

Quantum concentration inequalities and equivalence of the thermodynamical ensembles: an optimal mass transport approach

Giacomo De Palma^{*1} and Davide Pastorello^{†1,2}

¹University of Bologna, Department of Mathematics, Piazza di Porta San Donato 5, 40126 Bologna, Italy

²TIFPA-INFN, via Sommarive 14, 38123 Povo (Trento), Italy

March 28, 2024

Abstract

We prove new concentration inequalities for quantum spin systems which apply to any local observable measured on any product state or on any state with exponentially decaying correlations. Our results do not require the spins to be arranged in a regular lattice, and cover the case of observables that contain terms acting on spins at arbitrary distance. Moreover, we introduce a local W_1 distance, which quantifies the distinguishability of two states with respect to local observables. We prove a transportation-cost inequality stating that the local W_1 distance between a generic state and a state with exponentially decaying correlations is upper bounded by a function of their relative entropy. Finally, we apply such inequality to prove the equivalence between the canonical and microcanonical ensembles of quantum statistical mechanics and the weak eigenstate thermalization hypothesis for the Hamiltonians whose Gibbs states have exponentially decaying correlations.

1 Introduction

Concentration inequalities provide an upper bound to the probability of the deviations of a random variable from its mean and have become a fundamental tool in mathematics, physics and computer science [1]. Concentration inequalities have been recently proved in

^{*}giacomo.depalma@unibo.it

[†]davide.pastorello3@unibo.it

the setting of local observables measured on quantum spin systems with weak correlations and have explained important physical effects. Local observables have the form

$$H = \sum_{\Lambda \subseteq [n]} h_{\Lambda}, \quad (1.1)$$

where each Λ is a subset of the set of all the spins and each h_{Λ} only acts on the spins in the corresponding Λ . The topic has originally been investigated for product states [2–5] and later for high-temperature Gibbs states [6, 7] and states of spin lattices whose correlations decay exponentially with the distance [8]. Concentration inequalities for Gibbs states have been the key tool to prove the equivalence between the canonical and the microcanonical ensembles of quantum statistical mechanics and the weak version of the eigenstate thermalization hypothesis [6, 7, 9–12].

The first part of the present paper is devoted to proving new concentration inequalities for product states and states of a quantum spin lattice with correlations that decay exponentially with the distance. Such concentration inequalities are based on a local norm for observables, which is given by the maximum over the spins of twice the sum of the operator norms of the addends in (1.1) that act on that spin (Definition 3.1).

The best available concentration inequality for product states is [8, Theorem 5.3], and states that the concentration of local observables is exponential in the number of spins whenever each region Λ in the sum (1.1) intersects only $O(1)$ other regions. Our Theorem 3.2 extends this result to the setting where each region Λ contains $O(1)$ spins and the local norm of the observable is $O(1)$. We stress that our result covers also the case of observables with all-to-all interactions, *e.g.*, where the sum in (1.1) runs over all the couples of spins, provided that the sum of the operator norms of the addends that act on any given spin is $O(1)$.

The best available concentration inequalities for the states of a quantum spin lattice whose correlations decay exponentially with the distance are [8, Theorem 4.2] and [12, Eq. (S.17)], and state that the probability of deviations from the average decays as $\exp\left(-\Omega\left(n^{\frac{1}{2d+1}}\right)\right)$, where n is the number of spins and $d > 0$ depends on the geometry of the lattice, whenever the diameter of all the regions in the sum (1.1) is $O(1)$. Our Corollary 3.2 extends these results to the setting where each region Λ contains $O(1)$ spins and the local norm of the observable is $O(1)$. Furthermore, while [8, Theorem 4.2] and [12, Eq. (S.17)] require a regular lattice, our results do not require any regular structure nor the notion of neighboring spins, but only require a distance on the set of the spins such that the size of each metric ball is $O(r^d)$, where r is the radius (Assumption 2.1).

The second part of the paper is devoted to the equivalence between the ensembles of quantum statistical mechanics. It has long been known that a concentration inequality for an observable measured on a Gibbs state implies an upper bound in the difference between the expectation value of that observable on the Gibbs and on any microcanonical states with sufficiently close average energy [6, 7, 9–12], thus bounding the distinguishability between the Gibbs and the microcanonical states and proving the equivalence between the canonical and microcanonical ensembles for the observables that exhibit concentration. Such link between concentration and equivalence of the ensembles has been connected to the quantum

theory of optimal mass transport [5, 13–15]. In this context, a quantum generalization of the Lipschitz constant has been defined to quantify the maximum dependence of an observable on any given spin, and concentration inequalities have been proved for observables with $O(1)$ Lipschitz constant measured on weakly correlated states [6]. The distinguishability between quantum states with respect to observables with unit Lipschitz constant has been quantified by the quantum Wasserstein distance of order 1, or quantum W_1 distance [5, 16]. Such concentration has been proved to be linked to transportation-cost inequalities that provide an upper bound to the W_1 distance between a generic state and a Gibbs state in terms of the square root of their relative entropy, and the equivalence between the canonical and microcanonical ensembles has been proved via such transportation-cost inequalities [6, Proposition 12]. It has further been proved that, whenever a Gibbs state satisfies such transportation-cost inequality, it is close with respect to the W_1 distance to any state that has approximately the same entropy, *i.e.*, such that the difference between the entropy of the two states is $o(n)$, where n is the number of spins.

In this paper, we consider the distinguishability with respect to k -local observables, *i.e.*, the local observables such that each region Λ in the sum (1.1) contains at most $k = O(1)$ spins, and we quantify such distinguishability with the k -local quantum W_1 distance (Definition 4.1). First, we employ our concentration inequality for the states of a quantum spin lattice with exponentially decaying correlations to prove transportation-cost inequalities for the same states (Theorem 4.1). Then, we apply this result to prove that the Gibbs states with exponentially decaying correlations are close in the k -local quantum W_1 distance to any state with approximately the same entropy, *i.e.*, such that the difference between the entropy of the two states is $o\left(n^{\frac{1}{2d+1}}\right)$ (Proposition 5.1), where $d > 0$ depends on the growth of the size of the metric balls as explained above. In particular, we prove the equivalence between the canonical and microcanonical ensembles for such Gibbs states with respect to all the k -local observables whose local norm is $O(1)$ (Theorem 5.1).

The paper is organized as follows. In section 2, we set the notation for the paper. In section 3, we introduce the local norm and prove the concentration inequalities for product states and states with exponentially decaying correlations. In section 4, we introduce the k -local quantum W_1 distance and prove that any quantum state with exponentially decaying correlations satisfies a transportation-cost inequality with respect to such distance. In section 5, we introduce the problem of the equivalence of the ensembles of quantum statistical mechanics and prove such equivalence for the above states. We conclude in section 6. Appendix A contains the proof of the auxiliary lemmas.

2 Setup

Let us start by setting the notation for the paper:

Definition 2.1. For any $k \in \mathbb{N}$ we define

$$[k] = \{1, \dots, k\}. \quad (2.1)$$

We consider a quantum system made by n spins, which we label with the integers from 1 to n . The local Hilbert space of each spin is $\mathbb{C}^q$, such that the Hilbert space of the system is $(\mathbb{C}^q)^{\otimes n}$.

Definition 2.2. For any subset of the spins $\Lambda \subseteq [n]$, let

$$\mathcal{H}_\Lambda = \bigotimes_{x \in \Lambda} \mathbb{C}^q \quad (2.2)$$

be the Hilbert space associated with the spins in Λ , let $\mathcal{O}_\Lambda$ be the set of the self-adjoint linear operators acting on $\mathcal{H}_\Lambda$, let $\mathcal{O}_\Lambda^T$ be the set of the traceless operators in $\mathcal{O}_\Lambda$, and let $\mathcal{S}_\Lambda$ be the set of the quantum states acting on $\mathcal{H}_\Lambda$.

Definition 2.3 (k -local operator). A linear operator X acting on $(\mathbb{C}^q)^{\otimes n}$ is k -local if it is a linear combination of linear operators acting on at most k spins each, *i.e.*, if

$$X = \sum_{\Lambda \subseteq [n]: |\Lambda| \leq k} X_\Lambda, \quad (2.3)$$

where each X_Λ acts only on the corresponding $\mathcal{H}_\Lambda$. For any $\Lambda \subseteq [n]$, we denote with $\mathcal{O}_\Lambda^{(k)}$ the set of the k -local self-adjoint linear operators acting on $\mathcal{H}_\Lambda$.

When dealing with non-product states, we will need to introduce a distance $\mathbf{d}$ on the set of the spins $[n]$. For example, if the spin are located at the vertices of a graph, $\mathbf{d}$ can be the graph distance given by the length of the minimum path connecting two vertices. Otherwise, if the spins are located at points of $\mathbb{R}^d$, the distance can be the Euclidean distance or any ℓ^p distance. We stress that we do not require any regular structure such as a square lattice, nor a notion of neighboring spins.

Assumption 2.1 (Dimensionality of the spin lattice). We assume that the size of the balls of $\mathbf{d}$ with radius r are $O(r^d)$, *i.e.*, that there exist $d \geq 1$ and $A > 0$ such that for any $v \in [n]$ and any $r > 0$ we have

$$|B_{\mathbf{d}}(v, r)| \leq A r^d, \quad (2.4)$$

where d plays the role of the dimensionality of the lattice.

When dealing with non-product states, we will also need the notion of correlations decaying exponentially with the distance $\mathbf{d}$:

Definition 2.4 (Correlation length). We say that the state $\omega \in \mathcal{S}_{[n]}$ has correlation length $\xi > 0$ if there exist $C > 0$ independent of n such that for any two disjoint regions $\Lambda_1, \Lambda_2 \subset [n]$ and any two observables $h_{\Lambda_1} \in \mathcal{O}_{\Lambda_1}$ and $h_{\Lambda_2} \in \mathcal{O}_{\Lambda_2}$ we have

$$|\langle h_{\Lambda_1} h_{\Lambda_2} \rangle - \langle h_{\Lambda_1} \rangle \langle h_{\Lambda_2} \rangle| \leq C \exp \left(-\frac{\mathbf{d}(\Lambda_1, \Lambda_2)}{\xi} \right), \quad (2.5)$$

where

$$\mathbf{d}(\Lambda_1, \Lambda_2) = \min \{ \mathbf{d}(x_1, x_2) : x_1 \in \Lambda_1, x_2 \in \Lambda_2 \} \quad (2.6)$$

and the angle brackets denote the expectation with respect to ω .

3 Concentration inequalities

In this section we introduce the local norm and derive the concentration inequalities for two relevant kinds of states of a spin lattice system: product states and states with exponentially decaying correlations. More precisely, we bound from below the cumulative distribution functions of local observables with a function of their local norm.

3.1 The local norm

Let us introduce the local norm on which all our results are based:

Definition 3.1 (Local norm). The local norm of an observable $H \in \mathcal{O}_{[n]}$ is

$$\|H\|_{\text{loc}} = 2 \min \left\{ \max_{x \in [n]} \sum_{\Lambda \ni x} \|h_\Lambda\| : H = \sum_{\Lambda \subseteq [n]} h_\Lambda, h_\Lambda \in \mathcal{O}_\Lambda \right\}, \quad (3.1)$$

where $\|\cdot\|$ denotes the operator norm. In words, we consider all the decompositions of H as a sum of local operators, and define the dependence of any such decomposition on a spin x as twice the sum of the operator norm of each local operator that acts on x . We then define the local norm of such decomposition as the maximum dependence on a spin, and the local norm of H as the minimum local norm of all its possible decompositions.

Remark 3.1. The local norm (3.1) is similar to the norm considered in [7, Eq. (4)]. The novelty of (3.1) with respect to [7, Eq. (4)] is that, while [7, Eq. (4)] considers only a fixed decomposition of H of the form (1.1), (3.1) optimizes over all the possible decompositions.

Remark 3.2. The local norm (3.1) can be recovered as a special case of the local norm defined in [17], where the norm of each h_Λ is multiplied by a penalty $c_{|\Lambda|}$ depending on $|\Lambda|$:

$$\|H\|_{\widetilde{\text{loc}}} = 2 \min \left\{ \max_{x \in [n]} \sum_{\Lambda \ni x} c_{|\Lambda|} \|h_\Lambda\| : H = \sum_{\Lambda \subseteq [n]} h_\Lambda, h_\Lambda \in \mathcal{O}_\Lambda \right\}. \quad (3.2)$$

The local norm (3.1) can be recovered from (3.2) by setting to one all the penalties.

Remark 3.3. $\|\cdot\|_{\text{loc}}$ is a seminorm such that $\|I\|_{\text{loc}} = 0$.

3.2 Product states

We start by proving a bound on the cumulative distribution function of the local observable H measured on a product state as a function of the local norm of H .

In order to derive a concentration inequality involving the quantum local norm, we will first prove an upper bound to the expectation value of e^H . The concentration inequality will then follow from the Markov inequality.

Theorem 3.1. Let ω be a product state of n spins. Let $H \in \mathcal{O}_{[n]}$ be an observable, and let

$$H = \sum_{\Lambda \subseteq [n]} \tilde{h}_\Lambda \quad (3.3)$$

be the decomposition of H that achieves the local norm as in (3.1). For any $\emptyset \neq \Lambda \subseteq [n]$, let

$$h_\Lambda := \tilde{h}_\Lambda - \langle \tilde{h}_\Lambda \rangle, \quad (3.4)$$

where the angle brackets denote the expectation with respect to ω . Then,

$$\langle e^H \rangle \leq \exp \left[\langle H \rangle + \sum_{\Lambda \subseteq [n]} \left(e^{|\Lambda| \|H\|_{\text{loc}}} - 1 - \frac{|\Lambda| \|H\|_{\text{loc}}}{2} \right) \|h_\Lambda\| \right]. \quad (3.5)$$

Proof. By definition of h_Λ , for every $\emptyset \neq \Lambda \subseteq [n]$ we have $\langle h_\Lambda \rangle = 0$ and

$$H = \sum_{\Lambda \subseteq [n]} \tilde{h}_\Lambda = \tilde{h}_\emptyset + \sum_{\emptyset \neq \Lambda \subseteq [n]} \left(\langle \tilde{h}_\Lambda \rangle + h_\Lambda \right) = \langle H \rangle + \sum_{\emptyset \neq \Lambda \subseteq [n]} h_\Lambda. \quad (3.6)$$

Let us define $H_0 := \sum_{\emptyset \neq \Lambda \subseteq [n]} h_\Lambda$ and observe that:

$$\langle H_0^2 \rangle = \sum_{\Lambda_1 \cap \Lambda_2 \neq \emptyset} \langle h_{\Lambda_1} h_{\Lambda_2} \rangle \leq \sum_{\Lambda_1 \cap \Lambda_2 \neq \emptyset} \|h_{\Lambda_1}\| \|h_{\Lambda_2}\|, \quad (3.7)$$

since $\langle h_{\Lambda_1} h_{\Lambda_2} \rangle = \langle h_{\Lambda_1} \rangle \langle h_{\Lambda_2} \rangle = 0$ for $\Lambda_1 \cap \Lambda_2 = \emptyset$. Therefore, in order to bound $\langle H_0^r \rangle$, we take the sum over the collections of r regions $\Lambda_1, \dots, \Lambda_r$ such that any region does not have empty intersection with all the others:

$$\langle H_0^r \rangle \leq \sum_{(\Lambda_1, \dots, \Lambda_r) \in \mathfrak{R}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|, \quad (3.8)$$

where $\mathfrak{R}_r = \{(\Lambda_1, \dots, \Lambda_r) : \Lambda_i \cap (\cup_{j \neq i} \Lambda_j) \neq \emptyset, \forall i = 1, \dots, r\}$. For any graph $G = (V, E)$ with $V = \{1, \dots, r\}$ let

$$C_G := \sum_{\Lambda_1, \dots, \Lambda_r} \prod_{(i,j) \in E} |\Lambda_i \cap \Lambda_j| \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|. \quad (3.9)$$

Since $|\Lambda_i \cap \Lambda_j| \geq 1$ whenever $\Lambda_i \cap \Lambda_j \neq \emptyset$, we have $\langle H_0^r \rangle \leq \sum_{G \in \mathcal{G}_r} C_G$ where $\mathcal{G}_r = \{(V, E) : |V| = r, \deg(v) \geq 1 \forall v \in V\}$. In words, we associate any non-empty intersection among two regions of the collections $\Lambda_1, \dots, \Lambda_r$ to an edge of a graph G with r vertices, in this sense we take the sum only over the graphs with no isolated vertices. Furthermore, the bound over $\langle H_0^r \rangle$ can be improved considering the minimal graphs. A *minimal graph* is defined as a graph such that the removal of any edge leaves at least one isolated vertex. For any $(\Lambda_1, \dots, \Lambda_r) \in \mathfrak{R}_r$ there is a graph $G \in \mathcal{G}_r$ of r vertices such that $(i, j) \in E$ if and only if $\Lambda_i \cap \Lambda_j \neq \emptyset$. Let G' be a minimal subgraph of G , then:

$$\|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| \leq \prod_{(i,j) \in E'} |\Lambda_i \cap \Lambda_j| \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|. \quad (3.10)$$

Now, let $\mathfrak{G}_r := \{\text{minimal graphs with } r \text{ vertices}\}$, for bounding $\langle H_0^r \rangle$ we can take the sum only on the minimal graphs:

$$\langle H_0^r \rangle \leq \sum_{G \in \mathfrak{G}_r} C_G. \quad (3.11)$$

Therefore:

$$\langle e^{H_0} \rangle = \sum_{r=0}^{+\infty} \frac{\langle H_0^r \rangle}{r!} \leq \sum_{r=0}^{+\infty} \sum_{G \in \mathfrak{G}_r} \frac{C_G}{r!} = \sum_{G \in \mathfrak{G}_{min}} \frac{C_G}{V_G!}, \quad (3.12)$$

where $\mathfrak{G}_{min}$ is the set of all the minimal graph with a finite set of vertices and V_G is the number of vertices of G . We claim that any connected minimal graph is a star graph, *i.e.*, has a vertex that belongs to all the edges. Indeed, let v be a vertex with maximum degree. If such degree is 1, then all vertices have degree 1, and the graph is made by 2 vertices connected by an edge. Let the maximum degree be at least 2. Any vertex w that is connected with v by an edge cannot be connected to any other vertex, otherwise the edge (v, w) could be removed without disconnecting any vertex. Therefore, v is connected by an edge to all the other vertices and the graph is a star graph. Then, any minimal graph is the union of star graphs. Given a minimal graph with V_G vertices, let us denote the number of connected components with r vertices as N_r , then $V_G = \sum_{r \geq 2} r N_r$. The contribution C_r in (3.11) of the star graphs with r vertices is:

$$C_r := \sum_{\Lambda_1, \dots, \Lambda_r} |\Lambda_1 \cap \Lambda_2| \cdots |\Lambda_1 \cap \Lambda_r| \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|, \quad (3.13)$$

within the choice that Λ_1 corresponds to the core vertex. Let us consider a minimal graph G formed by N_r connected components, for $r \geq 2$. By definitions (3.9) and (3.13), we have:

$$C_G = \prod_{r \geq 2} C_r^{N_r}. \quad (3.14)$$

As a consequence, the last term in (3.12) can be re-written as a sum over the connected components:

$$\begin{aligned} \sum_{G \in \mathfrak{G}_{min}} \frac{C_G}{V_G!} &= \sum_{N_2, N_3, \dots} \frac{1}{(2N_2 + 3N_3 + \dots)!} \times \frac{(2N_2 + 3N_3 + \dots)!}{N_2! 2^{N_2} N_3! 2^{N_3} N_4! 3^{N_4} \dots} \times C_2^{N_2} C_3^{N_3} C_4^{N_4} \dots \\ &= \sum_{N_2} \frac{C_2^{N_2}}{2^{N_2} N_2!} \sum_{N_3, N_4, \dots} \frac{C_3^{N_3} C_4^{N_4} \dots}{2^{N_3} N_3! 3^{N_4} N_4! \dots} = \exp \left[\frac{C_2}{2} + \sum_{r=3}^{\infty} \frac{C_r}{(r-1)!} \right]. \end{aligned} \quad (3.15)$$

In (3.15), the factor $\frac{(2N_2+3N_3+\dots)!}{N_2! 2^{N_2} N_3! 2^{N_3} N_4! 3^{N_4} \dots}$ is the number of different graphs corresponding to the fixed values of N_i for every $i \geq 2$. In fact, a star graph with $r > 2$ vertices is invariant w.r.t. all the $(r-1)!$ permutations of the external vertices (and a graph with 2 vertices is trivially invariant under the swap of its vertices). Moreover, a graph with N_r connected

components of r vertices is invariant w.r.t. all the $N_r!$ permutations of these connected components. So, we take the quotient of the number of all the permutation of the $\sum_{r \geq 2} r N_r$ vertices over the total number of invariant operations. Therefore, (3.15) provides this bound:

$$\langle e^{H_0} \rangle \leq \exp \left[\frac{C_2}{2} + \sum_{r=3}^{\infty} \frac{C_r}{(r-1)!} \right]. \quad (3.16)$$

By definition (3.13), C_r can be written in this form:

$$C_r = \sum_{\Lambda_1, \dots, \Lambda_r} \sum_{v_2 \in \Lambda_1 \cap \Lambda_2} \cdots \sum_{v_r \in \Lambda_1 \cap \Lambda_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|, \quad (3.17)$$

because $\sum_{v_i \in \Lambda_1 \cap \Lambda_i} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| = |\Lambda_1 \cap \Lambda_i| \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|$. Furthermore, since:

$$\sum_{\Lambda_i} \sum_{v_i \in \Lambda_1 \cap \Lambda_i} \|h_{\Lambda_i}\| = \sum_{v_i \in \Lambda_1} \sum_{\Lambda_i \ni v_i} \|h_{\Lambda_i}\| \quad \forall i = 2, \dots, r, \quad (3.18)$$

by (3.17), we have:

$$C_r = \sum_{\Lambda_1} \sum_{v_2, \dots, v_r \in \Lambda_1} \sum_{\Lambda_2 \ni v_2} \cdots \sum_{\Lambda_r \ni v_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| = \sum_{\Lambda_1} \|h_{\Lambda_1}\| \left(\sum_{v_2 \in \Lambda_1} \sum_{\Lambda_2 \ni v_2} \|h_{\Lambda_2}\| \right)^{r-1}. \quad (3.19)$$

In view of Lemma A.2 we have $\sum_{\Lambda_2 \ni v_2} \|h_{\Lambda_2}\| \leq \|H\|_{\text{loc}}$ and observing that

$$\sum_{v_2 \in \Lambda_1} \|H\|_{\text{loc}} = |\Lambda_1| \|H\|_{\text{loc}}, \quad (3.20)$$

from (3.19) we obtain:

$$C_r \leq \sum_{\Lambda_1} \|h_{\Lambda_1}\| \left(\sum_{v_2 \in \Lambda_1} \sum_{\Lambda_2 \ni v_2} \|h_{\Lambda_2}\| \right)^{r-1} \leq \sum_{\Lambda \subseteq [n]} |\Lambda|^{r-1} \|H\|_{\text{loc}}^{r-1} \|h_{\Lambda}\| \quad \forall r \geq 2, \quad (3.21)$$

then, by substituting in (3.16):

$$\begin{aligned} \langle e^{H_0} \rangle &\leq \exp \left[\frac{C_2}{2} + \sum_{r=3}^{\infty} \sum_{\Lambda \subseteq [n]} \frac{|\Lambda|^{r-1} \|H\|_{\text{loc}}^{r-1} \|h_{\Lambda}\|}{(r-1)!} \right] = \\ &= \exp \left[\frac{C_2}{2} + \sum_{\Lambda \subseteq [n]} \left(e^{|\Lambda| \|H\|_{\text{loc}}} - 1 - |\Lambda| \|H\|_{\text{loc}} \right) \|h_{\Lambda}\| \right] \leq \\ &\leq \exp \left[\sum_{\Lambda \subseteq [n]} \left(e^{|\Lambda| \|H\|_{\text{loc}}} - 1 - \frac{|\Lambda| \|H\|_{\text{loc}}}{2} \right) \|h_{\Lambda}\| \right], \end{aligned} \quad (3.22)$$

where, in the last inequality, we have used the fact $C_2 \leq \sum_{\Lambda \subseteq [n]} |\Lambda| \|H\|_{\text{loc}} \|h_{\Lambda}\|$. The claim follows from $\langle e^H \rangle = e^{\langle H \rangle} \langle e^{H_0} \rangle$. $\square$

The inequality (3.5) stated by the theorem above provides a bound that is not a function of the local norm of H alone. However, such a bound can be obtained with the additional requirement that the size of the regions, on which H acts locally, satisfies $|\Lambda| \leq k$.

Corollary 3.1. *Under the hypotheses of Theorem 3.1, if H is k -local then*

$$\langle e^H \rangle \leq \exp \left[\langle H \rangle + \frac{n}{k} \left(e^{k\|H\|_{\text{loc}}} - 1 - \frac{k\|H\|_{\text{loc}}}{2} \right) \|H\|_{\text{loc}} \right] \quad (3.23)$$

Proof. By Theorem 3.1, we have:

$$\begin{aligned} \langle e^H \rangle &\leq \exp \left[\langle H \rangle + \sum_{\Lambda \subseteq [n]} \sum_{v \in \Lambda} \frac{1}{|\Lambda|} \left(e^{|\Lambda|\|H\|_{\text{loc}}} - 1 - \frac{|\Lambda|\|H\|_{\text{loc}}}{2} \right) \|h_\Lambda\| \right] \\ &\leq \exp \left[\langle H \rangle + \frac{1}{k} \left(e^{k\|H\|_{\text{loc}}} - 1 - \frac{k\|H\|_{\text{loc}}}{2} \right) \sum_{\Lambda \subseteq [n]} \sum_{v \in \Lambda} \|h_\Lambda\| \right] \\ &= \exp \left[\langle H \rangle + \frac{1}{k} \left(e^{k\|H\|_{\text{loc}}} - 1 - \frac{k\|H\|_{\text{loc}}}{2} \right) \sum_{v \in [n]} \sum_{\Lambda \ni v} \|h_\Lambda\| \right] \\ &\leq \exp \left[\langle H \rangle + \frac{1}{k} \left(e^{k\|H\|_{\text{loc}}} - 1 - \frac{k\|H\|_{\text{loc}}}{2} \right) \sum_{v \in [n]} \|H\|_{\text{loc}} \right] \\ &= \exp \left[\langle H \rangle + \frac{1}{k} \left(e^{k\|H\|_{\text{loc}}} - 1 - \frac{k\|H\|_{\text{loc}}}{2} \right) n\|H\|_{\text{loc}} \right]. \end{aligned} \quad (3.24)$$

The first inequality in (3.24) is implied by the hypothesis $|\Lambda| \leq k$ and the second one is a consequence of $\sum_{\Lambda \ni v} \|h_\Lambda\| \leq \|H\|_{\text{loc}}$ for every site v . $\square$

Corollary 3.1 can be used to derive an exponential concentration inequality via the inequality

$$\mathbb{P}(H \geq na) \leq e^{-tna} \langle e^{tH} \rangle \quad \forall t, a > 0 : \quad (3.25)$$

Theorem 3.2. *Let ω be a product state of n spins and H be a k -local Hamiltonian. Then, for any $a > 0$ we have*

$$\mathbb{P}(H \geq \langle H \rangle + na) \leq \exp \left[-\frac{n}{k^2} F \left(\frac{ak}{\|H\|_{\text{loc}}} \right) \right], \quad (3.26)$$

where $F(x) := \max_{s>0} s(x - e^s + 1 + \frac{s}{2})$.

Proof. Without loss of generality, we can assume $\langle H \rangle = 0$. First, we apply Corollary 3.1 inserting (3.23) into (3.25):

$$\mathbb{P}(H \geq na) \leq \exp \left[-tna + \frac{nt}{k} \left(e^{tk\|H\|_{\text{loc}}} - 1 - \frac{tk\|H\|_{\text{loc}}}{2} \right) \|H\|_{\text{loc}} \right]. \quad (3.27)$$

Let us apply the substitution $s = tk\|H\|_{\text{loc}}$:

$$\mathbb{P}(H \geq na) \leq \exp \left[-\frac{n}{k^2} s \left(\frac{ak}{\|H\|_{\text{loc}}} - e^s + 1 + \frac{s}{2} \right) \right] \quad (3.28)$$

and, in turn, perform the minimization over s in the right-hand side of (3.28) obtaining (3.26). $\square$

Remark 3.4. The function F is strictly increasing with $F(0) = 0$, then it is strictly positive for any $x > 0$.

3.3 Quantum states with finite correlation length

We now consider the quantum states with exponentially decaying correlations characterized by a finite correlation length ξ in the sense of Definition 2.4. Also in this case, we consider the decomposition of a local Hamiltonian H that achieves the local norm $H = \sum_{\Lambda \subseteq [n]} \tilde{h}_\Lambda$ and define $h_\Lambda := \tilde{h}_\Lambda - \langle \tilde{h}_\Lambda \rangle$ for any $\emptyset \neq \Lambda \subseteq [n]$. Moreover, under Assumption 2.1, for any region Λ , let us define its *enlargement* by $l > 0$ as $\tilde{\Lambda} := \{x : d(x, \Lambda) \leq l\}$. We can now apply to the enlarged regions the argument that we used for product states. In the new case, for $\tilde{\Lambda}_1 \cap \tilde{\Lambda}_2 = \emptyset$ we have

$$|\langle h_{\Lambda_1} h_{\Lambda_2} \rangle| \leq \|h_{\Lambda_1}\| \|h_{\Lambda_2}\| C e^{-\frac{d(\Lambda_1, \Lambda_2)}{\xi}} \leq \|h_{\Lambda_1}\| \|h_{\Lambda_2}\| C e^{-\frac{l}{\xi}}, \quad (3.29)$$

where $C > 0$ and ξ is the correlation length. Therefore, for a collections of regions $\Lambda_1, \dots, \Lambda_r$ such that $\tilde{\Lambda}_i \cap (\cup_{j \neq i} \tilde{\Lambda}_j) = \emptyset$ for some $i \in [r]$, we have

$$|\langle h_{\Lambda_1} \cdots h_{\Lambda_r} \rangle| \leq \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| C e^{-\frac{l}{\xi}}. \quad (3.30)$$

Theorem 3.3. *Let $\omega \in \mathcal{S}_{[n]}$ be a state of a quantum system of n spins with correlation length ξ as in Definition 2.4 with respect to a distance satisfying Assumption 2.1. Then, for any k -local Hamiltonian H we have*

$$\langle e^H \rangle \leq \exp \left[\langle H \rangle + \frac{n}{k} \left(e^{k(AI^d)^2 \|H\|_{\text{loc}}} - 1 - \frac{k(AI^d)^2 \|H\|_{\text{loc}}}{2} \right) \|H\|_{\text{loc}} \right] + C e^{\langle H \rangle + n \|H\|_{\text{loc}} - \frac{l}{\xi}}, \quad (3.31)$$

where the angle brackets denote the expectation with respect to ω .

Proof. Let $H = \sum_{\Lambda \subseteq [n]} \tilde{h}_\Lambda$ be the decomposition of H that achieves the local norm as in (3.1) and define $h_\Lambda := \tilde{h}_\Lambda - \langle \tilde{h}_\Lambda \rangle$ for any $\emptyset \neq \Lambda \subseteq [n]$ and $H_0 := \sum_{\emptyset \neq \Lambda \subseteq [n]} h_\Lambda$. Then, (3.6) is still valid and $\langle e^H \rangle = e^{\langle H \rangle} \langle e^{H_0} \rangle$. Adapting the inequality (3.8) to enlarged regions $\{\tilde{\Lambda}\}_{\Lambda \subseteq [n]}$ and considering (3.30), the expectation of H_0^r can be bounded as follows:

$$\langle H_0^r \rangle \leq \sum_{(\Lambda_1, \dots, \Lambda_r) \in \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| + \sum_{(\Lambda_1, \dots, \Lambda_r) \notin \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| C e^{-\frac{l}{\xi}}, \quad (3.32)$$

where $\tilde{\mathfrak{R}}_r = \{(\Lambda_1, \dots, \Lambda_r) : \tilde{\Lambda}_i \cap (\cup_{j \neq i} \tilde{\Lambda}_j) \neq \emptyset \forall i = 1, \dots, r\}$. Therefore:

$$\langle e^{H_0} \rangle \leq \sum_{r=0}^{+\infty} \frac{1}{r!} \sum_{(\Lambda_1, \dots, \Lambda_r) \in \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| + \sum_{r=0}^{+\infty} \frac{1}{r!} \sum_{(\Lambda_1, \dots, \Lambda_r) \notin \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| C e^{-\frac{l}{\xi}}. \quad (3.33)$$

Now, let us focus on the sum over $(\Lambda_1, \dots, \Lambda_r) \in \tilde{\mathfrak{R}}_r$ in (3.33) and define:

$$C_r := \sum_{\Lambda_1, \dots, \Lambda_r} |\tilde{\Lambda}_1 \cap \tilde{\Lambda}_2| \cdots |\tilde{\Lambda}_1 \cap \tilde{\Lambda}_r| \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\|. \quad (3.34)$$

so that we can adapt the inequality (3.21):

$$C_r \leq \sum_{\Lambda_1} \|h_{\Lambda_1}\| \left(\sum_{v_2 \in \tilde{\Lambda}_1} \sum_{\Lambda_2: v_2 \in \tilde{\Lambda}_2} \|h_{\Lambda_2}\| \right)^{r-1}. \quad (3.35)$$

In general, we have that $v \in \tilde{\Lambda}$ if and only if $\Lambda \cap B_d(v, l) \neq \emptyset$ where $B_d(v, l)$ is the ball centered in v with radius l referred to the distance $\mathbf{d}$ used to define the enlargement of the regions. Therefore:

$$\sum_{\Lambda: v \in \tilde{\Lambda}} \|h_{\Lambda}\| = \sum_{\Lambda: \Lambda \cap B_d(v, l) \neq \emptyset} \|h_{\Lambda}\| \leq \sum_{\Lambda \subseteq [n]} |\Lambda \cap B_d(v, l)| \|h_{\Lambda}\|, \quad (3.36)$$

the last term in (3.36) can be rewritten as:

$$\sum_{\Lambda \subseteq [n]} \sum_{w \in \Lambda \cap B_d(v, l)} \|h_{\Lambda}\| = \sum_{w \in B_d(v, l)} \sum_{\Lambda \ni w} \|h_{\Lambda}\| \leq |B_d(v, l)| \|H\|_{\text{loc}} \leq A l^d \|H\|_{\text{loc}}, \quad (3.37)$$

where, in the last inequality, we have taken into account **Assumption 2.1**. This results in the following variant of (3.21):

$$C_r \leq \sum_{\Lambda \subseteq [n]} \|h_{\Lambda}\| [(A l^d)^2 |\Lambda| \|H\|_{\text{loc}}]^{r-1} \quad \forall r = 2, \dots, r, \quad (3.38)$$

where we have also used the fact that $|\tilde{\Lambda}| \leq |\Lambda| A l^d$ for any $\Lambda \subseteq [n]$. The latter is a consequence of (2.4) since $\tilde{\Lambda} \subset \cup_{v \in \Lambda} B_d(v, l)$, so $|\tilde{\Lambda}| \leq \sum_{v \in \Lambda} |B_d(v, l)| \leq \sum_{v \in \Lambda} A l^d = |\Lambda| A l^d$. Therefore we can state:

$$\begin{aligned} \sum_{r=0}^{\infty} \frac{1}{r!} \sum_{(\Lambda_1, \dots, \Lambda_r) \in \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| &\leq \\ &\leq \exp \left[\sum_{\Lambda \subseteq [n]} \left(e^{(A l^d)^2 |\Lambda| \|H\|_{\text{loc}}} - 1 - \frac{(A l^d)^2 |\Lambda| \|H\|_{\text{loc}}}{2} \right) \|h_{\Lambda}\| \right], \end{aligned} \quad (3.39)$$

requiring that $|\Lambda| \leq k$ for any $\Lambda \in [n]$, the proof of [Corollary 3.1](#) can be repeated yielding:

$$\begin{aligned} \sum_{r=0}^{\infty} \frac{1}{r!} \sum_{(\Lambda_1, \dots, \Lambda_r) \in \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| &\leq \\ &\leq \exp \left[\frac{n}{k} \left(e^{k(Al^d)^2 \|H\|_{\text{loc}}} - 1 - \frac{k(Al^d)^2 \|H\|_{\text{loc}}}{2} \right) \|H\|_{\text{loc}} \right]. \end{aligned} \quad (3.40)$$

Now, let us bound the sum over $(\Lambda_1, \dots, \Lambda_r) \notin \tilde{\mathfrak{R}}_r$ in [\(3.32\)](#):

$$\sum_{(\Lambda_1, \dots, \Lambda_r) \notin \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| C e^{-\frac{l}{\xi}} \leq \left(\sum_{\Lambda \subseteq [n]} \|h_{\Lambda}\| \right)^r C e^{-\frac{l}{\xi}} \leq (n \|H\|_{\text{loc}})^r C e^{-\frac{l}{\xi}}, \quad (3.41)$$

then:

$$\sum_{r=0}^{\infty} \frac{1}{r!} \sum_{(\Lambda_1, \dots, \Lambda_r) \notin \tilde{\mathfrak{R}}_r} \|h_{\Lambda_1}\| \cdots \|h_{\Lambda_r}\| C e^{-\frac{l}{\xi}} \leq \sum_{r=0}^{+\infty} \frac{(n \|H\|_{\text{loc}})^r}{r!} C e^{-\frac{l}{\xi}} \leq e^{n \|H\|_{\text{loc}}} C e^{-\frac{l}{\xi}}. \quad (3.42)$$

Taking into account [\(3.33\)](#), [\(3.39\)](#), and [\(3.42\)](#) we have:

$$\langle e^{H_0} \rangle \leq \exp \left[\frac{n}{k} \left(e^{k(Al^d)^2 \|H\|_{\text{loc}}} - 1 - \frac{k(Al^d)^2 \|H\|_{\text{loc}}}{2} \right) \|H\|_{\text{loc}} \right] + C e^{n \|H\|_{\text{loc}}} e^{-\frac{l}{\xi}}, \quad (3.43)$$

that directly implies the claim. $\square$

As in the case of product states, [Theorem 3.3](#) implies the following concentration inequality for the quantum states with finite correlation length:

Corollary 3.2. *Let ω be a state of n quantum spins with correlation length ξ as in [Definition 2.4](#) with respect to a distance satisfying [Assumption 2.1](#). Then, for any k -local Hamiltonian H we have*

$$\mathbb{P}(H \geq \langle H \rangle + na) \leq (C + 1) \exp \left[-n^{\frac{1}{2d+1}} \frac{F\left(\frac{ak}{\|H\|_{\text{loc}}}\right)}{k^2 A^2 \left[\frac{\xi}{kA^2} \left(k s^*\left(\frac{ak}{\|H\|_{\text{loc}}}\right) + F\left(\frac{ak}{\|H\|_{\text{loc}}}\right) \right) \right]^{\frac{2d}{2d+1}}} \right], \quad (3.44)$$

where $F(x) := \max_s s(x - e^s + 1 + \frac{s}{2})$ and $s^*(x) := \arg \max_s s(x - e^s + 1 + \frac{s}{2})$.

Proof. Without loss of generality, we can assume $\langle H \rangle = 0$. Inserting the bound [\(3.31\)](#) into the Markov inequality [\(3.25\)](#), we have

$$\begin{aligned} \mathbb{P}(H \geq na) &\leq \exp \left[-tna + \frac{nt}{k} \left(e^{tkA^2 l^{2d} \|H\|_{\text{loc}}} - 1 - \frac{tkA^2 l^{2d} \|H\|_{\text{loc}}}{2} \right) \|H\|_{\text{loc}} \right] \\ &\quad + C \exp \left[nt \|H\|_{\text{loc}} - \frac{l}{\xi} \right]. \end{aligned} \quad (3.45)$$

Following the approach adopted in the proof of [Theorem 3.2](#), once introduced the variable $s = tkA^2l^{2d}\|H\|_{\text{loc}}$, the concentration inequality becomes

$$\mathbb{P}(H \geq na) \leq \exp \left[-\frac{n}{k^2 A^2 l^{2d}} s \left(\frac{ak}{\|H\|_{\text{loc}}} - e^s + 1 + \frac{s}{2} \right) \right] + C \exp \left[\frac{ns}{kA^2 l^{2d}} - \frac{l}{\xi} \right]. \quad (3.46)$$

In addition to the minimization over s the right-hand side of the above inequality can be optimized also w.r.t. l . Instead of performing such a double optimization, let us improve the bound on $\mathbb{P}(H \geq na)$ minimizing over s only the first exponential term in [\(3.46\)](#) with the same strategy of the proof of [Theorem 3.2](#). We have

$$\mathbb{P}(H \geq na) \leq \exp \left[-\frac{n}{k^2 A^2 l^{2d}} F \left(\frac{ak}{\|H\|_{\text{loc}}} \right) \right] + C \exp \left[\frac{n s^* \left(\frac{ak}{\|H\|_{\text{loc}}} \right)}{kA^2 l^{2d}} - \frac{l}{\xi} \right]. \quad (3.47)$$

Now, let us fix the value of l requiring the balance of the two exponential terms:

$$-nF \left(\frac{ak}{\|H\|_{\text{loc}}} \right) = kn s^* \left(\frac{ak}{\|H\|_{\text{loc}}} \right) - \frac{l^{2d+1}kA^2}{\xi}, \quad (3.48)$$

then:

$$l = \left[\frac{\xi}{kA^2} \left(kn s^* \left(\frac{ak}{\|H\|_{\text{loc}}} \right) + nF \left(\frac{ak}{\|H\|_{\text{loc}}} \right) \right) \right]^{\frac{1}{2d+1}}. \quad (3.49)$$

By the substitution into [\(3.47\)](#) we obtain:

$$\mathbb{P}(H \geq na) \leq (C+1) \exp \left[-\frac{n^{\frac{1}{2d+1}} F \left(\frac{ak}{\|H\|_{\text{loc}}} \right)}{k^2 A^2 \left[\frac{\xi}{kA^2} \left(k s^* \left(\frac{ak}{\|H\|_{\text{loc}}} \right) + F \left(\frac{ak}{\|H\|_{\text{loc}}} \right) \right) \right]^{\frac{2d}{2d+1}}} \right], \quad (3.50)$$

which is the claim. $\square$

Remark 3.5. In the proof of [Corollary 3.2](#), we have improved the bound of inequality [\(3.46\)](#) optimizing over s and fixing the value of l by the condition that the two exponential terms are balanced. However, even if these two steps are swapped, the obtained bound is the same as one can check by direct inspection.

4 A transportation-cost inequality

4.1 The k -local quantum W_1 distance

We define the k -local quantum W_1 distance as the distance that quantifies the distinguishability of quantum states with respect to the k -local observables with unit local norm:

Definition 4.1 (k -local quantum W_1 norm). We define the k -local quantum W_1 norm as the norm on $\mathcal{O}_{[n]}^T$ that is dual to the local norm on $\mathcal{O}_{[n]}^{(k)}$: For any $\Delta \in \mathcal{O}_{[n]}^T$,

$$\|\Delta\|_{k-W_1\text{loc}} = \max \left\{ \text{Tr} [\Delta H] : H \in \mathcal{O}_{[n]}^{(k)}, \|H\|_{\text{loc}} \leq 1 \right\}. \quad (4.1)$$

The k -local quantum W_1 distance is the distance on $\mathcal{S}_{[n]}$ induced by the k -local quantum W_1 norm: For any and for any $\rho, \sigma \in \mathcal{S}_{[n]}$

$$\|\rho - \sigma\|_{k-W_1\text{loc}} = \max \left\{ \text{Tr} [(\rho - \sigma) H] : H \in \mathcal{O}_{[n]}^{(k)}, \|H\|_{\text{loc}} \leq 1 \right\}. \quad (4.2)$$

Remark 4.1. Ref. [17] defines a local quantum W_1 norm as the dual of the local norm (3.2):

$$\|\Delta\|_{W_1\text{loc}} = \max \left\{ \text{Tr} [\Delta H] : H \in \mathcal{O}_{[n]}^{(k)}, \|H\|_{\widetilde{\text{loc}}} \leq 1 \right\}. \quad (4.3)$$

The k -local quantum W_1 norm can be recovered from (4.3) by setting to one the penalties associated with the regions containing at most k spins and to infinity the penalties associated with the regions containing more than k spins. Indeed, with such choice the observables $H \in \mathcal{O}_{[n]}$ with $\|H\|_{\widetilde{\text{loc}}} \leq 1$ are exactly the k -local observables with $\|H\|_{\text{loc}} \leq 1$, and the optimizations in (4.3) and (4.1) coincide.

The k -local quantum W_1 norm can be computed with a linear program. (4.1) constitutes the dual program, while the primal program is provided by the following:

Proposition 4.1. *For any $\rho, \sigma \in \mathcal{S}_{[n]}$ we have*

$$\|\rho - \sigma\|_{k-W_1\text{loc}} = \min \left\{ \sum_{x \in [n]} a_x : \|\rho_\Lambda - \sigma_\Lambda\|_1 \leq 2 \sum_{x \in \Lambda} a_x \forall \Lambda \subseteq [n] : |\Lambda| \leq k \right\}. \quad (4.4)$$

Proof. Analogous to [17, Proposition 2.4]. □

Remark 4.2. From Proposition 4.1, the k -local quantum W_1 distance between ρ and σ depends only on the marginals ρ_Λ and σ_Λ on regions Λ containing at most k spins. Therefore, if $\rho_\Lambda = \sigma_\Lambda$ for all such Λ , such distance vanishes even if ρ and σ do not coincide.

4.2 Properties of the k -local quantum W_1 distance

The k -local quantum W_1 distance inherits all the properties of the local quantum W_1 distance of [17]:

- For $n = 1$, it coincides with the trace distance:

Proposition 4.2. *For $n = 1$, $\|\cdot\|_{k-W_1\text{loc}} = \frac{1}{2} \|\cdot\|_1$.*

Proof. Follows from Proposition 4.1. □

- It is lower than the quantum W_1 distance of [5]:

Proposition 4.3. *We have $\|\cdot\|_{k-W_1\text{loc}} \leq \|\cdot\|_{W_1}$.*

Proof. Analogous to [17, Proposition 2.5]. $\square$

- It can be upper bounded by the trace distances between the marginal of the states:

Proposition 4.4. *For any $\rho, \sigma \in \mathcal{S}_{[n]}$ we have*

$$\|\rho - \sigma\|_{k-W_1\text{loc}} \leq \sum_{x \in [n]} \max_{\Lambda \ni x: |\Lambda| \leq k} \frac{\|\rho_\Lambda - \sigma_\Lambda\|_1}{2|\Lambda|} \leq n \max_{\Lambda \subseteq [n]: |\Lambda| \leq k} \frac{\|\rho_\Lambda - \sigma_\Lambda\|_1}{2|\Lambda|}. \quad (4.5)$$

Proof. Analogous to [17, Proposition 2.6]. $\square$

- It is superadditive with respect to the composition of quantum systems and additive with respect to the tensor product:

Proposition 4.5. *For any $\rho, \sigma \in \mathcal{S}_{[n]}$ and any $\Lambda \subset [n]$ we have*

$$\|\rho - \sigma\|_{k-W_1\text{loc}} \geq \|\rho_\Lambda - \sigma_\Lambda\|_{k-W_1\text{loc}} + \|\rho_{\Lambda^c} - \sigma_{\Lambda^c}\|_{k-W_1\text{loc}}. \quad (4.6)$$

Moreover, equality is achieved when $\rho = \rho_\Lambda \otimes \rho_{\Lambda^c}$ and $\sigma = \sigma_\Lambda \otimes \sigma_{\Lambda^c}$.

Proof. Analogous to [17, Proposition 2.7]. $\square$

- It is lower bounded by the sum of the trace distances between the single-spin marginals:

Proposition 4.6. *For any $\rho, \sigma \in \mathcal{S}_{[n]}$ we have*

$$\|\rho - \sigma\|_{k-W_1\text{loc}} \geq \frac{1}{2} \sum_{x \in [n]} \|\rho_x - \sigma_x\|_1. \quad (4.7)$$

Proof. Applying repeatedly [Proposition 4.5](#) we get

$$\|\rho - \sigma\|_{k-W_1\text{loc}} \geq \sum_{x \in [n]} \|\rho_x - \sigma_x\|_{k-W_1\text{loc}} = \frac{1}{2} \sum_{x \in [n]} \|\rho_x - \sigma_x\|_1, \quad (4.8)$$

where the last equality follows from [Proposition 4.2](#). The claim follows. $\square$

- It recovers the Hamming distance for the states of the computational basis:

Corollary 4.1. *For any $x, y \in [q]^n$,*

$$\| |x\rangle\langle x| - |y\rangle\langle y| \|_{k-W_1\text{loc}} = h(x, y) = |\{i \in [n] : x_i \neq y_i\}|. \quad (4.9)$$

Proof. Analogous to [17, Corollary 2.1]. $\square$

- It is contracting with respect to the action of single-spin quantum channels:

Proposition 4.7. *For any $\Delta \in \mathcal{O}_n^T$ and any quantum channel Φ acting on a single spin, we have*

$$\|\Phi(\Delta)\|_{k-W_1\text{loc}} \leq \|\Delta\|_{k-W_1\text{loc}}. \quad (4.10)$$

Proof. Analogous to [17, Proposition 2.8]. $\square$

4.3 A transportation-cost inequality

The concentration inequality [Theorem 3.3](#) can be applied to prove that any quantum state σ with exponentially decreasing correlations, as in [Definition 2.4](#), satisfies a transportation-cost inequality of the following form:

$$f(\|\rho - \sigma\|_{k-W_1\text{loc}}) \leq S_M(\rho\|\sigma) \quad \forall \rho \in \mathcal{S}_{[n]} \quad (4.11)$$

where f is an increasing function and S_M is the measured relative entropy.

Definition 4.2. Let ρ and σ be density matrices on a finite-dimensional Hilbert space $\mathcal{H}$. The *measured relative entropy* between ρ and σ is defined as:

$$S_M(\rho\|\sigma) := \sup_{\{A_a\}_{a \in X}} \sum_{a \in X} \text{Tr}(A_a \rho) \log \frac{\text{Tr}(A_a \rho)}{\text{Tr}(A_a \sigma)}, \quad (4.12)$$

where the supremum is taken over all the POVMs $\{A_a\}_{a \in X}$ on $\mathcal{H}$.

Let us recall that a POVM $\{A_a\}_{a \in X}$ on $\mathcal{H}$, where X is a finite set, is a collection of positive semidefinite operators such that $\sum_{a \in X} A_a = I$ and provides the general mathematical description of a quantum measurement process. Intuitively, the notion of measured relative entropy is given by the maximization, over all the measurement processes, of the Kullback-Leibler divergence between the classical probability distributions of the measurement outcomes associated to the considered states. The measured relative entropy admits the following characterization [\[18\]](#):

Lemma 4.1. *The following identity holds:*

$$S_M(\rho\|\sigma) = \sup_{A \geq 0} (\text{Tr}(\rho \log A) - \log \text{Tr}(\sigma A)). \quad (4.13)$$

The variational expression [\(4.13\)](#) can be used to bound from below the measured relative entropy by means of a function of the k -local quantum W_1 distance obtaining a transportation cost inequality.

Theorem 4.1. Let ρ and σ be states of n quantum spins. Let σ have correlation length ξ as in [Definition 2.4](#) with respect to a distance satisfying [Assumption 2.1](#). Then, the measured relative entropy between ρ and σ is bounded as follows:

$$S_M(\rho\|\sigma) \geq n^{\frac{1}{2d+1}} f(w) - \log(C+1), \quad (4.14)$$

where c is a positive constant, $w = \frac{1}{n} \|\rho - \sigma\|_{k-W_1\text{loc}}$, and

$$f(w) = \frac{F(kw)}{k^2 A^2 \|H\|_{\text{loc}} \left[\frac{\xi}{k A^2 \|H\|_{\text{loc}}} (k s^*(kw) + F(kw)) \right]^{\frac{2d}{2d+1}}}, \quad (4.15)$$

where $F(x) := \max_s s(x - e^s + 1 + \frac{s}{2})$ and $s^*(x) = \arg\max_s s(x - e^s + 1 + \frac{s}{2})$.

Proof. We apply the claim of [Lemma 4.1](#) for $E = e^H$ where $H = \sum_{\Lambda \subseteq [n]} h_\Lambda$ is a self-adjoint operator given by the sum of local terms h_Λ acting non trivially on $\Lambda \subseteq [n]$ such that $\langle h_\Lambda \rangle = 0$ and $|\Lambda| \leq k \forall \Lambda$, where the angle brackets denote the expectation with respect to σ . Therefore:

$$S_M(\rho\|\sigma) \geq \text{Tr}(\rho H) - \log \langle e^H \rangle. \quad (4.16)$$

We can choose the Hamiltonian H such that $\text{Tr}[(\rho - \sigma)H] = \|H\|_{\text{loc}} \|\rho - \sigma\|_{k-W_1\text{loc}}$ using the fact that $\|\cdot\|_{k-W_1\text{loc}}$ is the dual of $\|\cdot\|_{\text{loc}}$ as provided by [\(4.2\)](#). Such a Hamiltonian can be always chosen in such a way that $\langle \sigma \rangle = 0$, since $\|H\|_{\text{loc}} = \|H - \langle \sigma \rangle I\|_{\text{loc}}$ as $\|\cdot\|_{\text{loc}}$ is a semi-norm with $\|I\|_{\text{loc}} = 0$. Moreover, by the arbitrary choice of the Hamiltonian, we can control the value of $\|H\|_{\text{loc}}$. From [\(4.16\)](#), we have:

$$S_M(\rho\|\sigma) \geq \text{Tr}[(\rho - \sigma)H] - \log \langle e^H \rangle \geq \|H\|_{\text{loc}} \|\rho - \sigma\|_{k-W_1\text{loc}} - \log \langle e^H \rangle, \quad (4.17)$$

Let us set $\|H\|_{\text{loc}} = t$ and $\|\rho - \sigma\|_{k-W_1\text{loc}} = W$, then:

$$S_M(\rho\|\sigma) \geq tW - \log \langle e^H \rangle. \quad (4.18)$$

Taking the exponential of [\(4.18\)](#) we obtain:

$$e^{-S_M(\rho\|\sigma)} \leq e^{-tW} \langle e^H \rangle = e^{-tnw} \langle e^H \rangle. \quad (4.19)$$

This bound can be improved optimizing w.r.t. t following the argument used in the proof of [Corollary 3.2](#). More precisely, we can apply [Theorem 3.3](#) to bound $\langle e^H \rangle$ taking into account that $\langle H \rangle = 0$, thus:

$$e^{-S_M(\rho\|\sigma)} \leq \exp \left[-\frac{n}{k^2 A^2 l^{2d} \|H\|_{\text{loc}}} s \left(kw - e^s + 1 + \frac{s}{2} \right) \right] + C \exp \left[\frac{ns}{k A^2 l^{2d}} - \frac{l}{\xi} \right], \quad (4.20)$$

where $s = tk A^2 l^{2d} \|H\|_{\text{loc}}$. As in the proof of [Corollary 3.2](#), we can improve the bound minimizing over s and requiring the balance of the two exponential terms to fix l , this yields:

$$e^{-S_M(\rho\|\sigma)} \leq (C+1) \exp \left[-\frac{n^{\frac{1}{2d+1}} F(kw)}{k^2 A^2 \|H\|_{\text{loc}} \left[\frac{\xi}{k A^2 \|H\|_{\text{loc}}} (k s^*(kw) + F(kw)) \right]^{\frac{2d}{2d+1}}} \right]. \quad (4.21)$$

We then obtain the inequality

$$\exp[-S_M(\rho||\sigma)] \leq \exp\left[\log(C+1) - n^{\frac{1}{2d+1}} f(w)\right], \quad (4.22)$$

which implies the claim. □

5 Equivalence of the ensembles of quantum statistical mechanics

The three main ensembles employed in quantum statistical mechanics to compute the equilibrium properties of quantum systems are the canonical ensemble, the microcanonical ensemble and the diagonal ensemble. The quantum state associated to the canonical ensemble is the Gibbs state, which describes the physics of a system that is at thermal equilibrium with a large bath at a given temperature. The diagonal and microcanonical ensembles both describe the physics of an isolated quantum system, and the associated states are convex combinations of the eigenstates of the Hamiltonian. The microcanonical ensemble assumes a uniform probability distribution for the energy in a given energy shell. The diagonal ensemble includes all the states that are diagonal in the eigenbasis of the Hamiltonian, and in particular it includes the eigenstates themselves. For many quantum systems, the canonical and microcanonical ensembles give the same expectation values for local observables if the corresponding states have approximately the same average energy. A lot of effort has been devoted to determining conditions under which the two ensembles are equivalent [19–22]. The most prominent among such conditions are short-ranged interactions and a finite correlation length, but so far analytical proofs could be obtained only in the case of regular lattices [22]. The situation is more complex considering also the diagonal ensemble which includes all the states that are diagonal in the eigenbasis of the Hamiltonian. The condition under which this ensemble is equivalent to the microcanonical and canonical ensembles is called Eigenstate Thermalization Hypothesis (ETH) [23–27], and states that the expectation values of local observables on the eigenstates of the Hamiltonian are a smooth function of the energy, *i.e.*, for any given local observable, *any* two eigenstates with approximately the same energy yield approximately the same expectation value. The ETH is an extremely strong condition on the Hamiltonian and several quantum systems, including all integrable systems, do not satisfy it. A weak version of the ETH has been formulated [12, 28], stating that for any given local observable, *most* eigenstates in an energy shell yield approximately the same expectation value, or, more precisely, that the fraction of eigenstates yielding expectation values far from the Gibbs state with the same average energy vanishes in the thermodynamical limit. The weak ETH implies the equivalence between the canonical and microcanonical ensembles, but is not sufficient to prove their equivalence with the diagonal ensemble. Under the hypothesis of finite correlation length in the Gibbs state, an analytical proof of the weak ETH is available only for regular lattices [12].

In this paper, we consider the equivalence of the statistical mechanical ensembles and the weak ETH from the perspective of the k -local quantum W_1 distance. The closeness in such distance implies closeness of the expectation values of all the k -local observables with $O(1)$ local norm. We stress that such observables do not need to be geometrically local, since they can contain terms acting on spins at arbitrary distance. Our results are based on the transportation-cost inequality [Theorem 4.1](#) and are similar in spirit to the results of [\[6, Sec. 8\]](#), which considers the equivalence of the ensembles with respect to the quantum W_1 distance of [\[5\]](#) employing transportation-cost inequalities for such distance. The key difference between the present paper and Ref. [\[6\]](#) is that the k -local quantum W_1 distance that we employ is weaker than the quantum W_1 distance of [\[6\]](#). As a consequence, our [Theorem 4.1](#) requires only exponentially decaying correlations. On the contrary, the transportation-cost inequalities employed in [\[6\]](#) require a Hamiltonian that is the sum of local commuting terms (the subsequent Ref. [\[29\]](#) has proved transportation-cost inequalities for the quantum W_1 distance under the hypothesis of exponential decay of the conditional mutual information, which is however still stronger than exponential decay of correlations). Therefore, we are able to prove the equivalence of the ensembles for a significantly larger class of Hamiltonians than the class for which the results of [\[6, Sec. 8\]](#) apply, at the price of a weaker notion of equivalence. We also stress that, contrarily to the results of Refs. [\[12, 22\]](#), we do not need to restrict to regular lattices, and our results do not require any regularity nor the notion of neighboring spins, but only the upper bound [Assumption 2.1](#) to the growth of the size of the metric balls with the distance.

Let $H \in \mathcal{O}_{[n]}$ be the Hamiltonian of a quantum system of n spins. A *Gibbs state* for the Hamiltonian H is:

$$\omega := \frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})}, \quad (5.1)$$

where β is the inverse temperature. The following [Proposition 5.1](#) implies that any state $\rho \in \mathcal{S}_{[n]}$ is close in the k -local W_1 distance to the Gibbs state ω with the same average energy, provided that ω has finite correlation length and approximately the same entropy as ρ , *i.e.*,

$$S(\omega) - S(\rho) = o\left(n^{\frac{1}{2d+1}}\right). \quad (5.2)$$

Moreover, under the same hypothesis, the average reduced states over one spin of ρ and ω are close in trace distance.

Proposition 5.1. *Let $\omega \in \mathcal{S}_{[n]}$ be a Gibbs state for the Hamiltonian $H \in \mathcal{O}_{[n]}$ with correlation length ξ as in [Definition 2.4](#) with respect to a distance satisfying [Assumption 2.1](#). Then, any quantum state $\rho \in \mathcal{S}_{[n]}$ with the same average energy as ω satisfies*

$$\frac{1}{n} \|\rho - \omega\|_{k-W_1\text{loc}} \leq f^{-1}\left(\frac{S(\omega) - S(\rho)}{n^{\frac{1}{2d+1}}} + c\right), \quad (5.3)$$

with f as in [\(4.15\)](#). Moreover, let $\Lambda : \mathcal{S}_{[n]} \rightarrow \mathcal{S}(\mathbb{C}^d)$ be the quantum channel that computes

the average marginal state over one spin, i.e., for any $\rho \in \mathcal{S}_{[n]}$,

$$\Lambda(\rho) = \frac{1}{n} \sum_{x \in [n]} \rho_x. \quad (5.4)$$

Then,

$$\frac{1}{2} \|\Lambda(\rho) - \Lambda(\omega)\|_1 \leq f^{-1} \left(\frac{S(\omega) - S(\rho)}{n^{\frac{1}{2d+1}}} + c \right). \quad (5.5)$$

Proof. We have

$$S_M(\rho \parallel \omega) \leq S(\rho \parallel \omega) = S(\omega) - S(\rho), \quad (5.6)$$

where the last equality follows since ρ and ω have the same average energy. We then have from [Theorem 4.1](#)

$$\frac{1}{n} \|\rho - \omega\|_{k-W_1 \text{loc}} \leq f^{-1} \left(\frac{S_M(\rho \parallel \omega)}{n^{\frac{1}{2d+1}}} + c \right) \leq f^{-1} \left(\frac{S(\omega) - S(\rho)}{n^{\frac{1}{2d+1}}} + c \right). \quad (5.7)$$

We have from [Proposition 4.6](#) and [Proposition 5.1](#)

$$\|\Lambda(\rho) - \Lambda(\omega)\|_1 \leq \frac{1}{n} \sum_{x \in [n]} \|\rho_x - \omega_x\|_1 \leq \frac{2}{n} \|\rho - \omega\|_{k-W_1 \text{loc}} \leq 2 f^{-1} \left(\frac{S(\omega) - S(\rho)}{n^{\frac{1}{2d+1}}} + c \right), \quad (5.8)$$

and the claim follows. $\square$

Choosing ρ to be diagonal in the eigenbasis of the Hamiltonian, [Proposition 5.1](#) implies that any convex combination of a sufficiently large number of eigenstates is close in k -local W_1 distance to the Gibbs state with the same average energy. Such number of eigenstates can be a fraction $\exp\left(-o\left(n^{\frac{1}{2d+1}}\right)\right)$ of the total number of eigenstates appearing in a microcanonical state. Therefore, [Proposition 5.1](#) constitutes an exponential improvement over the weak ETH.

5.1 Width of the microcanonical energy shell

We can apply the transportation-cost inequality provided by [Theorem 4.1](#) to prove the convergence of the microcanonical ensemble to the canonical ensemble in the local W_1 distance as $n \rightarrow +\infty$. Let $H = \sum_E E P(E)$ be the spectral decomposition of H . The *microcanonical state* with energy E and width Δ is

$$\omega_{E,\Delta} := \frac{P(E, \Delta)}{\text{Tr}(P(E, \Delta))}, \quad (5.9)$$

where $P(E, \Delta)$ is the orthogonal projector onto the subspace spanned by the eigenvectors of H with eigenvalues in $(E - \Delta, E]$.

Theorem 5.1. Let ω be a Gibbs state for the Hamiltonian $H \in \mathcal{O}_{[n]}$ with correlation length ξ as in [Definition 2.4](#) with respect to a distance satisfying [Assumption 2.1](#) and let $\omega_{E^*, \Delta}$ be the microcanonical state with energy

$$E^* = \operatorname{argmax}_E e^{-\beta E} \operatorname{Tr} P(E, \Delta). \quad (5.10)$$

Then:

$$\frac{1}{n} \|\omega_{E^*, \Delta} - \omega\|_{k\text{-}W_1 \text{loc}} = o_{n \rightarrow \infty}(1), \quad (5.11)$$

whenever $\Delta = \exp\left(-o\left(n^{\frac{1}{2d+1}}\right)\right)$ and $\Delta = o\left(n^{\frac{1}{2d+1}}\right)$.

Proof. [Theorem 4.1](#) provides a bound on the measured relative entropy as a function of the k -local W_1 distance. Since $S(\omega_{E, \Delta} \| \omega) \geq S_M(\omega_{E, \Delta} \| \omega)$ let us control the relative entropy between the Gibbs and the microcanonical states:

$$S(\omega_{E^*, \Delta} \| \omega) = \ln \frac{\operatorname{Tr}[e^{-\beta H}]}{\operatorname{Tr}[P(E^*, \Delta)]} + \beta \ln \left[H \frac{P(E^*, \Delta)}{\operatorname{Tr}[P(E^*, \Delta)]} \right] \leq \ln \frac{\operatorname{Tr}[e^{-\beta H}]}{\operatorname{Tr}[P(E^*, \Delta)]} + \beta E^*. \quad (5.12)$$

Let us observe that:

$$\operatorname{Tr} \left[\sum_E \frac{e^{-\beta E}}{\operatorname{Tr}[e^{-\beta H}]} P(E) \right] = 1, \quad (5.13)$$

where the sum over E is taken over all the spectrum of H . Let us consider the partition function:

$$Z := \operatorname{Tr} \left[\sum_E e^{-\beta E} P(E) \right] = \operatorname{Tr}[e^{-\beta H}], \quad (5.14)$$

and define the interval $I = (-\|H\| - \Delta, \|H\|]$, then:

$$Z \leq \sum_{\nu \in \mathbb{Z} : \nu \Delta \in I} \operatorname{Tr}[P(\nu \Delta, \Delta)] e^{-\beta \Delta(\nu-1)}. \quad (5.15)$$

By definition of E^* , we have:

$$\operatorname{Tr}[P(\nu \Delta, \Delta)] e^{-\beta \Delta(\nu-1)} \leq e^{\beta(\Delta - E^*)} \operatorname{Tr}[P(E^*, \Delta)], \quad (5.16)$$

therefore:

$$Z \leq \frac{2\|H\| + \Delta}{\Delta} e^{\beta(\Delta - E^*)} \operatorname{Tr}[P(E^*, \Delta)]. \quad (5.17)$$

From equations [\(5.14\)](#) and [\(5.17\)](#), we obtain:

$$\frac{\operatorname{Tr}[e^{-\beta H}]}{\operatorname{Tr}[P(E^*, \Delta)]} \leq \left(\frac{2\|H\|}{\Delta} + 1 \right) e^{\beta(\Delta - E^*)}, \quad (5.18)$$

that is:

$$\ln \left[\frac{\operatorname{Tr}[e^{-\beta H}]}{\operatorname{Tr}[P(E^*, \Delta)]} \right] \leq \ln \left[\frac{2\|H\|}{\Delta} + 1 \right] + \beta(\Delta - E^*). \quad (5.19)$$

Without loss of generality, we can add a constant to H such that its minimum and maximum eigenvalues become opposite. Then, $\|H\| \leq \frac{n}{2}\|H\|_{\text{loc}}$ by [Lemma A.1](#). Inserting [\(5.19\)](#) into [\(5.12\)](#), we have:

$$S(\omega_{E^*,\Delta}||\omega) \leq \ln \left[\frac{n\|H\|_{\text{loc}}}{\Delta} + 1 \right] + \beta\Delta. \quad (5.20)$$

Applying [Theorem 4.1](#) to bound $S(\omega_{E^*,\Delta}||\omega)$ from below, we obtain:

$$\ln \left[\frac{n\|H\|_{\text{loc}}}{\Delta} + 1 \right] + \beta\Delta + c \geq n^{\frac{1}{2d+1}} f(w), \quad (5.21)$$

where $c > 0$ and f is the positive increasing function introduced in [\(4.14\)](#). Therefore:

$$w \leq f^{-1} \left(n^{-\frac{1}{2d+1}} \ln \left[\frac{n\|H\|_{\text{loc}}}{\Delta} + 1 \right] + n^{-\frac{1}{2d+1}} \beta\Delta + n^{-\frac{1}{2d+1}} c \right). \quad (5.22)$$

The right-hand side of the inequality above is $o(1)$ whenever $\ln(n/\Delta) = o\left(n^{\frac{1}{2d+1}}\right)$, that is $\Delta = n \exp\left(-o\left(n^{\frac{1}{2d+1}}\right)\right) = \exp\left(-o\left(n^{\frac{1}{2d+1}}\right)\right)$, and $\Delta = o\left(n^{\frac{1}{2d+1}}\right)$. The claim follows. $\square$

A similar result on the equivalence between the quantum canonical and microcanonical ensembles under the hypothesis of exponentially decaying correlations is proved in [\[12\]](#). Such results states that for any local observable $O = \frac{1}{n} \sum_{v \in [n]} O_v$, where each O_v is supported on the ball centered in v with fixed radius and $\|O_v\| \leq 1$, the expectations over the Gibbs and the microcanonical states converges to the same value as $n \rightarrow +\infty$ whenever $\Delta = \exp\left(-O\left(n^{\frac{1}{d+1}}\right)\right)$. While our result poses a stronger requirement on the width of the energy shell, we do not have any requirement on the radius of the regions where the local terms of O are supported, since we require such regions only to have bounded cardinality.

6 Conclusions

In this paper we have proved new concentration inequalities for quantum spin systems and we have applied such inequalities to prove the equivalence between the canonical and microcanonical ensembles of quantum statistical mechanics.

First, we have introduced the *local norm* [\(3.1\)](#) for observables of quantum spin systems. Then, we have proved a concentration inequality for local observables whose local norm is $O(1)$ measured on product states ([Theorem 3.2](#)), stating that the probability of deviations from the average decays exponentially with the number of spins. We have extended the result to the states with finite correlation length as in [Definition 2.4](#), proving that the probability of deviations from the average decays quasi exponentially with the number of spins ([Corollary 3.2](#)). Remarkably, our results do not require any regular structure nor any notion of neighboring spins.

Moreover, we have defined the k -local quantum W_1 distance as the distance on the states of a quantum spin system that quantifies the distinguishability with respect to k -local observables. We have proved a transportation-cost inequality stating that the k -local quantum

W_1 distance between a generic state and a state with exponentially decaying correlations is upper bounded by a function of the relative entropy between the states ([Theorem 4.1](#)). Such inequality implies that any Gibbs state ω with exponentially decaying correlations is close w.r.t. the k -local W_1 distance, to any state ρ with the same average energy and such that $S(\omega) - S(\rho) = o\left(n^{\frac{1}{2d+1}}\right)$ ([Proposition 5.1](#)), where d is the power of the radius governing the growth of the volume of the metric balls of the set of the spins. This result implies the equivalence between the canonical and the microcanonical ensembles of quantum statistical mechanics for all the Hamiltonians whose Gibbs states have exponentially decaying correlations and all the microcanonical states whose energy shell is wide enough. Furthermore, [Proposition 5.1](#) implies that any convex combination of eigenstates which constitute a fraction $\exp\left(-o\left(n^{\frac{1}{2d+1}}\right)\right)$ of the total number of eigenstates in a microcanonical shell is close in the k -local W_1 distance to the Gibbs state. This is an exponential improvement over the weak ETH. Our results do not require any regular structure nor any notion of neighboring spins, and hold at any temperature provided the correlations in the Gibbs state decay exponentially with the distance.

The main open problem we point out concerns the strong ETH, stating that the energy eigenstate themselves are close to the Gibbs states with the same average energy with respect to a suitable class of observables. From [\[6, Sec. 8\]](#), the strong ETH cannot be captured by the quantum W_1 distance of [\[5\]](#). Indeed, the von Neumann entropy per spin is continuous with respect to the W_1 distance per spin [\[30\]](#), therefore no pure state can be close to any state with large entropy such a Gibbs state. On the contrary, we do not expect such entropy continuity to hold for the k -local quantum W_1 distance, and therefore there is hope that this distance can capture the strong ETH.

Acknowledgements

This work was partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU.

GDP has been supported by the HPC Italian National Centre for HPC, Big Data and Quantum Computing - Proposal code CN00000013 - CUP J33C22001170001 and by the Italian Extended Partnership PE01 - FAIR Future Artificial Intelligence Research - Proposal code PE00000013 - CUP J33C22002830006 under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU. Funded by the European Union - NextGenerationEU under the National Recovery and Resilience Plan (PNRR) - Mission 4 Education and research - Component 2 From research to business - Investment 1.1 Notice Prin 2022 - DD N. 104 del 2/2/2022, from title “understanding the LEarning process of QUantum Neural networks (LeQun)”, proposal code 2022WHZ5XH – CUP J53D23003890006. GDP and DP are members of the “Gruppo Nazionale per la Fisica Matematica (GNFM)” of the “Istituto Nazionale di Alta Matematica “Francesco Severi” (INdAM)”.

A Auxiliary lemmas

Lemma A.1. *Let $H \in \mathcal{O}_{[n]}$ such that its minimum and maximum eigenvalues are opposite. Then, $2 \|H\| \leq n \|H\|_{\text{loc}}$.*

Proof. Let

$$H = \sum_{\Lambda \subseteq [n]} h_{\Lambda}, \quad h_{\Lambda} \in \mathcal{O}_{\Lambda} \quad (\text{A.1})$$

such that

$$\|H\|_{\text{loc}} = 2 \max_{x \in [n]} \sum_{\Lambda \ni x} \|h_{\Lambda}\|. \quad (\text{A.2})$$

We have

$$\begin{aligned} \|H\| &\leq \|H - h_{\emptyset}\| \leq \sum_{\Lambda \subseteq [n]: \Lambda \neq \emptyset} \|h_{\Lambda}\| \leq \sum_{\Lambda \subseteq [n]} |\Lambda| \|h_{\Lambda}\| = \sum_{\Lambda \subseteq [n]} \sum_{x \in \Lambda} \|h_{\Lambda}\| \\ &= \sum_{x=1}^n \sum_{\Lambda \ni x} \|h_{\Lambda}\| \leq \frac{n}{2} \|H\|_{\text{loc}}. \end{aligned} \quad (\text{A.3})$$

The claim follows. $\square$

Lemma A.2. *Let $H \in \mathcal{O}_{[n]}$ and let $H = \sum_{\Lambda \subseteq [n]} \tilde{h}_{\Lambda}$ be the decomposition of H that achieves the local norm as in (3.1). For any $\emptyset \neq \Lambda \subseteq [n]$, let $h_{\Lambda} := \tilde{h}_{\Lambda} - \langle \tilde{h}_{\Lambda} \rangle$, where the angle brackets denote the expectation with respect to a generic state. Then,*

$$\sum_{\Lambda \ni v} \|h_{\Lambda}\| \leq \|H\|_{\text{loc}} \quad \forall v \in [n]. \quad (\text{A.4})$$

Proof. By definition of h_{Λ} , we have $\|h_{\Lambda}\| \leq \|\tilde{h}_{\Lambda}\| + \left| \langle \tilde{h}_{\Lambda} \rangle \right|$ for every $\emptyset \neq \Lambda \subseteq [n]$ then

$$\sum_{\Lambda \ni v} \|h_{\Lambda}\| \leq \sum_{\Lambda \ni v} \|\tilde{h}_{\Lambda}\| + \sum_{\Lambda \ni v} \left| \langle \tilde{h}_{\Lambda} \rangle \right| \leq 2 \sum_{\Lambda \ni v} \|\tilde{h}_{\Lambda}\| \leq \|H\|_{\text{loc}} \quad \forall v \in [n], \quad (\text{A.5})$$

where we have respectively used: the triangle inequality, the operator norm inequality $\|\tilde{h}_{\Lambda}\| \geq \langle \tilde{h}_{\Lambda} \rangle$, and $\|H\|_{\text{loc}} = 2 \max_{v \in [n]} \sum_{\Lambda \ni v} \|\tilde{h}_{\Lambda}\|$. $\square$

References

- [1] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013.
- [2] D Goderis and P Vets. Central limit theorem for mixing quantum systems and the ccr-algebra of fluctuations. *Communications in Mathematical Physics*, 122:249–265, 1989.

- [3] Michael Hartmann, Günter Mahler, and Ortwin Hess. Existence of temperature on the nanoscale. *Physical review letters*, 93(8):080402, 2004.
- [4] Nilin Abrahamsen. Short proof of a spectral chernoff bound for local hamiltonians. *arXiv preprint arXiv:2009.04993*, 2020.
- [5] Giacomo De Palma, Milad Marvian, Dario Trevisan, and Seth Lloyd. The Quantum Wasserstein Distance of Order 1. *IEEE Transactions on Information Theory*, 67(10):6627–6643, 2021.
- [6] Giacomo De Palma and Cambyse Rouzé. Quantum Concentration Inequalities. *Annales Henri Poincaré*, 23:3391–3429, 2022.
- [7] Tomotaka Kuwahara and Keiji Saito. Gaussian concentration bound and Ensemble equivalence in generic quantum many-body systems including long-range interactions. *Annals of Physics*, 421:168278, 2020.
- [8] Anurag Anshu. Concentration bounds for quantum states with finite correlation length on quantum spin lattice systems. *New Journal of Physics*, 18(8):083011, 2016.
- [9] Fernando GSL Brandao and Marcus Cramer. Equivalence of statistical mechanical ensembles for non-critical quantum systems. *arXiv preprint arXiv:1502.03263*, 2015.
- [10] Hal Tasaki. On the local equivalence between the canonical and the microcanonical ensembles for quantum spin systems. *Journal of Statistical Physics*, 172:905–926, 2018.
- [11] Álvaro M Alhambra. Quantum many-body systems in thermal equilibrium. *PRX Quantum*, 4(4):040201, 2023.
- [12] Tomotaka Kuwahara and Keiji Saito. Eigenstate thermalization from the clustering property of correlation. *Physical review letters*, 124(20):200604, 2020.
- [13] Cambyse Rouzé and Nilanjana Datta. Concentration of quantum states from quantum functional and transportation cost inequalities. *Journal of Mathematical Physics*, 60(1):012202, 2019.
- [14] Eric A Carlen and Jan Maas. Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems. *Journal of Statistical Physics*, 178(2):319–378, 2020.
- [15] Li Gao, Marius Junge, and Nicholas LaRacuente. Fisher information and logarithmic sobolev inequality for matrix-valued functions. *Annales Henri Poincaré*, 21(11):3409–3478, 2020.
- [16] Giacomo De Palma and Dario Trevisan. Quantum optimal transport: Quantum channels and qubits. *arXiv:2307.16268*, 2023.

- [17] Giacomo De Palma, Tristan Klein, and Davide Pastorello. Classical shadows meet quantum optimal mass transport. *arXiv:2309.08426*, 2023.
- [18] Mario Berta, Omar Fawzi, and Marco Tomamichel. On variational expressions for quantum relative entropies. *Lett Math Phys*, 107:2239–2265, 2017.
- [19] R Lima. Equivalence of ensembles in quantum lattice systems. *Annales de l’IHP Physique théorique*, 15(1):61–68, 1971.
- [20] R Lima et al. Equivalence of ensembles in quantum lattice systems: states. *Communications in Mathematical Physics*, 24(3):180–192, 1972.
- [21] Markus P Müller, Emily Adlam, Lluís Masanes, and Nathan Wiebe. Thermalization and canonical typicality in translation-invariant quantum lattice systems. *Communications in Mathematical Physics*, 340(2):499–561, 2015.
- [22] Fernando G. S. L. Brandão and Marcus Cramer. Equivalence of Statistical Mechanical Ensembles for Non-Critical Quantum Systems. *arXiv:1502.03263 [cond-mat, physics:quant-ph]*, February 2015. arXiv: 1502.03263.
- [23] Josh M Deutsch. Quantum statistical mechanics in a closed system. *Physical Review A*, 43(4):2046, 1991.
- [24] Mark Srednicki. Chaos and quantum thermalization. *Physical Review E*, 50(2):888–901, August 1994.
- [25] Christian Gogolin and Jens Eisert. Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems. *Reports on Progress in Physics*, 79(5):056001, 2016.
- [26] Peter Reimann. Dynamical typicality approach to eigenstate thermalization. *Physical review letters*, 120(23):230601, 2018.
- [27] Giacomo De Palma, Alessio Serafini, Vittorio Giovannetti, and Marcus Cramer. Necessity of eigenstate thermalization. *Physical Review Letters*, 115(22):220401, 2015.
- [28] Giulio Biroli, Corinna Kollath, and Andreas M Läuchli. Effect of rare fluctuations on the thermalization of isolated quantum systems. *Physical review letters*, 105(25):250401, 2010.
- [29] Emilio Onorati, Cambyse Rouzé, Daniel Stilck França, and James D. Watson. Efficient learning of ground & thermal states within phases of matter. *arXiv:2301.12946*, 2023.
- [30] Giacomo De Palma and Dario Trevisan. The Wasserstein distance of order 1 for quantum spin systems on infinite lattices. *Annales Henri Poincaré*, 24(12):4237–4282, 6 2023.