Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

Tomotaka Kuwahara

WPI, Advanced Institute for Materials Research, Tohoku University, Sendai 980-8577, Japan
E-mail: tomotaka.kuwahara.e8@tohoku.ac.jp

Received 4 April 2016, revised 15 September 2016
Accepted for publication 15 October 2016
Published 16 November 2016
Online at stacks.iop.org/JSTAT/2016/113103
doi:10.1088/1742-5468/2016/11/113103

Abstract. This work aimed to explore the fundamental aspects of the spectral properties of few-body general operators. We first consider the following question: when we know the probability distributions of a set of observables, what do we know about the probability distribution of their summation? When considering arbitrary operators, we could not obtain useful information over the third-order moment, while under the assumption of $k$-locality, we can rigorously prove a much stronger bound on the moment generating function for arbitrary quantum states. Second, with the use of this bound, we generalize the Chernoff inequality (or the Hoeffding inequality), which characterizes the asymptotic decay of the probability distribution for the product states by Gaussian decay. In the present form, the Chernoff inequality can be applied to a summation of independent local observables (e.g. single-site operators). We extend the range of application of the Chernoff inequality to the generic few-body observables.

Keywords: Exact results, Rigorous results in statistical mechanics, Large deviation, Random/ordered microstructures
1. Introduction

In quantum mechanics, one of the most fundamental problems is the calculation of the probability distributions of observables, which is equivalent to determining the parameters that describe a quantum state. Usually, in condensed matter systems, the spins (or particles) are not independent of each other; hence such an analysis is generally quite a challenging problem [1] even for very simple quantum states [2]. This problem is deeply related to the entanglement structure of states [3, 4]. Here, we refer to the entanglement by quantum correlation, which results from the superposition principle; for example, a product state like $|0\rangle^\otimes N$ has no entanglement. Usually, we can treat only the class of slightly entangled states [5–7] because the number of parameters required to describe generic quantum states increases exponentially with the system size. This,
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

in turn, leads to the problem of characterizing slightly entangled states. In general, the complete analysis of the entanglement is extremely difficult for states with more than four qubits \[8\]; hence, we need to characterize it in a more coarse-grained manner.

For this purpose, we focus here on the fact that, in slightly entangled states, the spins (particles) are weakly dependent on each other; notice that, in the product states, there are no correlations among the spins, and each of the spins provides an independent probability distribution. In such cases, the asymptotic behaviors of the probability distribution can be determined by the Chernoff inequality \[9\] (figure 1), which is also called the Hoeffding inequality \[10\]. Let us consider an \(N\) spin system, a product state \(\rho_{\text{Prod}} = \bigotimes\rho\) with \(\|a\| = 1\) for \(i = 1, 2, \ldots, N\). The Chernoff inequality ensures that the probability distribution of \(A\) for \(\rho_{\text{Prod}}\) decays faster than the Gaussian decay with the variance \(O(N)\).

\[\text{Figure 1. Schematic of the Chernoff inequality. Let us consider a product state } \rho_{\text{Prod}} \text{ and an operator } A, \text{ which is given by the summation of single-site operators } \{a_i\}_{i=1}^N \text{ with } \|a_i\| = 1 \text{ for } i = 1, 2, \ldots, N. \text{ The Chernoff inequality ensures that the probability distribution of } A \text{ for } \rho_{\text{Prod}} \text{ decays faster than the Gaussian decay with the variance } O(N).\]

\(A = \sum_{i=1}^{N} a_i\)

\(\text{Distribution of } A\)

\(\mathcal{O}(\sqrt{N})\)

\(\mathcal{O}(\sqrt{N})\)

\(\text{Gaussian decay}\)

\(\langle A \rangle\)

\(\chi\)

\(\text{tr}(\rho_{\text{Prod}} A^X) \leq e^{-\frac{(x - \langle A \rangle)^2}{CN}}, \quad (1)\)

where \(\Pi^A_{\geq x}\) is the projector onto the eigenspace of \(A\) in which the eigenvalues are greater than \(x\), \(\langle A \rangle := \text{tr}(\rho_{\text{Prod}} A)\), and \(C\) is a constant. The Chernoff inequality is originally exploited to analyze the hypothesis testing problem \[9, 11, 12\]. Here, we consider exactly independent spins, i.e. the product states, but we expect that a similar inequality holds for weakly correlated states: for example, states with exponential clustering (the exponential decay of bi-partite correlations), ground states in non-critical regimes or with non-vanishing spectral gaps, and short-range entangled states. In these cases, the Chernoff inequality \(1\) is expected to hold, but it has not been proved yet.

A good place to start studying this question is in the short-range entanglement class \[13\], which is characterized by a constant-depth quantum circuit acting on a product state (figure 2). A constant-depth quantum circuit is a unitary operator that can be decomposed into a product of \(O(1)\) unitary operators; here, each of the unitary operators has the form \(U_{X_1} \otimes U_{X_2} \otimes \cdots \otimes U_{X_n}\), where \(\{X_i\}_{i=1}^n\) are non-overlapping supports. The short-range entangled states are often called trivial states as they do not show any non-local quantum behavior such as macroscopic superposition \[14\] and topological

\[\text{doi:10.1088/1742-5468/2016/11/113103}\]
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

Thus, we expect that the Chernoff inequality holds for this class of states. Now, the problem is equivalent to expressing the probability distribution of $U_{\text{loc}}^\dagger A U_{\text{loc}}$ with respect to a product state, where $U_{\text{loc}}$ is given by a constant-depth quantum circuit. Notice that the operator $U_{\text{loc}}^\dagger A U_{\text{loc}}$ is a few-body operator because $U_{\text{loc}}^\dagger a_i U_{\text{loc}}$ is still acting in the small region around site $i$. In the present paper, we aim to answer the following slightly general problem: for generic few-body operators, which may not be given by $U_{\text{loc}}^\dagger A U_{\text{loc}}$, can we prove the Chernoff inequality for arbitrary product states (see section 2 for the detailed definition of the few-body operators)?

Here, we compare the present analysis with previous works. The most relevant work recently appeared in [16] in which the Chernoff inequality has been proved in a different approach. There is a long history for the analysis of probability distributions in weakly or finitely correlated quantum many-body systems. One of the well-known examples is the central limit theorem, which states that the distribution of an observable converges to a Gaussian normal distribution in the thermodynamic limit, or for an infinite system size. In [17–20], the central limit theorem has been proved for several classes of weakly correlated quantum states. More recently, the related Berry–Essen theorem [21, 22] was proved for states under the assumption of exponential clustering [23]. However, in their analysis, the best convergence rate is of $O(1/\sqrt{N})$ [23], and a good estimation could not be provided for the asymptotic behavior of the probability distribution. The asymptotic behaviors of finite systems are often analyzed using the large deviation theorem [24], which has been also proven for special cases (e.g. finitely correlated spin chains [25] and 1D high-temperature Gibbs states [26, 27]). In another direction, for ground states with a non-vanishing spectral gap, a weaker version of the Chernoff bound $\text{tr}(\rho \Pi_{\omega^t}) \leq e^{-\frac{1}{2C} \lambda^2}$ was proved for systems with arbitrary dimensions [28, 29], where the constant $C$ is proportional to the inverse of the spectral gap. Thus far, even in simple setups, the Chernoff-type asymptotic decay (1) has not been proven for systems with higher or infinite dimensions.

In the analysis, the main difficulty lies in the fact that, when the observable $A$ is a generic few-body operator, the spectral analysis is usually non-trivial. To overcome this difficulty, we aim to answer the following questions: given the probability distributions of a set of observables $\{A_i\}_{i=1}^g$ for a quantum state $\rho$, what do we...
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

1D Heisenberg chain

![Diagram of 1D Heisenberg chain](image)

Figure 3. Decomposition of operators. Let an observable be given by the Hamiltonian $H$ of the 1D Heisenberg chain. Then, the operator is decomposed into two observables, namely $H_{\text{odd}}$ and $H_{\text{even}}$. Each of the operators can be given by the summation of the non-overlapping local operators; hence, the Chernoff inequality (1) holds for $H_{\text{odd}}$ and $H_{\text{even}}$, respectively. However, the analysis of $H_{\text{odd}} + H_{\text{even}}$ is usually non-trivial and it is not a simple question as to whether the Chernoff inequality also holds for $H$.

know about the distribution function of $\sum_{i=1}^{n} A_i / n$? We often encounter a situation in which some observables $\{A_i\}_{i=1}^{n}$ are easy to analyze, while their summation $\sum_{i=1}^{n} A_i / n$ is difficult to analyze. For example, let us consider the calculation of the probability distribution with respect to a product state $\otimes_{i=1}^{N} |0\rangle$ ($N$: system size), for the Heisenberg-type observable $H = \sum_{i=1}^{N} J (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + S_i^z S_{i+1}^z)$. Then as shown in figure 3, it is trivial to analyze each $H_{\text{even}} = \sum_{i=\text{even}}^{N} J (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + S_i^z S_{i+1}^z)$ and $H_{\text{odd}} = \sum_{i=\text{odd}}^{N} J (S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + S_i^z S_{i+1}^z)$ separately, while the analysis of $H$ is a non-trivial problem [30, 31]. In this case, the inequality (1) trivially holds for both $H_{\text{even}}$ and $H_{\text{odd}}$. Hence, if we can show that the asymptotic distribution behavior does not change because of the summation of $H_{\text{even}} + H_{\text{odd}}$, the probability distribution of $H$ is proved to follow the inequality (1). Indeed, we can generally connect the momentum-generating function of $\sum_{i=1}^{n} A_i / n$ to those of $\{A_i\}_{i=1}^{n}$, which plays a key role in the generalization of the Chernoff inequality.

The reminder of this paper is organized as follows. In section 2, we first show the fundamental setup of the system and definitions of the terms. We also show several preliminary lemmas for the analysis. In section 3, we first give the general statements to connect the probability distributions of different operators. As one application, we give an improved version of the main lemma in [32], which discusses the energy excitation by local perturbations. In section 4, we show the generalized version of the Chernoff inequality (1). In section 5, we prove all the theorems, lemmas and corollaries. Finally, section 6 concludes the paper.

2. Notation and general setup

We consider a spin system of finite volume with each spin having a $d$-dimensional Hilbert space, and we label each spin by $i = 1, 2, \ldots N$. We denote the set of all spins by $\Lambda = \{1, 2, \ldots N\}$, a partial set of sites by $X$, and the cardinality of $X$, that is, the size of this subset, by $|X|$ (e.g. $X = \{i_1, i_2, \ldots, i_{|X|}\}$).

Furthermore, for an arbitrary operator $A$, we define $\Pi_{x}^{A}$ ($\Pi_{<x}^{A}$) as the projection operator onto the eigenspace of $A$ in the range $\geq x$ ($< x$). Note that the probability
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

distribution of $A$ for a quantum state $\rho$ having values greater than $x$ is given by $\text{tr}(\rho \Pi^A_x)$. As quantum states, we consider an arbitrary $\rho$ in section 3, while in section 4 we restrict $\rho$ to the product states.

Throughout the paper, we consider few-body observables, which is characterized by ‘$q$-locality’ and ‘$g$-extensiveness’. We here refer to a operator with $q = O(1)$ and $g = O(1)$ as a few-body operator:

**Definition 1 ($q$-local).** An operator $O$ is $q$-local, with $q$ being a positive integer, if $O$ is given by the form

$$O = \sum_{|X| \leq q} o_X,$$

(2)

where each $\{o_X\}_{X \subset \Lambda}$ is an operator acting on a finite subset of spins $X$ having cardinality less than or equal to $q$. Here, the spins in subset $X$ may not exist next to each other in the lattice.

More explicitly, a $q$-local operator can be given in the form of

$$O = \sum_i h_i s_i + \sum_{i_1, i_2} J_{i_1 i_2} s_{i_1} s_{i_2} + \sum_{i_1, i_2, i_3} J_{i_1 i_2 i_3} s_{i_1} s_{i_2} s_{i_3} + \cdots + \sum_{i_1, i_2, \ldots, i_q} J_{i_1 i_2 \cdots i_q} s_{i_1} s_{i_2} \cdots s_{i_q},$$

where $\{s_i\}$ are operator bases on the $i$th spin; for example, it can be given by the Pauli matrices for $(1/2)$-spin systems, namely $\{s_i\} = \{\sigma^x_i, \sigma^y_i, \sigma^z_i\}$.

Second, we introduce $g$-extensiveness as the normalization of an operator:

**Definition 2 ($g$-extensive).** For a positive constant $g > 0$, an operator is $g$-extensive if

$$\sum_{X: X \ni i} \|o_X\| \leq g, \quad \text{for } \forall i \in \Lambda,$$

(3)

where we denote $O$ by $O = \sum_{X: X \ni i} o_X$ and $\|\cdots\|$ is the operator norm (i.e. the maximum singular value of operator) and $\sum_{X: X \ni i}$ denotes the summation with respect to the subsets $X$ containing the spin $i$.

This condition implies that a local norm of one spin is bounded by a finite constant $g$. A trivial algebra ensures that the norm $\|O\|$ does not exceed $gN$, where $N$ is the system size:

$$\|O\| = \left\| \sum_{X: X \subset \Lambda} o_X \right\| \leq \sum_{X: X \subset \Lambda} \|o_X\| \leq \sum_{i=1}^N \sum_{X: X \ni i} \|o_X\| \leq \sum_{i=1}^N g = gN,$$

(4)

where we note that the $q$-locality of $O$ is not assumed.

For an arbitrary quantum state $\rho$ and operator $A$, we define the moment-generating function $e^{M(\rho, A, \tau)}$ as follows:

$$M(\rho, A, \tau) := \log \left[ \text{tr}(e^{\tau A} \rho) \right].$$

(5)

Notice that $M(\rho, A, \tau)$ contains information on the asymptotic behavior of the probability distribution. For example, let us consider a case where $M(\rho, A, \tau)$ is bounded from above by
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ M(\rho, A, \tau) \leq c_1 \tau^2 + c_2 \]  

with \( \text{tr}(\rho A) = 0 \), where \( c_1 \) and \( c_2 \) are positive constants. We then have

\[ e^{x^2} \text{tr}(\Pi^d_{\rho} \rho) \leq \text{tr}(e^{x^2} \rho) \leq e^{x^2} c_1^2 + c_2 \text{ for } x \geq 0, \]

which yields

\[ \text{tr}(\Pi^d_{\rho} \rho) \leq \exp\left(-\frac{x^2}{4c_1} + c_2\right) \]  

by choosing \( \tau = x/(2c_1) \). The same inequality holds for \( \text{tr}(\Pi^d_{\rho} \rho) \).

2.1. Preliminaries

Here, we show three basic lemmas. First, for the analysis of generic few-body observables, we often need to treat multi-commutators. We give two lemmas as useful technical tools in the analysis. Second, we formulate the Chernoff inequality for operators that are given by the summation of independent local observables.

For the norm of multi-commutators, we can prove the following lemma (see lemma 3 in [33]):

**Lemma 1.** Let \( \{A_i\}_{i=1}^{n} \) be \( k \)-local and \( g \)-extensive, respectively, and \( O_X \) be an arbitrary operator supported in a subset \( X \). Then, the norm of the multi-commutator \([A_n, [A_{n-1}, \ldots, [A_1, O_X], \ldots]]\) is bounded from above by

\[ ||[A_n, [A_{n-1}, \ldots, [A_1, O_X], \ldots]]|| \leq \prod_{m=1}^{n} (2g_mK_m)||O_X||, \]

where \( K_m := |X| + \sum_{i \leq m-1} k_i \).

Then, we consider the decomposition of a few-body operator \( A \) into a summation of operators \( \{A_i\}_{i=1}^{n} \), each of which is easy to analyze. For example, when considering a spatially local Hamiltonian on a lattice with finite dimensions, a decomposition similar to that shown in figure 3 is always possible by taking \( \bar{n} = O(D) \), where \( D \) is the number of dimensions of the system. On the other hand, for generic few-body observables \( A \) (e.g., a local operator on infinite-dimensional graph), the existence of such a decomposition is non-trivial. Actually, we can also ensure the existence of the following decomposition for general \( k \)-local and \( g \)-extensive operators [34] as follows:

**Lemma 2.** Let \( A = \sum a_X \) be an arbitrary \( k \)-local and \( g \)-extensive operator and \( N \) be a fixed system size. We consider a decomposition of \( A \) into a summation of \( \bar{n} \) operators \( \{A_j\}_{j=1}^{\bar{n}} \) such that (figure 4)

\[ A_j = \tilde{a}_{X_j}^{(1)} + \tilde{a}_{X_j}^{(2)} + \cdots + \tilde{a}_{X_j}^{(\bar{n})}, \]

s.t. \( ||\tilde{a}_{X_j}^{(m)}|| \leq gk \) and \( X_j^{(m)} \cap X_j^{(m')} = \emptyset \) for \( \forall m, m' \in \{1, 2, \ldots, N_j\} \),

where \( \bar{n} \) is an arbitrary integer which is independent of the system size \( N \), each of \( \{\tilde{a}_{X_j}^{(1)}, \tilde{a}_{X_j}^{(2)}, \ldots, \tilde{a}_{X_j}^{(\bar{n})}\} \) is proportional to one component of \( \{a_X\}_X \) in \( A \) and \( N_j \leq N/k \) for...
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ A^c = a_X^{(1)} + a_X^{(2)} + a_X^{(3)} + \cdots \]

Figure 4. Decomposed operators in lemma 2. The lemma states that an arbitrary few-body operator can be decomposed into a set of simple operators \( \{ A_j^{(s)} \}_{j=1}^n \), each of which is given by the summation of non-overlapping local operators. The term ‘non-overlapping’ means that the local components in \( A_j^{(s)} \), namely \( \{ a_X \}_{m=1}^n \) in equation (9), are supported in subsets \( \{ X^{(s)}_m \}_{m=1}^N \), which do not overlap with each other.

\[ j = 1, 2, \ldots, n. \] Note that \( \{ A_j^{(s)} \}_{j=1}^n \) are now \( k \)-local and \((gk)\)-extensive. Then, there always exists a decomposition of \( A = \frac{1}{\bar{n}} \sum_{j=1}^n A_j^c \) with the error

\[ \left\| A - \frac{1}{\bar{n}} \sum_{j=1}^n A_j^c \right\| \leq C_0 \frac{1}{\bar{n}}, \tag{10} \]

where \( C_0 \) is a constant which only depends on \( k, g \) and \( N \). With the integer \( \bar{n} \) increasing, the precision becomes better as \( 1/\bar{n} \); that is, if we choose \( \bar{n} \) infinitely large, the operator \( A \) can be exactly decomposed into a summation of operators of the form (9). Such a decomposition has been explicitly given in appendix A of [34].

Lemma 2 says that if we take \( \bar{n} \) large enough, we can decompose \( k \)-local operator into the operators in the form of (9). As a concrete example, let us consider an operator \( A_{1D} \) on a one-dimensional lattice with nearest-neighbor couplings:

\[ A_{1D} = \sum_{i=1}^{N-1} a_{i,i+1}, \tag{11} \]

where \( \{ a_{i,i+1} \}_{i=1}^{N-1} \) are arbitrary operators defined on the sites \( \{(i, i+1)\}_{i=1}^{N-1} \), respectively, and satisfy \( \| a_{i,i+1} \| \leq g \) with \( g \) being a constant. Then, this can be decomposed as \( A_{1D} = (A_1^c + A_2^c)/2 \) with

\[ A_1^c = \sum_{i=\text{even}}^{N-1} 2a_{i,i+1}, \quad A_2^c = \sum_{i=\text{odd}}^{N-1} 2a_{i,i+1}. \tag{12} \]

The operators \( A_1^c \) and \( A_2^c \) are now 2-local (\( k = 2 \)) and \((2g)\)-extensive, respectively.

We formalize the quantum Chernoff inequality as the upper bound of the moment generating function \( M(A, \rho, \tau) \):

**Lemma 3 (The Chernoff inequality for product states).** Let \( A \) be a summation of non-overlapping local operators:

\[ A = \sum_{i=1}^{N_i} a_X, \quad \text{s.t.} \quad \| a_X \| \leq 1, \quad \| X \| \leq k \quad \text{and} \quad X_i \cap X_j = \emptyset \quad \forall i, j \in \{1, 2, \ldots, N_i \}, \tag{13} \]

\[ \text{doi:10.1088/1742-5468/2016/11/113103} \]
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

and \( \rho_{\text{Prod}} \) be a product state in the form of \( \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_N \). For simplicity, we set \( \text{tr}(\rho A) = 0 \). Then, the Chernoff inequality gives the Gaussian decay of the probability distribution of \( \rho_{\text{Prod}} \) with respect to \( A \), or equivalently,

\[
M(A, \rho_{\text{Prod}}, \tau) \leq \frac{4N}{k} \tau^2. \tag{14}
\]

**Proof.** The lemma is simply proved as follows:

\[
\text{tr}(e^{\sigma A} \rho_{\text{Prod}}) = \text{tr}\left(e^{\sum_{i=1}^{N_1} (a_i - \langle a_i \rangle) \rho_{\text{Prod}}} \right) = \prod_{i=1}^{N_1} \text{tr}(e^{(a_i - \langle a_i \rangle)^m \rho_{\text{Prod}}}) \tag{15}
\]

where \( \langle a_i \rangle := \text{tr}(a_i \rho_{\text{Prod}}) \) and \( \sum_{i=1}^{N_1} \langle a_i \rangle = 0 \). We have

\[
\text{tr}(e^{(a_i - \langle a_i \rangle)^m \rho_{\text{Prod}}}) = 1 + \sum_{m \geq 2} \frac{\tau^m}{m!} \text{tr}[(a_i - \langle a_i \rangle)^m \rho_{\text{Prod}}] \leq 1 + \sum_{m \geq 2} \frac{\tau^m}{m^2} 2^m = e^{2\tau} - 2\tau \leq e^{(2\tau)^2}, \tag{16}
\]

where we use \( \|a_i - \langle a_i \rangle\| \leq 2\|a_i\| \leq 2 \). By combining above two inequalities, we have

\[
\text{tr}(e^{\sigma A} \rho_{\text{Prod}}) \leq e^{(2\tau)^2 N_1} \leq e^{\tau^2 N/k}, \tag{17}
\]

where \( N_1 \leq N/k \). By taking the logarithm of this inequality, we reach the inequality (14).

From the lemma, we immediately obtain

\[
\text{tr}(\Pi_{\geq 2} \rho_{\text{Prod}}) \leq \exp\left(-\frac{k\sigma^2}{16N}\right) \tag{18}
\]

by following the same discussion as in the derivation of (7).

### 3. Upper bound on the moment-generating function

In this section, we consider the following problem. Let \( \rho \) be an arbitrary quantum state and \{ \( A_j \) \}_{j=1}^{\tilde{n}} \) be a set of arbitrary operators. We also assume that we know the probability distribution of each of \{ \( A_j \) \}_{j=1}^{\tilde{n}}. Here, we consider the relationship of the probability distributions of \{ \( A_j \) \}_{j=1}^{\tilde{n}} with that of their summation \( \sum_{j=1}^{\tilde{n}} A_j \).

We begin with the most general statement without any assumption on the operator \{ \( A_j \) \}_{j=1}^{\tilde{n}}. For this problem, we can trivially obtain information on the average and the standard variance. For simplicity, we discuss the case of \( \tilde{n} = 2 \) and consider two observables, \( A_1 \) and \( A_2 \). We denote the average and variance by \( \mu \) and \( \sigma \), respectively, and set the averages to zero (\( \mu_1 = \mu_2 = 0 \)). Then, for the summation \( A_1 + A_2 \), we have

doi:10.1088/1742-5468/2016/11/113103
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ \text{tr}[\rho(A_1 + A_2)] = \text{tr}(\rho A_1) + \text{tr}(\rho A_2) = \mu_1 + \mu_2 = 0 \]  \hspace{1cm} (19)

and

\[
\sqrt{\text{tr}[\rho(A_1 + A_2)^2]} = \sqrt{\text{tr}(\rho A_1^2) + \text{tr}(\rho A_2^2) + \text{tr}(\rho A_2 A_1) + \text{tr}(\rho A_1 A_2)}
\leq \sqrt{\left(\text{tr}(\rho A_1^2) + \text{tr}(\rho A_2^2)\right)^2} = \sigma_1 + \sigma_2, \]  \hspace{1cm} (20)

where we apply the Schwartz inequality \(\text{tr}(\rho A_2 A_1) \leq \sqrt{\text{tr}(\rho A_1^2) \text{tr}(\rho A_2^2)}\). By using the Chebyshev inequality, we can ensure that the probability distribution for \(A_1 + A_2\) decays faster than

\[ \text{tr}(\rho \Pi^{A_1 + A_2}_{\geq x}) \leq \left(\frac{\sigma_1 + \sigma_2}{|x|}\right)^2, \]  \hspace{1cm} (21)

where \(\Pi^{A_1 + A_2}_{\geq x}\) is the projection operator onto the eigenspace of \(A_1 + A_2\), which is in \([x, \infty)\).

From the bound (21), we cannot know information on moments above the third order, such as \(\text{tr}[\rho(A_1 + A_2)^k]\) with \(k \geq 3\). We aim to obtain a much better bound by considering additional assumptions on the observables \(A_1\) and \(A_2\). More specifically, we restrict the observables to be \(k\)-local and \(g\)-extensive. This condition is one of the most natural assumptions when we analyze realistic quantum many-body systems.

Nevertheless, recent studies have shown that this ‘few-body’ assumption gives us fruitful information on the fundamental properties. For arbitrary sets of few-body observables \(\{A_i\}_{i=1}^n\), we obtain strong upper bounds on the moment-generating function. We start from the case of \(n = 2\):

**Theorem 3.** Let \(A\) and \(B\) be \(k\)-local \(g\)-extensive operators, respectively. Then, for an arbitrary quantum state \(\rho\) and \(|\tau| < 1/(4\lambda)\), we have

\[ M(A + B, \rho, \tau) \leq \frac{M(2A, \rho, \tau) + M(2B, \rho, \tau)}{2} + \frac{3\tau^2}{2} \|[A, B]\|, \]  \hspace{1cm} (22)

where \(\lambda\) is defined by

\[ \lambda := 2gk. \]  \hspace{1cm} (23)

By using lemma 1 and the \(g\)-extensiveness of operators, the commutator norm \(\|[A, B]\|\) is bounded from above by \(\|[A, B]\| \leq \lambda g N\). Hence, we can generally obtain the inequality

\[ M(A + B, \rho, \tau) \leq \frac{M(2A, \rho, \tau) + M(2B, \rho, \tau)}{2} + \frac{3\lambda g N \tau^2}{2}. \]  \hspace{1cm} (24)

Here, we would like to emphasize that we need not make any assumption on the quantum state \(\rho\); for example, we need not assume that it is a product state or pure state.

We can easily extend the results to general summations:

\[ \text{tr}(\rho(A_1 + A_2)) = \text{tr}(\rho A_1) + \text{tr}(\rho A_2) = \mu_1 + \mu_2 = 0 \]  \hspace{1cm} (19)

and

\[
\sqrt{\text{tr}[\rho(A_1 + A_2)^2]} = \sqrt{\text{tr}(\rho A_1^2) + \text{tr}(\rho A_2^2) + \text{tr}(\rho A_2 A_1) + \text{tr}(\rho A_1 A_2)}
\leq \sqrt{\left(\text{tr}(\rho A_1^2) + \text{tr}(\rho A_2^2)\right)^2} = \sigma_1 + \sigma_2, \]  \hspace{1cm} (20)

where we apply the Schwartz inequality \(\text{tr}(\rho A_2 A_1) \leq \sqrt{\text{tr}(\rho A_1^2) \text{tr}(\rho A_2^2)}\). By using the Chebyshev inequality, we can ensure that the probability distribution for \(A_1 + A_2\) decays faster than

\[ \text{tr}(\rho \Pi^{A_1 + A_2}_{\geq x}) \leq \left(\frac{\sigma_1 + \sigma_2}{|x|}\right)^2, \]  \hspace{1cm} (21)

where \(\Pi^{A_1 + A_2}_{\geq x}\) is the projection operator onto the eigenspace of \(A_1 + A_2\), which is in \([x, \infty)\).

From the bound (21), we cannot know information on moments above the third order, such as \(\text{tr}[\rho(A_1 + A_2)^k]\) with \(k \geq 3\). We aim to obtain a much better bound by considering additional assumptions on the observables \(A_1\) and \(A_2\). More specifically, we restrict the observables to be \(k\)-local and \(g\)-extensive. This condition is one of the most natural assumptions when we analyze realistic quantum many-body systems.

Nevertheless, recent studies have shown that this ‘few-body’ assumption gives us fruitful information on the fundamental properties. For arbitrary sets of few-body observables \(\{A_i\}_{i=1}^n\), we obtain strong upper bounds on the moment-generating function. We start from the case of \(n = 2\):

**Theorem 3.** Let \(A\) and \(B\) be \(k\)-local \(g\)-extensive operators, respectively. Then, for an arbitrary quantum state \(\rho\) and \(|\tau| < 1/(4\lambda)\), we have

\[ M(A + B, \rho, \tau) \leq \frac{M(2A, \rho, \tau) + M(2B, \rho, \tau)}{2} + \frac{3\tau^2}{2} \|[A, B]\|, \]  \hspace{1cm} (22)

where \(\lambda\) is defined by

\[ \lambda := 2gk. \]  \hspace{1cm} (23)

By using lemma 1 and the \(g\)-extensiveness of operators, the commutator norm \(\|[A, B]\|\) is bounded from above by \(\|[A, B]\| \leq \lambda g N\). Hence, we can generally obtain the inequality

\[ M(A + B, \rho, \tau) \leq \frac{M(2A, \rho, \tau) + M(2B, \rho, \tau)}{2} + \frac{3\lambda g N \tau^2}{2}. \]  \hspace{1cm} (24)

Here, we would like to emphasize that we need not make any assumption on the quantum state \(\rho\); for example, we need not assume that it is a product state or pure state.

We can easily extend the results to general summations:
Corollary 4. Let \( \{A_j\}^{n}_{j=1} \) be arbitrary k-local g-extensive operators and \( A \) be defined by their summation:
\[
A = \frac{1}{\bar{n}} \sum_{j=1}^{n} A_j.
\]
(25)

Then, for an arbitrary quantum state \( \rho \) and \( |\tau| < 1/(4\lambda) \), we have
\[
M(A, \rho, \tau) \leq \frac{1}{2m_0} \sum_{j=1}^{\bar{n}} M\left(\frac{2m_0}{\bar{n}} A_j, \rho, \tau\right) + \frac{3\lambda g N\tau^2}{8} m_0,
\]
(26)
where \( m_0 = \lfloor \log_2 \bar{n} \rfloor \).

The error term is proportional to \( N\tau^2 \log \bar{n} \); hence as long as we consider a summation of a finite number of operators (\( \bar{n} = \mathcal{O}(1) \)), the term is not influential. This \( \log \bar{n} \) dependence becomes dominant when considering the Chernoff inequality for general few-body operators (see section 4.2), where \( \bar{n} \) is at least as large as \( \text{Poly}(N) \). This theorem plays central roles in deriving the Chernoff inequality (1) for general few-body operators.

We further show that a much stronger inequality holds if we add the assumption that the distributions are completely localized for \( \rho \):

Theorem 5. Let \( A \) be defined by
\[
A = \frac{1}{\bar{n}} \sum_{j=1}^{n} A^c_j,
\]
(27)
where each \( \{A^c_j\}^{n}_{j=1} \) is k-local, g-extensive, and given by a summation of non-overlapping local operators (see figure 4). We also assume that, for a quantum state \( \rho \), the distribution of \( \{A^c_j\}^{n}_{j=1} \) is exactly localized with a width \( 2\sigma \) and \( \text{tr}(\rho A_j) = 0 \) for \( j = 1, 2, \ldots, n \); that is, we assume
\[
\rho \Pi_{>\sigma}^A = \rho \Pi_{\leq \sigma}^A = 0
\]
(28)
for \( j = 1, 2, \ldots, n \). Then, the moment-generating function \( M(A, \rho, \tau) \) is upperbounded by
\[
M(A, \rho, \tau) \leq -\frac{\sigma}{\lambda} \log(1 - \lambda |\tau|),
\]
(29)
where \( \lambda \) is defined in equation (23).

This implies that the probability distribution of \( A \) is also localized in a finite range with exponentially decaying errors; this is ensured by
\[
\text{tr}(\Pi^A_{\geq x} \rho) \leq e^{-\tau x} \text{tr}(e^{\sigma A} \rho) \leq e^{-\tau x} (1 - \lambda |\tau|)^{-\sigma/\lambda},
\]
(30)
which yields
\[
\text{tr}(\Pi^A_{\geq x} \rho) \leq \left(\frac{x}{\sigma}\right)^{\sigma/\lambda} e^{-(x-\sigma)/\lambda} \text{ for } x \geq \sigma,
\]
(31)
where we choose \( \tau = (x - \sigma)/(|\lambda x|) \). We can obtain the same inequality for \( \text{tr}(\Pi^H_{\leq -x} \rho) \).
As one application of theorem 5, we show the improvement of the following inequality [32]: let \( H \) be a Hamiltonian that is \( k \)-local and \( g \)-extensive and \( A \) be an arbitrary \( q \)-local operator. Then, the energy excitation due to the operator \( A \) is exponentially suppressed as

\[
\| \Pi_{\geq x}^H A | \Omega \rangle \| \leq \| A \| \exp \left( -\frac{x}{5k\lambda} + q \right),
\]

where \( | \Omega \rangle \) is the ground state of the Hamiltonian \( H \) as \( H | \Omega \rangle = 0 \). This kind of inequality has been first introduced to analyze an effective Hamiltonian that governs low-energy regimes [32]. It plays central roles in connecting the spectral gap and the fundamental properties of ground states [28, 35, 36].

On the other hand, owing to the coefficient \( \| A \| \) in the right-hand side, an arbitrary large energy can be still locally excited with very low-probability. To demonstrate this point, we consider the case where the ground state \( | \Omega \rangle \) is decomposed as follows:

\[
| \Omega \rangle = \lambda_0 | 0 \rangle \otimes | \phi_0 \rangle + \lambda_1 | 1 \rangle \otimes | \phi_1 \rangle,
\]

where \( | \lambda_0 |^2 + | \lambda_1 |^2 = 1 \). We now choose \( A = | 0 \rangle \langle 0 | / \lambda_0 \) or \( A = | 1 \rangle \langle 1 | / \lambda_1 \), and then the inequality (32) reads

\[
\| \Pi_{\geq x}^H | 0 \rangle \otimes | \phi_0 \rangle \| \leq \frac{1}{\lambda_0} \exp \left( -\frac{x}{5k\lambda} + 1 \right), \quad \| \Pi_{\geq x}^H | 1 \rangle \otimes | \phi_1 \rangle \| \leq \frac{1}{\lambda_1} \exp \left( -\frac{x}{5k\lambda} + 1 \right),
\]

where \( A \) is 1-local \((q = 1)\) and \( \| A \| = 1/\lambda_0 \) \((1/\lambda_1)\). Thus, if \( \lambda_0 \) or \( \lambda_1 \) is as small as \( e^{-\alpha N} \), the macroscopic energy of \( O(N) \) can be locally excited with the probability \( e^{-\alpha N} \). This seems rather strange, and hence, we expect that the inequality can be improved in the following manner:

\[
\| \Pi_{\geq x}^H A | \Omega \rangle \| \leq \| A | \Omega \rangle \| \exp \left( -\frac{x}{5k\lambda} + q \right),
\]

which modifies the inequality (34) as\( \| \Pi_{\geq x}^H | 0 \rangle \otimes | \phi_0 \rangle \| \leq e^{-\frac{x}{5k\lambda} + 1} \) and\( \| \Pi_{\geq x}^H | 1 \rangle \otimes | \phi_1 \rangle \| \leq e^{-\frac{x}{5k\lambda} + 1} \), where the high-energy excitation is impossible even probabilistically.

We can indeed improve the inequality (32) in the case where the Hamiltonian is frustration free. A frustration-free Hamiltonian satisfies the following property: the Hamiltonian can be expressed as a sum of terms such that the lowest-energy states of the full Hamiltonian are also the lowest energy states of each individual term. In other words, the global ground states are also local ground states. Then, we can prove the following corollary:

**Corollary 6.** Let \( H \) be a frustration-free \( k \)-local Hamiltonian with \( g \)-extensiveness, i.e. \( h_X | \Omega \rangle = 0 \) for \( \forall h_X \) with \( H = \sum_{|X| \leq k} h_X \). We also denote the ground state of \( H \) by \( | \Omega \rangle \) \((H | \Omega \rangle = 0)\). Then, for any \( q \)-local operator \( A \), we have

\[
\| \Pi_{\geq x}^H A | \Omega \rangle \| \leq \| A | \Omega \rangle \| \left( \frac{x}{\lambda q} \right)^{q/2k} \exp \left( -\frac{x}{2k\lambda} + \frac{q}{2k} \right) \quad \text{for} \quad x \geq \lambda q.
\]
4. The Chernoff inequality

In this section, we extend the Chernoff inequality (14) to more general cases compared to the standard ones (13), which is restricted to the summation of non-overlapping local operators. Here, we consider the probability distribution of generic few-body observables for a product state $\rho_{\text{Prod}}$. We summarize the present results in figure 5.

4.1. Finite dimensional systems

When considering finite-dimensional systems, we have to define the structure of the system explicitly (e.g. a square lattice). For simplicity, we restrict ourselves here to $D$-dimensional regular lattices. In this lattice system, we define the distance $\text{dist}(X, Y)$ as the length of the shortest path connecting the two partial sets $X$ and $Y$. When we consider ‘spatially local operators’, we introduce the following additional assumption to the operator:

$$A = \sum_{|X| \leq k} a_X, \quad \text{s.t.} \quad \sum_{X : X \ni \text{diam}(X) \geq r} \|a_X\| = 0 \quad \text{for} \quad \forall i \in \Lambda,$$

where $r$ is an $O(1)$ constant and $\text{diam}(X) := \sup_{\{i,j\} \in X} \text{dist}(i, j)$.

First, we consider the case where the operator $A$ is given by operators that satisfy (37). This assumption ensures that we can always decompose the operator $A$ into

$$A = \frac{1}{\bar{n}} \sum_{j=1}^{\bar{n}} A_j^c$$

where each $\{A_j^c\}_{j=1}^{\bar{n}}$ is given by the summation of non-overlapping local operators (see equations (11) with (12) for example). Notice that the Chernoff inequality (14) holds for each of $\{A_j^c\}_{j=1}^{\bar{n}}$.

Here, from corollary 4, we give the following generalization of the Chernoff inequality in lemma 3:

**Lemma 4 (Chernoff inequality for finite dimensional systems).** Let $\rho_{\text{Prod}}$ be a product state $\rho_{\text{Prod}}$ and $A$ be an operator given by equation (38), where each of $\{A_j^c\}_{j=1}^{\bar{n}}$ is $k$-local and $g$-extensive. We also set $\text{tr}(\rho_{\text{Prod}}A_j^c) = 0$ for $j = 1, 2, \ldots, \bar{n}$. Then, the upper bound of the moment-generating function for $\rho_{\text{Prod}}$ is given by

$$M(A, \rho_{\text{Prod}}, \tau) \leq \tilde{C} N \tau^2$$

for $|\tau| \leq 1/(4 \lambda)$, where

$$\tilde{C} := \frac{8g^2}{k} \left[ \frac{3\lambda g}{2} \left\lfloor \log_2 \bar{n} \right\rfloor \right].$$

Now, the parameter $\tilde{C}$ increases logarithmically with the number of dimensions of the system $D$. 

doi:10.1088/1742-5468/2016/11/113103
We can immediately apply this lemma to the probability distribution. We have an inequality

\[ \text{tr}(\Pi^A_x \rho_{\text{Prod}}) \leq e^{-\tau x + C N \tau^2} \]  

for \( x \geq 0 \) and \( \tau \leq 1/(4 \lambda) \). By choosing the parameter \( \tau \) appropriately, the inequality (41) reduces to

\[ \text{tr}(\Pi^A_x \rho_{\text{Prod}}) \leq \max\left(e^{-x^2/4CN}, e^{-x/8}\right). \]  

The same inequality holds for \( \text{tr}(\Pi^A_{-x} \rho_{\text{Prod}}) \).

4.2. Infinite dimensional systems (generic few-body operators)

In the case where the operator \( A \) is given by a general \( k \)-local and \( g \)-extensive operator, the number of decomposed operators, namely \( \tilde{n} \), can be arbitrarily large as in lemma 2. Thus, we can no longer expect that \( \log_2 \tilde{n} \) is an \( \mathcal{O}(1) \) constant; \( \tilde{C}(\tau) \) in (40) is infinitely large in the limit of \( \tilde{n} \to \infty \). We can still prove the following weaker statement for an arbitrarily large \( \tilde{n} \).

**Theorem 7.** Let \( A \) be given in the form of equation (38) with each of \( \{A^i_j\}_{j=1}^\tilde{n} \) being \( k \)-local and \( g \)-extensive, where the number \( \tilde{n} \) may be infinitely large. Then, the probability distribution of \( A \) for \( \rho_{\text{Prod}} \) is bounded from above by

\[ \mathcal{O}(\sqrt{N}) \]  

for a product state \( \rho_{\text{Prod}} \) decays faster than the Gaussian decay, but with the variance of \( \mathcal{O}(N \log D) \) (lemma 4). On the other hand, for generic few-body operators or local operators on infinite-dimensional lattices, we obtain slightly weaker decays compared to the strict Gaussian, \( e^{-\tilde{C}(\tau)(N \log(D)/\sqrt{N})} \) (corollary 8). (a) Operators on finite-dimensional lattices. (b) Generic few-body operators.
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[
\text{tr}(\Pi^4_x \rho_{\text{Prod}}) \leq c_1 \exp \left( -c_0 \frac{x^2}{N \log(x/\sqrt{N})} \right)
\]

(43)

with \( x > 0 \), where we set \( \text{tr}(A_j^c \rho_{\text{Prod}}) = 0 \) for \( j = 1, 2, \ldots, \tilde{n} \) and \( \{ c_0, c_1 \} \) are positive constants of \( \mathcal{O}(1) \) that only depend on \( k \) and \( g \). The same upper bound is also given for \( \text{tr}(\Pi^4_x \rho_{\text{Prod}}) \).

This inequality implies that the probability distribution asymptotically decays as \( e^{-\mathcal{O}(N/\log N)} \) for \( x = \mathcal{O}(N) \). This decay is weaker than (42) for \( \tilde{n} = \mathcal{O}(1) \), whereas even for \( \tilde{n} \to \infty \), the inequality (43) gives us a meaningful bound. From lemma 2, we immediately generalize the Chernoff inequality to generic few-body observables.

**Corollary 8 (Chernoff bound for generic few-body operators).** For an arbitrary \( k \)-local and \( g \)-extensive operator \( A \) and a product state \( \rho_{\text{Prod}} \) with \( \text{tr}(\rho_{\text{Prod}} A) = 0 \), we have the upper bound of

\[
\text{tr}(\Pi^4_x \rho_{\text{Prod}}) \leq c'_1 \exp \left( -c'_0 \frac{x^2}{N \log(x/\sqrt{N})} \right)
\]

(44)

for \( x > 0 \), where \( c'_0 \) and \( c'_1 \) are positive constants of \( \mathcal{O}(1) \) that only depend on \( k \) and \( g \). We have the same inequality for \( \text{tr}(\Pi^4_x \rho_{\text{Prod}}) \).

**4.3. For quasi-local operators**

We often encounter the case where the operator is not given by a strict \( k \)-local operator but by a *quasi-local* operator. For example, we consider a 1-local operator \( \sum a_i \) which evolves with a spatially local Hamiltonian \( H \). Then, because of the Lieb–Robinson bound [37, 38], the time-evolved operator \( A(t) = \sum a_i e^{-iHt} \) is given by the form of

\[
a_i(t) = \sum X a_X,
\]

(45)

where \( \sum_{\text{diam}(X) \geq 1} \| a_X \| \leq \text{const} e^{-(t-v)\xi} \) with \( \xi \) and \( v \) \( \mathcal{O}(1) \) constants which depend only on the details of the Hamiltonian; that is, the non-local couplings between spins decay exponentially with respect to the spatial distance outside the light cone which grows linearly with the time. The operator \( A(t) \) is approximately \( \mathcal{O}(t^2) \)-local (\( D \): dimension of the system), but contains terms up to \( N \)-body coupling. For this quasi-local operators, we can also prove the Chernoff–Hoeffding inequality. For simplicity, we only consider small-dimensional systems, namely \( D = \mathcal{O}(1) \).

To this aim, we first decompose the system into a set of ‘supersites’ with multiple sites (figure 6); each of the supersites has a length of \( \mathcal{O}(t) \) (i.e. the width of the light cone) and consists of \( \mathcal{O}(t^D) \) sites. The number of supersite is \( \tilde{N} = \mathcal{O}(N/t^D) \). In the description of the supersite, the operator \( A(t) \) is given by

\[
doi:10.1088/1742-5468/2016/11/113103
\]

15
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\begin{equation}
A(t) = \sum_{l \geq 0} A_l,
\end{equation}

where \( A_l \) contains couplings between supersites which are distance \( l \) apart from each other.

Owing to the Lieb–Robinson bound, each of \( \{ A_l \} \) is \( \mathcal{O}(t^D) \)-local and \( g(l) \)-extensive with

\begin{equation}
g(l) = g_0 t^D e^{-\text{const} \cdot l},
\end{equation}

where \( g_0 \) is an \( \mathcal{O}(1) \) constant which does not depend on the system size \( N \) and the time \( t \). Note that because one supersite contains \( \mathcal{O}(t^D) \) sites, a local norm of one supersite is proportional to \( t^D \). Then, our task is to obtain the moment generating function \( M(A(t), \rho, \tau) \) for operator (46) with (47). It is given in the following theorem:

**Theorem 9.** Let \( A(t) \) be given in the form of equations (46) with (47). Then, the moment-generating function for a product state \( M(A(t), \rho_{\text{Prod}}, \tau) \) is bounded from above by

\begin{equation}
M(A(t), \rho_{\text{Prod}}, \tau) \leq \text{const} \cdot \tau^2 g_0^2 N t^D
\end{equation}

for \( \tau \leq \text{const} \cdot 1/(g_0 t^D) \), where the coefficient depends only on the system dimension \( D \) and details of the system parameters.

As in the derivation of the inequality (42), the probability distribution for \( A(t) \) can be also bounded from above by

\begin{equation}
\text{tr}(\Pi_x(t) \rho_{\text{Prod}}) \leq \exp \left\{ -\frac{x^2}{\text{const} \cdot g_0^2 t^D N} \right\}.
\end{equation}

This implies that as long as \( t \) is small the Chernoff–Hoeffding inequality holds for the time-evolved operator \( A(t) \).
5. Proof of results

5.1. Proof of theorem 3

We begin with the following inequality:

\[ \text{tr}(e^{(A+B)}\rho) = \sum_{j=1}^{n} p_j \langle \psi_j | e^{(A+B)} | \psi_j \rangle = \sum_{j=1}^{n} p_j \langle \psi_j | e^{\sum_{A} A} e^{-\tau A} e^{\sum_{B} B} | \psi_j \rangle \]

\[ \leq \|e^{\sum_{B} B} e^{\sum_{A} A} e^{-\tau A}\| \sum_{j=1}^{n} p_j \|e^{\tau A} | \psi_j \rangle|| \|e^{\sum_{B} B} | \psi_j \rangle|| \]

\[ = \|e^{\sum_{B} B} e^{\sum_{A} A} e^{-\tau A}\| \sum_{j=1}^{n} \sqrt{p_j \langle \psi_j | e^{2\tau A} | \psi_j \rangle} \sqrt{p_j \langle \psi_j | e^{2\tau B} | \psi_j \rangle} \]

\[ \leq \|e^{\sum_{B} B} e^{\sum_{A} A} e^{-\tau A}\| \sqrt{\text{tr}(e^{2\tau A} \rho)} \sqrt{\text{tr}(e^{2\tau B} \rho)}, \] (50)

where we define \( \rho = \sum_j p_j | \psi_j \rangle \langle \psi_j | \) and used the Schwartz inequality in the second inequality. Then, our task is to estimate the norm of \( \|e^{\sum_{B} B} e^{\sum_{A} A} e^{-\tau A}\| \). Here, the conditions for the operators, namely \( k \)-locality and \( g \)-extensiveness, play central roles in the estimation.

For this purpose, we define \( e^{-xB} e^{\sum_{A} A} e^{-xA} =: U(x) \) and obtain

\[ \frac{dU(x)}{dx} = -BU - UA + e^{-xB} e^{\sum_{A} A} (A+B) e^{-xA} \]

\[ = -BU + e^{-xB} e^{\sum_{A} A} Be^{-xA} e^{xB} U = [(\hat{B}(x) - B)] U, \] (51)

where \( \hat{B}(x) := e^{-xB} e^{\sum_{A} A} Be^{-xA} e^{xB} \). We thus obtain

\[ U(\tau) = T \left[ e^\int_0^\tau [(\hat{B}(x) - B)] dx \right], \] (52)

where \( T[\cdot] \) is the time-ordering operator. We then arrive at the inequality of

\[ \| U(\tau) \| \leq e^\int_0^\tau \| \hat{B}(x) - B \| dx. \] (53)

We notice that \( \hat{B}(x) \) is given by

\[ \hat{B}(x) - B = \sum_{m,n=0}^{\infty} \frac{x^{n+m}}{n! m!} \text{ad}_B^n[\text{ad}_{A+B}^m(B)] - B = \sum_{n=0}^{\infty} \frac{x^{n+m}}{n! m!} \text{ad}_B^n[\text{ad}_{A+B}^m(B)], \] (54)

where \( \text{ad}_B^n[\cdot] = [B, \text{ad}_B^{n-1}([B, \cdot])]. \)

By using the basic lemma 1, we obtain

\[ \| \text{ad}_{A+B}^m(B) \| \leq \| \text{ad}_{A+B}^{m-1}([A, B]) \| \leq (2\lambda)^{m-1} m! \|([A, B])\|, \] (55)

where we defined \( \lambda := 2g \) and use the fact that \( A + B \) is at most \( 2g \)-extensive. Similarly, we have

\[ \| \text{ad}_B^n[\text{ad}_{A+B}^m(B)] \| \leq 2^{m-1} \lambda^{n+m-1} (n+m)! \|([A, B])\|. \] (56)

Therefore, we calculate the upper bound of \( \| \hat{B}(x) - B \| \) as

doi:10.1088/1742-5468/2016/11/113103

17
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ \| B(x) - B \| \leq \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{|x|^{n+m}}{n! m!} 2^{n-1} \lambda^{n+m-1}(n + m)! \| [A, B] \| \]

\[ = \sum_{m=1}^{\infty} \frac{(\lambda |x|)^m}{2\lambda} 2^m (1 - \lambda |x|)^{-m-1} \| [A, B] \| \]

\[ = \frac{|x|}{(1 - 3\lambda |x|)(1 - \lambda |x|)} \| [A, B] \| \]  

(57)

for \( \lambda |x| < 1/3 \), where we use the Taylor expansion of \( (1 - x)^{-m-1} = \sum_{n=0}^{\infty} \frac{x^n}{n!} (m + 1)(m + 2) \cdots (m + n) \) in the first equality. Because we have assumed \( |\tau|\lambda < 1/4 \) for \( \tau \), we finally obtain

\[ \| U(\tau) \| \leq \exp \left[ \frac{\| [A, B] \|}{\lambda^2} \int_{0}^{\lambda |\tau|} \frac{t}{(1 - t)(1 - 3t)} dt \right] \]

\[ = \exp \left[ \frac{\| [A, B] \|}{\lambda^2} (2^{-1} \log(1 - \lambda |\tau|) - 6^{-1} \log(1 - 3\lambda |\tau|)) \right] \leq \exp \left[ \frac{3\tau^2}{2} \| [A, B] \| \right], \]  

(58)

where we use the inequality \( 2^{-1}\log(1 - x) - 6^{-1}\log(1 - 3x) \leq 3x^2/2 \) for \( 0 \leq x \leq 1/4 \).

Thus, from the inequalities (50) and (58), we arrive at the inequality (B.1) for \( |\tau| < 1/(4\lambda) \). This completes the proof. \( \square \)

5.2. Proof of corollary 4

Because of \( n \leq 2^{m_0} \), we first denote

\[ A = \frac{1}{n} \sum_{j=1}^{n} A_j = \sum_{j=1}^{2^{m_0}} \tilde{A}_j, \]  

(59)

where \( \tilde{A}_j = A_j/n \) for \( j \leq n \) and \( \tilde{A}_j = 0 \) for \( n < j \leq 2^{m_0} \). Note that each of \( \{\tilde{A}_j\}_{j=1}^{2^{m_0}} \) is now \( k \)-local and \( (g/n) \)-extensive. For the proof, we begin with

\[ A = \sum_{j=1}^{2^{m_0}-1} \tilde{A}_j + \sum_{j=2^{m_0}+1}^{2^{m_0}} \tilde{A}_j =: A_0 + A_1, \]  

(60)

where \( A_0 \) and \( A_1 \) are \( (g/2) \)-extensive. According to theorem 3, we obtain

\[ M(A, \rho, \tau) \leq \frac{M(2A_0, \rho, \tau) + M(2A_1, \rho, \tau)}{2} + \frac{3\tau^2}{2} \| [A_0, A_1] \| \]

\[ \leq \frac{M(2A_0, \rho, \tau) + M(2A_1, \rho, \tau)}{2} + \frac{3\lambda g N \tau^2}{8}, \]  

(61)

where we utilize the inequality \( \| [A_0, A_1] \| \leq \lambda gN/4 \). For \( M(2A_0, \rho, \tau) \) and \( M(2A_1, \rho, \tau) \), we apply the same procedure. We then obtain
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[
M(A, \rho, \tau) \leq \frac{M(4A_{00}, \rho, \tau) + M(4A_{01}, \rho, \tau) + M(4A_{10}, \rho, \tau) + M(4A_{11}, \rho, \tau)}{4} + \frac{3\lambda gN\tau^2}{4},
\]

where we decompose \( A_0 = A_{00} + A_{01} \) and \( A_1 = A_{10} + A_{11} \) in the same manner as in equation (60). By repeating this process \( m_0 \) times, we can prove corollary 4. \( \square \)

5.3. Proof of theorem 5

From the assumption of (28), we have

\[
\rho = \Pi^{(j)}_{|x| \leq \sigma} \rho
\]

for all of \( \{A^0_j\}_{j=1}^\sigma \). Now, \( \{A^0_j\}_{j=1}^\sigma \) are given by the summation of non-overlapping local operators. Hence, regarding the spectral properties of \( \{A^0_j\}_{j=1}^\sigma \), we obtain the following inequality [32] for any \( q \)-local operator \( \mathcal{O} \):

\[
\| \Pi^{(A^0_j)}_{|x| \geq h} A^m \rho \| \leq \| O \| \quad \text{for} \quad \Delta \epsilon \leq 2gq,
\]

\[
= 0 \quad \text{for} \quad \Delta \epsilon > 2gq.
\]

where each of \( \{\Pi^{(A^0_j)}_{|x| \geq h}\}_{j=1}^\sigma \) denotes the projection operator onto the eigenspace of \( A^0_j \) which is in \([\epsilon, \infty)\). By using inequalities (63) and (64), we have

\[
\Pi^{(A^0_j)}_{|x| \geq h} A^m \rho = \Pi^{(A^0_j)}_{|x| \geq h} A^m \Pi^{(A^0_j)}_{|x| \leq \sigma} \rho = 0 \quad \text{for} \quad h - \sigma > 2gkm,
\]

where we use the fact that \( A^m \) is at most \((mk)\)-local. This equality gives

\[
\text{tr}(A^m \rho) = \text{tr}(A_m A_{m-1} \cdots A_2 A_1 \rho)
\]

with

\[
A_s = \frac{1}{n} \sum_{j=1}^\sigma \Pi^{(A^0_j)}_{|x| \leq \sigma + 2gk(s-1)} A^0_j
\]

for \( s = 1, 2, \ldots, m \). Note that

\[
\|A_s\| \leq \frac{1}{n} \sum_{j=1}^\sigma \| \Pi^{(A^0_j)}_{|x| \leq \sigma + 2gk(s-1)} A^0_j \| \leq \frac{\sigma + 2gk(s-1)}{\sigma + 2gk(s-1) - 1} \sum_{j=1}^\sigma = \sigma + 2gk(s-1).
\]

From equation (66) and inequality (68), we obtain

\[
\text{tr}(A^m \rho) \leq \lambda^m [(m-1) + \sigma/\lambda][(m-2) + \sigma/\lambda] \cdots \sigma/\lambda.
\]

Thus, we arrive at

\[
\text{tr}(e^{\sigma A} \rho) \leq \sum_{m=0}^{\infty} \frac{(\lambda \tau)^m}{m!} [(m-1) + \sigma/\lambda][(m-2) + \sigma/\lambda] \cdots \sigma/\lambda = (1 - \lambda \tau)^{-\sigma/\lambda}
\]
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality for $0 \leq \tau \leq \lambda$. The last equality results from the Taylor expansion of $(1 - x)^z = \sum_{n=0}^{\infty} \frac{x^n}{n!} s(s + 1) \cdots (s + n - 1)$. This completes the proof. \hfill \Box

5.4. Proof of corollary 6

We can obtain inequality (36) immediately from lemma 2 and theorem 5. From lemma 2, the Hamiltonian can be decomposed into $k$-local and $(gk)$-extensive operators in the form of equation (9), say $\{H_j^g\}_{j=1}^{\tilde{n}}$, where $H_j^g|\Omega\rangle = 0$ for $j = 1, 2, 3, \ldots, \tilde{n}$ because $H$ is frustration-free. From inequality (64) and $q$-locality of the operator $A$, the energy distribution of $|\Omega\rangle$ with respect to each of the Hamiltonians $\{H_j^g\}_{j=1}^{\tilde{n}}$ is completely localized at most with the width $2gkq$, i.e. $\sigma = 2gkq$. Then, inequality (31) in theorem 5 gives

$$\frac{||\Pi_{x \geq x} A|\Omega\rangle||^2}{||A|\Omega\rangle||^2} \leq \left(\frac{x}{2gkq}\right)^{gk} \exp\left(-\frac{x - 2gkq}{k\lambda}\right), \quad \text{for } x \geq 2gkq,$$

(71)

which reduces to inequality (36). Notice that the Hamiltonian is now $(gk)$-extensive instead of $g$-extensive; hence, the parameter $\lambda$ in inequality (31) should be replaced by $k\lambda$. \hfill \Box

5.5. Proof of lemma 4

The proof is immediately given by applying corollary 4 and inequality (14). First, from corollary 4, we have

$$M(A, \rho_{Prod}, \tau) \leq \frac{1}{2^{m_0}} \sum_{j=1}^{\tilde{n}} M(2^{m_0} A_j^g/\tilde{n}, \rho_{Prod}, \tau) + \frac{3\lambda g N \tau^2}{8} [\log_2 \tilde{n}].$$

(72)

By replacing $\tau$ with $\frac{2^{m_0}}{\tilde{n}} g\tau$ in inequality (14), we obtain

$$M(2^{m_0} A_j^g/\tilde{n}, \rho_{Prod}, \tau) \leq \frac{4}{k} N \tau^2 \left(\frac{2^{m_0} g\tau}{\tilde{n}}\right)^2$$

(73)

for $j = 1, 2, \ldots, \tilde{n}$. This reduces inequality (72) to

$$M(A, \rho_{Prod}, \tau) \leq \frac{1}{2^{m_0}} \sum_{j=1}^{\tilde{n}} \frac{4}{k} N \left(\frac{2^{m_0} g\tau}{\tilde{n}}\right)^2 + \frac{3\lambda g N \tau^2}{8} [\log_2 \tilde{n}]$$

$$\leq N \tau^2 \left(\frac{8g^2}{k} + \frac{3\lambda g}{8} [\log_2 \tilde{n}]\right),$$

(74)

where we use $2^{m_0} \leq 2\tilde{n}$. Hence, lemma 4 is proved. \hfill \Box

5.6. Proof of theorem 7

For the proof, we calculate the moment $\langle A^m \rangle$ separately. The main idea for the calculations is to decompose $\langle A^m \rangle$ as follows. We begin with $m = 1$ and $m = 2$, and consider the decomposition of
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ \langle A^1 \rangle = \frac{1}{\bar{n}} \sum_{j=1}^{\bar{n}} \langle A_j^c \rangle, \]

\[ \langle A^2 \rangle = \frac{1}{\bar{n}^2} \sum_{i,j=1}^{\bar{n}} \langle A_i^c A_j^c \rangle = \frac{1}{\bar{n}^2} \sum_{i<j} \langle (A_i^c + A_j^c)^2 \rangle - \frac{\bar{n}-2}{\bar{n}^2} \sum_{j=1}^{\bar{n}} \langle (A_j^c)^2 \rangle. \]

(75)

The point of these decompositions is that we only have to calculate a set of the moment functions, which are constructed from at most two operators. For example, \( A_i^c + A_j^c \) corresponds to the case of \( \bar{n} = 2 \); hence, the upper bound of \( \langle (A_i^c + A_j^c)^2 \rangle \) can be efficiently estimated from inequality (42).

In the same manner, we aim to decompose \( \langle A^m \rangle \) so that we have to calculate a set of the moment functions constructed from at most \( m \) operators. For this purpose, we consider the following decomposition:

\[ A^m = \sum_{j=0}^{m-1} \theta_j S_{m-j}. \]

(76)

Here, \( \{\theta_j\}_{j=0}^{m-1} \) are integers and

\[ S_{m-j} := \frac{1}{\bar{n}^m} \sum_{|\Xi|=m-j} |\Xi|^m A^m_{\Xi}, \quad A_{\Xi} := \frac{1}{|\Xi|} \sum_{i \in \Xi} A_i^c, \]

where \( \Xi \) is a positive-integer set of size \( |\Xi| \) with elements from \( \{1, 2, \ldots, \bar{n}\} \), e.g. \( \Xi = \{i_1, i_2, i_3, \ldots, i_{|\Xi|}\} \). In appendix A, we show that the integers \( \{\theta_j\}_{j=0}^{m-1} \) are given by

\[ \theta_j = (-1)^j \binom{\bar{n} - m + j - 1}{j}. \]

(78)

By using this decomposition, we obtain

\[ |\langle A^m \rangle| \leq \sum_{j=0}^{m-1} \frac{\theta_j}{\bar{n}^m} \sum_{|\Xi|=m-j} |\Xi|^m |\langle A^m_{\Xi} \rangle|. \]

(79)

Here, the operator \( A_{\Xi} \) is a summation of \( |\Xi| \) operators; hence, the inequality (42) implies that \( \text{tr}(\Pi_{\Xi}^2 \rho_{\text{Prod}}) \) decays as \( e^{-\text{const.} |\Xi|^2/\bar{n}^2} \), namely Gaussian decay with the variance \( \mathcal{O}(N \log |\Xi|) \). Thus, we obtain

\[ |\langle A^m_{\Xi} \rangle| \leq c_1 \Gamma \left( \frac{|\Xi| + 1}{2} \right) (c_2 N \log |\Xi|)^{m/2} \leq c_1 \Gamma \left( \frac{m + 1}{2} \right) (c_2 N \log m)^{m/2}, \]

(80)

where \( \Gamma(x) \) is the gamma function, and \( c_1 \) and \( c_2 \) are constants of \( \mathcal{O}(1) \) that depend on \( g \) and \( k \). Note that \( A_{\Xi} \) is now the summation of at most \( m \) operators \( (|\Xi| \leq m) \).

The remaining problem is to count the number of summands in equation (79). First, the number of sets \( \Xi \) that satisfy \( |\Xi| = m - j \) is given by \( \binom{\bar{n}}{m-j} \). Hence, the total number of summands in equation (76) is
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ \sum_{j=0}^{m-1} |\theta_j| \left( \frac{\bar{n}}{m-j} \right) = \sum_{j=0}^{m-1} \left( \frac{\bar{n} - m + j - 1}{j} \right) \left( \frac{\bar{n}}{m-j} \right) \]
\[ = \sum_{j=0}^{m-1} \frac{1}{j!} \left( \frac{\bar{n} - m + 1}{(m-j)!} \right) \frac{\bar{n}!}{(\bar{n} - m + j)!} \]
\[ = \sum_{j=0}^{m-1} \frac{1}{j!} \left( \frac{\bar{n} - m}{(m-j)!} \right) \frac{\bar{n}!}{(\bar{n} - m + j)!} \]
\[ \leq \left( \frac{\bar{n}}{m} \right) \sum_{j=0}^{m-1} \left( \frac{m}{j} \right) \leq 2^m \left( \frac{\bar{n}}{m} \right). \]  

By combining inequalities (79)–(81), we obtain

\[ |\langle A^m \rangle| \leq \frac{2^m}{\bar{n}^m} \left( \frac{\bar{n}}{m} \right) c_1 \Gamma \left( \frac{m+1}{2} \right) \left( c_2 N \log m \right)^{m/2} \]
\[ \leq c_1 (2Nc_2m \log m)^{m/2}, \]  

where we use the inequalities \( \left( \frac{\bar{n}}{m} \right) \leq \bar{n}^m \) and \( \Gamma \left( \frac{m+1}{2} \right) \leq (m/2)^{m/2} \). To connect inequality (82) to the distribution function of \( A \), we use the following inequality:

\[ |\langle A^m \rangle| \geq x^m \text{tr} (\Pi_{\geq x} \rho_{\text{Prod}}) \quad \text{for} \quad x > 0, \]

which yields with (82) the following inequality:

\[ \text{tr} (\Pi_{\geq x} \rho_{\text{Prod}}) \leq c_1 \left( \frac{2Nc_2m \log m}{x^2} \right)^{m/2}. \]  

By defining \( m_0 \) as the minimum integer such that

\[ \frac{2Nc_2m \log m}{x^2} \leq \frac{1}{e}, \]  

we obtain

\[ \text{tr} (\Pi_{\geq x} \rho_{\text{Prod}}) \leq c_1 e^{-m_0/2}. \]  

Now, the integer \( m_0 \) satisfying (85) has the value of \( \frac{\log(x^2)}{N \log(N/x^2)} \) and inequality (86) reduces to (43). Hence, the theorem is proved. \( \square \)

5.7. Proof of theorem 9

We calculate the moment-generating function for \( A(t) \) with the form of equation (46). For this purpose, we first decompose \( A(t) \) as

\[ A(t) = \sum_{i \geq 0} A_i = \sum_{s \geq 1} Q_s \quad \text{with} \quad Q_s := \sum_{l = s}^{l+1} A_l, \]  

where
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

where \( l_1 = 0 \) and the set of \( \{ l_s \} \) is chosen appropriately such that \( Q_s \) is at most \([g_0 t^D 2^{-2s}]\)-extensive; from equation (47), we have to choose \( l_s \) as \( l_s = O(s) \).

We then define \( Q_{g_s} := \sum_{s \geq 1} Q_s \) (i.e. \( Q_{g_1} = A(t) \)) and decompose \( M(A(t), \rho_{\text{prod}}, \tau) \) as

\[
M(A(t), \rho_{\text{prod}}, \tau) = M(Q_{g_1}, \rho_{\text{prod}}, \tau) = \frac{M(2Q_1, \rho_{\text{prod}}, \tau) + M(2Q_{g_2}, \rho_{\text{prod}}, \tau)}{2} + E_{r1},
\]

where \( E_{r1} := M(Q_{g_1}, \rho_{\text{prod}}, \tau) - [M(2Q_1, \rho_{\text{prod}}, \tau) + M(2Q_{g_2}, \rho_{\text{prod}}, \tau)]/2. \) Note that \( Q_{g_1} = Q_1 + Q_{g_2} \). Similarly, we decompose \( M(2Q_{g_2}, \rho_{\text{prod}}, \tau) \) as

\[
M(2Q_{g_2}, \rho_{\text{prod}}, \tau) = \frac{M(2^2Q_2, \rho_{\text{prod}}, \tau) + M(2^2Q_{g_3}, \rho_{\text{prod}}, \tau)}{2} + E_{r2},
\]

where \( E_{r2} := M(2Q_{g_2}, \rho_{\text{prod}}, \tau) - [M(2^2Q_2, \rho_{\text{prod}}, \tau) + M(2^2Q_{g_3}, \rho_{\text{prod}}, \tau)]/2. \) By repeating this process, we reach the equality

\[
M(Q_{g_1}, \rho_{\text{prod}}, \tau) = \sum_{s \geq 1} M(2^sQ_s, \rho_{\text{prod}}, \tau) + 2E_r,
\]

where

\[
E_r := M(2^{s-1}Q_{g_s}, \rho_{\text{prod}}, \tau) - \frac{M(2^sQ_s, \rho_{\text{prod}}, \tau) + M(2^sQ_{g_{s+1}}, \rho_{\text{prod}}, \tau)}{2}.
\]

Our task is then to estimate the terms \( M(2^sQ_s, \rho_{\text{prod}}, \tau) \) and \( E_r \). We will treat the two terms separately.

We first estimate the function \( M(2^sQ_s, \rho_{\text{prod}}, \tau) \). From the definition of \( Q_s \) in equation (87), we apply the lemma 4 to \( M(2Q_1, \rho, \tau) \). Now, \( 2^{s-1}Q_{s-1} \) is \( O(l_s^D) \)-local and \( (g_0 t^D 2^{s-1}) \)-extensive; hence,

\[
M(2^{s-1}Q_{s-1}, \rho_{\text{prod}}, \tau) \leq \text{const} \cdot \tau^2 \bar{N}(g_0 t^D 2^{-s})^2 l_s^D = \text{const} \cdot \tau^2 g_0^2 N 2^{-2s} (st)^D,
\]

for \( \tau \leq \text{const} \cdot 2^s N/(g_0 t^D) \) and \( D = O(1) \), where we use \( l_s = O(s) \) and \( \bar{N} = O(N t^D) \). By using this inequality, we obtain

\[
\sum_{s \geq 1} \frac{M(2^sQ_s, \rho, \tau)}{2^s} \leq \text{const} \cdot \tau^2 g_0^2 N t^D
\]

for \( \tau \leq \text{const} \cdot 1/(g_0 t^D) \).

We second estimate the value \( E_r \). For the purpose, we have to extend the basic theorem 3 to the quasi-local operators:

**Lemma 5.** Let \( A \) and \( B \) be quasi-local operators such that

\[
A = \sum_{q \geq q_0} A_q, \quad B = \sum_{q \geq q_0} B_q
\]

where \( A_q \) and \( B_q \) are \( q \)-local and \( g(q) \)-extensive with
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ g(q) := \frac{q}{k_0} e^{-\left(q/k_0\right)^2}, \quad (95) \]

where \( \gamma \) is a positive constant. Then, for an arbitrary quantum state \( \rho \) and \( |\tau| < \text{const} \cdot 1/(g_0^2 \gamma^2) \), we have

\[ M(A + B, \rho, \tau) \leq \frac{M(2A, \rho, \tau) + M(2B, \rho, \tau)}{2} + \text{const} \cdot \gamma^3 g^2 k_0 \mathcal{N} e^{-\text{const} \cdot (g_0/k_0)^2 \tau^2}. \quad (96) \]

This lemma immediately gives the upper bound of \( \mathbb{E}_s \) by substituting \( A \rightarrow 2^{s-1} Q_s, \ B \rightarrow 2^{s-1} Q_{s+1}, \ g \rightarrow g_0^D 2^{-s-1}, \ k_0 \rightarrow 1, \ \gamma \rightarrow D \) and \( \mathcal{N} \rightarrow \tilde{\mathcal{N}} \) and \( g_0 \rightarrow \text{const} \cdot l^D \) in (96), respectively:

\[ \sum_{s \geq 1} 2^{-s+1} \mathbb{E}_s \leq \text{const} \cdot \sum_{s \geq 1} g_0^2 l^D \tilde{\mathcal{N}} 2^{-s-1} e^{-\text{const} \cdot l \tau^2} \]

\[ = \text{const} \cdot \tau^2 g_0^2 \mathcal{N} l^D \quad (97) \]

for \( \tau \leq \text{const} \cdot 1/(g_0 l^D) \), where we use \( \tilde{\mathcal{N}} = \mathcal{O}(N/l^D) \) and include the \( D \)-dependence into the coefficient since we here assume that \( D \) is an \( \mathcal{O}(1) \) constant.

6. Summary and future work

In the first half of this paper, we proved a general theorem on the moment-generating function, which connects a set of moment-generating functions \( \{ M(\rho, \tau, A_i) \}_{i=1}^n \) to \( M(\rho, \tau, \sum_{i=1}^n A_i) \). The crucial point of this theorem is that it is restricted to few-body operators instead of arbitrary operators. This result is quite general in the sense that we do not need any assumption on the quantum state \( \rho \) and few-body operators are the most general observables when we analyze realistic quantum many-body systems. Moreover, experimentally, this analysis may be helpful to infer probability distributions for other eigenbases that are difficult to measure experimentally; it is often the case that we can prepare projective measurements only onto simple eigenbases. As one of the applications, we utilize this upper bound to generalize the Chernoff inequality so that it may be applied to more general observables beyond the summation of single-site operators as in figure 1.

On the other hand, we have left several open problems. First, in corollary 4, the error term in (26) is proportional to \( \log \tilde{n} \). This term hampers us in theorem 7 from generalizing the Chernoff inequality for generic few-body operators in a complete manner, i.e. in the strict Gaussian form instead of the quasi-Gaussian form. Thus far, we have not clarified whether this estimation is qualitatively optimal, because the scaling of \( \log \tilde{n} \) is too subtle to observe numerically.

Second, as an important future direction, can we apply the present techniques to analyze more general quantum states? Our approach gives strong statements on the Chernoff inequality, whereas the range of application is now restricted to the cases where the quantum state \( \rho \) is given by the product states or short-range entangled states. One of the most prominent classes is the class of gapped ground states, or
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality equivalently ground states in non-critical phases. A Chernoff-type inequality has been proved [28] in such systems as well but in a weaker way as $e^{-|x|/\Theta(N)}$ instead of the Gaussian decay form of $e^{-x^2/2\Theta(N)}$. This kind of probability-distribution analysis provides us useful information for constructing the approximate projection of ground states, which has been a backbone in recent analysis of ground states [28, 35, 36]. It is quite intriguing to know whether we can refine the weak Chernoff inequality for the gapped ground states to the Gaussian form.

Acknowledgments

The author is grateful to Anurag Anshu for helpful discussions and comments on related topics. This work was supported by World Premier International Research Center Initiative (WPI), Mext, Japan. TK also acknowledges support from JSPS grant no. 2611111.

Appendix A. Derivation of equation (78)

For the derivation, we need to subtract the overcounted terms. To make the point clear, we begin with the case of $m = 2$:

$$A^2 = \left( \sum_{j=1}^{a} A_j \right)^2. \quad \text{(A.1)}$$

In this case, from equation (76), we have

$$S_0 = \sum_{i<j} (A_i^c + A_j^c)^2, \quad S_1 = \sum_{j=1}^{a} (A_j^c)^2. \quad \text{(A.2)}$$

Then, in $S_0$, the terms $A_i^c A_j^c (i \neq j)$ are not overcounted, but the terms $A_i^c A_j^c$ are counted $(\bar{n} - 1)$ times; that is, we overcount the terms $(\bar{n} - 2)$ times and have to subtract $(\bar{n} - 2)S_1$. We thus obtain $A^2 = S_0 - (\bar{n} - 2)S_1$, namely $\theta_0 = 1$ and $\theta_1 = -(\bar{n} - 2)$.

In the same manner, by subtracting the overcounted terms, we arrive at the following recurrence equation for general $m$:

$$\theta_{j_0} = -\sum_{j=0}^{j_0-1} \theta_j \left( \bar{n} - m + j_0 \right) + 1 \quad \text{for} \quad 0 \leq j_0 \leq m - 1 \quad \text{(A.3)}$$

with $\theta_0 = 1$. We obtain equation (78) by solving this equation. We prove it by the inductive method. For $j = 0$, it is clear that $\theta_0 = 1$, and we assume that the equality is true for $j \leq j_0 - 1$. Then, we obtain

$$\theta_{j_0} = -\sum_{j=0}^{j_0-1} (-1)^j \binom{\bar{n} - m + j_0 - 1}{j} \binom{\bar{n} - m + j_0}{j_0 - j} + 1. \quad \text{(A.4)}$$

Our task is to show the equality

doi:10.1088/1742-5468/2016/11/113103
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\[ - \sum_{j=0}^{n-1} (-1)^j \left( \binom{n - m + j - 1}{j} \binom{n - m + j_0}{j_0 - j} \right) + 1 = (-1)^b \binom{n - m + j_0 - 1}{j_0}, \tag{A.5} \]

which gives us equation (78).

For the derivation of equation (A.5), we use the inductive method again. First, by denoting \( M = n - m \), the inequality reduces to

\[ - \sum_{j=0}^{n-1} (-1)^j \left( \binom{M + j - 1}{j} \binom{M + j_0}{j_0 - j} \right) + 1 = (-1)^b \binom{M + j_0 - 1}{j_0}. \tag{A.6} \]

For \( j_0 = 1 \), we have

\[ - \binom{M - 1}{0} \binom{M + 1}{1} + 1 = -M = (-1)^0 \binom{M + 1 - 1}{1}, \tag{A.7} \]

and hence the equality is true. We then assume that the equality is true for the case of \( j_0 \) and prove it for the case of \( j_0 + 1 \). We begin with the decomposition of

\[ - \sum_{j=0}^{n-1} (-1)^j \left( \binom{M + j - 1}{j} \binom{M + j_0 + 1}{j_0 + 1 - j} \right) + 1 \]

\[ = - \sum_{j=0}^{n-1} (-1)^j \left( \binom{M + j - 1}{m} \binom{M + j_0 + 1}{j_0 - j + 1} \right) + 1 - (-1)^b \binom{M + j_0 - 1}{j_0} \binom{M + j_0 + 1}{j_0 + 1 - j_0}. \tag{A.8} \]

For the calculation of the first term in equation (A.8), we utilize the Pascal’s rule as

\[ \binom{M + j_0 + 1}{j_0 - j + 1} = \binom{M + j + (j_0 - j + 1)}{j_0 - j} = \binom{M + j_0}{j_0 - j} + \binom{M + j_0}{j_0 - j + 1}, \tag{A.9} \]

which yields

\[ - \sum_{j=0}^{n-1} (-1)^j \left( \binom{M + j - 1}{j} \binom{M + 1 + j_0}{j_0 - j + 1} \right) + 1 \]

\[ = - \sum_{j=0}^{n-1} (-1)^j \left( \binom{M + j - 1}{j} \left( \binom{M + j_0}{j_0 - j} + \binom{M + j_0}{j_0 - j + 1} \right) \right) + 1 \]

\[ = (-1)^b \binom{M + j_0 - 1}{j_0} - \sum_{j=0}^{n-1} (-1)^j \binom{M + j - 1}{j_0 - j + 1} \left( \binom{M + j_0}{j_0} \right) \]

\[ = (-1)^b \binom{M + j_0 - 1}{j_0} - \sum_{j=0}^{n-1} (-1)^j \binom{j_0 + 1}{j} \binom{M + j_0}{j_0 + 1} \]

\[ = (-1)^b \binom{M + j_0 - 1}{j_0} + (-1)^b \binom{j_0 + 1}{j_0} \binom{M + j_0}{j_0 + 1} + (-1)^{b+1} \binom{M + j_0}{j_0 + 1}. \tag{A.10} \]
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

The second term in equation (A.8) is also given by

$$(-1)^j \binom{M + j_0 - 1}{j_0} \binom{M + j_0 + 1}{j_0 + 1} = (-1)^j \binom{M + j_0 - 1}{j_0} (M + j_0 + 1). \quad (A.11)$$

By combining equations (A.10) and (A.11) and substituting the combination in equation (A.8), we arrive at

$$\sum_{j=0}^{k} (-1)^m \binom{M + j - 1}{j} \binom{M + j_0 + 1}{j_0 + 1} + 1$$

$$= (-1)^{j_0} \binom{M + j_0 - 1}{j_0} + (-1)^{j_0} \binom{M + j_0 }{j_0 + 1} - (-1)^{j_0} \binom{M + j_0 - 1}{j_0} (M + j_0 + 1)$$

$$= (-1)^{j_0} \binom{M + j_0 }{j_0 + 1} - (-1)^{j_0} \binom{M + j_0 - 1}{j_0 } (M + j_0) = (-1)^{j_0 + 1} \binom{M + j_0}{j_0 + 1}, \quad (A.12)$$

where we use the equality

$$(-1)^{j_0} \binom{M + j_0 - 1}{j_0 } (M + j_0) = (-1)^{j_0} \frac{(M + j_0)!}{j_0! (M - 1)!} (M + j_0)$$

$$= (-1)^{j_0} \frac{j_0 + 1}{(j_0 + 1)! (M - 1)!}. \quad (A.13)$$

Thus, we prove equation (A.5) for the case of $j_0 + 1$. This completes the proof. \qed

Appendix B. Proof of lemma 5

For the proof, we reconsider the norm of $\| \text{ad}_B^n \left[ \text{ad}_A^n \right] \|$ which were given by (56) for strict $k$-local operators. We notice that from the inequalities (50) and (53) the moment-generating function $M(A + B, \rho, \tau)$ is bounded from above by

$$M(A + B, \rho, \tau) \leq \frac{M(2A, \rho, \tau) + M(2B, \rho, \tau)}{2} + \int_0^{\tau} \| \tilde{B}(x) - B \| dx. \quad (B.1)$$

We thus need to calculate

$$\| \tilde{B}(x) - B \| = \| e^{-xB} e^{x(A+B)} B e^{-x(A+B)} e^{xB} - B \|

= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{|x|^{n+m}}{n! m!} \| \text{ad}_B^n \left[ \text{ad}_A^n \right] \|. \quad (B.2)$$

The difficulty lies in that the operators $A$ and $B$ contain non-local terms up to $N$-body operators. To calculate the norm of the multi-commutator norm, we have to prove the following lemma.
Lemma 6. Let $A$ be an operator that is given by the form

$$A = \sum_{q \geq q_0} A_q$$

where $A_q$ is $q$-local and $g(q)$-extensive with $\log[g(q)]$ a concave function. Then, for an arbitrary operator $O_L$ which is supported in a subset $L$, the norm of the multi-commutator $\text{ad}^m_A(O_L)$ is bounded from above by

$$\| \text{ad}^m_A(O_L) \| \leq \frac{2|L| \cdot \| O_L \|}{(m - 1)!} G(m, |L|, q_0)$$

where the function $G(m, |L|, q_0)$ is defined as

$$G(m, |L|, q_0) := \sum_{M = q_0(m+1)}^{\infty} \left[ g\left(\frac{M}{m}\right) \right]^m [M(|L| + M)]^{m-1}.$$  

In particular, for an operator $B$ such that

$$B = \sum_{q \geq q_0} B_q$$

with each of $\{B_q\}$ a $q$-local and $g(q)$-extensive operator with $\log[g(q)]$ defined above, we obtain

$$\| \text{ad}^m_B(B) \| \leq \frac{N}{m!} \frac{G'(m, q_0)}{2}$$

where the function $G'(m,q_0)$ is defined as

$$G'(m, q_0) := \sum_{M = q_0(m+1)}^{\infty} \left[ 2g\left(\frac{M}{m}\right) \right]^{m+1} M^{2m}.$$  

We defer the proof to section B.1.

By applying lemma 6 to $\text{ad}^m_B[\text{ad}^m_A + B(B)]$ with the definition of $g(q)$ in equation (95), we obtain

$$\| \text{ad}^m_B[\text{ad}^m_A + B(B)] \| \leq \frac{N/2}{\tilde{m}!} \sum_{M = q_0(\tilde{m}+1)}^{\infty} \left[ \frac{4g\left(\frac{M}{m}\right)}{k_0^2} e^{-2[M(\tilde{m}k_0)]^{\gamma}} \right]^{\tilde{m}+1} M^{2\tilde{m}}$$

$$\leq \left( c_1 g(k_0)^{\tilde{m}+1} N \right) \frac{(\tilde{m}^{\gamma})^{2\tilde{m}+1} e^{-c_2 g(k_0)^{\gamma}}}{k_0 \cdot \tilde{m}!},$$

where $c_1$ and $c_2$ are $O(1)$ constants and $\tilde{m} = n + m$; note that $A + B = \sum_{q \geq q_0} (A_q + B_q)$ and $A_q + B_q$ is at most $[2g(q)]$-extensive. Then, the left-hand side of the inequality (B.2) is bounded from above by
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

\begin{align}
\sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{|x|^{n+m}}{n! \cdot m!} \| \text{ad}_A^n[\text{ad}_A^n + \text{ad}_B^n(B)] \| \\
\leq e^{-c(q_0/k_0)^{1/\gamma}} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{|x|^{n+m}}{n! \cdot m!} \frac{(c_1gk_0)^{m+1}N}{k_0 \cdot m!} (m\gamma)^{2\hat{m}+1} \\
\leq g \gamma^2 N e^{-c(q_0/k_0)^{1/\gamma}} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} (c^*|x|)^{n+m} = \text{const} \cdot \gamma^2 g^2 k_0 N e^{-c(q_0/k_0)^{1/\gamma}} |x|, \tag{B.10}
\end{align}

for $|x| < 1/c^*$, where $c^* = \text{const} \cdot gk_0\gamma^2$. By applying the inequality (B.10) to (B.1) with equation (B.2), we obtain the inequality (96). This completes the proof.

### B.1. Proof of lemma 6

For simplicity, we set $\|O_L\| = 1$. We start from the expression

\[ \text{ad}_A^m(O_L) = \sum_{q_1, q_2, \ldots, q_m} [A_{q_1}, [A_{q_{n-1}}, \ldots [A_{q_2}, O_L] \ldots]]. \]

By using lemma 1, we obtain

\[ \| [[A_{q_n}, [A_{q_{n-1}}, \ldots [A_{q_2}, O_L] \ldots]]] \| \leq 2^m g(q_1)g(q_2) \cdots g(q_m)|L|(|L|+q_1)(|L|+q_1+q_2) \cdots (|L|+q_1+q_2+\cdots+q_{m-1}). \tag{B.11} \]

The norm of $\text{ad}_A^m(O_L)$ is bounded from above by

\[ \| \text{ad}_A^m(O_L) \| \leq \sum_{q_1, q_2, \ldots, q_m} 2^m g(q_1)g(q_2) \cdots g(q_m)|L|(|L|+q_1)(|L|+q_1+q_2) \cdots \times (|L|+q_1+q_2+\cdots+q_{m-1}) \\
= \sum_{M \geq q_0m} \sum_{q_1+q_2+\ldots+q_m=M} 2^m g(q_1)g(q_2) \cdots g(q_m)|L|(|L|+q_1) \times (|L|+q_1+q_2) \cdots (|L|+q_1+q_2+\cdots+q_{m-1}), \tag{B.12} \]

where the summation over $M$ starts from $q_0m$ because each of $\{q_i\}_{i=1}^m$ is larger than $q_0$ from the definition (B.3). Here, because $\log[g(q)]$ is a concave function, we have $\sum_{s=1}^m \log[g(q_s)] \leq m \log[g(M/m)]$ for $q_1+q_2+\cdots+q_m=M$; hence, the summmand in the inequality (B.12) is upperbounded by

\[ g(q_1)g(q_2) \cdots g(q_m)|L|(|L|+q_1)(|L|+q_1+q_2) \cdots (|L|+q_1+q_2+\cdots+q_{m-1}) \leq \left[ g \left( \frac{M}{m} \right) \right]^m |L| \cdot |L|+M-q_0(m-1)|L|+M-q_0(m-2)|L|+M-q_0 \cdots |L|+M-q_0, \tag{B.13} \]

where we use the inequality $(q_1+q_2+\cdots+q_s=M-(q_{s+1}+q_{s+2}+\cdots+q_m) \leq M-q_0(m-s)$ for $s=1,2,\ldots,m-1$. By combining the two inequalities (B.12) and (B.13), we obtain
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeüfding inequality

\[ \| \text{ad}_A^n(O_L) \| \]

\[ \leq \sum_{M \geq qm} \sum_{q_1 + q_2 + \ldots + q_n = M} \left[ 2g\left( \frac{M}{m} \right) \right]^m |L| \cdot \left[ |L| + M - q_0(m - 1) \right] \]

\[ \cdot \left[ |L| + M - q_0(m - 2) \right] \cdots \left[ |L| + M - q_0 \right] \]

\[ \leq \sum_{M \geq qm} \left( \frac{m + M - q_0m - 1}{m - 1} \right)^m \left[ 2g\left( \frac{M}{m} \right) \right]^m |L| \cdot \left[ |L| + M - q_0(m - 1) \right] \]

\[ \cdot \left[ |L| + M - q_0(m - 2) \right] \cdots \left[ |L| + M - q_0 \right] \]

\[ = \frac{|L|}{(m - 1)!} \sum_{M \geq qm} \left[ 2g\left( \frac{M}{m} \right) \right]^m \prod_{s=1}^{m-1} (M - q_0m + s)(|L| + M - q_0s) \]

\[ \leq \frac{2|L|}{(m - 1)!} \sum_{M \geq qm} \left[ g\left( \frac{M}{m} \right) \right]^m [M(|L| + M)]^{m-1} = \frac{2|L|}{(m - 1)!} G(m, |L|, q_0), \]

(B.14)

where, in the second inequality, we use the fact that the number of summand from \( \sum_{q_1 + q_2 + \ldots + q_n = M} \) is at most \( (m + M - q_0m - 1) \) for \( q_i \geq q_0 \) for \( s = 1, 2, \ldots, m \). Hence, the first inequality (B.4) is proved.

The inequality (B.7) is proved in the same way. The inequality (B.12) is now replaced by

\[ \| \text{ad}_A^{-1}(B) \| \]

\[ \leq \sum_{q \geq q_0} 2^{m-1} g(q) g(q_2) \cdots g(q_{m-1})(q(q + q_1)(q + q_1 + q_2) \cdots (q + q_1 + q_2 + \ldots + q_{m-2})) \| B \| \]

\[ \leq N \sum_{q \geq q_0} 2^{m-1} g(q) g(q_2) \cdots g(q_{m-1})(q(q + q_1)(q + q_1 + q_2) \cdots (q + q_1 + q_2 + \ldots + q_{m-2})) \]

\[ = \sum_{M \geq qm} \sum_{q + q_1 + q_2 + \ldots + q_{m-1} = M} 2^{m-1} g(q) g(q_2) \cdots g(q_{m-1})(q(q + q_1)(q + q_1 + q_2) \cdots (q + q_1 + q_2 + \ldots + q_{m-2})). \]

(B.15)

where in the second inequality we use \( \| B \| \leq \sum_{q \geq q_0} g(q)N \). This inequality, in the same way as the derivation of (B.14), yields

\[ \| \text{ad}_A^{-1}(B) \| \leq \frac{N}{2} \sum_{M \geq qm} \sum_{q + q_1 + q_2 + \ldots + q_{m-1} = M} \left[ 2g\left( \frac{M}{m} \right) \right]^m [M - q_0(m - 1)] \]

\[ \cdot [M - q_0(m - 2)] \cdots [M - q_0] \]

\[ \leq \frac{N}{2(m - 1)!} \sum_{M \geq qm} \left[ 2g\left( \frac{M}{m} \right) \right]^m \prod_{s=1}^{m-1} (M - q_0m + s)(M - q_0s) \]

\[ \leq \frac{N}{2(m - 1)!} \sum_{M \geq qm} \left[ 2g\left( \frac{M}{m} \right) \right]^m M^{2m-2}. \]

(B.16)

Hence, the inequality (B.7) is proved. \( \square \)

We thus prove lemma 6.
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

References

[1] Kempe J, Kitaev A and Regev O 2006 The complexity of the local hamiltonian problem SIAM J. Comput. 35 1070–97
[2] Osborne T J 2012 Hamiltonian complexity Rep. Prog. Phys. 75 022001
[3] Eisert J, Cramer M and Plenio M B 2010 Colloquium: area laws for the entanglement entropy Rev. Mod. Phys. 82 277–306
[4] Hastings M B 2007 An area law for one-dimensional quantum systems J. Stat. Mech. P08024
[5] Perez-Garcia D, Verstraete F, Wolf M M and Cirac J I 2006 Matrix product state representations (arXiv:quant-ph/0608197)
[6] Verstraete F, Murg V and Cirac J I 2008 Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems Adv. Phys. 57 143–224
[7] Eisert J 2013 Entanglement and tensor network states (arXiv:1308.3318)
[8] Horodecki R, Horodecki P, Horodecki M and Horodecki K 2009 Quantum entanglement Rev. Mod. Phys. 81 865–942
[9] Chernoff H 1952 A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations Ann. Math. Stat. 23 493–507
[10] Hoeffding W 1963 Probability inequalities for sums of bounded random variables J. Am. Stat. Assoc. 58 13–30
[11] Michael Nussbaum A S 2009 The chernoff lower bound for symmetric quantum hypothesis testing Ann. Stat. 37 1040–57
[12] Audenaert K M R, Calsamiglia J, Muñoz Tapia R, Bagan E, Li. Masanes, Acín A and Verstraete F 2007 Discriminating states: the quantum chernoff bound Phys. Rev. Lett. 98 160501
[13] Chen X, Gu Z-C and Wen X-G 2010 Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order Phys. Rev. B 82 155138
[14] Frowis F and Dürr W 2012 Measures of macroscopicity for quantum spin systems New J. Phys. 14 093039
[15] Hastings M B 2011 Topological order at nonzero temperature Phys. Rev. Lett. 107 210501
[16] Anshu A 2016 Concentration bounds for quantum states with finite correlation length on quantum spin lattice systems New J. Phys. 18 083011
[17] Hartmann M, Mahler G and Hess O 2004 Gaussian quantum fluctuations in interacting many particle systems Lett. Math. Phys. 68 103–12
[18] Goderis D and Vets P 1989 Central limit theorem for mixing quantum systems and the CCR-algebra of fluctuations Commun. Math. Phys. 122 249–65
[19] Goderis D, Verbeure A and Vets P 1989 Non-commutative central limits Probab. Theory Relat. Fields 82 527–44
[20] Matsui T 2002 Bosonic central limit theorem for the one-dimensional xy model Rev. Math. Phys. 14 675–700
[21] Berry A C 1941 The accuracy of the gaussian approximation to the sum of independent variates Trans. Am. Math. Soc. 49 122–36
[22] Esseen C-G 1945 Fourier analysis of distribution functions. A mathematical study of the laplace-gaussian law Acta Math. 77 1–125
[23] Brandao F G and Cramer M 2015 Equivalence of statistical mechanical ensembles for non-critical quantum systems (arXiv:1502.03263)
[24] Touchette H 2009 The large deviation approach to statistical mechanics Phys. Rep. 478 1–69
[25] Ogata Y 2010 Large deviations in quantum spin chains Commun. Math. Phys. 296 35–68
[26] Hiai F, Mosonyi M and Ogawa T 2007 Large deviations and chernoff bound for certain correlated states on a spin chain J. Math. Phys. 48 123301
[27] Lenci M and Rey-Bellet L 2005 Large deviations in quantum lattice systems: one-phase region J. Stat. Phys. 119 715–46
[28] Kuwahara T, Arad I, Amico L and Vedral V 2015 Local reversibility and entanglement structure of many-body ground states (arXiv:1502.05330)
[29] Kuwahara T 2016 Asymptotic behavior of macroscopic observables in generic spin systems J. Stat. Mech. 053103
[30] Haldane F D M 1983 Nonlinear field theory of large-spin heisenberg antiferromagnets: semiclassically quantized solitons of the one-dimensional easy-axis nêel state Phys. Rev. Lett. 50 1153–6
[31] Haldane F 1983 Continuum dynamics of the 1-d heisenberg antiferromagnet: identification with the o(3) nonlinear sigma model Phys. Lett. A 93 461–8
[32] Arad I, Kuwahara T and Landau Z 2016 Connecting global and local energy distributions in quantum spin models on a lattice J. Stat. Mech. 033301

doi:10.1088/1742-5468/2016/11/113103
Connecting the probability distributions of different operators and generalization of the Chernoff–Hoeffding inequality

[33] Kuwahara T, Mori T and Saito K 2016 Floquet-magnus theory and generic transient dynamics in periodically driven many-body quantum systems Ann. Phys. 367 96–124
[34] Kuwahara T 2016 Exponential bound on information spreading induced by quantum many-body dynamics with long-range interactions New J. Phys. 18 053034
[35] Arad I, Kitaev A, Landau Z and Vazirani U 2013 An area law and sub-exponential algorithm for 1d systems (arXiv:1301.1162)
[36] Arad I, Landau Z, Vazirani U and Vidick T 2016 Rigorous rg algorithms and area laws for low energy eigenstates in 1d (arXiv:1602.08828)
[37] Lieb E and Robinson D 1972 The finite group velocity of quantum spin systems Commun. Math. Phys. 28 251–7
[38] Bravyi S, Hastings M B and Verstraete F 2006 Lieb-robinson bounds and the generation of correlations and topological quantum order Phys. Rev. Lett. 97 050401

doi:10.1088/1742-5468/2016/11/113103