta} \right) \mathcal{N}_{\nu,\delta,\nu} e^{-\beta \nu} = 2e \left( 1 + \frac{\gamma n}{\delta} \right) \mathcal{N}_{U,\delta} e^{-\beta U},
\]
where we use the inequality (5), \( \delta = \min(\Delta, 1/\beta) \) and the definition of \( U \) in Eq. (26). By combining the inequalities (49) and (50), we obtain the inequality (43). \( \square \)

4. Summary and future perspective
In summary, we have worked on the concentration bound and the ensemble equivalence between the canonical and the micro-canonical distributions in long-range interacting systems. Our first theorem 1 (or Corollary 1) ensures that the Gaussian concentration inequality holds above a threshold temperature \( \beta_c = 1/(4e^3 g k) \) with \( g \) and \( k \) the parameters of Hamiltonian. We have then connected the concentration bound to the ensemble equivalence...
in Theorem 2. The theorem itself is not restricted to the high-temperature Gibbs states and can be applied to more general cases. In applying it to the high-temperature Gibbs states, the Gaussian concentration bound implies \( \frac{1}{n}|(\Omega)_{\Omega,\Delta} - \langle \Omega \rangle_{\beta}| \leq C_{2} \left( n^{-1} \log(\Delta^{-1} n^{3/2}) \right)^{1/2} \) for arbitrary few-body operators \( \Omega \in \mathcal{B}(\mathcal{H}) \) with \( \Delta (\leq 1/\beta) \) the width of the energy shell. As shown in Corollary 4, we have also proved the weak eigenstate thermalization, namely that almost all the eigenstates in the energy shell have the similar value as the micro-canonical average. Our results have given the first theoretical step to quantitatively treat the ensemble equivalence as well as the weak eigenstate thermalization in generic quantum systems including long-range interaction.

We have left several open questions. First, the Gaussian concentration bound in Corollary 1 is applied to only the few-body observables as Eq. (8). This class of observables covers almost all the interesting thermodynamic properties. However, in order to discuss the trace distance between the reduced density matrices of the canonical and the micro-canonical states, we need to consider summation of non-local operators (see Refs. [28, 29, 65]):

\[
\hat{\Omega} = \sum_{i=1}^{\tilde{n}} \hat{\Omega}_{i}, \quad \|\tilde{\Omega}_{i}\| \leq 1, \tag{51}
\]

where \( \{\tilde{\Omega}_{i}\}_{i=1}^{\tilde{n}} \) are supported on large subsystems \( B_{i} \subset V \) with \( |B_{i}| \gg 1 \) which are not overlapped with each other (i.e., \( B_{i} \cap B_{j} = \emptyset \)). In this case, we still expect the Gaussian concentration inequality in the form of \( \exp[-x^{2}/(c\tilde{n})] \) above a threshold temperature. Unfortunately, our present proof cannot be directly extended to this case.

The second question is whether we can prove the ensemble equivalence in long-range interacting systems at low temperatures. We have already shown that the concentration bound (29) in Gibbs states is a sufficient condition that the ensemble equivalence holds. We note that the bound should not universally holds since the violation of the ensemble equivalence has been reported in Refs. [32, 33]. Therefore, our task is to identify the condition where the concentration bound (29) holds at low temperatures. We have left several open questions. First, the Gaussian concentration bound in Corollary 1 is applied to only the few-body observables as Eq. (8). This class of observables covers almost all the interesting thermodynamic properties. However, in order to discuss the trace distance between the reduced density matrices of the canonical and the micro-canonical states, we need to consider summation of non-local operators (see Refs. [28, 29, 65]):

\[
\hat{\Omega} = \sum_{i=1}^{\tilde{n}} \hat{\Omega}_{i}, \quad \|\tilde{\Omega}_{i}\| \leq 1, \tag{51}
\]

where \( \{\tilde{\Omega}_{i}\}_{i=1}^{\tilde{n}} \) are supported on large subsystems \( B_{i} \subset V \) with \( |B_{i}| \gg 1 \) which are not overlapped with each other (i.e., \( B_{i} \cap B_{j} = \emptyset \)). In this case, we still expect the Gaussian concentration inequality in the form of \( \exp[-x^{2}/(c\tilde{n})] \) above a threshold temperature. Unfortunately, our present proof cannot be directly extended to this case.

The second question is whether we can prove the ensemble equivalence in long-range interacting systems at low temperatures. We have already shown that the concentration bound (29) in Gibbs states is a sufficient condition that the ensemble equivalence holds. We note that the bound should not universally holds since the violation of the ensemble equivalence has been reported in Refs. [32, 33]. Therefore, our task is to identify the condition where the concentration bound (29) holds at low temperatures.

Finally, further applications of the concentration bounds to other problems in the statistical mechanics are important future challenge.

5. Proof of Theorem 1

5.1. Cluster notation

We first define several basic terminologies. We define \( E_{k} \) as the set of \( X \subset V \) such that \( |X| \leq k \), namely

\[
E_{k} := \{ X \subset V | |X| \leq k \}. \tag{52}
\]

We call a multiset \( w = \{X_{1}, X_{2}, \ldots, X_{|w|}\} \) with \( X_{j} \in E_{k} \) for \( j = 1, 2, \ldots, |w| \) as “cluster”, where \(|w|\) is the cardinality of \( w \) (i.e., the number of subsets in \( w \)). We denote \( C_{m} \) by the set of \( w \) with \(|w| = m \). We define \( V_{w} \subseteq V \) as

\[
V_{w} := X_{1} \cup X_{2} \cup \cdots \cup X_{|w|}. \tag{53}
\]

Also, we define the connected cluster as follows:

**Definition 1. (Connected cluster)** For a cluster \( w \in C_{|w|} \), we say that \( w \) is a connected cluster if there are no decompositions of \( w = w_{1} \oplus w_{2} \) such that \( V_{w_{1}} \cap V_{w_{2}} = \emptyset \). We denote by \( \mathcal{G}_{m} \) the set of the connected clusters with \(|w| = m \).

**Definition 2. (Connected cluster to a region, FIG. 2)** Similarly, we say that \( w \) is a connected cluster to \( L \) if there are no decompositions of \( w = w_{1} \oplus w_{2} \) such that \( (L \cup V_{w_{1}}) \cap V_{w_{2}} = \emptyset \). We denote by \( \mathcal{G}_{m}^{L} \) the set of the connected clusters to \( L \) with \(|w| = m \).

5.2. Generalized cluster Expansion

We here introduce the generalized cluster expansion, which we distinguish from the standard cluster expansion. We first parametrize \( \hat{H} \) by a parameter set \( \hat{\alpha} := \{a_{X}\}_{X \in E_{k}} \) as

\[
H_{\hat{\alpha}} = \sum_{X \in E_{k}} a_{X} h_{X}, \tag{54}
\]
where $H_\tilde{\imath} = H$ with $\tilde{\imath} = \{1, 1, \ldots, 1\}$. By using Eq. (54), we define a parametrized Gibbs state $\rho_\tilde{\alpha}$ as

$$\rho_\tilde{\alpha} := \frac{e^{-\beta H_\tilde{\alpha}}}{Z_\tilde{\alpha}}$$

where $Z_\tilde{\alpha} := \text{tr}(e^{-\beta H_\tilde{\alpha}})$. Similarly, we parametrize $\Omega$ by $\Omega_\tilde{\alpha}$ as in Eq. (54); that is, $\Omega_\tilde{\alpha} = \sum_{X \in E_k} b_X \omega_X$. In the standard cluster expansion, we consider the Taylor expansion of $e^{-\beta H_\tilde{\alpha}}$ with respect to the parameters $\tilde{\alpha}$. Instead, we here utilize the Taylor expansion of $\log[\text{tr}(e^{-\beta \Omega_\tilde{\alpha}} \rho_\tilde{\alpha})]$ with respect to the parameters $\tilde{\alpha}$ and $\tilde{\beta}$.

First, the Taylor expansion of $\rho_\tilde{\alpha}$ with respect to $\tilde{\alpha}$ reads

$$\rho_\tilde{\alpha} = \sum_{m=0}^\infty \frac{1}{m!} \left[ \left( \sum_{X \in E_k} \frac{\partial}{\partial x_X} \right)^m \rho_\tilde{\alpha} \right]_{\tilde{\alpha} = \tilde{0}} = \sum_{m=0}^\infty \frac{1}{m!} \sum_{X_1, X_2, \ldots, X_m \in E_k} \prod_{j=1}^m \frac{\partial}{\partial x_{X_j}} \rho_\tilde{\alpha} \bigg|_{\tilde{\alpha} = \tilde{0}},$$

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

$$\sum_{X_1, X_2, \ldots, X_m \in E_k} = \sum_{w \in C_m} n_w$$

which yields

$$\rho_\tilde{\alpha} = \sum_{m=0}^\infty \frac{1}{m!} \sum_{w \in C_m} n_w D_w \rho_\tilde{\alpha} \bigg|_{\tilde{\alpha} = \tilde{0}} \quad \text{with} \quad D_w := \prod_{j=1}^m \frac{\partial}{\partial x_{X_j}},$$

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

We second expand the moment generating function $\log[\text{tr}(e^{-\beta \Omega_\tilde{\alpha}} \rho_\tilde{\alpha})]$ with respect to $\tilde{\beta}$. In the same way as the derivation of Eq. (58), we obtain

$$\log[\text{tr}(e^{-\beta \Omega_\tilde{\alpha}} \rho_\tilde{\alpha})] = \sum_{m=0}^\infty \frac{1}{m!} \sum_{w \in C_m} n_w D_w \log[\text{tr}(e^{-\beta \Omega_\tilde{\alpha}} \rho_\tilde{\alpha})] \bigg|_{\tilde{\beta} = \tilde{0}}$$

where we defined

$$D_w := \prod_{X \in w} \frac{\partial}{\partial \beta_X}.$$

In the summation (59), let us consider the cases of $m = 0$ and $m = 1$ in

$$D_w \log[\text{tr}(e^{-\beta \Omega_\tilde{\alpha}} \rho_\tilde{\alpha})] \bigg|_{\tilde{\beta} = \tilde{0}} = D_w \left[ \log[\text{tr}(e^{-\beta \Omega_\tilde{\alpha}} e^{-\beta H_\tilde{\alpha}})] - \log Z_\tilde{\alpha} \right] \bigg|_{\tilde{\beta} = \tilde{0}}.$$

Figure 2: Schematic pictures of clusters of $w \in C^L$ and $w \notin C^L$. Each of the elements $\{X_i | X_e \in E_k\}$ is a subset of the total set $V$ (i.e., $X \subset V$). In (a), there are no decompositions of $w = w_1 \oplus w_2$ 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 $w' = w'_1 \oplus w'_2$ with $w'_i = \{X'_1, X'_2\}$ and $w'_2 = \{X'_1, X'_4\}$ satisfies $(L \cup V_{w'_1}) \cap V_{w'_2} = \emptyset$.
For $m = 0$ (i.e., $|w| = 0$), the derivation (61) vanishes. For $m = 1$, we have
\[
\tilde{D}_w \left( \log \left[ \frac{\tr(e^{-\tau \Omega \omega e^{-\beta H_t}})}{\tr(e^{-\beta H_t})} \right] - \log Z_t \right) \bigg|_{b=0} = D_w \log \left[ \frac{\tr(e^{-\tau \Omega \omega e^{-\beta H_t}})}{\tr(e^{-\beta H_t})} \right] = -\tau \tr(\omega X \rho_t) - \tau \tr(\omega X \rho_t),
\]
which yields
\[
\sum_{w \in C_1} n_w \tilde{D}_w \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = -\tau \sum_{\omega X \in E_k} \tr(\omega X \rho_t) = -\tau \tr(\Omega \rho_t).
\]
By taking out the terms of $m = 0, 1$ in Eq. (59), we have
\[
\log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] + \tau \tr(\Omega \rho_t) = \sum_{m=2}^{\infty} \sum_{w \in C_m} \frac{n_w}{m!} \tilde{D}_w \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = 0,
\]
which reduces to
\[
\log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] + \tau \tr(\Omega \rho_t) = \sum_{m=2}^{\infty} \sum_{w \in C_m} \frac{n_w}{m!} \tilde{D}_w \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = 0.
\]
where we use the expansion (58) for $\rho_t$ in the second equation.

The generalized cluster expansion (65) is only the multi-parameter Taylor expansion in itself. However, we can prove that the summation with respect to $\sum_{w \in C_m}$ reduces to quite a simple form as follows:

Proposition 1. The cluster expansion (65) reduces to the summation of connected clusters as follows:
\[
\log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] + \tau \tr(\Omega \rho_t) = \sum_{m=2}^{\infty} \sum_{w \in C_m} \frac{n_w}{m!} \tilde{D}_w \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = 0.
\]
Here, the summation $\sum_{w \in C_m}$ in Eq. (65) is replaced by $\sum_{w \in G_m}$ in Eq. (66).

From this proposition, we need to estimate the contribution of clusters in $G_m$ to upperbound the moment generating function. In the following, we rewrite the summation as
\[
\sum_{w \in G_m} \frac{n_w}{m!} \sum_{w' \in G_m} \frac{D_{w2} \tilde{D}_{w1}}{D_w} \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = 0,
\]
which yields,
\[
\log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] + \tau \tr(\Omega \rho_t) = \sum_{m=1}^{\infty} \sum_{w \in C_m} \frac{n_w + 1}{(m+1)!} \sum_{w' \in C_{m-1}} \frac{n_{w'}}{m!} \tilde{D}_{w2} \tilde{D}_{w1} \frac{\partial}{\partial b_X} \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = 0,
\]
where $G_m^X$ was defined in Def. 2.

5.3. Summation of the expansion

In order to upperbound the summation with respect to connected clusters, the estimation of the upper bound of
\[
D_{w2} \tilde{D}_{w1} \frac{\partial}{\partial b_X} \log \left[ \frac{\tr(e^{-\tau \Omega \omega \rho_t})}{\tr(\omega X \rho_t)} \right] \bigg|_{b=0} = 0.
\]
is crucial. Because of $\frac{\partial}{\partial b_X} \log Z_a = 0$, we first obtain

$$\frac{\partial}{\partial b_X} \log \left[ \text{tr} \left( e^{-\Omega_b \rho_a} \right) \right] = -\tau \frac{\text{tr} \left( \omega_X e^{-\tau \Omega_b e^{-\beta H_a}} \right)}{\text{tr} \left( e^{-\tau \Omega_b e^{-\beta H_a}} \right)} = -\tau \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right), \quad (70)$$

where $\Phi_{\bar{a},\bar{b}}$ is defined as

$$\Phi_{\bar{a},\bar{b}} := \frac{e^{-\tau \Omega_b e^{-\beta H_a}}}{\text{tr} \left( e^{-\tau \Omega_b e^{-\beta H_a}} \right)}. \quad (71)$$

In the following proposition, we give an explicit form of $D_{w_2} D_{w_1} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) |_{\bar{a}=\bar{0},\bar{b}=\bar{0}}$.

**Proposition 2.** Let us take $m$ copies of the total Hilbert space $\mathcal{H}$ and consider $\mathcal{H} \otimes m+1 = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_{m+1}$. Then, for an arbitrary cluster $w = w_1 \otimes w_2$, we have

$$D_{w_2} D_{w_1} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) |_{\bar{a}=\bar{0},\bar{b}=\bar{0}} = (\beta)^{m_1} (\beta)^{m_2} \frac{P_{m_1} P_{m_2}}{D^{m+1}_H} \text{tr}_{H \otimes m+1} \left( \omega_X^{(1)} \omega_X^{(2)} \cdots \omega_X^{(m)} \right) h^{(m)} X_{1,1} X_{1,2} \ldots X_{1,m_1},$$

$$\text{where } w_1 = \{ X_{1,j} \}_{j=1}^{m_1} \text{ and } w_2 = \{ X_{2,j} \}_{j=1}^{m_2}, \text{ where for an arbitrary operator } O \in \mathcal{B}(\mathcal{H}) \text{ we define}$$

$$O_{H_s} := \frac{1}{s} \text{ tr} \left( \omega_X^{(1)} \omega_X^{(2)} \cdots \omega_X^{(m)} \right) h^{(m)} X_{1,1} X_{1,2} \ldots X_{1,m_1},$$

$$O^{(s)} := O_{H_s} + O_{H_{s+1}} + \cdots + O_{H_n} - s O_{H_1}, \quad (73)$$

for $s = 1, 2, \ldots, m$. Also, $P_{m_1}$ and $P_{m_2}$ denote the symmetrization operators for $h_X$ and $\omega_X$, respectively:

$$P_{m_1} \omega_X^{(1)} X_{1,1} X_{1,2} \cdots X_{1,m_1} = \sum_{\sigma} \omega_X^{(1)} X_{1,1} X_{1,2} \cdots X_{1,m_1} \left( \sigma \right),$$

$$P_{m_2} h^{(m)} X_{1,1} X_{1,2} \cdots X_{1,m_1} = \sum_{\sigma} h^{(m)} X_{1,1} X_{1,2} \cdots X_{1,m_1} \left( \sigma \right). \quad (74)$$

with $\sum_{\sigma}$ the summation of $m!$ terms which come from all the permutations.

We then obtain an upper bound of $|D_{w_2} D_{w_1} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) |_{\bar{a}=\bar{0},\bar{b}=\bar{0}}$ from Proposition 2, we have

$$\left| D_{w_2} D_{w_1} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) \right|_{\bar{a}=\bar{0},\bar{b}=\bar{0}} \leq \frac{2^{m_1} \tau^{m_2}}{D^{m+1}_H} \text{tr}_{H \otimes m+1} \left( \omega_X^{(0)} \omega_X^{(1)} \cdots \omega_X^{(m)} \right) h^{(m)} X_{1,1} X_{1,2} \ldots X_{1,m_1}. \quad (75)$$

In order to bound the right-hand side of (75), we utilize the following proposition:

**Proposition 3.** Let $O_0 \in \mathcal{B}(\mathcal{H})$ be an operator supported on a subset $L \subseteq V$ such that $\text{tr}(O_0) = 0$. We then consider arbitrary $m$ operators $\{ O_s \}_{s=1}^{m} \in \mathcal{B}(\mathcal{H})$ which are supported on $w := \{ X_s \}_{s=1}^{m}$ respectively and satisfy $\text{tr}(O_s) = 0$ for $s = 1, 2, \ldots, m$. For each of $\{ O_s \}_{s=1}^{m}$, we define $O^{(s)}$ as in Eq. (73). We then prove

$$\frac{1}{D^{m+1}_H} \left| \text{tr}_{H \otimes m+1} \left( O^{(0)}_0 O^{(1)}_1 O^{(2)}_2 \cdots O^{(m)}_m \right) \right| \leq \left| O_0 \right| \left| \prod_{s=1}^{m} 2 N_{X_s \mid w_L} \right| \left| O_s \right|. \quad (76)$$

where $w_L := \{ L, X_1, X_2, \ldots, X_m \}$ and $N_{X_s \mid w_L}$ is a number of subsets in $w_L$ that have overlap with $X_s$ (Fig. 3):

$$N_{X_s \mid w_L} = \# \left\{ X \in w_L | X \neq X_s \right\}.$$  \quad (77)
where $w$ by analysis, and Proposition 3 plays a crucial role in proving the convergence of the generalized cluster expansion. However, this estimation is too loose and cannot ensure the convergence of Eq. (68). We thus need more refined analysis, and Proposition 3 plays a crucial role in proving the convergence of the generalized cluster expansion.

Then, for an arbitrary subset $L$, we obtain

\[
\sum_{w \in G_L^m} \frac{n_w}{m_1} \prod_{s=1}^{m} N_{X_s \mid w_L} \|\phi_{X_s}\| \leq \frac{1}{2} e^{\lfloor L \rfloor k (2e^3 g k)^m}.
\]

By using Proposition 4 with $L = X$, we can derive an upper bound of

\[
\sum_{w \in G_L^m} \frac{n_w}{m_1} \prod_{s=1}^{m} N_{X_s \mid w_L} \|\phi_{X_s}\| \leq \frac{1}{2} e^{\lfloor L \rfloor k (2e^3 g k)^m} (\beta \tau)^{m_1} (2\beta)^{m_1} \left( \begin{array}{c} m_1 \\ m_1 \end{array} \right) \leq \frac{1}{2} e^{(4e^3 g k)^m ([\beta + \tau]^m - \beta^m)}.
\]
This reduces the inequality (79) to
\[
\log \left[ \text{tr} \left( e^{-\tau \Omega_1} \rho_1 \right) \right] + \tau \text{tr} (\Omega_1 \rho_1) \leq \sum_{X \in E_k}^\infty \frac{\tau}{m+1} \| \omega_X \| \left( \frac{\beta + \tau}{\beta_c} - \frac{\beta/\beta_c}{1 - (\beta + \tau)/\beta_c} \right)
\]
\[
\leq \frac{\beta^2}{\beta - \tau} \frac{\Omega}{\beta_c - \beta - \tau}, \tag{83}
\]
where we use the notation of \( \beta_c = 1/(4e^3 g k) \). This completes the proof of Theorem 1.

6. acknowledgments

The work of T. K. was supported by the RIKEN Center for AIP and JSPS KAKENHI Grant No. 18K13475. K.S. was supported by JSPS Grants-in-Aid for Scientific Research (JP16H02211).

Appendix A. Proof of Proposition 1

For the proof, we first rewrite
\[
D_{w_2} \hat{D}_{w_1} \log \left[ \text{tr} \left( e^{-\tau \Omega_2} \rho_2 \right) \right] \bigg|_{\tilde{a} = \tilde{b} = 0} = D_{w_2} \hat{D}_{w_1} \log \left[ \text{tr} \left( e^{-\tau \Omega_2} \rho_{\tilde{a}_{w_2}} \right) \right] \bigg|_{\tilde{a}_{w_2} = \tilde{b}_{w_1} = 0}, \tag{A.1}
\]
where we define \( \tilde{a}_w \) as a parameter vector such that only the elements in \( w \) are non-zero, namely
\[
(\tilde{a}_w)_X \begin{cases} 
\neq 0 & \text{for} \ X \in w, \\
= 0 & \text{for} \ X \notin w.
\end{cases} \tag{A.2}
\]
An element of \( a_X \) in \( \tilde{a} \) is denoted by \( (\tilde{a})_X \). In the same way, we define \( \tilde{b}_w \).

We now need to prove
\[
D_{w_2} \hat{D}_{w_1} \log \left[ \text{tr} \left( e^{-\tau \Omega_2} \rho_{\tilde{a}_{w_2}} \right) \right] \bigg|_{\tilde{a}_{w_2} = \tilde{b}_{w_1} = 0} = 0 \tag{A.3}
\]
for \( w_1 \oplus w_2 \notin \mathcal{G}_{|w_1| + |w_2|} \). The unconnected condition of the cluster \( w_1 \oplus w_2 \) implies the existence of the decomposition of
\[
w_1 \oplus w_2 = (\tilde{w}_1 \oplus \tilde{w}_2) \oplus (\tilde{w}_1 \oplus \tilde{w}_2), \quad V_{\tilde{w}_1 \oplus \tilde{w}_2} \cap V_{\tilde{w}_1 \oplus \tilde{w}_2} = \emptyset, \tag{A.4}
\]
where \( \tilde{w}_1 = \tilde{w}_1 \oplus \tilde{w}_1 \) and \( \tilde{w}_2 = \tilde{w}_2 \oplus \tilde{w}_2 \) with \( |\tilde{w}_1 \oplus \tilde{w}_2| > 0 \) and \( |\tilde{w}_1 \oplus \tilde{w}_2| > 0 \). Then, the operator \( \Omega_{\tilde{w}_1} = \Omega_{\tilde{w}_1 \oplus \tilde{w}_1} \) can be decomposed as
\[
\Omega_{\tilde{w}_1} = \Omega_{\tilde{w}_1} + \Omega_{\tilde{w}_1}, \tag{A.5}
\]
which yields
\[
e^{-\tau \Omega_2} = e^{-\tau \Omega_{\tilde{w}_1}} \otimes e^{-\tau \Omega_{\tilde{w}_1}}, \tag{A.6}
\]
because of \( V_{\tilde{w}_1} \cap V_{\tilde{w}_1} = \emptyset \). Notice that the operators \( \Omega_{\tilde{w}_1} \) and \( \Omega_{\tilde{w}_1} \) are supported on the subsets \( V_{\tilde{w}_1} \) and \( V_{\tilde{w}_1} \), respectively. Also, the density matrix \( \rho_{\tilde{a}_{w_2}} = \rho_{\tilde{a}_{w_2} \oplus \tilde{a}_{w_2}} \) is now defined by using the Hamiltonian
\[
H_{\tilde{a}_{w_2}} = H_{\tilde{a}_{w_2} \oplus \tilde{a}_{w_2}} = H_{\tilde{a}_{w_2}} + H_{\tilde{a}_{w_2}}, \tag{A.7}
\]
and hence $e^{-\beta H_{w_2}} = e^{-\beta H_{\bar{a}_2}} \otimes e^{-\beta H_{\bar{a}_2}}$. Therefore, from Eqs. (A.4) and (A.6), we obtain
\[
\text{tr} \left( e^{-\tau \Omega_{w_1}} \rho_{\bar{a}_2} \right) = \text{tr} \left( e^{-\tau \Omega_{\bar{a}_2}} \rho_{\bar{a}_2} \right) \text{tr} \left( e^{-\tau \Omega_{\bar{a}_2}} \rho_{\bar{a}_2} \right).
\] (A.8)

The equation (A.8) implies
\[
\mathcal{D}_{w_2} \mathcal{D}_{\bar{a}_2} \log \left[ \text{tr} \left( e^{-\tau \Omega_{w_1}} \rho_{\bar{a}_2} \right) \right] = \mathcal{D}_{w_2} \mathcal{D}_{\bar{a}_2} \mathcal{D}_{\bar{a}_2} \log \left[ \text{tr} \left( e^{-\tau \Omega_{w_1}} \rho_{\bar{a}_2} \right) \right] + \log \left[ \text{tr} \left( e^{-\tau \Omega_{\bar{a}_2}} \rho_{\bar{a}_2} \right) \right].
\] (A.9)

Finally, because of
\[
\mathcal{D}_{\bar{a}_2} \mathcal{D}_{\bar{a}_2} \log \left[ \text{tr} \left( e^{-\tau \Omega_{w_1}} \rho_{\bar{a}_2} \right) \right] = 0, \quad \mathcal{D}_{w_2} \mathcal{D}_{\bar{a}_2} \log \left[ \text{tr} \left( e^{-\tau \Omega_{w_1}} \rho_{\bar{a}_2} \right) \right] = 0
\] (A.10)
for $|\bar{w}_1 + \bar{w}_2| > 0$ and $|\bar{w}_1 + \bar{w}_2| > 0$, Eq. (A.9) reduces to
\[
\mathcal{D}_{w_2} \mathcal{D}_{w_2} \log \left[ \text{tr} \left( e^{-\tau \Omega_{w_1}} \rho_{\bar{a}_2} \right) \right] = 0.
\] (A.11)

This gives the equation (A.3) and completes the proof of Proposition 1. □

**Appendix B. Proof of Proposition 2**

We here show the proof of Proposition 2, which gives the explicit form of $\mathcal{D}_{w_2} \mathcal{D}_{w_1} \text{tr}(\omega_X \Phi_{\bar{a},\bar{b}})|_{\bar{a}=\bar{b}=\bar{g}}$. For the proof, we first consider the Taylor expansion with respect to $\beta$ and $\tau$:
\[
\text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) = \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \sum_{m_1}^{\infty} \frac{m_1 \beta m_2 \beta m_2}{m_1! m_2!} \partial_{\beta m_1} \partial_{\beta m_2} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) \bigg| \beta = \tau = 0
\] (B.1)

The following lemma gives the explicit form of the derivatives:

**Lemma 2.** By using the notation (73), we obtain derivatives of $\text{log} \left[ \text{tr} \left( e^{-\tau \Omega_{w_1}} e^{-\beta H_{\bar{a}}} \right) \right]$ with respect to $\beta$ and $\tau$ as
\[
\frac{\partial^{m_2} \partial^{m_1}}{\partial \beta^{m_2} \partial \beta^{m_1}} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) = (-1)^m \text{tr}_{H^{m+1}} \left( \omega_X \Omega_0^{(1)} \ldots \Omega_{m_1}^{(m_1)} \Phi_{\bar{a},\bar{b}}^{m+1} H_{\bar{a}}^{(m+1)} H_{\bar{a}}^{(m+2)} \ldots H_{\bar{a}}^{(m)} \right).
\] (B.2)

This yields for $\beta = \tau = 0$
\[
\frac{\partial^{m_2} \partial^{m_1}}{\partial \beta^{m_2} \partial \beta^{m_1}} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) \bigg| \beta = \tau = 0 = (-1)^m \text{tr}_{H^{m+1}} \left( \omega_X \Omega_0^{(1)} \ldots \Omega_{m_1}^{(m_1)} H_{\bar{a}}^{(m+1)} H_{\bar{a}}^{(m+2)} \ldots H_{\bar{a}}^{(m)} \right),
\] (B.3)

where we use $\Phi_{\bar{a},\bar{b}} = 1/D_{\bar{a}}$ for $\beta = 0$ and $\tau = 0$ (see the definition (71)).

By taking $\mathcal{D}_{w_2} \mathcal{D}_{w_1}$ with $|w_1| = m_1$ and $|w_2| = m_2$, only the $m_1$th order terms of $\tau$ and $m_2$th order terms of $\beta$ survives in Eq. (B.1), respectively. Hence, we have
\[
\mathcal{D}_{w_2} \mathcal{D}_{w_1} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) \bigg|_{\bar{a}=\bar{b}=\bar{g}} = \frac{\tau^{m_1} \beta^{m_2}}{m_1! m_2!} \mathcal{D}_{w_2} \mathcal{D}_{w_1} \left( \frac{\partial^{m_2} \partial^{m_1}}{\partial \beta^{m_2} \partial \beta^{m_1}} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) \bigg| \beta = \tau = 0 \right)
\]
\[
= \frac{(\tau^{m_1} \beta^{m_2})}{D_{\bar{a}}^{m_1} m_1! m_2!} \mathcal{D}_{w_2} \mathcal{D}_{w_1} \text{tr} \left( \omega_X \Phi_{\bar{a},\bar{b}} \right) \bigg|_{\beta = \tau = 0}
\] (B.4)

where the second equation comes from Lemma 2. By using the notations of $\mathcal{P}_{m_2}$ and $\mathcal{P}_{m_1}$ as in Eq. (74), we obtain
\[
\mathcal{D}_{w_2} \Omega_{\bar{a}}^{(1)} \ldots \Omega_{\bar{a}}^{(m_2)} |_{\bar{a}=\bar{b}=\bar{g}} = \mathcal{P}_{m_2} \omega_X^{(1)} \ldots \omega_X^{(m_2)}
\]
\[
\mathcal{D}_{w_2} \Omega_{\bar{a}}^{(m_1+1)} H_{\bar{a}}^{(m_1+2)} \ldots H_{\bar{a}}^{(m)} |_{\bar{a}=\bar{b}=\bar{g}} = \mathcal{P}_{m_2} h_{X_{1,1}}^{(m_1+1)} h_{X_{1,2}}^{(m_1+2)} \ldots h_{X_{1,m_1},m_1}^{(m)}.
\] (B.5)
By combining Eqs. (B.4) and (B.5), we obtain Eq. (72). This completes the proof of Proposition 2. □

Proof of Lemma 2. For the proof, we rely on the induction method. For $m_1 + m_2 = 1$, we can obtain Eq. (B.2) because of

$$
\frac{\partial}{\partial \beta} \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right) = \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} H_{\vec{a}} \right) - \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right) \cdot \text{tr} \left( \Phi_{\vec{a}, \vec{b}} H_{\vec{a}} \right) = \text{tr}_{H^\otimes 2} \left( \omega_X^{(0)} \Phi_{\vec{a}, \vec{b}} \Phi_{\vec{a}, \vec{b}} \right)
$$

(B.6)

and

$$
\frac{\partial}{\partial \tau} \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right) = \text{tr} \left( \omega X \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}} \right) - \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right) \cdot \text{tr} \left( \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}} \right) = \text{tr}_{H^\otimes 2} \left( \omega_X^{(0)} \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}} \right)
$$

(B.7)

We then assume Eq. (B.2) for $m_1 + m_2 = m - 1$ and prove the case of $m_1 + m_2 = m$. We first consider the case of $\frac{\partial^{m_1+1}}{\partial \beta^{m_1+1}} \frac{\partial^{m_1}}{\partial \beta^{m_1}} \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right)$ with $m_1 + m_2 = m - 1$. The assumption for $m_1 + m_2 = m - 1$ gives

$$
\frac{\partial^{m_1+1}}{\partial \beta^{m_1+1}} \frac{\partial^{m_1}}{\partial \beta^{m_1}} \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right) = \frac{\partial}{\partial \beta} \left( -1 \right)^{m-1} \text{tr}_{H^\otimes m+1} \left( \omega_X^{(0)} \Omega_{\vec{b}}^{(1)} \cdots \Omega_{\vec{b}}^{(m_1)} \Phi_{\vec{a}, \vec{b}} \otimes \Pi_{m_1+1} H_{\vec{a}}^{(m_1+1)} H_{\vec{a}}^{(m_1+2)} \cdots H_{\vec{a}}^{(m-1)} \right).
$$

(B.8)

Then, we have

$$
\frac{\partial}{\partial \beta} \Phi_{\vec{a}, \vec{b}}^{\otimes m+1} = \frac{\partial}{\partial \beta} \left( e^{-\tau \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}}^{\otimes m}} \right) \otimes m

= -\Phi_{\vec{a}, \vec{b}}^{\otimes m+1} \sum_{s=1}^{m} H_{\vec{a}, \vec{b}} + m \left( e^{-\tau \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}}^{\otimes m}} \right) \otimes m \frac{\text{tr} \left( e^{-\tau \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}}^{\otimes m} H_{\vec{a}}^{(m+1)} H_{\vec{a}}^{(m+2)} \cdots H_{\vec{a}}^{(m-1)} \right)}{\text{tr} \left( e^{-\tau \Omega_{\vec{b}} \Phi_{\vec{a}, \vec{b}}^{\otimes m} H_{\vec{a}}^{(m+1)} H_{\vec{a}}^{(m+2)} \cdots H_{\vec{a}}^{(m-1)} \right)}

= \frac{\partial}{\partial \beta} \Phi_{\vec{a}, \vec{b}}^{\otimes m+1} \sum_{s=1}^{m} H_{\vec{a}, \vec{b}} + m \left( \Phi_{\vec{a}, \vec{b}}^{\otimes m+1} H_{\vec{a}, \vec{b}} \right) = -\text{tr}_{H_{\vec{a}, \vec{b}}} \left( \Phi_{\vec{a}, \vec{b}}^{\otimes m+1} H_{\vec{a}, \vec{b}}^{(m)} \right)
$$

(B.9)

By combining Eqs. (B.8) and (B.9), we obtain

$$
\frac{\partial^{m_1+1}}{\partial \beta^{m_1+1}} \frac{\partial^{m_1}}{\partial \beta^{m_1}} \text{tr} \left( \omega X \Phi_{\vec{a}, \vec{b}} \right) = \left( -1 \right)^{m-1} \text{tr}_{H^\otimes m+1} \left( \omega_X^{(0)} \Omega_{\vec{b}}^{(1)} \cdots \Omega_{\vec{b}}^{(m_1)} \Phi_{\vec{a}, \vec{b}}^{\otimes m+1} H_{\vec{a}}^{(m+1)} H_{\vec{a}}^{(m+2)} \cdots H_{\vec{a}}^{(m-1)} \right).
$$

(B.10)

We thus obtain Eq. (B.2) by using $[H_{\vec{a}}^{(s)} : H_{\vec{a}}^{(s')}] = 0$ for all $s, s'$. The same analysis can be applied to the case of $\frac{\partial^{m_2}}{\partial \beta^{m_2}} \frac{\partial^{m_1}}{\partial \beta^{m_1}}$. This completes the proof of Eq. (B.2) for $m_1 + m_2 = m$. □

Appendix C. Proof of Proposition 3

For the proof, we first notice that $\text{tr}_{H^\otimes m+1} \left( O_{0}^{(0)} O_{1}^{(1)} O_{2}^{(2)} \cdots O_{m}^{(m)} \right) / D_{H}^{m+1}$ consists of a summation of multiplications as follows:

$$
\frac{\text{tr}(O_{0})}{D_{H}} \cdot \frac{\text{tr}(O_{2})}{D_{H}} \cdots \frac{\text{tr}(O_{m})}{D_{H}}
$$

(C.1)

where each of $\{u_j\}_{j=1}^{l}$ is an integer subset in $\{0, 1, 2, 3, \ldots, m\}$ with $|u_j| \geq 2$ and $u_1 \oplus u_2 \oplus \cdots \oplus u_q = \{0, 1, 2, 3, \ldots, m\}$; also for $u = \{i_1, i_2, \ldots, i_{|u|}\}$, we define $O_u := O_{i_1} O_{i_2} \cdots O_{i_{|u|}}$ with $0 \leq i_1 < i_2 < \cdots < i_{|u|} \leq m$. We notice that each of the sets $\{u_j\}_{j=1}^{l}$ in (C.1) is irreducible in the sense that $\{X_{i_1}, X_{i_2}, \ldots, X_{i_{|u|}}\} \in G_{|u|} \left( u = \{i_1, i_2, \ldots, i_{|u|}\} \right)$. We then obtain the following decomposition:

$$
\frac{1}{D_{H}^{m+1}} \text{tr}_{H^\otimes m+1} \left( O_{0}^{(0)} O_{1}^{(1)} O_{2}^{(2)} \cdots O_{m}^{(m)} \right) = \sum_{q=1}^{n/2} \sum_{u_1 \oplus u_2 \oplus \cdots \oplus u_q = \{0, 1, 2, 3, \ldots, m\}} N_{u_1, u_2, \ldots, u_q} \frac{\text{tr}(O_{u_1})}{D_{H}} \cdot \frac{\text{tr}(O_{u_2})}{D_{H}} \cdots \frac{\text{tr}(O_{u_q})}{D_{H}},
$$

(C.2)

where $N_{u_1, u_2, \ldots, u_q} \in \mathbb{Z}$ is a non-trivial coefficient which can be calculated from (73).
For example, let us consider the case of $m = 4$ as shown in Fig. C.4. Because of $\text{tr}(O_s) = 0$ for $s = 0, 1, 2, 3$, we have

$$\frac{1}{D_H^4} \text{tr}_{H^{\otimes 4}} \left( O_0^{(0)} O_1^{(1)} O_2^{(2)} O_3^{(3)} \right) = \frac{\text{tr}(O_0 O_1 O_2 O_3)}{D_H}$$

Then, due to $L \cap X_1 = L \cap X_2 = \emptyset$ from Fig. C.4, only the sets of $\{0, 1, 2, 3\}$, $\{0, 3\}$ and $\{1, 2\}$ are irreducible. We thus obtain

$$\frac{1}{D_H} \text{tr}_{H^{\otimes 4}} \left( O_0^{(0)} O_1^{(1)} O_2^{(2)} O_3^{(3)} \right) = \frac{\text{tr}(O_0 O_1 O_2 O_3)}{D_H} - \frac{\text{tr}(O_0 O_2 O_3)}{D_H} - \frac{\text{tr}(O_0 O_3)}{D_H} - \frac{\text{tr}(O_1 O_2)}{D_H}. \quad (C.3)$$

This yields $N_{\{0,1,2,3\}} = 1$ and $N_{\{0,3\},\{1,2\}} = -1$.

Our task is now to estimate the upper bound of

$$N \left( O_0^{(0)} O_1^{(1)} O_2^{(2)} \cdots O_m^{(m)} \right) := \sum_{q=1}^{n/2} \sum_{u_1, u_2, \ldots, u_q} |N_{x_1, u_2, \ldots, u_q}|. \quad (C.5)$$

Because of

$$\frac{\text{tr}(O_{u_1})}{D_H} \cdot \frac{\text{tr}(O_{u_2})}{D_H} \cdots \frac{\text{tr}(O_{u_q})}{D_H} \leq \|O_0\| \prod_{s=1}^{m} \|O_s\|, \quad (C.6)$$

we obtain the upper bound

$$\frac{1}{D_H^{m+1}} \text{tr}_{H^{\otimes m+1}} \left( O_0^{(0)} O_1^{(1)} O_2^{(2)} \cdots O_m^{(m)} \right) \leq N \left( O_0^{(0)} O_1^{(1)} O_2^{(2)} \cdots O_m^{(m)} \right) \|O_0\| \prod_{s=1}^{m} \|O_s\|. \quad (C.7)$$

In order to estimate the upper bound of $N \left( O_0^{(0)} O_1^{(1)} O_2^{(2)} \cdots O_m^{(m)} \right)$, we consider a more general form as follows:

$$\frac{1}{D_H^{m+1}} \text{tr}_{H^{\otimes m+1}} \left( O_0^{(0)} O_1^{(l_1)} O_2^{(l_2)} \cdots O_m^{(l_m)} \right), \quad (C.8)$$

where $1 \leq l_1 < l_2 < \cdots < l_m < \infty$. We then aim to prove

$$N \left( O_0^{(0)} O_1^{(l_1)} O_2^{(l_2)} \cdots O_m^{(l_m)} \right) \leq \prod_{s=1}^{m} 2N_{x_s|w_L}. \quad (C.9)$$

By applying the above inequality with $\{l_1, l_2, \ldots, l_m\} = \{1, 2, \ldots, m\}$ to Ineq. (C.7), we obtain the inequality (76).

In the following, we give the proof of (C.9) by using mathematical induction. For $m = 1$, we have

$$\text{tr}_{H^{\otimes l_1+1}} \left( O_0^{(0)} O_1^{(l_1)} \right) = \text{tr}_{H^{\otimes l_1+1}} \left[ O_{0,H_1} \left( -l_1 O_{1,H_{l_1+1}} + \sum_{j=1}^{l_1} O_{1,H_j} \right) \right] = \text{tr}(O_0 O_1) - \text{tr}(O_0) \cdot \text{tr}(O_1) = \text{tr}(O_0 O_1), \quad (C.10)$$
which gives $N(O_0^{(0)}O_1^{(1)}) = 1$ as long as $X_0 \cap X_1 \neq \emptyset$, where we define $O_{1,H_j}$ as in Eq. (73). We thus prove the inequality (C.9) for $m = 1$.

We then prove the case of $m = M$ by assuming the inequality (76) for $m = M - 1$. For the purpose, we introduce $C_{s,s'}$ as a operation which applies to $O_{u,H_j}$ ($j = 1, 2, \ldots, m + 1$) as follows:

$$C_{s,s'}O_{u,H_j} := \begin{cases} 0 & \text{for } \{s, s'\} \subset u, \\ O_{u,H_j} & \text{otherwise} \end{cases} \quad (C.11)$$

where the operator $C_{s,s'}$ acts on the index subspace $u \subset \{1, 2, \ldots, m\}$. Roughly speaking, the operator $C_{s,s'}$ prohibits the two operator $O_s$ and $O_{s'}$ to be in the same Hilbert space.

From the definition (C.11), we obtain

$$N \left( C_{s,s'}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(m)} \right) \leq N \left( O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(m)} \right) \quad (C.12)$$

and

$$C_{s,s'}O_s^{(l_s)}O_{s'}^{(l_{s'})} = C_{s,s'}O_{s'}^{(l_{s'})}O_s^{(l_s)} \quad \text{or} \quad C_{s,s'} \left[ O_s^{(l_s)}, O_{s'}^{(l_{s'})} \right] = 0, \quad (C.13)$$

where $[\cdot, \cdot]$ is the commutator. Also, the definition (C.11) implies

$$O_s^{(l_s)}O_{s'}^{(l_{s'})} = (O_sO_{s'}^{(l_{s'})})^{(l_s)} + C_{s,s'}O_s^{(l_s)}O_{s'}^{(l_{s'})} \quad (C.14)$$

for $l_s < l_{s'}$.

We here denote by $O_{s_0}$ the operator that has the minimum $N_{X_s}[w_L]$, namely $N_{X_{s_0}[w_L]} \leq N_{X_s}[w_L]$ ($s = 0, 1, 2, \ldots, M$). We also define $\{s_1, s_2, \ldots, s_{\tilde{m}}\}$ ($s_1 \leq s_2 \leq \cdots \leq s_{\tilde{m}}$) as the indices which satisfy $X_{s_j} \cap X_{s_0} \neq \emptyset$; notice that $\tilde{m} = N_{X_{s_0}[w_L]}$. In the following, we assume $s_0 \leq s_1$ for the simplicity, but the same discussion is applied to the general cases. Because $O_{s_0}$ commutes with $O_s$ if $s \notin \{s_1, s_2, \ldots, s_{\tilde{m}}\}$, we have

$$O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(m)} = O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_{s_1-1}^{(l_{s_1-1})}O_{s_0}^{(l_{s_0})}O_{s_1}^{(l_{s_1})} \cdots O_M^{(l_M)}$$

$$= O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_1})^{(l_{s_0})}O_{s_1}^{(l_{s_1})} \cdots O_M^{(l_M)} + C_{s_0,s_1}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(l_M)}, \quad (C.15)$$

where in the second equality we use Eq. (C.14). In the same way, from Eq. (C.13), we have

$$C_{s_0,s_1}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(m)} = C_{s_0,s_1}O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_2})^{(l_{s_0})}O_{s_2}^{(l_{s_2})} \cdots O_M^{(l_M)} + C_{s_0,s_1}C_{s_0,s_2}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(l_M)}$$

By repeating this process, we finally obtain

$$O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(m)} = O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_1})^{(l_{s_0})}O_{s_1}^{(l_{s_1})} \cdots O_M^{(l_M)}$$

$$+ C_{s_0,s_1}O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_2})^{(l_{s_0})}O_{s_2}^{(l_{s_2})} \cdots O_M^{(l_M)}$$

$$+ C_{s_0,s_1}C_{s_0,s_2}O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_3})^{(l_{s_0})}O_{s_3}^{(l_{s_3})} \cdots O_M^{(l_M)}$$

$$+ \cdots + C_{s_0,s_1}C_{s_0,s_2} \cdots C_{s_0,s_{\tilde{m}-1}}O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_{\tilde{m}}})^{(l_{s_0})}O_{s_{\tilde{m}}}^{(l_{s_{\tilde{m}}})} \cdots O_M^{(l_M)} + C_{s_0,s_1}C_{s_0,s_2} \cdots C_{s_0,s_{\tilde{m}}}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(l_M)}. \quad (C.16)$$

From $\text{tr}(O_s) = 0$ for $s = 0, 1, \ldots, M$, we obtain $\text{tr} \left( C_{s_0,s_1}C_{s_0,s_2} \cdots C_{s_0,s_{\tilde{m}}}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(l_M)} \right) = 0$, and hence

$$N \left( C_{s_0,s_1}C_{s_0,s_2} \cdots C_{s_0,s_{\tilde{m}}}O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(l_M)} \right) = 0. \quad (C.17)$$

By combining the equation (C.16) with the relations (C.12) and (C.17), we have

$$N \left( O_0^{(0)}O_1^{(1)}O_2^{(2)} \cdots O_M^{(l_M)} \right) \leq N \left( O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_1})^{(l_{s_0})}O_{s_1}^{(l_{s_1})} \cdots O_M^{(l_M)} \right)$$

$$+ N \left( O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_2})^{(l_{s_0})}O_{s_2}^{(l_{s_2})} \cdots O_M^{(l_M)} \right)$$

$$+ \cdots + N \left( O_0^{(0)}O_1^{(1)}(O_{s_0}O_{s_{\tilde{m}}})^{(l_{s_0})}O_{s_{\tilde{m}}}^{(l_{s_{\tilde{m}}})} \cdots O_M^{(l_M)} \right). \quad (C.18)$$
In order to estimate the summation, we first decompose

where

combining the two inequalities (C.18) and (C.19), we finally obtain

where we use

We now upperbound each of the terms in the right-hand side of (C.18). The assumption of Ineq. (C.9) for

implies the upper bound of

We then obtain

This completes the proof of the inequality (C.9). □

Appendix D. Proof of Proposition 4

We here obtain an upper bound of

In order to estimate the summation, we first decompose $w_L$ as follows:

where $w_0 = \{L\}$ and $w_j \subset w_L$ satisfy $\text{dist}(w_j, w_0) = j$ for $j = 1, 2, \ldots, l$. Here, we define $\text{dist}(w_j, w_0)$ as the shortest path length in the cluster $w_0 \oplus w_1 \oplus \cdots \oplus w_{j-1}$ which connects from $w_j$ to $w_0$. We also define $q_j := |w_j|$ with $q_j \geq 1$. We then obtain

where $\sum_{w_j|w_0, w_1, \ldots, w_{j-1}}$ denotes the summation with respect to $w_j$ such that $\text{dist}(w_j, w_0) = j$ when $\{w_1, w_2, \ldots, w_{j-1}\}$ are fixed. Note that $w_0$ has been already fixed as $w_0 = \{L\}$. Also, because the subsets in $w_j$ have overlaps with those only in $w_{j-1}$, $w_j$ and $w_{j+1}$, we have

$$N_{X_j(w_j)} = \sum_{w_j \in G_L^m} m! \prod_{s=1}^{m} N_{X_s(w_L)} \|a_{X_s}\|.$$
where we denote \( w_j = \{X_j^{(j)}\}_{j=1}^g \).

By using the above decomposition, we can reduce the inequality (D.1) to

\[
\sum_{w \in \mathcal{G}_n} \frac{n_w}{m!} \prod_{s=1}^m N_{X_s} |w| \|o_{X_s}\| \leq \frac{1}{m!} \sum_{s=1}^m \frac{m!}{\prod_{l \in \{q_1, q_2, \ldots, q_i\}} q_l!} \sum_{w_1} \cdots \sum_{w_m} \prod_{s=1}^m n_{w_1} \prod_{s=1}^m N_{X_s}^{(s)} |w_1 \oplus w_2 \oplus \ldots \oplus w_m| \|o_{X_s}^{(s)}\| \times
\]

\[
\sum_{w_2} \cdots \sum_{w_m} \prod_{s_2=1}^m N_{X_s^{(s_2)\oplus w_2 \oplus w_3}} \|o_{X_s^{(s_2)\oplus w_2 \oplus w_3}}\| \times \cdots \times \sum_{w_{m-1}} \cdots \sum_{w_m} \prod_{s_{m-1}=1}^m N_{X_s^{(s_{m-1})\oplus w_{m-1}}} |w_m| \|o_{X_s^{(s_{m-1})\oplus w_{m-1}}}\|, \tag{D.5}
\]

where we used

\[
\frac{n_w}{n_{w_1} n_{w_2} \cdots n_{w_m}} = \frac{m!}{q_1! q_2! \cdots q_i!}. \tag{D.6}
\]

In the following, we aim to calculate the upper bound of

\[
\sum_{w_j |w_0, \ldots, w_{j-1}} n_{w_j} \prod_{s_j=1}^q N_{X_s}^{(s) |w_{j-1} \oplus w_j} \|o_{X_s}^{(s)}\| \tag{D.7}
\]

only by using \( q_{j-1}, q_j, q_{j+1} \), which does not depend on the details of \( w_{j-1}, w_j, w_{j+1} \). For the purpose, we start from the summation with respect to \( w_1 \):

\[
\sum_{w_1 |w_0, \ldots, w_{l-1}} n_{w_1} \prod_{s_1=1}^{q_1-1} N_{X_s^{(s_1)}} |w_{j-1} \oplus w_j \oplus w_{l-1} \oplus w_1| \|o_{X_s^{(s_1)}}\| \sum_{w_2 |w_0, \ldots, w_{l-1}} \cdots \sum_{w_{l-1}} \prod_{s_{l-1}=1}^{q_{l-1}} N_{X_s^{(s_{l-1}) \oplus w_{l-1}}} |w_1| \|o_{X_s^{(s_{l-1}) \oplus w_{l-1}}}\| \times \prod_{w_1 |w_0, \ldots, w_{l-1}} \prod_{s_1=1}^{q_1} N_{X_s^{(s_1) \oplus w_1}} |w_{l-1} \oplus w_1| \|o_{X_s^{(s_1) \oplus w_1}}\| \tag{D.8}
\]

For fixed \( \{w_1, w_2, \ldots, w_{l-1}\} \), we aim to obtain an upper bound of

\[
\sum_{w_1 |w_0, \ldots, w_{l-1}} n_{w_1} \prod_{s_1=1}^{q_1} N_{X_s^{(s_1) \oplus w_1}} \|o_{X_s^{(s_1) \oplus w_1}}\|, \tag{D.9}
\]

only by using \( q_1 \) and \( q_{l-1} \).

First, we have

\[
\sum_{w_1 |w_0, \ldots, w_{l-1}} n_{w_1} \prod_{s_1=1}^{q_1} N_{X_s^{(s_1) \oplus w_1}} \|o_{X_s^{(s_1) \oplus w_1}}\| \leq \prod_{s_1=1}^{q_1} \left( \sum_{X_{s_1} |w_{l-1} \oplus w_1} N_{X_{s_1}^{(s_1) \oplus w_1}} \|o_{X_{s_1}^{(s_1) \oplus w_1}}\| \right) \leq \left( \max_{w_1 |w_0, \ldots, w_{l-1}} \sum_{X_{s_1} \in \mathcal{C}_{w_1 \oplus w_0 \oplus \ldots \oplus w_{l-1}}} N_{X_{s_1}} \|o_{X_{s_1}}\| \right)^{q_1}. \tag{D.10}
\]

We then prove the following inequality for arbitrary \( w \in \mathcal{C}_w \):

\[
\sum_{X: X \cap V_w \neq \emptyset} N_{X} \|o_{X}\| \leq g \sum_{j=1}^{|w|} |X_j| \quad \text{for } \forall w, \tag{D.11}
\]

where we denote \( w = \{X_j^{(j)}\}_{j=1}^{|w|} \). For the proof of (D.11), we start from the inequality as

\[
\sum_{X: X \cap V_w \neq \emptyset} N_{X} \|o_{X}\| \leq \sum_{v \in V_w} \sum_{X: X \cap v \neq \emptyset} n_v \|o_{X}\| \tag{D.12}
\]
where \( n_{v|w} \) is a number of subsets in \( w \) which have overlaps with the spin \( v \). From the condition (80), we obtain

\[
\sum_{v \in V_w} \sum_{X: X \ni v} n_{v|w} \|o_X\| \leq g \sum_{v \in V_w} n_{v|w} = g \sum_{j=1}^{|w|} |X_j|.
\]

Finally, by combining the inequalities (D.12) and (D.13), we obtain the inequality (D.11). By using the inequality (D.11) with \( w = w_{l-1} \oplus \tilde{w}_l \), the inequality (D.14) reduces to

\[
\sum_{w_l|w_{l-1}} \prod_{q=1}^{q_l} N_{X^{(l)}_{wj}} \|o_{X^{(l)}_{wj}}\| \leq [g_k(q_{l-1} + q_l)]^{q_l},
\]

where we use the fact that the cardinality of \( X \in E_k \) satisfies \( |X| \leq k \).

After the summation with respect to \( w_l \), we can apply the same calculation for the summation with respect to \( w_{l-1} \) for a fixed \( w_{l-2} \):

\[
\sum_{w_{l-2}|w_{l-1}|w_{l-3}} n_{w_{l-2}} \prod_{s_{l-2}=1}^{q_{l-2}} N_{X^{(l-2)}_{w_{l-2}}} \|o_{X^{(l-2)}_{w_{l-2}}}\| \times \max_{\tilde{w}_l \in C_{q_l}} \sum_{w_{l-2}|w_{l-1}|w_{l-3}} n_{w_{l-2}} \prod_{s_{l-1}=1}^{q_{l-1}} N_{X^{(l-1)}_{w_{l-1}}} \|o_{X^{(l-1)}_{w_{l-1}}}\|
\]

\[
\leq \left( \prod_{s_{l-2}=1}^{q_{l-2}} N_{X^{(l-2)}_{w_{l-2}}} \|o_{X^{(l-2)}_{w_{l-2}}}\| \right) \times \max_{\tilde{w}_l \in C_{q_l}} \sum_{w_{l-2}|w_{l-1}|w_{l-3}} n_{w_{l-2}} \prod_{s_{l-1}=1}^{q_{l-1}} N_{X^{(l-1)}_{w_{l-1}}} \|o_{X^{(l-1)}_{w_{l-1}}}\|.
\]

In the same way as the derivation of Ineq. (D.14), we obtain

\[
\sum_{w_l|w_{l-1}|w_{l-2}} n_{w_{l-1}} \prod_{s_{l-1}=1}^{q_{l-1}} N_{X^{(l-1)}_{w_{l-1}}} \|o_{X^{(l-1)}_{w_{l-1}}}\| \leq [g_k(q_{l-2} + q_{l-1} + q_l)]^{q_{l-1}}.
\]

By repeating this process, we finally obtain

\[
\sum_{w \in \mathcal{G}_m} \frac{n_w}{m!} \prod_{s=1}^{m} N_{X_{s|w_l}} \|o_{X_s}\| \leq \sum_{l=1}^{m} \sum_{\{q_1, q_2, \ldots, q_l\}} \frac{1}{q_1! q_2! \ldots q_l!} \left[ g(|L| + kq_1 + kq_2) \right]^{q_1} \prod_{j=2}^{l} \left[ g(k(q_{j-1} + q_j + q_{j+1})) \right]^{q_j},
\]

where we set \( q_{l+1} = 0 \). By using \( n! \geq (n/e)^n \) and \( e^{q_j/|q_j+q_j+q_{j+1}|} \), we have

\[
\frac{1}{q_1! q_2! \ldots q_l!} \left[ g(|L| + kq_1 + kq_2) \right]^{q_1} \prod_{j=2}^{l} \left[ g(k(q_{j-1} + q_j + q_{j+1})) \right]^{q_j}
\]

\[
\leq (gk)^m e^{(|L|/k) \exp \left( \sum_{j=1}^{l} (q_{j-1} + q_j + q_{j+1}) \right)} \leq e^{(|L|/k) (e^3 gk)^m},
\]

which yields

\[
\sum_{w \in \mathcal{G}_m} \frac{n_w}{m!} \prod_{s=1}^{m} N_{X_{s|w_l}} \|o_{X_s}\| \leq \sum_{l=1}^{m} \sum_{\{q_1, q_2, \ldots, q_l\}} e^{(|L|/k) (e^3 gk)^m}.
\]
The summation with respect to \( \{q_1, q_2, \ldots, q_l\} \) is equal to the \((m - l)\)-multicombination from a set of \( l \) elements:

\[
\sum_{\{q_1, q_2, \ldots, q_l\}} = \binom{l}{m - l} = \binom{m - 1}{l - 1},
\]

and hence

\[
\sum_{w \in G_L^m} \frac{n_w}{m!} \prod_{s=1}^{m} N_{X_s | w_L} \|\sigma_{X_s}\| \leq \sum_{l=1}^{m} \binom{m - 1}{l - 1} e^{[L]/k} e^{3gk} = \frac{1}{2} e^{[L]/k} (2e^{3gk})^m.
\]

This completes the proof of Proposition 4. \( \square \)

References

[1] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, C. Monroe, Non-local propagation of correlations in quantum systems with long-range interactions, Nature 511 (7508) (2014) 198.

[2] I. Bloch, J. Dalibard, W. Zwerger, Many-body physics with ultracold gases, Rev. Mod. Phys. 80 (2008) 885–964. doi:10.1103/RevModPhys.80.885. URL https://link.aps.org/doi/10.1103/RevModPhys.80.885

[3] J. W. Britton, B. C. Sawyer, A. C. Keith, C.-C. J. Wang, J. K. Freericks, H. Uys, M. J. Biercuk, J. J. Bollinger, Engineered two-dimensional ising interactions in a trapped-ion quantum simulator with hundreds of spins, Nature 484 (7395) (2012) 489.

[4] M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sen(De), U. Sen, Ultracold atomic gases in optical lattices: mimicking condensed matter physics and beyond, Advances in Physics 56 (2) (2007) 243–379. arXiv: https://doi.org/10.1080/00018730701223200, doi:10.1080/00018730701223200. URL https://doi.org/10.1080/00018730701223200

[5] M. Saffman, T. G. Walker, K. Mølmer, Quantum information with rydberg atoms, Rev. Mod. Phys. 82 (2010) 2313–2363. doi:10.1103/RevModPhys.82.2313. URL https://link.aps.org/doi/10.1103/RevModPhys.82.2313

[6] B. Yan, S. A. Moses, B. Gadway, J. P. Covey, K. R. Hazzard, A. M. Rey, D. S. Jin, J. Ye, Observation of dipolar spin-exchange interactions with lattice-confined polar molecules, Nature 501 (7468) (2013) 521.

[7] V. Bendkowsky, B. Butscher, J. Nipper, J. P. Shaffer, R. Löw, T. Pfau, Observation of ultralong-range rydberg molecules, Nature 458 (7241) (2009) 1005.

[8] M. Kastner, Diverging equilibration times in long-range quantum spin models, Phys. Rev. Lett. 106 (2011) 130601. doi:10.1103/PhysRevLett.106.130601. URL https://link.aps.org/doi/10.1103/PhysRevLett.106.130601

[9] J. Eisert, M. van den Worm, S. R. Manmana, M. Kastner, Breakdown of quasilocality in long-range quantum lattice models, Phys. Rev. Lett. 111 (2013) 260401. doi:10.1103/PhysRevLett.111.260401. URL http://link.aps.org/doi/10.1103/PhysRevLett.111.260401

[10] N. Y. Yao, A. V. Gorshkov, C. R. Laumann, A. M. Läuchli, J. Ye, M. D. Lukin, Realizing fractional chern insulators in dipolar spin systems, Phys. Rev. Lett. 110 (2013) 185302. doi:10.1103/PhysRevLett.110.185302. URL https://link.aps.org/doi/10.1103/PhysRevLett.110.185302

[11] M. F. Maghrebi, Z.-X. Gong, A. V. Gorshkov, Continuous symmetry breaking in 1d long-range interacting quantum systems, Phys. Rev. Lett. 119 (2017) 023001. doi:10.1103/PhysRevLett.119.023001. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.023001

24
[12] B. Neyenhuis, J. Zhang, P. W. Hess, J. Smith, A. C. Lee, P. Richerme, Z.-X. Gong, A. V. Gorshkov, C. Monroe, Observation of prethermalization in long-range interacting spin chains, Science Advances 3 (8). arXiv:http://advances.sciencemag.org/content/3/8/e1700672.full.pdf, doi:10.1126/sciadv.1700672. URL http://advances.sciencemag.org/content/3/8/e1700672

[13] Z.-X. Gong, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, Persistence of locality in systems with power-law interactions, Phys. Rev. Lett. 113 (2014) 030602. doi:10.1103/PhysRevLett.113.030602. URL http://link.aps.org/doi/10.1103/PhysRevLett.113.030602

[14] M. Foss-Feig, Z.-X. Gong, C. W. Clark, A. V. Gorshkov, Nearly linear light cones in long-range interacting quantum systems, Phys. Rev. Lett. 114 (2015) 157201. doi:10.1103/PhysRevLett.114.157201. URL http://link.aps.org/doi/10.1103/PhysRevLett.114.157201

[15] D. Métivier, R. Bachelard, M. Kastner, Spreading of perturbations in long-range interacting classical lattice models, Phys. Rev. Lett. 112 (2014) 210601. doi:10.1103/PhysRevLett.112.210601. URL http://link.aps.org/doi/10.1103/PhysRevLett.112.210601

[16] M. Pino, Entanglement growth in many-body localized systems with long-range interactions, Phys. Rev. B 90 (2014) 174204. doi:10.1103/PhysRevB.90.174204. URL http://link.aps.org/doi/10.1103/PhysRevB.90.174204

[17] T. Kuwahara, Exponential bound on information spreading induced by quantum many-body dynamics with long-range interactions, New Journal of Physics 18 (5) (2016) 053034. doi:10.1088/1367-2630/18/5/053034. URL https://doi.org/10.1088%2F1367-2630%2F18%2F5%2F053034

[18] S. Bravyi, M. B. Hastings, F. Verstraete, Lieb-robinson bounds and the generation of correlations and topological quantum order, Phys. Rev. Lett. 97 (2006) 050401. doi:10.1103/PhysRevLett.97.050401. URL http://link.aps.org/doi/10.1103/PhysRevLett.97.050401

[19] E. H. Lieb, D. W. Robinson, The finite group velocity of quantum spin systems, Communications in mathematical physics 28 (3) (1972) 251–257.

[20] Z. Eldredge, Z.-X. Gong, J. T. Young, A. H. Moosavian, M. Foss-Feig, A. V. Gorshkov, Fast quantum state transfer and entanglement renormalization using long-range interactions, Phys. Rev. Lett. 119 (2017) 170503. doi:10.1103/PhysRevLett.119.170503. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.170503

[21] M. Kliesch, A. Riera, Properties of thermal quantum states: locality of temperature, decay of correlations, and more, arXiv preprint arXiv:1803.09754 arXiv:arXiv:1803.09754

[22] H.-O. Georgii, The equivalence of ensembles for classical systems of particles, Journal of Statistical Physics 80 (5) (1995) 1341–1378. doi:10.1007/BF02179874. URL https://doi.org/10.1007/BF02179874

[23] H. Touchette, Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels, Journal of Statistical Physics 159 (5) (2015) 987–1016. doi:10.1007/s10955-015-1212-2. URL https://doi.org/10.1007/s10955-015-1212-2

[24] R. Lima, Equivalence of ensembles in quantum lattice systems, Ann. Inst. H. Poincaré Sect. A 15 (1971) 61–68.

[25] R. Lima, Equivalence of ensembles in quantum lattice systems: States, Communications in Mathematical Physics 24 (3) (1972) 180–192. doi:10.1007/BF01877711. URL https://doi.org/10.1007/BF01877711

[26] W. De Roeck, C. Maes, K. Netočný, Quantum macrostates, equivalence of ensembles, and an h-theorem, Journal of mathematical physics 47 (7) (2006) 073303.

[27] M. P. Müller, E. Adlam, L. Masanes, N. Wiebe, Thermalization and canonical typicality in translation-invariant quantum lattice systems, Communications in Mathematical Physics 340 (2) (2015) 499–561. doi:10.1007/s00220-015-2473-y. URL https://doi.org/10.1007/s00220-015-2473-y
28. F. G. Brandao, M. Cramer, Equivalence of statistical mechanical ensembles for non-critical quantum systems, arXiv preprint arXiv:1502.03263.

29. H. Tasaki, On the local equivalence between the canonical and the microcanonical ensembles for quantum spin systems, Journal of Statistical Physics 172 (4) (2018) 905–926. doi:10.1007/s10955-018-2077-y. URL https://doi.org/10.1007/s10955-018-2077-y

30. L. D’Alessio, Y. Kafri, A. Polkovnikov, M. Rigol, From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, Advances in Physics 65 (3) (2016) 239–362. arXiv:https://doi.org/10.1080/00018732.2016.1198134, doi:10.1080/00018732.2016.1198134.

31. R. Nandkishore, D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, Annual Review of Condensed Matter Physics 6 (1) (2015) 15–38. arXiv:https://doi.org/10.1146/annurev-conmatphys-031214-014726, doi:10.1146/annurev-conmatphys-031214-014726.

32. J. Barré, D. Mukamel, S. Ruffo, Inequivalence of ensembles in a system with long-range interactions, Phys. Rev. Lett. 87 (2001) 030601. doi:10.1103/PhysRevLett.87.030601. URL https://link.aps.org/doi/10.1103/PhysRevLett.87.030601

33. L. Casetti, M. Kastner, Partial equivalence of statistical ensembles and kinetic energy, Physica A: Statistical Mechanics and its Applications 384 (2) (2007) 318 – 334. doi:https://doi.org/10.1016/j.physa.2007.05.043.

34. L. Gross, Decay of correlations in classical lattice models at high temperature, Communications in Mathematical Physics 68 (1) (1979) 9–27. doi:10.1007/BF01562538. URL https://doi.org/10.1007/BF01562538

35. H. Araki, Gibbs states of a one dimensional quantum lattice, Communications in Mathematical Physics 14 (2) (1969) 120–157. doi:10.1007/BF01645134. URL https://doi.org/10.1007/BF01645134

36. Y. M. Park, H. J. Yoo, Uniqueness and clustering properties of gibbs states for classical and quantum unbounded spin systems, Journal of Statistical Physics 80 (1) (1995) 223–271. doi:10.1007/BF02178359. URL https://doi.org/10.1007/BF02178359

37. D. Ueltschi, Cluster expansions and correlation functions, Moscow Mathematical Journal 4 (2) (2004) 511–522.

38. M. Kliesch, C. Gogolin, M. J. Kastoryano, A. Riera, J. Eisert, Locality of temperature, Phys. Rev. X 4 (2014) 031019. doi:10.1103/PhysRevX.4.031019. URL https://link.aps.org/doi/10.1103/PhysRevX.4.031019

39. J. Fröhlich, D. Ueltschi, Some properties of correlations of quantum lattice systems in thermal equilibrium, Journal of Mathematical Physics 56 (5) (2015) 053302.

40. H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, The Annals of Mathematical Statistics 23 (4) (1952) 493–507. URL http://www.jstor.org/stable/2236576

41. W. Hoeffding, Probability inequalities for sums of bounded random variables, Journal of the American Statistical Association 58 (301) (1963) 13–30. arXiv:https://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1963.10500830, doi:10.1080/01621459.1963.10500830.

42. X. Chen, Z.-C. Gu, X.-G. Wen, Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order, Phys. Rev. B 82 (2010) 155138. doi:10.1103/PhysRevB.82.155138. URL http://link.aps.org/doi/10.1103/PhysRevB.82.155138
[43] T. Kuwahara, Connecting the probability distributions of different operators and generalization of the chernoff–hoeffding inequality, Journal of Statistical Mechanics: Theory and Experiment 2016 (11) (2016) 113103. doi:10.1088/1742-5468/2016/11/113103. URL https://doi.org/10.1088%2F1742-5468%2F2016%2F11%2F113103

[44] A. Anshu, Concentration bounds for quantum states with finite correlation length on quantum spin lattice systems, New Journal of Physics 18 (8) (2016) 083011. doi:10.1088/1367-2630/18/8/083011. URL https://doi.org/10.1088%2F1367-2630%2F18%2F8%2F083011

[45] T. Kuwahara, Asymptotic behavior of macroscopic observables in generic spin systems, Journal of Statistical Mechanics: Theory and Experiment 2016 (5) (2016) 053103. doi:10.1088/1742-5468/2016/05/053103. URL https://doi.org/10.1088%2F1742-5468%2F2016%2F05%2F053103

[46] T. Kuwahara, I. Arad, L. Amico, V. Vedral, Local reversibility and entanglement structure of many-body ground states, Quantum Science and Technology 2 (1) (2017) 015005. doi:10.1088/2058-9565/aa523d. URL https://doi.org/10.1088%2F2058-9565%2Faa523d

[47] I. Arad, T. Kuwahara, Z. Landau, Connecting global and local energy distributions in quantum spin models on a lattice, Journal of Statistical Mechanics: Theory and Experiment 2016 (3) (2016) 033301. URL http://stacks.iop.org/1742-5468/2016/i=3/a=033301

[48] C. Külske, Concentration inequalities for functions of gibbs fields with application to diffraction and random gibbs measures, Communications in Mathematical Physics 239 (1) (2003) 29–51. doi:10.1007/s00220-003-0841-5. URL https://doi.org/10.1007/s00220-003-0841-5

[49] J. R. Chazottes, P. Collet, C. Külske, F. Redig, Concentration inequalities for random fields via coupling, Probability Theory and Related Fields 137 (1) (2007) 201–225. doi:10.1007/s00440-006-0026-1. URL https://doi.org/10.1007/s00440-006-0026-1

[50] F. Redig, J. R. Chazottes, Concentration inequalities for markov processes via coupling, Electron. J. Probab. 14 (2009) 1162–1180. doi:10.1214/EJP.v14-657. URL https://doi.org/10.1214/EJP.v14-657

[51] J.-R. Chazottes, P. Collet, F. Redig, On concentration inequalities and their applications for gibbs measures in lattice systems, Journal of Statistical Physics 169 (3) (2017) 504–546. doi:10.1007/s10955-017-1884-x. URL https://doi.org/10.1007/s10955-017-1884-x

[52] T. Squartini, J. de Mol, F. den Hollander, D. Garlaschelli, Breaking of ensemble equivalence in networks, Phys. Rev. Lett. 115 (2015) 268701. doi:10.1103/PhysRevLett.115.268701. URL https://link.aps.org/doi/10.1103/PhysRevLett.115.268701

[53] M. Hartmann, G. mahler, O. Hess, Gaussian quantum fluctuations in interacting many particle systems, Letters in Mathematical Physics 68 (2) 103–112. doi:10.1023/B:MATH.0000043321.00896.86. URL http://dx.doi.org/10.1023/B:MATH.0000043321.00896.86

[54] D. Goderis, P. Vets, Central limit theorem for mixing quantum systems and the ccr-algebra of fluctuations, Communications in Mathematical Physics 122 (2) 249–265. doi:10.1007/BF01257415. URL http://dx.doi.org/10.1007/BF01257415

[55] D. Goderis, A. Verbeure, P. Vets, Non-commutative central limits, Probability Theory and Related Fields 82 (4) 527–544. doi:10.1007/BF00341282. URL http://dx.doi.org/10.1007/BF00341282

[56] T. Matsui, Bosonic central limit theorem for the one-dimensional xy model, Reviews in Mathematical Physics 14 (07n08) (2002) 675–700.

[57] A. C. Berry, The accuracy of the gaussian approximation to the sum of independent variates, Transactions of the American Mathematical Society 49 (1) (1941) 122–136. URL http://www.jstor.org/stable/1990053
[58] C.-G. Esseen, Fourier analysis of distribution functions. a mathematical study of the laplace-gaussian law, Acta Mathematica 77 (1) 1–125. doi:10.1007/BF02392223.
URL http://dx.doi.org/10.1007/BF02392223

[59] H. Touchette, The large deviation approach to statistical mechanics, Physics Reports 478 (1) (2009) 1 – 69. doi:https://doi.org/10.1016/j.physrep.2009.05.002.
URL http://www.sciencedirect.com/science/article/pii/S0370157309001410

[60] Y. Ogata, Large deviations in quantum spin chains, Communications in Mathematical Physics 296 (1) (2010) 35–68. doi:10.1007/s00220-010-0986-y.
URL http://dx.doi.org/10.1007/s00220-010-0986-y

[61] F. Hiai, M. Mosonyi, T. Ogawa, Large deviations and chernoff bound for certain correlated states on a spin chain, Journal of Mathematical Physics 48 (12). doi:10.1063/1.2812417.
URL http://scitation.aip.org/content/aip/journal/jmp/48/12/10.1063/1.2812417

[62] M. Lenci, L. Rey-Bellet, Large deviations in quantum lattice systems: One-phase region, Journal of Statistical Physics 119 (3) 715–746. doi:10.1007/s10955-005-3015-3.
URL http://dx.doi.org/10.1007/s10955-005-3015-3

[63] K. Netočný, F. Redig, Large deviations for quantum spin systems, Journal of Statistical Physics 117 (3) (2004) 521–547. doi:10.1007/s10955-004-3452-4.
URL https://doi.org/10.1007/s10955-004-3452-4

[64] J. P. Keating, N. Linden, H. J. Wells, Spectra and eigenstates of spin chain hamiltonians, Communications in Mathematical Physics 338 (1) (2015) 81–102. doi:10.1007/s00220-015-2366-0.
URL https://doi.org/10.1007/s00220-015-2366-0

[65] T. Kuwahara, K. Saito, Ensemble equivalence and eigenstate thermalization from clustering of correlation, arXiv preprint arXiv:1905.01886.

[66] G. Biroli, C. Kollath, A. M. Läuchli, Effect of rare fluctuations on the thermalization of isolated quantum systems, Phys. Rev. Lett. 105 (2010) 250401. doi:10.1103/PhysRevLett.105.250401.
URL https://link.aps.org/doi/10.1103/PhysRevLett.105.250401

[67] E. Iyoda, K. Kaneko, T. Sagawa, Fluctuation theorem for many-body pure quantum states, Phys. Rev. Lett. 119 (2017) 100601. doi:10.1103/PhysRevLett.119.100601.
URL https://link.aps.org/doi/10.1103/PhysRevLett.119.100601

[68] V. Alba, Eigenstate thermalization hypothesis and integrability in quantum spin chains, Phys. Rev. B 91 (2015) 155123. doi:10.1103/PhysRevB.91.155123.
URL https://link.aps.org/doi/10.1103/PhysRevB.91.155123