Annales Henri Poincaré

The Wasserstein Distance of Order 1 for Quantum Spin Systems on Infinite Lattices

Giacomo De Palma◉ and Dario Trevisan

Abstract. We propose a generalization of the Wasserstein distance of order 1 to quantum spin systems on the lattice $\mathbb{Z}^d$, which we call specific quantum W_1 distance. The proposal is based on the W_1 distance for qudits of De Palma et al. (IEEE Trans Inf Theory 67(10):6627–6643, 2021) and recovers Ornstein's $\bar{d}$-distance for the quantum states whose marginal states on any finite number of spins are diagonal in the canonical basis. We also propose a generalization of the Lipschitz constant to quantum interactions on $\mathbb{Z}^d$ and prove that such quantum Lipschitz constant and the specific quantum W_1 distance are mutually dual. We prove a new continuity bound for the von Neumann entropy for a finite set of quantum spins in terms of the quantum W_1 distance, and we apply it to prove a continuity bound for the specific von Neumann entropy in terms of the specific quantum W_1 distance for quantum spin systems on $\mathbb{Z}^d$. Finally, we prove that local quantum commuting interactions above a critical temperature satisfy a transportation-cost inequality, which implies the uniqueness of their Gibbs states.

1. Introduction

Let $\mathcal{X}$ be a finite set endowed with the distance D and let μ and ν be probability distributions on $\mathcal{X}$. A *coupling* between μ and ν is a probability distribution on two copies of $\mathcal{X}$ with marginal distributions equal to μ and ν, respectively. The theory of optimal mass transport considers μ and ν as distributions of a unit amount of mass, and any coupling π prescribes a plan to transform the distribution μ into the distribution ν, i.e., $\pi(x, y)$ is the amount of mass that is moved from x to y for any $x, y \in \mathcal{X}$. Assuming that the cost of moving a unit of mass from x to y is equal to $D(x, y)$, the cost of the coupling π is equal to $\mathbb{E}_{(X,Y)\sim\pi} D(X, Y)$. The Monge–Kantorovich distance between μ and ν is given by the minimum cost among all the couplings between μ and ν [1–3]. Such distance is also called earth mover's distance or Wasserstein distance

of order 1, often shortened to W_1 distance. The exploration of the theory of optimal mass transport has led to the creation of an extremely fruitful field in mathematical analysis, with applications ranging from differential geometry and partial differential equations to machine learning [3–6].

The Hamming distance constitutes a natural choice for the distance D when $\mathcal{X}$ is a set of finite strings over an alphabet. The W_1 distance with respect to the Hamming distance is called Ornstein's $\bar{d}$-distance and was first considered in [7], together with its extension to stationary stochastic processes. Originally introduced as a tool for the classification of stationary processes in ergodic theory, since then it has found further applications in probability theory, such as the statistical estimation of processes [8–12], information theory, such as coding theorems for a large class of discrete noisy channels with memory and rate distortion theory [13,14] and recently also machine learning, as a peculiar case of Wasserstein auto-encoders [15,16].

Ref. [17] proposed a generalization of the W_1 distance to the space of the quantum states of a finite set of qudits or spins, called quantum W_1 distance. The generalization is based on the notion of neighboring quantum states. Two quantum states of a finite set of qudits are neighboring if they coincide after discarding one qudit. The quantum W_1 distance proposed in Ref. [17] is the distance induced by the maximum norm that assigns distance at most 1 to any couple of neighboring states. Such quantum W_1 distance recovers Ornstein's $\bar{d}$-distance in the case of quantum states diagonal in the canonical basis and inherits most of its properties. The quantum W_1 distance has found several applications in quantum information theory. In the context of statistical mechanics of quantum spin systems, a connection with quantum speed limits [18] has been found. Furthermore, transportation-cost inequalities have been proved, which upper bound the square of the quantum W_1 distance between a generic quantum state and the Gibbs state of a local quantum commuting Hamiltonian with the relative entropy between the same states [19]. Such inequalities have been applied to prove the equivalence between the microcanonical and the canonical ensembles of quantum statistical mechanics [19] and to prove limitations of variational quantum algorithms [20,21]. Moreover, the quantum W_1 distance has been applied to quantify the complexity of quantum circuits [22]. In the context of quantum state tomography, the quantum W_1 distance has been employed as quantifier of the quality of the learned quantum state, and the transportation-cost inequalities have led to an efficient algorithm to learn Gibbs states of local quantum commuting Hamiltonians [23,24]. In the context of quantum machine learning, the quantum W_1 distance has been employed as cost function of the quantum version of generative adversarial networks [25–28]. Furthermore, the quantum W_1 distance has been applied in the context of differential privacy of a quantum computation [29,30]. Finally, the quantum W_1 distance has been extended to general composite systems [31], which include the case of a finite tensor product of C^* algebras, but also provides a way to define a quantum W_1 distance between quantum channels.

1.1. Our Contribution

In this paper, we propose a generalization of the W_1 distance to quantum spin systems on the lattice $\mathbb{Z}^d$ [32–35] based on the quantum W_1 distance of Ref. [17]. Quantum spin systems on infinite lattices play a key role in quantum statistical mechanics since they provide a model to study the thermodynamic limit of infinite size of the system. Such limit is necessary to define phase transitions and to identify the properties of the system that are independent on boundary effects and boundary conditions, and more generally to make a clear distinction between the local and the global properties of the system.

We define the specific quantum W_1 distance between two translation-invariant states as the limit of the distance between their marginal states on an hypercube divided by the volume of the hypercube for the volume of the hypercube tending to infinity (Definition 4.1). Contrarily to the trace distance, the specific quantum W_1 distance has an intensive nature that makes it suitable to capture the closeness of states that are locally similar but become perfectly distinguishable globally, such as Gibbs states at close but different temperatures. We provide in Definition 4.2 an equivalent definition of the specific quantum W_1 distance that does not require the limit. We propose a generalization of the Lipschitz constant to quantum interactions on $\mathbb{Z}^d$ (Definition 5.2), and we prove in Theorem 6.1 that the specific quantum W_1 distance and the Lipschitz constant are mutually dual.

We prove in Proposition 7.1 that the specific quantum W_1 distance recovers Ornstein's $\bar{d}$-distance in the case of quantum states whose marginal states on a finite number of spins are all diagonal in the canonical basis. We prove in Proposition 8.1 a Poincaré inequality stating that for any product state, the variance of the local Hamiltonians associated with an interaction grows linearly with the volume. In Theorem 8.1, we prove a Gaussian concentration inequality for the maximally mixed state of a finite set of spins, and we apply it in Corollary 8.1 to prove an upper bound to the pressure of a quantum interaction on $\mathbb{Z}^d$ in terms of its Lipschitz constant.

In Theorem 9.1, we prove a continuity bound for the von Neumann entropy in terms of the W_1 distance. The bound applies to quantum systems made by a finite number of spins or qudits and states that the difference between the von Neumann entropy of any two quantum states divided by the number of spins is upper bounded by a universal function of the ratio between the W_1 distance and the number of spins. The bound of Theorem 9.1 contains only intensive quantities, and thanks to this property we apply it to prove a continuity bound for the specific von Neumann entropy in terms of the specific quantum W_1 distance (Theorem 10.1). Theorem 9.1 improves [17, Theorem 1], which is a weaker continuity bound for the von Neumann entropy in terms of the W_1 distance. Contrarily to the bound of Theorem 9.1, the bound of [17, Theorem 1] cannot be expressed in terms of only intensive quantities, and therefore such bound would not be sufficient to prove a continuity bound for the specific von Neumann entropy. Besides the applications to quantum spin

systems, Theorem 9.1 can be useful in quantum Shannon theory in the context of rate-distortion theory, which addresses the problem of determining the maximum compression rate of a quantum state if a certain level of distortion in the recovered state is allowed [36–43].

In the remainder of the paper, we apply our W_1 distance to study the statistical mechanics of quantum spin systems on infinite lattices. We propose a definition of w_1-Gibbs state as a translation-invariant state such that the W_1 distance between its marginal state on a hypercube and the Gibbs state of the local Hamiltonian on the same hypercube scales sublinearly with the volume of the hypercube (Definition 11.1). If an interaction admits a w_1-Gibbs state, then such state is unique (Proposition 11.1) and is an equilibrium state of the interaction (Proposition 11.2) in the sense of Kubo–Martin–Schwinger [33]. In Sect. 12, we consider transportation-cost inequalities for interactions on the quantum spin lattice $\mathbb{Z}^d$. Such inequalities imply the uniqueness of the Gibbs state of the interaction (Theorem 12.1) and a continuity bound for the specific entropy in terms of the specific relative entropy with respect to the Gibbs state (Proposition 12.2). Finally, we prove that transportation-cost inequalities are satisfied by interactions that contain only terms acting on single spins (Corollary 12.1) and geometrically local commuting interactions above a critical temperature (Theorem 12.2 and Theorem 12.3).

The paper is structured as follows. In Sect. 2, we introduce quantum spin systems on the lattice $\mathbb{Z}^d$ and in Sect. 3 we present the quantum W_1 distance and the quantum Lipschitz constant of Ref. [17]. In Sect. 4 and Sect. 5, we generalize the quantum W_1 distance and the quantum Lipschitz constant, respectively, to quantum spin systems on the lattice $\mathbb{Z}^d$. In Sect. 6, we prove the duality between the specific quantum W_1 distance and the Lipschitz constant and in Sect. 7 we prove that the specific quantum W_1 distance that we propose recovers Ornstein's $\bar{d}$-distance. In Sect. 8, we prove the quantum Poincaré and Gaussian concentration inequalities for product states. In Sect. 9, we prove the continuity bound for the von Neumann entropy in terms of the W_1 distance, and in Sect. 10, we prove the continuity bound for the specific entropy in terms of the specific quantum W_1 distance. In Sect. 11, we introduce the notion of w_1-Gibbs state. In Sect. 12, we present and prove the transportation-cost inequalities for Gibbs states. We conclude in Sect. 13 presenting some perspective applications of this work. "Appendix A" recalls some relevant properties of the quantum W_1 distance. "Appendix B" contains some auxiliary proofs, and "Appendix C" contains the proof of the auxiliary lemmas.

1.2. Related Approaches

Several quantum generalizations of optimal transport distances have been proposed. One line of research by Carlen, Maas, Datta and Rouzé [44–50] defines a quantum Wasserstein distance of order 2 from a Riemannian metric on the space of quantum states based on a quantum analog of a differential structure. Exploiting their quantum differential structure, Refs. [46,47,51] also define a quantum generalization of the Lipschitz constant and of the Wasserstein distance of order 1. Alternative definitions of quantum Wasserstein distances of

order 1 based on a quantum differential structure are proposed in Refs. [52–55]. Refs. [56–58] propose quantum Wasserstein distances of order 1 based on a distance between the vectors of the canonical basis.

Another line of research by Golse, Mouhot, Paul and Caglioti [59–69] arose in the context of the study of the semiclassical limit of quantum mechanics and defines a family of quantum Wasserstein distances of order 2 built on a quantum generalization of couplings. Such distances have been generalized to von Neumann algebras [70–72].

Ref. [73] proposes another quantum Wasserstein distance of order 2 based on couplings, with the property that each quantum coupling is associated to a quantum channel. The relation between quantum couplings and quantum channels in the framework of von Neumann algebras has been explored in [74]. The problem of defining a quantum Wasserstein distance of order 1 through quantum couplings has been explored in Ref. [75].

The quantum Wasserstein distance between two quantum states can be defined as the classical Wasserstein distance between the probability distributions of the outcomes of an informationally complete measurement performed on the states, which is a measurement whose probability distribution completely determines the state. This definition has been explored for Gaussian quantum systems with the heterodyne measurement in Refs. [76–78].

2. Quantum Spin Systems on Infinite Lattices

In this section, we introduce the setting of quantum spin systems on infinite lattices and fix the notation for the paper. For more details, the reader is encouraged to consult the books [32–35].

2.1. Algebra and States

We associate to each $x \in \mathbb{Z}^d$ the single-spin Hilbert space $\mathcal{H}_x = \mathbb{C}^q$. Let $\mathcal{F}_{\mathbb{Z}^d}$ be the collection of all the finite subsets of $\mathbb{Z}^d$. We associate to each $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ the Hilbert space

$$\mathcal{H}_\Lambda = \bigotimes_{x \in \Lambda} \mathcal{H}_x. \tag{2.1}$$

For each $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, we denote with $\mathfrak{U}_\Lambda$ the algebra of the linear operators acting on $\mathcal{H}_\Lambda$ equipped with the operator norm, which we denote with $\|\cdot\|_\infty$. For any $\Lambda' \subseteq \Lambda$, $\mathfrak{U}_{\Lambda'}$ can be canonically identified with a subalgebra of $\mathfrak{U}_\Lambda$. This identification will always be implicit.

We denote with $\mathcal{O}_\Lambda \subset \mathfrak{U}_\Lambda$ the set of the self-adjoint linear operators acting on $\mathcal{H}_\Lambda$, and with $\mathcal{O}_\Lambda^T \subset \mathcal{O}_\Lambda$ the set of the traceless self-adjoint linear operators acting on $\mathcal{H}_\Lambda$. We denote with $\mathcal{S}_\Lambda \subset \mathcal{O}_\Lambda$ the set of the quantum states acting on $\mathcal{H}_\Lambda$, i.e., the positive semidefinite linear operators with unit trace, and with Tr_Λ the partial trace over $\mathcal{H}_\Lambda$. We say that $\rho \in \mathcal{S}_\Lambda$ is a product state if there exists a collection of states $\{\rho_x \in \mathcal{S}_x\}_{x \in \Lambda}$ such that

$$\rho = \bigotimes_{x \in \Lambda} \rho_x. \tag{2.2}$$

Some results of this paper do not require the lattice structure of $\mathbb{Z}^d$ and apply to generic finite spin systems. If Λ is a generic finite set, we still define $\mathcal{H}_\Lambda$, Tr_Λ, $\mathfrak{U}_\Lambda$, $\mathcal{O}_\Lambda$, $\mathcal{O}_\Lambda^T$ and $\mathcal{S}_\Lambda$ as above.

The strictly local algebra of the spin lattice $\mathbb{Z}^d$ is

$$\mathfrak{U}_{\mathbb{Z}^d}^{loc} = \bigcup_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \mathfrak{U}_\Lambda, \tag{2.3}$$

and is equipped with the norm inherited from the operator norm of each $\mathfrak{U}_\Lambda$. The quasi-local algebra $\mathfrak{U}_{\mathbb{Z}^d}$ is the completion of $\mathfrak{U}_{\mathbb{Z}^d}^{loc}$ with respect to such norm, which we still denote with $\| \cdot \|_\infty$. For any (not necessarily finite) $\Gamma \subseteq \mathbb{Z}^d$, we define

$$\mathfrak{U}_\Gamma = \overline{\bigcup_{X \in \mathcal{F}_{\mathbb{Z}^d},\, X \subseteq \Gamma} \mathfrak{U}_X} \subseteq \mathfrak{U}_{\mathbb{Z}^d}, \tag{2.4}$$

where the closure is with respect to the $\| \cdot \|_\infty$ norm in $\mathfrak{U}_{\mathbb{Z}^d}$. When Γ is finite or $\Gamma = \mathbb{Z}^d$, (2.4) is consistent with the previous definitions. We denote with $\mathcal{O}_\Gamma$ the set of the self-adjoint elements of $\mathfrak{U}_\Gamma$.

A quantum state ρ of the spin lattice $\mathbb{Z}^d$ is a positive linear functional on $\mathfrak{U}_{\mathbb{Z}^d}$ with $\rho(\mathbb{I}) = 1$. We denote the set of the quantum states of $\mathbb{Z}^d$ with $\mathcal{S}_{\mathbb{Z}^d}$. Analogously, for any (not necessarily finite) $\Gamma \subseteq \mathbb{Z}^d$, a quantum state ρ of Γ is a positive linear functional on $\mathfrak{U}_\Gamma$ with $\rho(\mathbb{I}) = 1$. We denote with $\mathcal{S}_\Gamma$ the set of the quantum states of Γ. If Γ is finite, this definition is consistent with the previous one by setting for any $A \in \mathfrak{U}_\Gamma$

$$\rho(A) = \mathrm{Tr}_\Gamma [\rho\, A]. \tag{2.5}$$

Let $\Gamma' \subseteq \Gamma \subseteq \mathbb{Z}^d$ and let $\rho \in \mathcal{S}_\Gamma$. The marginal state $\rho_{\Gamma'}$ of ρ on Γ' is the restriction of ρ on $\mathfrak{U}_{\Gamma'}$. If Γ is finite, we have

$$\rho_{\Gamma'} = \mathrm{Tr}_{\Gamma \setminus \Gamma'} \rho. \tag{2.6}$$

Since $\mathfrak{U}_{\mathbb{Z}^d}^{loc}$ is dense in $\mathfrak{U}_{\mathbb{Z}^d}$ by construction, any $\rho \in \mathcal{S}_{\mathbb{Z}^d}$ is completely determined by the collection of its marginal states $(\rho_\Lambda)_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}}$. We say that $\rho \in \mathcal{S}_{\mathbb{Z}^d}$ is a product state if for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ the marginal state ρ_Λ is a product state.

We associate to each $x \in \mathbb{Z}^d$ the translation operator τ_x, which is the automorphism of $\mathfrak{U}_{\mathbb{Z}^d}$ that sends $\mathfrak{U}_\Gamma$ to $\mathfrak{U}_{\Gamma+x}$ for any $\Gamma \subseteq \mathbb{Z}^d$. With some abuse of notation, we denote with τ_x also the automorphism of $\mathcal{S}_{\mathbb{Z}^d}$ such that for any $\rho \in \mathcal{S}_{\mathbb{Z}^d}$ and any $A \in \mathfrak{U}_{\mathbb{Z}^d}$ we have

$$(\tau_x \rho)(\tau_x A) = \rho(A). \tag{2.7}$$

With some further abuse of notation, for any $\Gamma \subseteq \mathbb{Z}^d$ we denote with τ_x also the isomorphism between $\mathcal{S}_\Gamma$ and $\mathcal{S}_{\Gamma+x}$ such that (2.7) holds for any $\rho \in \mathcal{S}_\Gamma$ and any $A \in \mathfrak{U}_\Gamma$. We say that $\rho \in \mathcal{S}_{\mathbb{Z}^d}$ is translation invariant if $\tau_x \rho = \rho$ for any $x \in \mathbb{Z}^d$. We denote with $\mathcal{S}_{\mathbb{Z}^d}^I \subset \mathcal{S}_{\mathbb{Z}^d}$ the set of the translation-invariant quantum states of $\mathbb{Z}^d$.

For each $a \in \mathbb{N}_+^d$, we denote with Λ_a the box

$$\Lambda_a = \left\{ x \in \mathbb{Z}^d : -a \le x < a \right\}, \qquad |\Lambda_a| = \prod_{i=1}^d 2a_i, \tag{2.8}$$

where inequalities between vectors hold for each component. Given a sequence $\left(a^{(n)}\right)_{n\in\mathbb{N}} \subset \mathbb{N}_+^d$, we say that $a^{(n)} \to \infty$ if $a_i^{(n)} \to \infty$ for any $i = 1, \ldots, d$.

Definition 2.1 (*Trace distance*). The *trace distance* on $\mathcal{S}_{\mathbb{Z}^d}$ is the distance induced by the norm on $\mathfrak{U}_{\mathbb{Z}^d}$: For any $\rho,\, \sigma \in \mathcal{S}_{\mathbb{Z}^d}$,

$$T(\rho, \sigma) = \frac{1}{2} \sup \left\{ |\rho(A) - \sigma(A)| : A \in \mathfrak{U}_{\mathbb{Z}^d},\, \|A\|_\infty \leq 1 \right\}. \qquad (2.9)$$

Proposition 2.1. *The trace distance on $\mathcal{S}_{\mathbb{Z}^d}$ is the supremum of the trace distances between the marginal states: For any $\rho,\, \sigma \in \mathcal{S}_{\mathbb{Z}^d}$,*

$$T(\rho, \sigma) = \frac{1}{2} \sup_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \|\rho_\Lambda - \sigma_\Lambda\|_1, \qquad (2.10)$$

where $\| \cdot \|_1$ denotes the trace norm on $\mathfrak{U}_\Lambda$ given by

$$\|A\|_1 = \mathrm{Tr}_\Lambda \sqrt{A^\dagger A}, \qquad A \in \mathfrak{U}_\Lambda. \qquad (2.11)$$

Proof. See Sect. B.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

Definition 2.2 (*Specific entropy* [33, Proposition 6.2.38]). The *von Neumann entropy* of a quantum state ρ acting on a finite-dimensional Hilbert space is [79–81]

$$S(\rho) = -\mathrm{Tr}\left[\rho \ln \rho\right]. \qquad (2.12)$$

The *specific entropy* of $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ is the entropy per site in the limit of infinite volume:

$$s(\rho) = \lim_{a\to\infty} \frac{S(\rho_{\Lambda_a})}{|\Lambda_a|} = \inf_{a\in\mathbb{N}_+^d} \frac{S(\rho_{\Lambda_a})}{|\Lambda_a|}. \qquad (2.13)$$

Definition 2.3 (*Specific relative entropy* [82]). The *relative entropy* [79–81] between the quantum states ρ and σ acting on a finite-dimensional Hilbert space is

$$S(\rho\|\sigma) = -\mathrm{Tr}\left[\rho\left(\ln\rho - \ln\sigma\right)\right]. \qquad (2.14)$$

The *specific relative entropy* between the states $\rho,\, \sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$ is the relative entropy per site in the limit of infinite volume:

$$s(\rho\|\sigma) = \lim_{a\to\infty} \frac{S(\rho_{\Lambda_a}\|\sigma_{\Lambda_a})}{|\Lambda_a|}, \qquad (2.15)$$

whenever the limit exists.

Remark 2.1. $s(\rho\|\sigma) = 0$ does not imply $\rho = \sigma$. Indeed, let $\rho_\Lambda = |0\rangle\langle0|^{\otimes\Lambda}$ and $\sigma_\Lambda = \frac{1}{2}\left(|0\rangle\langle0|^{\otimes\Lambda} + |1\rangle\langle1|^{\otimes\Lambda}\right)$ for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$. Then, $S(\rho_\Lambda\|\sigma_\Lambda) = \ln 2$ and $s(\rho\|\sigma) = 0$.

2.2. Interactions

An interaction Φ is a collection of observables $(\Phi(\Lambda))_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}}$, where $\Phi(\Lambda) \in \mathcal{O}_\Lambda$ for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$. We can associate to Φ the formal Hamiltonian

$$H_{\mathbb{Z}^d}^\Phi = \sum_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \Phi(\Lambda). \tag{2.16}$$

In general, the series (2.16) does not converge not even weakly, and $H_{\mathbb{Z}^d}^\Phi$ cannot be defined as an element of $\mathfrak{U}_{\mathbb{Z}^d}$.

We can define for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ the local Hamiltonian on Λ with open boundary conditions

$$H_\Lambda^\Phi = \sum_{X \subseteq \Lambda} \Phi(X) \in \mathcal{O}_\Lambda. \tag{2.17}$$

An interaction Φ is translation invariant if $\Phi(\Lambda + x) = \tau_x(\Phi(\Lambda))$ for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ and any $x \in \mathbb{Z}^d$. For $r > 0$, we denote with $\mathcal{B}_{\mathbb{Z}^d}^r$ the set of translation-invariant interactions satisfying

$$\|\Phi\|_r = \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} e^{r(|\Lambda|-1)} \|\Phi(\Lambda)\|_\infty < \infty. \tag{2.18}$$

The specific energy observable of $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ is

$$E_\Phi = \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \frac{\Phi(\Lambda)}{|\Lambda|} \in \mathcal{O}_{\mathbb{Z}^d} \tag{2.19}$$

and satisfies [33, Proposition 6.2.39]

$$\lim_{a \to \infty} \frac{1}{|\Lambda_a|} \left\| H_{\Lambda_a}^\Phi - \sum_{x \in \Lambda_a} \tau_x E_\Phi \right\|_\infty = 0. \tag{2.20}$$

Therefore, for any $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ we have that $\rho(E_\Phi)$ is equal to the average energy per site of ρ in the limit of infinite volume:

$$\lim_{a \to \infty} \frac{\rho\left(H_{\Lambda_a}^\Phi\right)}{|\Lambda_a|} = \rho\left(E_\Phi\right). \tag{2.21}$$

2.3. Gibbs States

Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$. For any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, the local Gibbs state of Φ on Λ with open boundary conditions is the Gibbs state of H_Λ^Φ:

$$\omega_\Lambda^\Phi = \frac{e^{-H_\Lambda^\Phi}}{\mathrm{Tr}_\Lambda e^{-H_\Lambda^\Phi}} \in \mathcal{S}_\Lambda. \tag{2.22}$$

Since the temperature can always be reabsorbed in the interaction, in the whole paper we set it to one.

Remark 2.2. The states $\left(\omega_\Lambda^\Phi\right)_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}}$ defined in (2.22) are not necessarily the marginal states of a single global state $\omega \in \mathcal{S}_{\mathbb{Z}^d}$.

The pressure of $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ is [33, Theorem 6.2.40]

$$P(\Phi) = \lim_{a \to \infty} \frac{\ln \operatorname{Tr}_{\Lambda_a} e^{-H_{\Lambda_a}^{\Phi}}}{|\Lambda_a|} = \sup_{\rho \in \mathcal{S}_{\mathbb{Z}^d}^I} \left(s(\rho) - \rho(E_{\Phi}) \right). \tag{2.23}$$

The states $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ that achieve the supremum in (2.23) are called equilibrium states of Φ and satisfy the Kubo–Martin–Schwinger condition [33]. We denote with $\mathcal{S}_{eq}(\Phi)$ the set of such states. For any $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$, $\mathcal{S}_{eq}(\Phi)$ is nonempty, convex and compact with respect to the trace distance.

3. The Quantum W_1 Distance

Ref. [17] introduced the following generalization of the Wasserstein distance of order 1 and of the Lipschitz constant to quantum systems made by a finite number of spins. Since Ref. [17] does not require the lattice structure of $\mathbb{Z}^d$, here Λ denotes a fixed generic finite set. The quantum W_1 distance is based on the notion of neighboring quantum states. The states $\rho,\ \sigma \in \mathcal{S}_\Lambda$ are neighboring if there exists $x \in \Lambda$ such that $\operatorname{Tr}_x \rho = \operatorname{Tr}_x \sigma$. The quantum W_1 norm is the maximum norm that assigns distance at most 1 to each couple of neighboring states:

Definition 3.1 (W_1 *norm*). Let Λ be a finite set. We define for any $\Delta \in \mathcal{O}_\Lambda^T$

$$\|\Delta\|_{W_1} = \frac{1}{2} \min \left\{ \sum_{x \in \Lambda} \left\| \Delta^{(x)} \right\|_1 : \Delta^{(x)} \in \mathcal{O}_\Lambda^T,\ \operatorname{Tr}_x \Delta^{(x)} = 0,\ \sum_{x \in \Lambda} \Delta^{(x)} = \Delta \right\}. \tag{3.1}$$

The *quantum W_1 distance* on $\mathcal{S}_\Lambda$ is the distance induced by the quantum W_1 norm: For any $\rho,\ \sigma \in \mathcal{S}_\Lambda$,

$$W_1(\rho, \sigma) = \|\rho - \sigma\|_{W_1}. \tag{3.2}$$

Definition 3.2 (*Lipschitz constant*). Let Λ be a finite set. We define the dependence of $H \in \mathcal{O}_\Lambda$ on the site $x \in \Lambda$ as

$$\partial_x H = 2 \min_{H_{\Lambda \setminus x} \in \mathcal{O}_{\Lambda \setminus x}} \left\| H - H_{\Lambda \setminus x} \right\|_\infty. \tag{3.3}$$

The *quantum Lipschitz constant* of $H \in \mathcal{O}_\Lambda$ is

$$\|H\|_L = \max_{x \in \Lambda} \partial_x H. \tag{3.4}$$

Proposition 3.1 ([17, Proposition 8]). *The quantum W_1 norm and the quantum Lipschitz constant are mutually dual, i.e., for any $\Delta \in \mathcal{O}_\Lambda^T$ we have*

$$\|\Delta\|_{W_1} = \max \left\{ \operatorname{Tr}_\Lambda [\Delta H] : H \in \mathcal{O}_\Lambda,\ \|H\|_L \le 1 \right\}. \tag{3.5}$$

4. The Quantum W_1 Distance for Infinite Lattices

In this section, we extend the quantum Wasserstein distance of order 1 of Ref. [17] to the quantum states of the spin lattice $\mathbb{Z}^d$. As for the entropy and the relative entropy, we define a specific quantum W_1 distance, which we denote with w_1, as the W_1 distance per site in the limit of infinite volume:

Definition 4.1 (*Specific quantum W_1 distance*). For any ρ, $\sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$ we define

$$w_1(\rho, \sigma) = \lim_{a \to \infty} \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|}. \tag{4.1}$$

Remark 4.1. We define the specific quantum W_1 distance only for translation-invariant states since the limit (4.1) may not exist for generic states in $\mathcal{S}_{\mathbb{Z}^d}$.

Proposition 4.1. *The limit in (4.1) exists for any ρ, $\sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$ and is equal to*

$$w_1(\rho, \sigma) = \sup_{a \in \mathbb{N}_+^d} \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|}. \tag{4.2}$$

Moreover, w_1 is a distance on $\mathcal{S}_{\mathbb{Z}^d}^I$.

Proof. For any $a \in \mathbb{N}_+^d$, let

$$f(a) = \|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}. \tag{4.3}$$

Given $a \in \mathbb{N}_+^d$, $k \in \mathbb{N}$ and $i \in \{1, \ldots, d\}$, let

$$b = (a_1, \ldots, a_i + k, \ldots, a_d), \qquad c = (a_1, \ldots, k, \ldots, a_d). \tag{4.4}$$

We have

$$\Lambda_b = (\Lambda_a - k\, e_i) \cup (\Lambda_c + a_i\, e_i), \tag{4.5}$$

where e_i is the i-th vector of the canonical basis of $\mathbb{R}^d$. Then, we get from Proposition A.3 and from the translation invariance of ρ and σ

$$f(b) \geq \|\rho_{\Lambda_a - ke_i} - \sigma_{\Lambda_a - ke_i}\|_{W_1} + \|\rho_{\Lambda_c + a_i e_i} - \sigma_{\Lambda_c + a_i e_i}\|_{W_1} = f(a) + f(c). \tag{4.6}$$

Then, f is superadditive in each variable, and we have from the multidimensional Fekete's lemma Lemma C.1

$$w_1(\rho, \sigma) = \lim_{a \to \infty} \frac{f(a)}{|\Lambda_a|} = \sup_{a \in \mathbb{N}_+^d} \frac{f(a)}{|\Lambda_a|}. \tag{4.7}$$

The nonnegativity and the triangle inequality for w_1 follow from the nonnegativity and the triangle inequality for W_1, respectively. Let $w_1(\rho, \sigma) = 0$. Then, (4.2) implies

$$\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1} = 0 \qquad \forall\, a \in \mathbb{N}_+^d, \tag{4.8}$$

i.e., $\rho_{\Lambda_a} = \sigma_{\Lambda_a}$ for any $a \in \mathbb{N}_+^d$. Let $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, and let us choose $a \in \mathbb{N}_+^d$ such that $\Lambda \subseteq \Lambda_a$. Then,

$$\rho_\Lambda = \mathrm{Tr}_{\Lambda_a \setminus \Lambda} \rho_{\Lambda_a} = \mathrm{Tr}_{\Lambda_a \setminus \Lambda} \sigma_{\Lambda_a} = \sigma_\Lambda, \tag{4.9}$$

hence $\rho = \sigma$. $\qquad\square$

The specific quantum W_1 distance is always upper bounded by the trace distance:

Proposition 4.2. *For any $\rho, \sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$, we have*

$$w_1(\rho, \sigma) \leq T(\rho, \sigma). \tag{4.10}$$

Moreover, for any $a \in \mathbb{N}_+^d$ we have

$$\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_1 \leq 2 |\Lambda_a| \, w_1(\rho, \sigma). \tag{4.11}$$

Proof. We have

$$w_1(\rho, \sigma) = \sup_{a \in \mathbb{N}_+^d} \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|} \overset{(a)}{\leq} \sup_{a \in \mathbb{N}_+^d} \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_1}{2} \overset{(b)}{\leq} T(\rho, \sigma), \tag{4.12}$$

where (a) follows from Proposition A.1 and (b) follows from Proposition 2.1. From Proposition A.1, we have for any $a \in \mathbb{N}_+^d$

$$w_1(\rho, \sigma) \geq \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|} \geq \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_1}{2 |\Lambda_a|}. \tag{4.13}$$

The claim follows. □

The specific quantum W_1 distance admits an equivalent definition, which directly generalizes Definition 3.1 to infinite lattices and does not require the limit over hypercubes.

Definition 4.2. We define for any $\sigma, \rho \in \mathcal{S}_{\mathbb{Z}^d}^I$,

$$\|\rho - \sigma\|_{w_1} = \inf \left\{ c \geq 0 : \exists \, \rho', \sigma' \in \mathcal{S}_{\mathbb{Z}^d} : \rho'_{\mathbb{Z}^d \setminus 0} = \sigma'_{\mathbb{Z}^d \setminus 0}, \right.$$

$$\left. \rho_\Lambda - \sigma_\Lambda = c \sum_{x \in \Lambda} (\tau_x \rho')_\Lambda - (\tau_x \sigma')_\Lambda \quad \forall \, \Lambda \in \mathcal{F}_{\mathbb{Z}^d} \right\}. \tag{4.14}$$

Let us collect some basic properties of the quantity defined above.

Proposition 4.3. *The infimum in (4.14) is attained for some $c \geq 0$, $\rho', \sigma' \in \mathcal{S}_{\mathbb{Z}^d}$. Moreover, given sequences $(\rho_n)_{n \in \mathbb{N}}$, $(\sigma_n)_{n \in \mathbb{N}} \subseteq \mathcal{S}_{\mathbb{Z}^d}^I$ weakly converging, respectively, toward ρ and σ, then*

$$\|\rho - \sigma\|_{w_1} \leq \liminf_{n \to \infty} \|\rho_n - \sigma_n\|_{w_1}. \tag{4.15}$$

Proof. Both statements follow from the weak sequential compactness of $\mathcal{S}_{\mathbb{Z}^d}$, together with the fact that for every $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, the restriction map on states $\rho \mapsto \rho_\Lambda$ is weakly continuous. Considering a sequence $(c_n, \rho'_n, \sigma'_n)_{n \in \mathbb{N}}$ such that $\lim_n c_n = \|\rho - \sigma\|_{w_1}$ and, using compactness to extract converging subsequences, assume that $\lim_n \rho'_n = \rho'$, $\lim_n \sigma'_n = \sigma'$. By continuity, we deduce that c, ρ', σ' satisfy the conditions in (4.14), hence they are minimizers. A similar argument gives (4.15). □

The rest of this section is devoted to showing the equivalence between Definition 4.1 and Definition 4.2.

Theorem 4.1. *For σ, $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$, we have*

$$w_1(\rho, \sigma) = \|\rho - \sigma\|_{w_1} . \tag{4.16}$$

We split the proof into several intermediate results. We begin with the following upper bound.

Lemma 4.1 (Upper bound). *For σ, $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$, we have*

$$w_1(\rho, \sigma) \leq \|\rho - \sigma\|_{w_1} . \tag{4.17}$$

Proof. Given $c \geq 0$, ρ', $\sigma' \in \mathcal{S}_{\mathbb{Z}^d}$ as in the right-hand side of (4.14), for any $a \in \mathbb{N}_+^d$, we write the identity

$$\rho_{\Lambda_a} - \sigma_{\Lambda_a} = \sum_{x \in \Lambda_a} c \left(\rho_{\Lambda_a}^{(x)} - \sigma_{\Lambda_a}^{(x)} \right) , \tag{4.18}$$

where we define, for $x \in \Lambda_a$, the states $\rho_{\Lambda_a}^{(x)} = (\tau_x \rho')_{\Lambda_a}$ and $\sigma_{\Lambda_a}^{(x)} = (\tau_x \sigma')_{\Lambda_a}$. Using (2.6), it follows that

$$\mathrm{Tr}_x \rho_{\Lambda_a}^{(x)} = \mathrm{Tr}_x \left[(\tau_x \rho')_{\Lambda_a} \right] = (\tau_x \rho')_{\Lambda_a \setminus x} = \tau_x \rho'_{(\Lambda_a - x) \setminus 0} , \tag{4.19}$$

and similarly with $\sigma_{\Lambda_a}^{(x)}$, so that $\mathrm{Tr}_x \rho_{\Lambda_a}^{(x)} = \mathrm{Tr}_x \sigma_{\Lambda_a}^{(x)}$ for every $x \in \Lambda_a$. Therefore, by definition of W_1 on $\mathcal{S}_{\Lambda_a}$, we have the inequality

$$\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1} \leq c |\Lambda_a| . \tag{4.20}$$

Dividing both sides by $|\Lambda_a|$ and letting $a \to \infty$, we deduce $w_1(\rho, \sigma) \leq c$, hence (4.17). $\qquad\square$

To establish the lower bound, we consider a periodic approximation of the marginal states over a box Λ_a. We write, for any $a \in \mathbb{N}_+^d$ and $k \in \mathbb{Z}^d$,

$$2ak = (2a_i k_i)_{i=1}^d . \tag{4.21}$$

Proposition 4.4 (Periodic approximation). *For σ, $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$, and $a \in \mathbb{N}_+^d$, define $\tilde{\rho}^a$, $\tilde{\sigma}^a \in \mathcal{S}_{\mathbb{Z}^d}^I$ as follows:*

$$\tilde{\rho}^a = \frac{1}{|\Lambda_a|} \sum_{x \in \Lambda_a} \tau_x \bigotimes_{k \in \mathbb{Z}^d} \tau_{2ak} \rho_{\Lambda_a} , \qquad \tilde{\sigma}^a = \frac{1}{|\Lambda_a|} \sum_{x \in \Lambda_a} \tau_x \bigotimes_{k \in \mathbb{Z}^d} \tau_{2ak} \sigma_{\Lambda_a} . \tag{4.22}$$

Then, we have

$$\|\tilde{\rho}^a - \tilde{\sigma}^a\|_{w_1} \leq \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|} . \tag{4.23}$$

Proof. We introduce first some notation. For disjoint sets $R, S \subseteq \mathbb{Z}^d$, write

$$\alpha_{R,S} = \left(\bigotimes_{k \in R} \tau_{2ak} \rho_{\Lambda_a} \right) \otimes \left(\bigotimes_{k \in S} \tau_{2ak} \sigma_{\Lambda_a} \right) , \tag{4.24}$$

which we further simplify to $\alpha_R = \alpha_{R,S}$ whenever $S = \mathbb{Z}^d \setminus R$. With the above notation, we have

$$\tilde{\rho}^a = \frac{1}{|\Lambda_a|} \sum_{x \in \Lambda_a} \tau_x \alpha_{\mathbb{Z}^d} , \qquad \tilde{\sigma}^a = \frac{1}{|\Lambda_a|} \sum_{x \in \Lambda_a} \tau_x \alpha_\emptyset . \tag{4.25}$$

For $x \in \Lambda_a$, let $c_x \geq 0$ and $\rho_{\Lambda_a}^{(x)}, \sigma_{\Lambda_a}^{(x)} \in \mathcal{S}_{\Lambda_a}$ be such that

$$\mathrm{Tr}_x \rho_{\Lambda_a}^{(x)} = \mathrm{Tr}_x \sigma_{\Lambda_a}^{(x)} , \qquad \rho_{\Lambda_a} - \sigma_{\Lambda_a} = \sum_{x \in \Lambda_a} c_x \left(\rho_{\Lambda_a}^{(x)} - \sigma_{\Lambda_a}^{(x)} \right) , \qquad (4.26)$$

and introduce the states

$$\tilde{\rho}^{(x)} = \alpha_{R_d, S_d} \otimes \rho_{\Lambda_a}^{(x)} , \qquad \tilde{\sigma}^{(x)} = \alpha_{R_d, S_d} \otimes \sigma_{\Lambda_a}^{(x)} , \qquad (4.27)$$

where R_d, S_d are disjoint sets with $R_d \cup S_d = \mathbb{Z}^d \backslash 0$, to be specified in (4.41) below (their precise definition will be relevant only later). Notice that

$$\mathrm{Tr}_x \tilde{\rho}^{(x)} = \alpha_{R_d, S_d} \otimes \mathrm{Tr}_x \rho_{\Lambda_a}^{(x)} = \alpha_{R_d, S_d} \otimes \mathrm{Tr}_x \sigma_{\Lambda_a}^{(x)} = \mathrm{Tr}_x \tilde{\sigma}^{(x)} , \qquad (4.28)$$

hence, for every $z \in \mathbb{Z}^d$, $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ with $x + z \notin \Lambda$,

$$\left(\tau_z \tilde{\rho}^{(x)} \right)_\Lambda = \left(\tau_z \tilde{\sigma}^{(x)} \right)_\Lambda , \qquad (4.29)$$

Moreover,

$$\sum_{x \in \Lambda_a} c_x \left(\tilde{\rho}^{(x)} - \tilde{\sigma}^{(x)} \right) = \alpha_{R_d, S_d} \otimes \sum_{x \in \Lambda_a} c_x \left(\rho_{\Lambda_a}^{(x)} - \sigma_{\Lambda_a}^{(x)} \right) = \alpha_{R_d, S_d} \otimes (\rho_{\Lambda_a} - \sigma_{\Lambda_a})$$

$$= \alpha_{R_d \cup 0} - \alpha_{R_d} . \qquad (4.30)$$

We assume that $c = \sum_{x \in \Lambda_a} c_x > 0$, otherwise (4.26) yields $\rho_{\Lambda_a} = \sigma_{\Lambda_a}$ hence $\tilde{\rho}^a = \tilde{\sigma}^a$ and (4.23) holds since $\|\tilde{\rho}^a - \tilde{\sigma}^a\|_{w_1} = 0$, by choosing $c = 0$ and any $\rho' = \sigma' \in \mathcal{S}_{\mathbb{Z}^d}$. Therefore, letting $p_x = c_x / c$, we define the states

$$\rho' = \sum_{x \in \Lambda} p_x \tau_{-x} \tilde{\rho}^{(x)} , \quad \sigma' = \sum_{x \in \Lambda} p_x \tau_{-x} \tilde{\sigma}^{(x)} . \qquad (4.31)$$

By (4.28), we have the identity

$$\mathrm{Tr}_0 \rho' = \sum_{x \in \Lambda} p_x \mathrm{Tr}_0 \tau_{-x} \tilde{\rho}^{(x)} = \sum_{x \in \Lambda} p_x \tau_{-x} \mathrm{Tr}_x \tilde{\rho}^{(x)} = \sum_{x \in \Lambda} p_x \tau_{-x} \mathrm{Tr}_x \tilde{\sigma}^{(x)} = \mathrm{Tr}_0 \sigma' .$$

$$(4.32)$$

If we prove that, for every $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$,

$$\tilde{\rho}_\Lambda^a - \tilde{\sigma}_\Lambda^a = c \sum_{y \in \Lambda} \left((\tau_y \rho')_\Lambda - (\tau_y \sigma')_\Lambda \right) , \qquad (4.33)$$

then

$$\|\tilde{\rho}^a - \tilde{\sigma}^a\|_{w_1} \leq c , \qquad (4.34)$$

and (4.23) follows. To show (4.33), we write explicitly

$$c \sum_{y \in \Lambda} \left((\tau_y \rho')_\Lambda - (\tau_y \sigma')_\Lambda \right) = \sum_{y \in \Lambda} \sum_{x \in \Lambda_a} c_x \left((\tau_{y-x} \tilde{\rho}^{(x)})_\Lambda - (\tau_{y-x} \tilde{\sigma}^{(x)})_\Lambda \right)$$

$$= \sum_{x \in \Lambda_a} c_x \sum_{y \in \Lambda} (\tau_{y-x} \tilde{\rho}^{(x)})_\Lambda - (\tau_{y-x} \tilde{\sigma}^{(x)})_\Lambda$$

$$= \sum_{x \in \Lambda_a} c_x \sum_{z \in \Lambda - x} (\tau_z \tilde{\rho}^{(x)})_\Lambda - (\tau_z \tilde{\sigma}^{(x)})_\Lambda . \qquad (4.35)$$
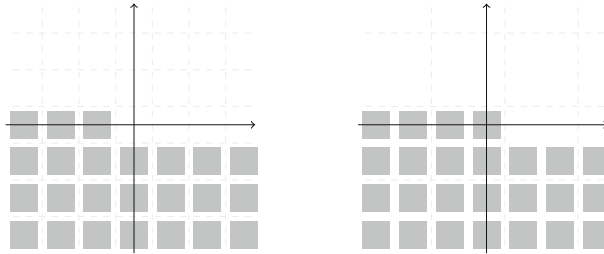
FIGURE 1. Representation of $R_2 \subseteq \mathbb{Z}^2$ (left) and its translated $R_2 + e_1 = R_2 \cup 0$ (right)

where the last line follows letting $z = y - x$. Using (4.29), we extend the summation over $z \in \Lambda'$, for any $\Lambda' \in \mathcal{F}_{\mathbb{Z}^d}$ such that

$$\Lambda' \supseteq \bigcup_{x \in \Lambda_a} (\Lambda - x). \tag{4.36}$$

Then, exchanging again the order of summation and using (4.30),

$$\sum_{x \in \Lambda_a} c_x \sum_{z \in \Lambda'} (\tau_z \tilde{\rho}^{(x)})_\Lambda - (\tau_z \tilde{\sigma}^{(x)})_\Lambda = \sum_{z \in \Lambda'} \sum_{x \in \Lambda_a} c_x \left((\tau_z \tilde{\rho}^{(x)})_\Lambda - (\tau_z \tilde{\sigma}^{(x)})_\Lambda) \right) \tag{4.37}$$

$$= \sum_{z \in \Lambda'} (\tau_z \alpha_{R_d \cup 0})_\Lambda - (\tau_z \alpha_{R_d})_\Lambda \tag{4.38}$$

We now specify the sets Λ' and R_d in such a way that the above summation is telescopic and yields (4.33). First, we let

$$\Lambda' = \bigcup_{k \in \Lambda_b} (\Lambda_a + 2ak), \tag{4.39}$$

with $b \in \mathbb{N}_+^d$ sufficiently large so that (4.36) holds. Then, recalling (4.25), to obtain (4.33) it is sufficient to prove that

$$\sum_{k \in \Lambda_b} (\tau_{2ak} \alpha_{R_d \cup 0})_\Lambda - (\tau_{2ak} \alpha_{R_d})_\Lambda = (\alpha_{\mathbb{Z}^d})_\Lambda - (\alpha_\emptyset)_\Lambda. \tag{4.40}$$

The following recursive definition for the subsets $R_d \subseteq \mathbb{Z}^d \setminus 0$ serves exactly this purpose. We let

$$R_1 = \mathbb{Z}_-, \ R_d = \left(\mathbb{Z}^{d-1} \times \mathbb{Z}_- \right) \cup (R_{d-1} \times 0), \tag{4.41}$$

so that

$$R_d \cup \{0\} = R_d + e_1, \tag{4.42}$$

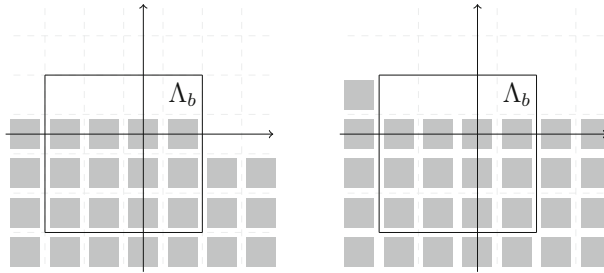where we write $e_i \in \mathbb{Z}^d$ for the natural basis vectors, for $i = 1, \ldots, d$ (see Fig. 1).

FIGURE 2. The rectangle Λ_b with $b = (2,2)$ is highlighted, providing an example of the general identity (4.44): in this case, $(R_2 + 2e_1) \cap \Lambda_b$ (left) equals $(R_2 - 2e_1 + e_2) \cap \Lambda_b$ (right)

We decompose the right-hand side in (4.40) as a double summation, over k_1 and $k_{\backslash 1} = (-b_1, k_2, \ldots, k_d)$, so that, for fixed $k_{\backslash 1}$, we find a telescopic sum

$$\sum_{0 \leq k_1 < 2b_1} \tau_{2ak_{\backslash 1}} \alpha_{R_d + e_1} - \tau_{2ak_{\backslash 1}} \alpha_{R_d} = \sum_{0 \leq k_1 < 2b_1} \tau_{2ak_{\backslash 1}} \alpha_{R_d + (k_1+1)e_1} - \tau_{2ak_{\backslash 1}} \alpha_{R_d + k_1 e_1}$$

$$= \tau_{2ak_{\backslash 1}} \alpha_{R_d + 2b_1 e_1} - \tau_{2ak_{\backslash 1}} \alpha_{R_d} . \qquad (4.43)$$

Since $\Lambda \subseteq \Lambda'$, and

$$(R_d + 2b_1 e_1 + k_{\backslash 1}) \cap \Lambda_b = (R_d + e_2 + k_{\backslash 1}) \cap \Lambda_b , \qquad (4.44)$$

(see Fig. 2), it follows that

$$(\tau_{2ak_{\backslash 1}} \alpha_{R_d + 2b_1 e_1})_\Lambda - (\tau_{2ak_{\backslash 1}} \alpha_{R_d})_\Lambda = (\tau_{2ak_{\backslash 1}} \alpha_{R_d + e_2})_\Lambda - (\tau_{2ak_{\backslash 1}} \alpha_{R_d})_\Lambda . \qquad (4.45)$$

We further proceed decomposing the summation upon k_2 and

$$k_{\backslash 1,2} = (-b_1, -b_2, k_3, \ldots, k_d), \qquad (4.46)$$

so that we obtain a similar telescopic sum. By iterating the same argument up to summation over k_d, we eventually conclude that

$$\sum_{k \in \Lambda_b} (\tau_{2ak} \alpha_{R_d \cup 0})_\Lambda - (\tau_{2ak} \alpha_{R_d})_\Lambda = (\tau_{-2ab} \alpha_{R_d + 2b_d e_d})_\Lambda - (\tau_{-2ab} \alpha_{R_d})_\Lambda , \qquad (4.47)$$

which gives (4.40) since

$$(R_d + 2b_d e_d - 2ab) \cap \Lambda_b = \Lambda_b , \ (R_d - 2ab) \cap \Lambda_b = \emptyset , \qquad (4.48)$$

thus the proof is completed. $\qquad \square$

Using the above construction, we establish the following lower bound, hence completing the proof of Theorem 4.1.

Lemma 4.2 (Lower bound). *For $\sigma, \rho \in \mathcal{S}^I_{\mathbb{Z}^d}$, we have*

$$w_1(\rho, \sigma) \geq \|\rho - \sigma\|_{w_1} . \qquad (4.49)$$

Proof. With the notation of Proposition 4.4, we argue that the states $\tilde{\rho}^a$ weakly converge to ρ. For any fixed $b \in \mathbb{N}_+^d$, if $x \in \Lambda_a$ is such that $\Lambda_b \subseteq \Lambda_a + x$, then

$$\left(\tau_x \bigotimes_{k \in \mathbb{Z}^d} \tau_{2ak} \rho_{\Lambda_a} \right)_{\Lambda_b} = \rho_{\Lambda_b} . \tag{4.50}$$

Therefore, denoting by $G_a \subseteq \Lambda_a$ the set of such x's, we write

$$\tilde{\rho}^a_{\Lambda_b} = \frac{|G_a|}{|\Lambda_a|} \rho_{\Lambda_b} + \sum_{x \in \Lambda_a \setminus G_a} \left(\tau_x \bigotimes_{k \in \mathbb{Z}^d} \tau_{2ak} \rho_{\Lambda_a} \right)_{\Lambda_b} \tag{4.51}$$

Since $|G_a| = |\Lambda_a| - |\Lambda_b|$, it follows that, as $a \to \infty$, $\tilde{\rho}^a_{\Lambda_b}$ converge in $\mathcal{S}_{\Lambda_b}$ toward ρ_{Λ_b}. This holds for any $b \in \mathbb{N}_+^d$, hence we obtain the desired weak convergence in $\mathcal{S}_{\mathbb{Z}^d}$ of $\tilde{\rho}^a$ toward ρ. By (4.15), Proposition 4.4 and Proposition 4.1 we have the inequalities

$$\|\rho - \sigma\|_{w_1} \leq \liminf_{a \to \infty} \|\tilde{\rho}^a - \tilde{\sigma}^a\|_{w_1}$$

$$\leq \limsup_{a \to \infty} \|\tilde{\rho}^a - \tilde{\sigma}^a\|_{w_1} \leq \sup_{a \in \mathbb{N}_+^d} \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|} = w_1(\rho, \sigma), \tag{4.52}$$

and the proof of (4.49) is completed.   □

As a consequence of the above argument, we also obtain that the periodic approximations always converge with respect to the specific quantum W_1 distance.

Corollary 4.1. *With the notation of Proposition 4.4, we have*

$$\lim_{a \to \infty} w_1(\tilde{\rho}^a, \tilde{\sigma}^a) = w_1(\rho, \sigma). \tag{4.53}$$

Proof. In (4.52), all inequalities must be equalities, hence the limit.   □

5. The Quantum Lipschitz Constant for Infinite Lattices

In this section, we extend the definition of quantum Lipschitz constant of Ref. [17] to interactions on the quantum spin lattice $\mathbb{Z}^d$.

The first step toward defining a Lipschitz constant for interactions is to extend to $\mathcal{O}_{\mathbb{Z}^d}$ the definition of dependence on a site:

Definition 5.1. For any $H \in \mathcal{O}_{\mathbb{Z}^d}$ and any $x \in \mathbb{Z}^d$, we define

$$\partial_x H = 2 \inf_{A \in \mathcal{O}_{\mathbb{Z}^d \setminus x}} \|H - A\|_\infty . \tag{5.1}$$

Proposition 5.1. *For any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, any $H \in \mathcal{O}_\Lambda$ and any $x \in \Lambda$, (3.3) and (5.1) are equivalent.*

Proof. See Sect. B.2.   □

We can now define the Lipschitz constant for interactions:

Definition 5.2 (*Lipschitz constant*)**.** We define the *Lipschitz constant* of $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ as the dependence of the formal Hamiltonian $H_{\mathbb{Z}^d}^\Phi$ on the site 0:

$$\|\Phi\|_L = \partial_0 \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \Phi(\Lambda), \tag{5.2}$$

where the series converges absolutely in the $\|\cdot\|_\infty$ norm.

The Lipschitz constant of Φ is also equal to the dependence on a fixed site of the local Hamiltonian on a region in the limit of infinite volume:

Proposition 5.2. *For any $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ and any $x \in \mathbb{Z}^d$, we have*

$$\lim_{a \to \infty} \partial_x H_{\Lambda_a}^\Phi = \|\Phi\|_L. \tag{5.3}$$

Proof. We have

$$\left| \|\Phi\|_L - \partial_x H_{\Lambda_a}^\Phi \right| \stackrel{(a)}{=} \left| \|\Phi\|_L - \partial_0 H_{\Lambda_a - x}^\Phi \right| = \left| \|\Phi\|_L - \partial_0 \sum_{0 \in \Lambda \subseteq \Lambda_a - x} \Phi(\Lambda) \right|$$

$$\stackrel{(b)}{\leq} \partial_0 \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d} : \Lambda \not\subseteq \Lambda_a - x} \Phi(\Lambda) \leq 2 \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d} : \Lambda \not\subseteq \Lambda_a - x} \|\Phi(\Lambda)\|_\infty, \tag{5.4}$$

where (a) follows from the translation invariance of Φ and (b) follows since ∂_0 is a seminorm. Since

$$\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \|\Phi(\Lambda)\|_\infty \leq \|\Phi\|_r < \infty, \tag{5.5}$$

we have

$$\lim_{a \to \infty} \left| \|\Phi\|_L - \partial_x H_{\Lambda_a}^\Phi \right| \leq 2 \lim_{a \to \infty} \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d} : \Lambda \not\subseteq \Lambda_a - x} \|\Phi(\Lambda)\|_\infty = 0. \tag{5.6}$$

The claim follows. □

5.1. Physical Equivalence

Different interactions may give rise to the same formal Hamiltonian. Such interactions are called physically equivalent. The concept of physical equivalence between interactions was formally introduced in [83–85]. We adopt the definition of [86, Section 4.7] and [82,87]. The reader can find more details in [88, Section 2.4.6].

Definition 5.3. The interaction Φ, $\Psi \in \mathcal{B}_{\mathbb{Z}^d}^r$ are *physically equivalent* if any of the following equivalent conditions holds:

1. The local Hamiltonians of Φ and Ψ normalized by the number of sites differ only by a constant in the limit of infinite volume:

$$\lim_{a \to \infty} \left\| \frac{H_\Lambda^\Phi - H_\Lambda^\Psi}{|\Lambda_a|} - \omega(E_{\Phi - \Psi}) \mathbb{I} \right\|_\infty = 0, \tag{5.7}$$

where $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$ is the uniform distribution, i.e., $\omega_\Lambda = \frac{\mathbb{I}_\Lambda}{q^{|\Lambda|}}$ for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$.

2. Φ and Ψ generate the same time evolution: For any $A \in \mathfrak{U}_{\mathbb{Z}^d}$ and any $t \in \mathbb{R}$, we have

$$\lim_{a \to \infty} \left\| e^{iH_{\Lambda_a}^{\Phi} t} A \, e^{-iH_{\Lambda_a}^{\Phi} t} - e^{iH_{\Lambda_a}^{\Psi} t} A \, e^{-iH_{\Lambda_a}^{\Psi} t} \right\|_{\infty} = 0. \tag{5.8}$$

3. For any $A \in \mathfrak{U}_{\mathbb{Z}^d}^{loc}$, we have

$$\sum_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} [\Phi(\Lambda) - \Psi(\Lambda), \, A] = 0. \tag{5.9}$$

Proposition 5.3. *The interaction* $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ *is physically equivalent to the null interaction iff* $\|\Phi\|_L = 0$.

Proof. • Let Φ be physically equivalent to the null interaction. Let

$$K = \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \Phi(\Lambda), \tag{5.10}$$

where the series converges absolutely in the $\|\cdot\|_{\infty}$ norm. We have for any $A \in \mathfrak{U}_0$

$$0 = \sum_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} [\Phi(\Lambda), \, A] = \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} [\Phi(\Lambda), \, A] = [K, \, A], \tag{5.11}$$

therefore $K \in \mathcal{O}_{\mathbb{Z}^d} \setminus 0$ and $\|\Phi\|_L = \partial_0 K = 0$.
• Let $\|\Phi\|_L = 0$. Let K be as in (5.10). We have

$$2 \inf_{H \in \mathcal{O}_{\mathbb{Z}^d \setminus 0}} \|K - H\|_{\infty} = \partial_0 K = \|\Phi\|_L = 0, \tag{5.12}$$

therefore there exists a sequence $\left(K^{(n)} \right)_{n \in \mathbb{N}} \subset \mathcal{O}_{\mathbb{Z}^d \setminus 0}$ such that

$$\lim_{n \to \infty} \left\| K - K^{(n)} \right\|_{\infty} = 0. \tag{5.13}$$

Then, $K \in \mathcal{O}_{\mathbb{Z}^d \setminus 0}$, and for any $A_0 \in \mathfrak{U}_0$ we have

$$\sum_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} [\Phi(\Lambda), \, A_0] = [K, \, A_0] = 0. \tag{5.14}$$

Let us prove that for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ and any $A_{\Lambda} \in \mathfrak{U}_{\Lambda}$ we have

$$\sum_{X \in \mathcal{F}_{\mathbb{Z}^d}} [\Phi(X), \, A_{\Lambda}] = 0 \tag{5.15}$$

by induction on the size of Λ. We have already proved the claim for $|\Lambda| = 1$. Let us fix $x \in \Lambda$. By linearity, we can assume that $A_{\Lambda} = A_x \otimes A_{\Lambda \setminus x}$ with $A_{\Lambda \setminus x} \in \mathfrak{U}_{\Lambda \setminus x}$. We have from the inductive hypothesis

$$\sum_{X \in \mathcal{F}_{\mathbb{Z}^d}} [\Phi(X), \, A_x \otimes A_{\Lambda \setminus x}] = \sum_{X \in \mathcal{F}_{\mathbb{Z}^d}} ([\Phi(X), \, A_x] \, A_{\Lambda \setminus x}$$
$$+ A_x \, [\Phi(X), \, A_{\Lambda \setminus x}]) = 0. \tag{5.16}$$

The claim follows.

$\square$

6. Duality for the w_1 Distance

Using Theorem 4.1, we prove the following dual formulation for the specific quantum W_1 distance.

Theorem 6.1 (Duality). *For ρ, $\sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$, we have*

$$w_1(\rho, \sigma) = \sup \left\{ \rho(E_\Phi) - \sigma(E_\Phi) : \Phi \in \mathcal{B}_{\mathbb{Z}^d}^r, \|\Phi\|_L \le 1 \right\}. \tag{6.1}$$

The result extends Proposition 3.1 to infinite spin systems, but unlike the finite dimensional case, in general there is no reason for the supremum in (6.1) to be attained in $\mathcal{B}_{\mathbb{Z}^d}^r$.

Proof. Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ with $\|\Phi\|_L \le 1$, and let $c \ge 0$, ρ', $\sigma' \in \mathcal{S}_{\mathbb{Z}^d}$ be as in (4.14). For $a \in \mathbb{N}_+^d$ and $\Lambda \subseteq \Lambda_a \in \mathcal{F}_{\mathbb{Z}^d}$, we have

$$\rho(\Phi(\Lambda)) - \sigma(\Phi(\Lambda)) = c \sum_{x \in \Lambda_a} \tau_x \rho'(\Phi(\Lambda)) - \tau_x \rho'(\Phi(\Lambda))$$

$$= c \sum_{x \in \Lambda_a} \rho'(\Phi(\Lambda - x)) - \sigma'(\Phi(\Lambda - x)). \tag{6.2}$$

If $x \notin \Lambda$, then $0 \notin \Lambda - x$, hence $\rho'(\Phi(\Lambda - x)) = \sigma'(\Phi(\Lambda - x))$ because $\rho'_{\Lambda-x} = \sigma'_{\Lambda-x}$. Therefore, the sum above can be restricted upon $x \in \Lambda$, i.e., $0 \in \Lambda - x$. We then sum upon $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ and make a change of variable $\Lambda' = \Lambda - x$, obtaining

$$\sum_{0 \in \Lambda \subseteq \Lambda_a} \frac{\rho(\Phi(\Lambda)) - \sigma(\Phi(\Lambda))}{|\Lambda|} = c \sum_{0 \in \Lambda \subseteq \Lambda_a} \sum_{x \in \Lambda} \frac{\rho'(\Phi(\Lambda - x)) - \sigma'(\Phi(\Lambda - x))}{|\Lambda|}$$

$$= c \sum_{0 \in \Lambda' \in \mathcal{F}_{\mathbb{Z}^d}} (\rho'(\Phi(\Lambda')) - \sigma'(\Phi(\Lambda'))) \frac{g_a(\Lambda')}{|\Lambda'|}, \tag{6.3}$$

where $g_a(\Lambda')$ denotes the number of pairs (x, Λ) with $x \in \Lambda \subseteq \Lambda_a$, $0 \in \Lambda$, such that $\Lambda' = \Lambda - x$. Notice that the above is a finite sum, since we may restrict upon $\Lambda' \subseteq \Lambda_{2a}$, otherwise $g_a(\Lambda') = 0$. Moreover, for every such pair (x, Λ), it must be $x \in -\Lambda'$, since $0 \in \Lambda$. Therefore, for every $\Lambda' \in \mathcal{F}_{\mathbb{Z}^d}$ with $0 \in \Lambda'$,

$$0 \le \frac{g_a(\Lambda')}{|\Lambda'|} \le 1, \quad \text{and} \quad \lim_{a \to \infty} \frac{g_a(\Lambda')}{|\Lambda'|} = 1, \tag{6.4}$$

since every pair (x, Λ) with $x \in -\Lambda'$ and $\Lambda = x + \Lambda'$ satisfies $x \in \Lambda \subseteq \Lambda_a$ if a is sufficiently large. Therefore, by the dominated convergence theorem for series, we deduce that

$$\lim_{a \to \infty} \sum_{0 \in \Lambda' \in \mathcal{F}_{\mathbb{Z}^d}} \rho'(\Phi(\Lambda')) \frac{g_a(\Lambda')}{|\Lambda'|} = \sum_{0 \in \Lambda' \in \mathcal{F}_{\mathbb{Z}^d}} \rho'(\Phi(\Lambda')), \tag{6.5}$$

and similarly for σ'. The left-hand side in (6.3) converges to $\rho(E_\Phi) - \sigma(E_\Phi)$ as $a \to \infty$, hence we obtain the identity

$$\rho(E_\Phi) - \sigma(E_\Phi) = c \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \rho'(\Phi(\Lambda)) - \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \sigma'(\Phi(\Lambda)). \tag{6.6}$$

Given $A \in \mathcal{O}_{\mathbb{Z}^d \setminus 0}$, since $\rho'_{\mathbb{Z}^d \setminus 0} = \sigma'_{\mathbb{Z}^d \setminus 0}$, we have $\rho'(A) = \sigma'(A)$, thus

$$\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \rho'(\Phi(\Lambda)) - \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \sigma'(\Phi(\Lambda)) = \rho' \left(\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \Phi(\Lambda) - A \right)$$

$$- \sigma' \left(\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \Phi(\Lambda) - A \right) \quad (6.7)$$

$$\leq 2 \left\| \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \Phi(\Lambda) - A \right\|_\infty . \quad (6.8)$$

Being $A \in \mathcal{O}_{\mathbb{Z}^d \setminus 0}$ arbitrary, we deduce the inequality

$$\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \rho'(\Phi(\Lambda)) - \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \sigma'(\Phi(\Lambda)) \leq \|\Phi\|_L \leq 1 , \quad (6.9)$$

which from (6.6) gives

$$\sup \left\{ \rho(E_\Phi) - \sigma(E_\Phi) : \Phi \in \mathcal{B}^r_{\mathbb{Z}^d}, \|\Phi\|_L \leq 1 \right\} \leq c . \quad (6.10)$$

Recalling that c and ρ', σ' are chosen as in (4.14), we deduce

$$\sup \left\{ \rho(E_\Phi) - \sigma(E_\Phi) : \Phi \in \mathcal{B}^r_{\mathbb{Z}^d}, \|\Phi\|_L \leq 1 \right\} \leq \|\rho - \sigma\|_{w_1} , \quad (6.11)$$

i.e., inequality $\geq$ holds in (6.1).

For the converse inequality, given any $H \in \mathcal{O}_{\Lambda_a}$ with $\|H\|_L \leq 1$, we define the translation-invariant interaction

$$\Phi^H(\Lambda) = \frac{\tau_x H}{|\Lambda_a|} \quad (6.12)$$

if $\Lambda = \Lambda_a + x$ for some $x \in \mathbb{Z}^d$, $\Phi^H(\Lambda) = 0$ otherwise. Notice that $\Phi^H \in \mathcal{B}^r_{\mathbb{Z}^d}$ and

$$\|\Phi^H\|_L \leq \sum_{x \in \Lambda_a} \frac{\partial_0 \tau_{-x} H}{|\Lambda_a|} \leq \sum_{x \in \Lambda_a} \frac{\partial_x H}{|\Lambda_a|} \leq \|H\|_L \leq 1 . \quad (6.13)$$

Since $\rho, \sigma \in \mathcal{S}^I_{\mathbb{Z}^d}$, we have $\rho(\tau_x H) = \rho(H)$, $\sigma(\tau_x H) = \sigma(H)$, hence

$$\rho(E_{\Phi^H}) - \sigma(E_{\Phi^H}) = \sum_{x \in \Lambda_a} \frac{\rho(\tau_{-x} H) - \sigma(\tau_{-x} H)}{|\Lambda_a|^2} = \frac{\rho(H) - \sigma(H)}{|\Lambda_a|} . \quad (6.14)$$

The duality for the quantum W_1 distance on the finite lattice Λ_a yields

$$\frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|} = \sup \left\{ \rho(E_{\Phi^H}) - \sigma(E_{\Phi^H}) : H \in \mathcal{O}_{\Lambda_a}, \|H\|_L \leq 1 \right\}$$

$$\leq \sup \left\{ \rho(E_\Phi) - \sigma(E_\Phi) : \Phi \in \mathcal{B}^r_{\mathbb{Z}^d}, \|\Phi\|_L \leq 1 \right\} . \quad (6.15)$$

Letting $a \to \infty$, we obtain inequality $\leq$ in (6.1), hence the thesis. $\qquad \square$

7. Recovery of Ornstein's $\bar{d}$-Distance

As in the finite dimensional case, the specific quantum W_1 distance between states recovers Ornstein's $\bar{d}$-distance, when restricted to diagonal states in the canonical basis, i.e., $\rho \in \mathcal{S}_{\mathbb{Z}^d}$ such that, for every $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, ρ_Λ is diagonal in the basis $(|x\rangle\langle x|)_{x \in [q]^\Lambda}$.

There is indeed a correspondence between probability measures μ on $[q]^{\mathbb{Z}}$ and such states, defined by mapping μ to the diagonal state $\rho \in \mathcal{S}_{\mathbb{Z}^d}$ such that, for every $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$,

$$\rho_\Lambda = \sum_{x \in [q]^\Lambda} \mu_\Lambda(x) |x\rangle\langle x|, \tag{7.1}$$

where μ_Λ denotes the marginal of μ on Λ. Since states are determined by their collection of marginals, (7.1) completely determines ρ.

The correspondence is clearly invertible, arguing similarly on the space of probability measures $[q]^{\mathbb{Z}}$. With a slight abuse of notation, we write $|x\rangle\langle x| \in \mathcal{S}_{\mathbb{Z}^d}$ for the diagonal state corresponding to the Dirac probability measure concentrated at $x \in [q]^{\mathbb{Z}^d}$, so that one can also write

$$\rho = \int_{[q]^{\mathbb{Z}^d}} |x\rangle\langle x| \, \mathrm{d}\mu(x), \tag{7.2}$$

where integration is in the sense of Pettis (also called weak integral).

Given two shift-invariant (i.e., stationary) probability measures μ, ν on the infinite product space $[q]^{\mathbb{Z}^d}$, Ornstein's $\bar{d}$-distance [7,13] is defined as

$$\bar{d}(\mu,\nu) = \sup_{a \in \mathbb{N}_+^d} \frac{W_1\left(\mu_{\Lambda_a}, \nu_{\Lambda_a}\right)}{|\Lambda_a|}, \tag{7.3}$$

where, for $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, W_1 denotes the classical optimal transport distance with Hamming cost on $[q]^\Lambda$, i.e.,

$$W_1\left(\mu_\Lambda, \nu_\Lambda\right) = \min_{\pi \in \mathcal{C}(\mu_\Lambda, \nu_\Lambda)} \sum_{x,y \in [q]^\Lambda} h(x,y)\pi(x,y), \tag{7.4}$$

with $\mathcal{C}(\mu_\Lambda, \nu_\Lambda)$ being the set of couplings between the probability distributions μ_Λ, ν_Λ, and

$$h(x,y) = |\{i \in \Lambda \,:\, x_i \neq y_i\}| \,. \tag{7.5}$$

This distance is usually defined only in the case $d = 1$, but the extension to $d \geq 1$ is straightforward and informally discussed already in [7, Appendix 4].

Proposition 7.1. *Given stationary probability measures μ, ν on $[q]^{\mathbb{Z}^d}$, let ρ, $\sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$ denote the associated diagonal states,*

$$\rho = \int_{[q]^{\mathbb{Z}^d}} |x\rangle\langle x| \mathrm{d}\mu(x), \quad \sigma = \int_{[q]^{\mathbb{Z}^d}} |x\rangle\langle x| \mathrm{d}\nu(x). \tag{7.6}$$

Then, we have

$$w_1(\rho, \sigma) = \bar{d}(\mu, \nu). \tag{7.7}$$

Proof. For every $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$, we have, by [17, Proposition 5],

$$\|\rho_\Lambda - \sigma_\Lambda\|_{W_1} = W_1(\mu_\Lambda, \nu_\Lambda). \tag{7.8}$$

Choosing $\Lambda = \Lambda_a$, for $a \in \mathbb{N}_+^d$, dividing by $|\Lambda_a|$ and letting $a \to \infty$ yields the thesis. □

Ornstein's $\bar{d}$-distance (7.3) can be equivalently defined [13, Theorem 1] as

$$\bar{d}(\mu, \nu) = \min_{\pi \in \mathcal{C}^I(\mu,\nu)} \sum_{x,y \in [q]} h(x,y)\pi_0(x,y), \tag{7.9}$$

where $\mathcal{C}^I(\mu, \nu)$ denotes the set of stationary couplings between the probability distributions μ, ν, and π_0 is the marginal density of π at 0. Definition 4.2 together with Theorem 4.1 provides a similar characterization for quantum spin systems, where stationary couplings are replaced in (4.14) by representations of the difference the states as series of translates. In fact, if the states ρ, σ are diagonal, we can also restrict minimization in (4.14) to diagonal states ρ', σ', corresponding to probability measures μ', ν', obtaining the following further equivalent representation of Ornstein's distance.

Corollary 7.1. *Given stationary probability measures μ, ν on $[q]^{\mathbb{Z}^d}$, we have*

$$\bar{d}(\mu, \nu) = \min \left\{ c \geq 0 : \exists \, \mu', \nu' \text{ probability measures on } [q]^{\mathbb{Z}^d} : \mu'_{\mathbb{Z}^d \setminus 0} = \nu'_{\mathbb{Z}^d \setminus 0}, \right.$$

$$\left. \mu_\Lambda - \nu_\Lambda = c \sum_{x \in \Lambda} (\tau_x \mu')_\Lambda - (\tau_x \nu')_\Lambda \quad \forall \, \Lambda \in \mathcal{F}_{\mathbb{Z}^d} \right\}. \tag{7.10}$$

To our knowledge, duality for Orstein's $\bar{d}$-distance is not explicitly discussed in the literature. A result can be obtained directly from Theorem 6.1 for diagonal states, simply noticing that the supremum may run among interactions Φ such that each $\Phi(\Lambda)$ is also diagonal, i.e., corresponding to a function

$$f(\Lambda) : [q]^{\mathbb{Z}^d} \to \mathbb{R}, \tag{7.11}$$

depending only on the coordinates in Λ. Let us denote by $\mathcal{B}_{\mathbb{Z}^d}^{r,\text{diag}}$ the set of translation invariant diagonal interactions satisfying (2.18). The Lipschitz constant of $f \in \mathcal{B}_{\mathbb{Z}^d}^{r,\text{diag}}$ coincides with the oscillation of the function on $[q]^{\mathbb{Z}^d}$,

$$x \mapsto \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} f(\Lambda)(x) \tag{7.12}$$

with respect to the 0-th coordinate, i.e.,

$$\|f\|_L = \sup \left\{ \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} (f(\Lambda)(x) - f(\Lambda)(y)) : x, y \in [q]^{\mathbb{Z}^d}, x_k = y_k \text{ for every } k \in \mathbb{Z}^d \setminus 0 \right\}. \tag{7.13}$$

The specific energy of f is identified with the function on $[q]^{\mathbb{Z}^d}$,

$$x \mapsto e_f(x) = \sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \frac{f(\Lambda)(x)}{|\Lambda|} . \tag{7.14}$$

With this notation, Theorem 6.1 yields the following result.

Corollary 7.2. *Given stationary probability measures μ, ν on $[q]^{\mathbb{Z}}$, we have*

$$\bar{d}(\mu, \nu) = \sup \left\{ \int_{[q]^{\mathbb{Z}^d}} e_f(x) d\mu(x) - \int_{[q]^{\mathbb{Z}^d}} e_f(x) d\nu(x) : f \in \mathcal{B}_{\mathbb{Z}^d}^{r,\mathrm{diag}}, \|f\|_L \leq 1 \right\} . \tag{7.15}$$

8. Quantum Concentration Inequalities

8.1. Poincaré Inequality

In this section, we prove the following quantum Poincaré inequality stating that for any interaction Φ, the variance of the local Hamiltonian on Λ_a on a product state scales at most linearly with the volume of Λ_a in the limit $a \to \infty$, and the proportionality constant is upper bounded by the square of the Lipschitz constant of Φ:

Proposition 8.1 (Poincaré inequality). *Let $\omega \in \mathcal{S}_{\mathbb{Z}^d}$ be a product state. Then, for any interaction $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ we have*

$$\limsup_{a \to \infty} \frac{\mathrm{Var}_{\omega_{\Lambda_a}} H_{\Lambda_a}^{\Phi}}{|\Lambda_a|} \leq \|\Phi\|_L^2 . \tag{8.1}$$

8.1.1. Proof of Proposition 8.1. The proof of Proposition 8.1 is based on its counterpart for quantum spin systems on finite lattices:

Proposition 8.2 (Quantum Poincaré inequality [20, Lemma F.1]). *Let Λ be a finite set, and let $\omega \in \mathcal{S}_\Lambda$ be a product state. Then, for any $H \in \mathcal{O}_\Lambda$ we have*

$$\mathrm{Var}_\omega H = \mathrm{Tr} \left[\omega \left(H - \mathrm{Tr} \left[\omega H \right] \mathbb{I} \right)^2 \right] \leq \sum_{x \in \Lambda} (\partial_x H)^2 . \tag{8.2}$$

The key step in the proof of Proposition 8.1 is the following proposition, which states that the Lipschitz constant of Φ is equal to the quadratic average of the dependence of the local Hamiltonians of Φ on the sites in the limit of infinite volume:

Proposition 8.3. *Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$. Then,*

$$\lim_{a \to \infty} \frac{1}{|\Lambda_a|} \sum_{x \in \Lambda_a} \left(\partial_x H_{\Lambda_a}^{\Phi} \right)^2 = \|\Phi\|_L^2 . \tag{8.3}$$

Proof. We fix $\epsilon > 0$. Let $N \in \mathbb{N}$ and $0 \in \Lambda_1 \in \mathcal{F}_{\mathbb{Z}^d}, \ldots, 0 \in \Lambda_N \in \mathcal{F}_{\mathbb{Z}^d}$ such that

$$\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}, \Lambda \neq \Lambda_1, \ldots, \Lambda_N} \|\Phi(\Lambda)\|_\infty < \epsilon . \tag{8.4}$$

Let $a \in \mathbb{N}_+^d$ such that $\Lambda_1 \cup \ldots \cup \Lambda_N \subseteq \Lambda_a$ and

$$\left| \partial_0 H_{\Lambda_a}^\Phi - \|\Phi\|_L \right| < 2\,\epsilon. \tag{8.5}$$

We have

$$\sum_{0 \in \Lambda \in \mathcal{F}_{\mathbb{Z}^d}, \, \Lambda \nsubseteq \Lambda_a} \|\Phi(\Lambda)\|_\infty < \epsilon. \tag{8.6}$$

Let us fix $x \in \mathbb{Z}^d$. For any $b \in \mathbb{N}_+^d$ with $b \geq a \pm x$, we have $\Lambda_a \subseteq \Lambda_b - x$ and

$$\left| \partial_x H_{\Lambda_b}^\Phi - \partial_0 H_{\Lambda_a}^\Phi \right| \overset{(a)}{=} \left| \partial_0 H_{\Lambda_b - x}^\Phi - \partial_0 H_{\Lambda_a}^\Phi \right| \overset{(b)}{\leq} \partial_0 \left(H_{\Lambda_b - x}^\Phi - H_{\Lambda_a}^\Phi \right) = \partial_0 \sum_{\Lambda \subseteq \Lambda_b - x, \, \Lambda \nsubseteq \Lambda_a} \Phi(\Lambda)$$

$$= \partial_0 \sum_{0 \in \Lambda \subseteq \Lambda_b - x, \, \Lambda \nsubseteq \Lambda_a} \Phi(\Lambda) \leq 2 \sum_{0 \in \Lambda \subseteq \Lambda_b - x, \, \Lambda \nsubseteq \Lambda_a} \|\Phi(\Lambda)\|_\infty < 2\,\epsilon,$$

$$\tag{8.7}$$

where (a) follows from the translation invariance of Φ and (b) follows since ∂_0 is a seminorm. We then have

$$\left| \partial_x H_{\Lambda_b}^\Phi - \|\Phi\|_L \right| < 4\,\epsilon. \tag{8.8}$$

Let $b > a$, such that any $x \in \Lambda_{b-a}$ satisfies $b \geq a \pm x$. We have from Lemma C.4

$$\sum_{x \in \Lambda_b} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 = \sum_{x \in \Lambda_{b-a}} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 + \sum_{x \in \Lambda_b \setminus \Lambda_{b-a}} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2$$

$$\leq |\Lambda_{b-a}| \left(\|\Phi\|_L + 4\,\epsilon \right)^2 + 4 \left(|\Lambda_b| - |\Lambda_{b-a}| \right) \|\Phi\|_r^2, \tag{8.9}$$

therefore

$$\limsup_{b \to \infty} \frac{1}{|\Lambda_b|} \sum_{x \in \Lambda_b} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 \leq \limsup_{b \to \infty} \left(\frac{|\Lambda_{b-a}|}{|\Lambda_b|} \left(\|\Phi\|_L + 4\,\epsilon \right)^2 + 4 \left(1 - \frac{|\Lambda_{b-a}|}{|\Lambda_b|} \right) \|\Phi\|_r^2 \right)$$

$$= \left(\|\Phi\|_L + 4\,\epsilon \right)^2. \tag{8.10}$$

Since ϵ is arbitrary, we get

$$\limsup_{b \to \infty} \frac{1}{|\Lambda_b|} \sum_{x \in \Lambda_b} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 \leq \|\Phi\|_L^2. \tag{8.11}$$

If $\|\Phi\|_L = 0$, the claim follows from (8.11). If $\|\Phi\|_L > 0$, we can choose $\epsilon < \|\Phi\|_L / 4$. We then have

$$\sum_{x \in \Lambda_b} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 \geq \sum_{x \in \Lambda_{b-a}} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 \geq |\Lambda_{b-a}| \left(\|\Phi\|_L - 4\,\epsilon \right)^2, \tag{8.12}$$

and

$$\liminf_{b \to \infty} \frac{1}{|\Lambda_b|} \sum_{x \in \Lambda_b} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 \geq \liminf_{b \to \infty} \frac{|\Lambda_{b-a}|}{|\Lambda_b|} \left(\|\Phi\|_L - 4\,\epsilon \right)^2 = \left(\|\Phi\|_L - 4\,\epsilon \right)^2. \tag{8.13}$$

Since ϵ is arbitrary, we get

$$\liminf_{b \to \infty} \frac{1}{|\Lambda_b|} \sum_{x \in \Lambda_b} \left(\partial_x H_{\Lambda_b}^\Phi \right)^2 \geq \|\Phi\|_L^2. \tag{8.14}$$

The claim follows. □

We can now conclude the proof of Proposition 8.1. We have from Proposition 8.2 and Proposition 8.3

$$\limsup_{a \to \infty} \frac{\mathrm{Var}_{\omega_{\Lambda_a}} H^\Phi_{\Lambda_a}}{|\Lambda_a|} \le \limsup_{a \to \infty} \frac{1}{|\Lambda_a|} \sum_{x \in \Lambda_a} \left(\partial_x H^\Phi_{\Lambda_a} \right)^2 = \|\Phi\|_L^2 . \tag{8.15}$$

The claim follows.

8.2. Gaussian Concentration Inequality

In this section, we prove the following Gaussian concentration inequality (Theorem 8.1) for quantum spin systems on finite lattices and apply it to prove an upper bound to the pressure of an interaction in terms of its Lipschitz constant (Corollary 8.1).

Theorem 8.1 (Gaussian concentration inequality). *Let Λ be a finite set and let $\omega \in \mathcal{S}_\Lambda$ be a product state with full support. Then, for any $H \in \mathcal{O}_\Lambda$ we have*

$$\ln \mathrm{Tr}_\Lambda e^{H + \ln \omega} \le \mathrm{Tr}_\Lambda [\omega H] + \frac{1}{2} \sum_{x \in \Lambda} (\partial_x H)^2 . \tag{8.16}$$

Remark 8.1. Ref. [17] proved the following Gaussian concentration inequality:

Theorem 8.2 ([17, Theorem 3]). *Let Λ be a finite set and let $\omega = \frac{\mathbb{I}_\Lambda}{q^{|\Lambda|}} \in \mathcal{S}_\Lambda$ be the uniform distribution. Then, for any $H \in \mathcal{O}_\Lambda$ we have*

$$\ln \mathrm{Tr}_\Lambda e^{H + \ln \omega} \le \mathrm{Tr}_\Lambda [\omega H] + \frac{|\Lambda|}{8} \|H\|_L^2 . \tag{8.17}$$

Upon replacing the constant $\frac{1}{8}$ by $\frac{1}{2}$, the inequality (8.17) is implied by (8.16).

Proof. We will prove the claim by induction on the size of Λ. For $\Lambda = \emptyset$ equality holds in (8.16). Let us fix $x \in \Lambda$, and let $\Lambda_0 = \Lambda \setminus x$. Let $H_{\Lambda_0} \in \mathcal{O}_{\Lambda_0}$ such that

$$\partial_x H = 2 \|H - H_{\Lambda_0}\|_\infty . \tag{8.18}$$

We have

$$\|H - \mathrm{Tr}_x [\omega_x H]\|_\infty = \|H - H_{\Lambda_0} - \mathrm{Tr}_x [\omega_x (H - H_{\Lambda_0})]\|_\infty$$
$$\le 2 \|H - H_{\Lambda_0}\|_\infty = \partial_x H. \tag{8.19}$$

Using the inequality

$$e^t \le \frac{\sinh a}{a} t + e^{\frac{a^2}{2}}, \qquad |t| \le a, \tag{8.20}$$

we get

$$e^{H - \mathrm{Tr}_x [\omega_x H]} \le \frac{\sinh \partial_x H}{\partial_x H} (H - \mathrm{Tr}_x [\omega_x H]) + e^{\frac{(\partial_x H)^2}{2}}, \tag{8.21}$$

therefore

$$\mathrm{Tr}_x \left[\omega_x e^{H - \mathrm{Tr}_x [\omega_x H]} \right] \le e^{\frac{(\partial_x H)^2}{2}} . \tag{8.22}$$

For any $y \in \Lambda_0$, let $H_{\Lambda \setminus y} \in \mathcal{O}_{\Lambda \setminus y}$ such that

$$\partial_y H = 2 \left\| H - H_{\Lambda \setminus y} \right\|_\infty . \tag{8.23}$$

We have

$$\partial_y \mathrm{Tr}_x \left[\omega_x H \right] \le 2 \left\| \mathrm{Tr}_x \left[\omega_x H \right] - \mathrm{Tr}_x \left[\omega_x H_{\Lambda \setminus y} \right] \right\|_\infty \le 2 \left\| H - H_{\Lambda \setminus y} \right\|_\infty = \partial_y H. \tag{8.24}$$

We then have

$$\ln \mathrm{Tr}_\Lambda e^{H + \ln \omega} = \ln \mathrm{Tr}_\Lambda \exp \left(H - \mathrm{Tr}_x \left[\omega_x H \right] + \ln \omega_x + \mathrm{Tr}_x \left[\omega_x H \right] + \ln \omega_{\Lambda_0} \right)$$

$$\overset{(a)}{\le} \ln \int_0^\infty \mathrm{Tr}_\Lambda \left[e^{\mathrm{Tr}_x [\omega_x H] + \ln \omega_{\Lambda_0}} \left(\omega_x^{-1} + t \right)^{-1} e^{H - \mathrm{Tr}_x [\omega_x H]} \left(\omega_x^{-1} + t \right)^{-1} \right] \mathrm{d}t$$

$$= \ln \mathrm{Tr}_{\Lambda_0} \left[e^{\mathrm{Tr}_x [\omega_x H] + \ln \omega_{\Lambda_0}} \mathrm{Tr}_x \left[e^{H - \mathrm{Tr}_x [\omega_x H]} \int_0^\infty \left(\omega_x^{-1} + t \right)^{-2} \mathrm{d}t \right] \right]$$

$$= \ln \mathrm{Tr}_{\Lambda_0} \left[e^{\mathrm{Tr}_x [\omega_x H] + \ln \omega_{\Lambda_0}} \mathrm{Tr}_x \left[\omega_x e^{H - \mathrm{Tr}_x [\omega_x H]} \right] \right]$$

$$\overset{(b)}{\le} \frac{(\partial_x H)^2}{2} + \ln \mathrm{Tr}_{\Lambda_0} e^{\mathrm{Tr}_x [\omega_x H] + \ln \omega_{\Lambda_0}}$$

$$\overset{(c)}{\le} \frac{(\partial_x H)^2}{2} + \mathrm{Tr}_\Lambda \left[\omega H \right] + \frac{1}{2} \sum_{y \in \Lambda_0} \left(\partial_y \mathrm{Tr}_x \left[\omega_x H \right] \right)^2$$

$$\overset{(d)}{\le} \mathrm{Tr}_\Lambda \left[\omega H \right] + \frac{1}{2} \sum_{y \in \Lambda} \left(\partial_y H \right)^2 , \tag{8.25}$$

where (a) follows from the Golden–Thompson inequality with three matrices [89], (b) from (8.22), (c) from the inductive hypothesis and (d) from (8.24). The claim follows. $\qquad\square$

Corollary 8.1. *Let* $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ *and let* $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$ *be the uniform distribution, i.e.,* $\omega_\Lambda = \frac{\mathbb{I}_\Lambda}{q^{|\Lambda|}}$ *for any* $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$. *Then,*

$$P(\Phi) \le \ln q + \frac{\|\Phi\|_L^2}{2} - \omega(E_\Phi). \tag{8.26}$$

Proof. We have

$$P(\Phi) = \lim_{a \to \infty} \frac{\ln \mathrm{Tr}_{\Lambda_a} e^{-H_{\Lambda_a}^\Phi}}{|\Lambda_a|} \overset{(a)}{\le} \ln q + \lim_{a \to \infty} \left(\frac{1}{2 |\Lambda_a|} \sum_{x \in \Lambda_a} \left(\partial_x H_{\Lambda_a}^\Phi \right)^2 - \frac{\omega \left(H_{\Lambda_a}^\Phi \right)}{|\Lambda_a|} \right)$$

$$\overset{(b)}{=} \ln q + \frac{\|\Phi\|_L^2}{2} - \omega(E_\Phi) , \tag{8.27}$$

where (a) follows from Theorem 8.1 and (b) from Proposition 8.3 and (2.21). The claim follows. $\qquad\square$

9. W_1 Continuity of the von Neumann Entropy

In this section, we prove the following continuity bound of the von Neumann entropy with respect to the quantum W_1 distance:

Theorem 9.1 (W_1 *continuity of the von Neumann entropy*). *Let Λ be a finite set. For any ρ, $\sigma \in \mathcal{S}_\Lambda$, we have*

$$\frac{|S(\rho) - S(\sigma)|}{|\Lambda|} \leq h_2 \left(\frac{\|\rho - \sigma\|_{W_1}}{|\Lambda|} \right) + \frac{\|\rho - \sigma\|_{W_1}}{|\Lambda|} \ln \left(q^2 - 1 \right). \qquad (9.1)$$

Theorem 9.1 generalizes to the quantum setting the following continuity bound of the Shannon entropy with respect to the classical W_1 distance:

Theorem 9.2 (W_1 *continuity of the Shannon entropy* [90, Proposition 8]). *Let Λ be a finite set. For any two probability distributions μ, ν on $[q]^\Lambda$, we have*

$$|S(\mu) - S(\nu)| \leq |\Lambda| \, h_2 \left(\frac{W_1(\mu, \nu)}{|\Lambda|} \right) + W_1(\mu, \nu) \ln \left(q - 1 \right). \qquad (9.2)$$

The quantum continuity bound (9.1) is identical to the classical bound (9.2) upon replacing q by q^2. Such replacement is necessary, since the von Neumann entropy does not always satisfy the classical bound (9.2) [17].

Remark 9.1. Ref. [17] proved the following weaker continuity bound for the von Neumann entropy in terms of the W_1 distance:

Theorem 9.3 ([17, Theorem 1]). *Let Λ be a finite set. For any ρ, $\sigma \in \mathcal{S}_\Lambda$,*

$$|S(\rho) - S(\sigma)| \leq g \left(\|\rho - \sigma\|_{W_1} \right) + \|\rho - \sigma\|_{W_1} \ln \left(q^2 \, |\Lambda| \right), \qquad (9.3)$$

where for any $t \geq 0$

$$g(t) = (t + 1) \ln (t + 1) - t \ln t. \qquad (9.4)$$

Due to the term $\ln |\Lambda|$, the bound (9.3) does not have the right scaling with respect to $|\Lambda|$ to prove a continuity bound for the specific entropy in terms of the specific quantum W_1 distance. On the contrary, Theorem 9.1 will be crucial in the proof of such a bound, which will be the subject of Sect. 10.

9.1. Proof of Theorem 9.1

The proof of Theorem 9.1 is based on the following notion of distance operator:

Definition 9.1 (*Distance operator* [91, Section 2], [92, Definition 15]). Let $\mathcal{V}$ be a subspace of $\mathcal{H}_\Lambda$. For any $k = 0, \ldots, |\Lambda|$, we define the *fattening* $\mathcal{V}_k$ of $\mathcal{V}$ of radius k as the span of the linear operators acting on at most k sites applied to a vector in $\mathcal{V}$:

$$\mathcal{V}_k = \mathrm{span} \left\{ O|\psi\rangle : |\psi\rangle \in \mathcal{V}, \ O \in \mathfrak{U}_X : X \subseteq \Lambda, \ |X| \leq k \right\}, \qquad (9.5)$$

such that

$$\mathcal{V} = \mathcal{V}_0 \subseteq \ldots \subseteq \mathcal{V}_{|\Lambda|} = \mathcal{H}_\Lambda. \qquad (9.6)$$

We define the *distance operator* of $\mathcal{V}$ as the linear operator $H_\mathcal{V} \in \mathcal{O}_\Lambda$ that has eigenvalue k on $\mathcal{V}_k \cap \mathcal{V}_{k-1}^\perp$ for each $k = 0, \ldots, |\Lambda|$.

The following Proposition 9.1 provides the link between the distance operator and the W_1 distance:

Proposition 9.1. *Let $\mathcal{V}$ be a subspace of $\mathcal{H}_\Lambda$, and let ρ, $\sigma \in \mathcal{S}_\Lambda$ such that the support of σ is contained in $\mathcal{V}$. Then,*

$$\|\rho - \sigma\|_{W_1} \geq \mathrm{Tr}_\Lambda\left[\rho\, H_\mathcal{V}\right]. \tag{9.7}$$

Proof. Since $\mathrm{Tr}_\Lambda\left[\sigma\, H_\mathcal{V}\right] = 0$, it is sufficient to prove that $\|H_\mathcal{V}\|_L \leq 1$. For any $k = 0, \ldots, |\Lambda|$, let Π_k be the orthogonal projector onto $\mathcal{V}_k$, such that

$$H_\mathcal{V} = \sum_{k=0}^{|\Lambda|} \left(\mathbb{I} - \Pi_k\right). \tag{9.8}$$

For any $x \in \Lambda$, let

$$\mathcal{V}_{k,x} = \mathrm{span}\left\{O|\psi\rangle : |\psi\rangle \in \mathcal{V},\ O \in \mathfrak{U}_X : X \subseteq \Lambda,\ |X| \leq k,\ x \in X\right\}, \tag{9.9}$$

and let $\Pi_{k,x}$ be the orthogonal projector onto $\mathcal{V}_{k,x}$. We have $\mathcal{V}_{k-1} \subseteq \mathcal{V}_{k,x} \subseteq \mathcal{V}_k$, therefore

$$0 \leq \Pi_k - \Pi_{k,x} \leq \Pi_k - \Pi_{k-1}. \tag{9.10}$$

The subspace $\mathcal{V}_{k,x}$ is invariant with respect to the action of any unitary operator $U \in \mathfrak{U}_x$. Then, $\Pi_{k,x}$ commutes with any such U, and therefore $\Pi_{k,x} \in \mathcal{O}_{\Lambda \setminus x}$. Then,

$$\partial_x H_\mathcal{V} = \partial_x \sum_{k=0}^{|\Lambda|} \left(\mathbb{I} - \Pi_k\right) = \partial_x \sum_{k=0}^{|\Lambda|} \Pi_k = \partial_x \sum_{k=0}^{|\Lambda|} \left(\Pi_k - \Pi_{k,x}\right)$$

$$\overset{(a)}{\leq} \left\|\sum_{k=0}^{|\Lambda|} \left(\Pi_k - \Pi_{k,x}\right)\right\|_\infty \left\|\sum_{k=0}^{|\Lambda|} \left(\Pi_k - \Pi_{k-1}\right)\right\|_\infty \overset{(b)}{=} 1, \tag{9.11}$$

where (a) follows from Lemma C.2 and (b) follows observing that $\Pi_k - \Pi_{k-1}$ is the orthogonal projector onto $\mathcal{V}_k \cap \mathcal{V}_{k-1}^\perp$. The claim follows. $\qquad\square$

We first prove Theorem 9.1 when σ is proportional to an orthogonal projector:

Proposition 9.2. *Let $\mathcal{V}$ be a subspace of $\mathcal{H}_\Lambda$, let Π be the associated orthogonal projector and let*

$$\sigma = \frac{\Pi}{\dim \mathcal{V}}. \tag{9.12}$$

Then, for any $\rho \in \mathcal{S}_\Lambda$ we have

$$S(\rho) - \ln \dim \mathcal{V} \leq |\Lambda|\, h_2\left(\frac{\|\rho - \sigma\|_{W_1}}{|\Lambda|}\right) + \|\rho - \sigma\|_{W_1} \ln\left(q^2 - 1\right). \tag{9.13}$$

Proof. For any $0 \leq t \leq 1$, let

$$\phi(t) = h_2(t) + t \ln\left(q^2 - 1\right), \tag{9.14}$$

and let

$$\|\rho - \sigma\|_{W_1} = |\Lambda|\, w. \tag{9.15}$$

The claim becomes

$$S(\rho) \leq \ln \dim \mathcal{V} + |\Lambda|\, \phi(w). \tag{9.16}$$

ϕ is increasing in $\left[0, 1 - \frac{1}{q^2}\right]$ and decreasing in $\left[1 - \frac{1}{q^2}, 1\right]$ with

$$\phi(0) = 0, \qquad \phi\left(1 - \frac{1}{q^2}\right) = \ln q^2, \qquad \phi(1) = \ln\left(q^2 - 1\right). \tag{9.17}$$

Let $0 < w^* < 1 - \frac{1}{q^2}$ satisfy

$$\phi(w^*) = \ln q. \tag{9.18}$$

If $w \geq w^*$, the claim is trivial. Indeed, if $w^* \leq w \leq 1 - \frac{1}{q^2}$ we have

$$\phi(w) \geq \phi(w^*) = \ln q, \tag{9.19}$$

while if $1 - \frac{1}{q^2} < w \leq 1$ we have

$$\phi(w) \geq \phi(1) = \ln\left(q^2 - 1\right) \geq \ln q. \tag{9.20}$$

We can then assume $w < w^*$.

Let $H_{\mathcal{V}}$ be the distance operator of $\mathcal{V}$. For any $k = 0, \ldots, |\Lambda|$, let $\mathcal{W}_k$ be the eigenspace of $H_{\mathcal{V}}$ with eigenvalue k, let P_k be the orthogonal projector onto $\mathcal{W}_k$, and let $p_k = \mathrm{Tr}_{\Lambda}\left[\rho \, P_k\right]$ be the probability that a measurement of $H_{\mathcal{V}}$ on ρ has outcome k. Let

$$\tilde{\rho} = \sum_{k=0}^{|\Lambda|} p_k \frac{P_k}{\dim \mathcal{W}_k}. \tag{9.21}$$

We have

$$0 \leq S(\rho \| \tilde{\rho}) = S(\tilde{\rho}) - S(\rho). \tag{9.22}$$

For any $x \in \left\{0, \ldots, q^2 - 1\right\}^{\Lambda}$, let $H(x)$ be the number of components of x that are different from 0, and for any $k = 0, \ldots, |\Lambda|$, let

$$D_k = \left|H^{-1}(k)\right| = \left|\left\{x \in \left\{0, \ldots, q^2 - 1\right\}^{\Lambda} : H(x) = k\right\}\right|. \tag{9.23}$$

Let X be a random variable with values in $\left\{0, \ldots, q^2 - 1\right\}^{\Lambda}$ distributed as follows. Let the probability distribution of $H(X)$ be p, and for any $k = 0, \ldots, |\Lambda|$, let the probability distribution of X conditioned on $H(X) = k$ be uniform, such that the probability of $x \in \left\{0, \ldots, q^2 - 1\right\}^{\Lambda}$ is

$$\mathbb{P}(X = x) = \frac{p_{H(x)}}{D_{H(x)}}. \tag{9.24}$$

Since $H(X)$ has the same probability distribution as $H_{\mathcal{V}}$ measured on ρ, we have

$$\mathbb{E}\, H(X) = \mathrm{Tr}_{\Lambda}\left[\rho \, H_{\mathcal{V}}\right] =: |\Lambda|\, u. \tag{9.25}$$

By the maximum entropy principle, the Shannon entropy of X is upper bounded by the Shannon entropy of the Gibbs distribution of H with average energy $|\Lambda|\, u$:

$$S(X) \leq |\Lambda|\, \phi(u). \tag{9.26}$$

We then have

$$S(\rho) \overset{(a)}{\le} S(\tilde{\rho}) = \sum_{k=0}^{|\Lambda|} p_k \ln \frac{\dim \mathcal{W}_k}{p_k} \overset{(b)}{\le} \sum_{k=0}^{|\Lambda|} p_k \ln \frac{D_{\tilde{k}} \dim \mathcal{V}}{p_k} = \ln \dim \mathcal{V} + S(X)$$

$$\overset{(c)}{\le} \ln \dim \mathcal{V} + |\Lambda|\, \phi(u)\,, \tag{9.27}$$

where (a) follows from (9.22), (b) from Lemma C.3 and (c) from (9.26). We have from Proposition 9.1

$$w \ge \frac{\mathrm{Tr}_\Lambda [\rho\, H_\mathcal{V}]}{|\Lambda|} = u, \tag{9.28}$$

hence

$$\phi(u) \le \phi(w). \tag{9.29}$$

The claim follows.   □

Without loss of generality, we can assume $S(\rho) \ge S(\sigma)$. For any $k \in \mathbb{N}$ and any $\delta > 0$, let $P_{k,\delta}$ be the δ-typical projector of $\sigma^{\otimes k}$, i.e., the orthogonal projector on the sum of the eigenspaces of $\sigma^{\otimes k}$ with eigenvalues contained in $\left[e^{-k(S(\sigma)+\delta)}, e^{-k(S(\sigma)-\delta)}\right]$. $P_{k,\delta}$ satisfies [79–81, Section 5.5]

$$\sigma^{\otimes k} \ge e^{-k(S(\sigma)+\delta)}\, P_{k,\delta}, \tag{9.30a}$$

$$\liminf_{k \to \infty} \frac{\ln \mathrm{Tr}_\Lambda P_{k,\delta}}{k} \ge S(\sigma) - \delta. \tag{9.30b}$$

The property (9.30a) implies

$$\frac{\ln \mathrm{Tr}_\Lambda P_{k,\delta}}{k} \le S(\sigma) + \delta, \tag{9.31a}$$

$$\frac{1}{k} S\left(\frac{P_{k,\delta}}{\mathrm{Tr}_\Lambda P_{k,\delta}} \,\middle\|\, \sigma^{\otimes k}\right) \le S(\sigma) + \delta - \frac{\ln \mathrm{Tr}_\Lambda P_{k,\delta}}{k}. \tag{9.31b}$$

We have from (9.31a) and Proposition 9.2

$$S(\rho) - S(\sigma) \le \frac{S\left(\rho^{\otimes k}\right) - \ln \mathrm{Tr}_\Lambda P_{k,\delta}}{k} + \delta \le |\Lambda|\, \phi\left(\frac{\left\|\rho^{\otimes k} - \frac{P_{k,\delta}}{\mathrm{Tr}_\Lambda P_{k,\delta}}\right\|_{W_1}}{k\,|\Lambda|}\right) + \delta. \tag{9.32}$$

We have

$$\frac{\left\|\rho^{\otimes k} - \frac{P_{k,\delta}}{\mathrm{Tr}_\Lambda P_{k,\delta}}\right\|_{W_1}}{k\,|\Lambda|} \le \frac{\left\|\rho^{\otimes k} - \sigma^{\otimes k}\right\|_{W_1}}{k\,|\Lambda|} + \frac{\left\|\sigma^{\otimes k} - \frac{P_{k,\delta}}{\mathrm{Tr}_\Lambda P_{k,\delta}}\right\|_{W_1}}{k\,|\Lambda|}$$

$$\overset{(a)}{\le} w + \sqrt{\frac{2}{k} S\left(\frac{P_{k,\delta}}{\mathrm{Tr}_\Lambda P_{k,\delta}} \,\middle\|\, \sigma^{\otimes k}\right)}$$

$$\overset{(b)}{\le} w + \sqrt{2\left(S(\sigma) + \delta - \frac{\ln \mathrm{Tr}_\Lambda P_{k,\delta}}{k}\right)}, \tag{9.33}$$

where (a) follows from Proposition A.3 and Proposition C.1, and (b) follows from (9.31b). We get from (9.30b)

$$\limsup_{k\to\infty} \frac{\left\| \rho^{\otimes k} - \frac{P_{k,\delta}}{\mathrm{Tr}_\Lambda P_{k,\delta}} \right\|_{W_1}}{k\,|\Lambda|} \le w + 2\sqrt{\delta}. \tag{9.34}$$

We then get from (9.32)

$$S(\rho) - S(\sigma) \le |\Lambda|\,\phi\left(w + 2\sqrt{\delta}\right) + \delta, \tag{9.35}$$

and the claim follows taking the limit $\delta \to 0$.

10. w_1 Continuity of the Specific Entropy

A fundamental consequence of Theorem 9.1 is the following continuity bound for the specific entropy in terms of the specific quantum W_1 distance:

Corollary 10.1 (w_1 continuity of the specific entropy). *The specific entropy satisfies the following continuity bound with respect to the specific quantum W_1 distance: For any ρ, $\sigma \in \mathcal{S}_{\mathbb{Z}^d}^I$, we have*

$$|s(\rho) - s(\sigma)| \le h_2\left(w_1(\rho,\sigma)\right) + w_1(\rho,\sigma)\ln\left(q^2 - 1\right), \tag{10.1}$$

where

$$h_2(t) = -t\ln t - (1-t)\ln(1-t), \qquad 0 \le t \le 1 \tag{10.2}$$

is the binary entropy function.

Proof. The claim follows from Theorem 9.1: We have

$$
\begin{aligned}
|s(\rho) - s(\sigma)| &= \lim_{a\to\infty} \frac{|S\left(\rho_{\Lambda_a}\right) - S\left(\sigma_{\Lambda_a}\right)|}{|\Lambda_a|} \\
&\le \lim_{a\to\infty} \left(h_2\left(\frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|}\right) + \frac{\|\rho_{\Lambda_a} - \sigma_{\Lambda_a}\|_{W_1}}{|\Lambda_a|}\ln\left(q^2 - 1\right) \right) \\
&= h_2\left(w_1(\rho,\sigma)\right) + w_1(\rho,\sigma)\ln\left(q^2 - 1\right) .
\end{aligned} \tag{10.3}
$$

$\square$

11. w_1-Gibbs States

We define the w_1-Gibbs states of the interaction Φ as the translation-invariant states whose marginal states have a W_1 distance from the local Gibbs states of Φ that scales sublinearly with the volume:

Definition 11.1 (w_1-*Gibbs state*). Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$. We define for any $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ the specific quantum W_1 distance between ρ and Φ as the limit of the W_1 distance per site between the marginals of ρ and the local Gibbs states of Φ:

$$w_1(\rho, \Phi) = \limsup_{a\to\infty} \frac{\left\| \rho_{\Lambda_a} - \omega_{\Lambda_a}^\Phi \right\|_{W_1}}{|\Lambda_a|}. \tag{11.1}$$

We say that the state $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$ is a w_1-*Gibbs state* of Φ if $w_1(\omega, \Phi) = 0$.

We also define for any $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ and any $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ the specific relative entropy between ρ and Φ as the limit of the relative entropy per site between the marginals of ρ and the local Gibbs states of Φ:

$$s(\rho\|\Phi) = \lim_{a\to\infty} \frac{S\left(\rho_{\Lambda_a}\|\omega_{\Lambda_a}^\Phi\right)}{|\Lambda_a|} = P(\Phi) - s(\rho) + \rho(E_\Phi). \qquad (11.2)$$

We have $s(\rho\|\Phi) \geq 0$, with equality iff $\rho \in \mathcal{S}_{eq}(\Phi)$.

An interaction can have at most one w_1-Gibbs state:

Proposition 11.1 (Uniqueness of the w_1-Gibbs state). *Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ have a w_1-Gibbs state $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$. Then, for any $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ we have*

$$w_1(\rho, \Phi) = w_1(\rho, \omega). \qquad (11.3)$$

In particular, Φ can have at most one w_1-Gibbs state.

Proof. We have

$$|w_1(\rho, \Phi) - w_1(\rho, \omega)| = \left|\limsup_{a\to\infty} \frac{\left\|\rho_{\Lambda_a} - \omega_{\Lambda_a}^\Phi\right\|_{W_1} - \left\|\rho_{\Lambda_a} - \omega_{\Lambda_a}\right\|_{W_1}}{|\Lambda_a|}\right|$$

$$\leq \limsup_{a\to\infty} \frac{\left\|\omega_{\Lambda_a} - \omega_{\Lambda_a}^\Phi\right\|_{W_1}}{|\Lambda_a|} = w_1(\omega, \Phi) = 0. \qquad (11.4)$$

If also ρ is a w_1-Gibbs state of Φ, we have

$$w_1(\rho, \omega) = w_1(\rho, \Phi) = 0, \qquad (11.5)$$

hence $\rho = \omega$. The claim follows. $\qquad \square$

If an interaction admits a w_1-Gibbs state, then such state is also an equilibrium state:

Proposition 11.2. *We have for any $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ and any $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$*

$$s(\rho\|\Phi) \leq h_2(w_1(\rho, \Phi)) + w_1(\rho, \Phi)\left(\ln\left(q^2 - 1\right) + 2\left\|\Phi\right\|_r\right). \qquad (11.6)$$

In particular, if Φ has a w_1-Gibbs state $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$, then $\omega \in \mathcal{S}_{eq}(\Phi)$.

Proof. We have

$$s(\rho\|\Phi) = \lim_{a\to\infty} \frac{S\left(\rho_{\Lambda_a}\|\omega_{\Lambda_a}^\Phi\right)}{|\Lambda_a|} = \lim_{a\to\infty} \frac{S\left(\omega_{\Lambda_a}^\Phi\right) - S(\rho_{\Lambda_a}) + \mathrm{Tr}_{\Lambda_a}\left[\left(\rho_{\Lambda_a} - \omega_{\Lambda_a}^\Phi\right)H_{\Lambda_a}^\Phi\right]}{|\Lambda_a|}$$

$$\overset{(a)}{\leq} \lim_{a\to\infty}\left(h_2\left(\frac{\left\|\rho_{\Lambda_a} - \omega_{\Lambda_a}^\Phi\right\|_{W_1}}{|\Lambda_a|}\right) + \frac{\left\|\rho_{\Lambda_a} - \omega_{\Lambda_a}^\Phi\right\|_{W_1}}{|\Lambda_a|}\left(\ln\left(q^2 - 1\right) + 2\left\|\Phi\right\|_r\right)\right)$$

$$= h_2(w_1(\rho, \Phi)) + w_1(\rho, \Phi)\left(\ln\left(q^2 - 1\right) + 2\left\|\Phi\right\|_r\right), \qquad (11.7)$$

where (a) follows from Theorem 9.1 and (C.16). The claim follows. $\qquad \square$

12. Quantum Transportation-Cost Inequalities

Definition 12.1 (*TCI*). The interaction $\Phi \in \mathcal{B}^r_{\mathbb{Z}^d}$ satisfies a *Transportation-Cost Inequality* (TCI) with constant $c > 0$ if the square of the specific quantum W_1 distance with respect to Φ is upper bounded by $\frac{c}{2}$ times the specific relative entropy with respect to Φ, i.e., if for any $\rho \in \mathcal{S}^I_{\mathbb{Z}^d}$ we have

$$w_1(\rho, \Phi)^2 \leq \frac{c}{2}\, s(\rho \| \Phi)\,. \tag{TCI}$$

A fundamental consequence of (TCI) is the uniqueness of the equilibrium state of Φ:

Proposition 12.1 (Uniqueness of the equilibrium state). *Let* $\Phi \in \mathcal{B}^r_{\mathbb{Z}^d}$ *satisfy* (TCI). *Then,* Φ *has a unique equilibrium state, which is a* w_1-*Gibbs state.*

Proof. Let $\omega \in \mathcal{S}_{eq}(\Phi)$. From (TCI), we have

$$w_1(\omega, \Phi)^2 \leq \frac{c}{2}\, s(\omega \| \Phi) = 0, \tag{12.1}$$

therefore ω is a w_1-Gibbs state of Φ. Since the w_1-Gibbs state is unique, the equilibrium state is unique, too. $\qquad\square$

Another property of the interactions satisfying (TCI) is the following upper bound to the variation of the specific entropy in terms of the specific relative entropy:

Proposition 12.2. *Let* $\Phi \in \mathcal{B}^r_{\mathbb{Z}^d}$ *satisfy* (TCI) *and let* $\omega \in \mathcal{S}^I_{\mathbb{Z}^d}$ *be its unique equilibrium state. Let* w^* *be as in* (9.18). *Then, for any* $\rho \in \mathcal{S}^I_{\mathbb{Z}^d}$ *such that*

$$s(\rho \| \Phi) \leq \frac{2\, w^{*2}}{c} \tag{12.2}$$

we have

$$|s(\rho) - s(\omega)| \leq h_2\left(\sqrt{\frac{c}{2}\, s(\rho\|\Phi)}\right) + \sqrt{\frac{c}{2}\, s(\rho\|\Phi)}\, \ln\left(q^2 - 1\right). \tag{12.3}$$

Proof. From Theorem 12.1, we have that ω is a w_1-Gibbs state of Φ. We then have

$$|s(\rho) - s(\omega)| \overset{(a)}{\leq} h_2(w_1(\rho, \omega)) + w_1(\rho, \omega) \ln\left(q^2 - 1\right)$$
$$\overset{(b)}{\leq} h_2\left(\sqrt{\frac{c\, s(\rho\|\Phi)}{2}}\right) + \sqrt{\frac{c\, s(\rho\|\Phi)}{2}}\, \ln\left(q^2 - 1\right), \tag{12.4}$$

where (a) follows from Theorem 10.1 and (b) from Proposition 11.1 and (TCI). The claim follows. $\qquad\square$

In the following, we will prove that (TCI) is satisfied by interactions containing only single-site terms (Sect. 12.1) and local commuting interactions at high temperature (Sect. 12.2).

12.1. Product States

The simplest setting where (TCI) holds is when the interaction contains only terms acting on single spins and the associated Gibbs state is a product state. Ref. [17] proved the following TCI for product states on finite lattices:

Theorem 12.1 (Quantum Marton's Transportation Inequality [17, Theorem 2]). *Let Λ be a finite set and let $\sigma \in \mathcal{S}_\Lambda$ be a product state. Then, for any $\rho \in \mathcal{S}_\Lambda$ we have*

$$\|\rho - \sigma\|_{W_1}^2 \leq \frac{|\Lambda|}{2} S(\rho \| \sigma). \tag{12.5}$$

Theorem 12.1 implies the following TCI for product states on $\mathbb{Z}^d$:

Corollary 12.1 (TCI for product states). *Let $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$ be a product state. Then, for any $\rho \in \mathcal{S}_{\mathbb{Z}^d}^I$ we have*

$$w_1(\rho, \omega)^2 \leq \frac{1}{2} s(\rho \| \omega). \tag{12.6}$$

Therefore, any $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ that contains only terms acting on single spins (i.e., such that $\Phi(\Lambda) = 0$ for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ with $|\Lambda| \geq 2$) satisfies (TCI) with $c = 1$.

Proof. The claim (12.6) follows from Theorem 12.1: We have

$$w_1(\rho, \omega)^2 = \lim_{a \to \infty} \frac{\|\rho_{\Lambda_a} - \omega_{\Lambda_a}\|_{W_1}^2}{|\Lambda_a|^2} \leq \lim_{a \to \infty} \frac{S(\rho_{\Lambda_a} \| \omega_{\Lambda_a})}{2 |\Lambda_a|} = \frac{s(\rho \| \omega)}{2}. \tag{12.7}$$

Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ contain only terms acting on single sites. We have for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$

$$\omega_\Lambda^\Phi = \bigotimes_{x \in \Lambda} \frac{e^{-\Phi(x)}}{\text{Tr}_x e^{-\Phi(x)}} = \bigotimes_{x \in \Lambda} \omega_x^\Phi, \tag{12.8}$$

therefore there exists a product state $\omega \in \mathcal{S}_{\mathbb{Z}^d}^I$ such that $\omega_\Lambda = \omega_\Lambda^\Phi$ for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$. We have $w_1(\rho, \omega) = w_1(\rho, \Phi)$ and $s(\rho \| \omega) = s(\rho \| \Phi)$, therefore (12.6) implies (TCI) with $c = 1$. The claim follows. □

12.2. Local Commuting Interactions at High Temperature

A more general setting where (TCI) can be proved is the case of geometrically local commuting interactions, where each spin interacts with a finite number of spins and all the terms of the interaction mutually commute.

Ref. [19] proved the following TCI for Gibbs states of local commuting interactions on a finite lattice employing a quantum generalization of Ollivier's coarse Ricci curvature [93,94]:

Theorem 12.2 (High-temperature TCI for local commuting interactions [19, Theorem 4 and Proposition 9]). *Let $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ be geometrically local and commuting, i.e., each spin interacts with at most N spins where*

$$N = \left| \bigcup_{0 \in X \in \mathcal{F}_{\mathbb{Z}^d} : \Phi(X) \neq 0} X \right| < \infty, \tag{12.9}$$

and $[\Phi(X), \Phi(Y)] = 0$ *for any* $X, Y \in \mathcal{F}_{\mathbb{Z}^d}$. *Let*

$$M = \inf_{t \geq 0} \left(\left(e^{\|\Phi\|_r} + 1 \right) \sqrt{1 + t^2} \, \|\Phi\|_r \, q^{\frac{3 + \sqrt{1 + t^2}}{2}} e^{\|\Phi\|_r \left(2 + \frac{\sqrt{1 + t^2}}{2} \right)} \right.$$

$$\left. + 2 \, \|\Phi\|_r \, e^{2\|\Phi\|_r} + 4 \, e^{-\pi t} \right), \tag{12.10}$$

and let us assume that

$$\kappa = 1 - (2N - 1)(N - 1) M > 0. \tag{12.11}$$

Then, for any $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ *and any* $\rho \in \mathcal{S}_\Lambda$ *we have*

$$\left\| \rho - \omega_\Lambda^\Phi \right\|_{W_1}^2 \leq \frac{2 \, N^2 \, |\Lambda|}{(1 - e^{-\kappa})^2} \, S\left(\rho \, \big\| \, \omega_\Lambda^\Phi \right). \tag{12.12}$$

In particular, Φ *satisfies* (TCI) *with*

$$c = \frac{4 \, N^2}{(1 - e^{-\kappa})^2}. \tag{12.13}$$

Remark 12.1. Choosing in (12.10)

$$t = \frac{\ln \frac{1}{\|\Phi\|_r}}{\pi + \frac{\ln q}{2}}, \tag{12.14}$$

we get $M \leq O\left(\|\Phi\|_r \right)$ for $\|\Phi\|_r \to 0$.

Another strategy to prove TCIs for quantum spin systems on a finite lattice is to prove that suitable local quantum Markov semigroups that have the Gibbs state as unique fixed point satisfy a modified logarithmic Sobolev inequality [19, 95–97], which states that the semigroup contracts exponentially the relative entropy with respect to the Gibbs state. Ref. [97] proved that above a critical temperature, the Gibbs states of commuting nearest-neighbor interactions satisfy a modified logarithmic Sobolev inequality. Exploiting this result, Ref. [19] proved the following TCI for such Gibbs states:

Theorem 12.3 (High-temperature TCI for nearest-neighbor interactions [19, Theorem 5]). *Let* $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ *be a nearest-neighbor interaction, i.e.,* $\Phi(\Lambda) = 0$ *for any* $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ *that contains at least two sites that are not neighboring. Then, there exists a critical inverse temperature* $\beta_c > 0$ *such that for any* $0 \leq \beta < \beta_c$ *there exists* $c_\beta > 0$ *such that for any* $a \in \mathbb{N}_+^d$ *we have*

$$\left\| \rho_{\Lambda_a} - \omega_{\Lambda_a}^{\beta\Phi} \right\|_{W_1}^2 \leq \frac{c_\beta \, |\Lambda_a|}{2} \, S\left(\rho_{\Lambda_a} \, \big\| \, \omega_{\Lambda_a}^{\beta\Phi} \right). \tag{12.15}$$

In particular, $\beta \, \Phi$ *satisfies* (TCI) *with* $c = c_\beta$.

Corollary 12.2. *All the interactions satisfying the hypotheses either of Theorem 12.2 or of Theorem 12.3 have a unique equilibrium state.*

Remark 12.2. The uniqueness of the equilibrium states for all the interactions $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$ such that $r > \log q$ and $\|\Phi\|_r < \frac{1}{2q}$ has been proved in [98].

13. Perspectives

In this paper, we have proposed a specific Wasserstein distance of order 1 for quantum spin systems on infinite lattices. We expect the proposed distance to be a powerful tool in the study of the statistical mechanics of quantum spin systems, quantum dynamical systems, and tomography of quantum states:

1. The specific quantum W_1 distance can be employed to study the diameter of the set of the equilibrium states of an interaction close to a thermal phase transition. Above the critical temperature, the equilibrium state is unique and the diameter is zero, while below the critical temperature the diameter is strictly positive. The limit of the diameter of the set of the equilibrium states as the temperature tends from below to the critical value, and in particular whether such limit is zero or strictly positive, can be employed to characterize the phase transition.

2. In [7], Ornstein proposed the $\bar{d}$-distance as a natural metric for the classification of stochastic processes and singled out a large class (the so-called B-processes), containing, e.g., all mixing Markov processes, such that a fundamental isomorphism theorem holds: two processes are isomorphic if and only if their entropies coincide. The isomorphism here is in the sense of dynamical systems, i.e., a measurable and invertible transformation mapping one probability measure to the other. The specific quantum W_1 distance could provide a useful analytical tool toward establishing analogous results in the setting of quantum dynamical systems [99].

3. The statistical problem of estimating a stationary ergodic process, in a given family, from the observation of a single sample path of length n can be quantitatively addressed using Ornstein's $\bar{d}$-distance. In [8], it is proved that an empirical block scheme, i.e., the product probability naturally obtained from the observed frequencies on sliding window of length k, converges in the $\bar{d}$-distance, as n grows, toward the target process, provided that it is a B-process and k grows at least logarithmically with respect to n. Exploring the quantum analogue of this and related results, e.g., for discrimination between two sampled processes [9], may extend the scope of the recent works [23, 24] on tomography of a quantum state and stimulate novel approaches, particularly when the number of accessible independent copies is extremely constrained.

A. Properties of the Quantum W_1 Distance

Proposition A.1 ([17, Proposition 2]). *For any finite set Λ and any $\Delta \in \mathcal{O}_\Lambda^T$, we have*

$$\frac{1}{2} \|\Delta\|_1 \leq \|\Delta\|_{W_1} \leq \frac{|\Lambda|}{2} \|\Delta\|_1. \tag{A.1}$$

Proposition A.2 ([17, Proposition 5]). *Let $\Lambda' \subseteq \Lambda$ be finite sets. Then, for any $\Delta \in \mathcal{O}_\Lambda^T$ such that $\mathrm{Tr}_{\Lambda'}\Delta = 0$ we have*

$$\|\Delta\|_{W_1} \leq \frac{q^2 - 1}{q^2} |\Lambda'| \|\Delta\|_1. \tag{A.2}$$

Proposition A.3 (Superadditivity [17, Proposition 4]). *The quantum W_1 distance is superadditive in general and additive for product states, i.e., for any two disjoint finite sets Λ, Λ' and any ρ, $\sigma \in \mathcal{S}_{\Lambda\Lambda'}$ we have*

$$\|\rho - \sigma\|_{W_1} \geq \|\rho_\Lambda - \sigma_\Lambda\|_{W_1} + \|\rho_{\Lambda'} - \sigma_{\Lambda'}\|_{W_1}, \tag{A.3}$$

and for any ρ_Λ, $\sigma_\Lambda \in \mathcal{S}_\Lambda$ and any $\rho_{\Lambda'}$, $\sigma_{\Lambda'} \in \mathcal{S}_{\Lambda'}$ we have

$$\|\rho_\Lambda \otimes \rho_{\Lambda'} - \sigma_\Lambda \otimes \sigma_{\Lambda'}\|_{W_1} = \|\rho_\Lambda - \sigma_\Lambda\|_{W_1} + \|\rho_{\Lambda'} - \sigma_{\Lambda'}\|_{W_1}. \tag{A.4}$$

B. Auxiliary Proofs

B.1. Proof of Proposition 2.1

Proposition. (2.1) *The trace distance on $\mathcal{S}_{\mathbb{Z}^d}$ is the supremum of the trace distances between the marginal states: For any ρ, $\sigma \in \mathcal{S}_{\mathbb{Z}^d}$,*

$$T(\rho, \sigma) = \frac{1}{2} \sup_{\Lambda \in \mathcal{F}_{\mathbb{Z}^d}} \|\rho_\Lambda - \sigma_\Lambda\|_1, \tag{B.1}$$

where $\|\cdot\|_1$ denotes the trace norm on $\mathfrak{U}_\Lambda$ given by

$$\|A\|_1 = \mathrm{Tr}_\Lambda \sqrt{A^\dagger A}, \qquad A \in \mathfrak{U}_\Lambda. \tag{B.2}$$

Proof. Since $\mathfrak{U}_{\mathbb{Z}^d}^{loc}$ is dense in $\mathfrak{U}_{\mathbb{Z}^d}$, we have

$$2\,T(\rho,\sigma) = \sup_{A\in\mathfrak{U}_{\mathbb{Z}^d}^{loc}:\|A\|_\infty\leq 1} |\rho(A)-\sigma(A)| = \sup_{\Lambda\in\mathcal{F}_{\mathbb{Z}^d}}\ \sup_{A\in\mathfrak{U}_\Lambda:\|A\|_\infty\leq 1} |\rho(A)-\sigma(A)|$$

$$= \sup_{\Lambda\in\mathcal{F}_{\mathbb{Z}^d}}\ \sup_{A\in\mathfrak{U}_\Lambda:\|A\|_\infty\leq 1} |\mathrm{Tr}_\Lambda\left[(\rho_\Lambda-\sigma_\Lambda)\,A\right]| = \sup_{\Lambda\in\mathcal{F}_{\mathbb{Z}^d}} \|\rho_\Lambda-\sigma_\Lambda\|_1\ . \tag{B.3}$$

The claim follows. $\qquad\square$

B.2. Proof of Proposition 5.1

Proposition. (5.1) *For any $\Lambda\in\mathcal{F}_{\mathbb{Z}^d}$, any $H\in\mathcal{O}_\Lambda$ and any $x\in\Lambda$, (3.3) and (5.1) are equivalent.*

Proof. Let

$$\partial_x H = 2\min_{A\in\mathcal{O}_{\Lambda\setminus x}} \|H-A\|_\infty, \qquad \tilde{\partial}_x H = 2\inf_{A\in\mathcal{O}_{\mathbb{Z}^d\setminus x}} \|H-A\|_\infty. \tag{B.4}$$

We clearly have $\tilde{\partial}_x H \leq \partial_x H$. Let $\omega_{\mathbb{Z}^d\setminus\Lambda}\in\mathcal{S}_{\mathbb{Z}^d\setminus\Lambda}$ be the uniform distribution on $\mathbb{Z}^d\setminus\Lambda$, and let $\Psi_\Lambda:\mathfrak{U}_{\mathbb{Z}^d}\to\mathfrak{U}_\Lambda$ be the completely positive unital linear map such that for any $A\in\mathfrak{U}_{\mathbb{Z}^d}$ and any $\rho_\Lambda\in\mathcal{S}_\Lambda$

$$\mathrm{Tr}_\Lambda\left[\rho_\Lambda\,\Psi_\Lambda(A)\right] = (\omega_{\mathbb{Z}^d\setminus\Lambda}\otimes\rho_\Lambda)(A). \tag{B.5}$$

Let $A\in\mathcal{O}_{\mathbb{Z}^d\setminus x}$. We have for any $\rho_\Lambda\in\mathcal{S}_\Lambda$ and any unitary operator $U_x\in\mathfrak{U}_x$

$$\mathrm{Tr}_\Lambda\left[\rho_\Lambda\,U_x^\dagger\,\Psi_\Lambda(A)\,U_x\right] = \mathrm{Tr}_\Lambda\left[U_x\,\rho_\Lambda\,U_x^\dagger\,\Psi_\Lambda(A)\right] = \left(\omega_{\mathbb{Z}^d\setminus\Lambda}\otimes U_x\,\rho_\Lambda\,U_x^\dagger\right)(A)$$

$$= (\omega_{\mathbb{Z}^d\setminus\Lambda}\otimes\rho_\Lambda)\left(U_x^\dagger\,A\,U_x\right) = (\omega_{\mathbb{Z}^d\setminus\Lambda}\otimes\rho_\Lambda)(A)$$

$$= \mathrm{Tr}_\Lambda\left[\rho_\Lambda\,\Psi_\Lambda(A)\right], \tag{B.6}$$

therefore $U_x^\dagger\,\Psi_\Lambda(A)\,U_x = \Psi_\Lambda(A)$, hence $\Psi_\Lambda(A)\in\mathcal{O}_{\Lambda\setminus x}$. We then have

$$\partial_x H \leq 2\,\|H-\Psi_\Lambda(A)\|_\infty = 2\,\|\Psi_\Lambda(H-A)\|_\infty \leq 2\,\|H-A\|_\infty, \tag{B.7}$$

where the last inequality follows since Ψ_Λ is completely positive and unital. We then have $\partial_x H \leq \tilde{\partial}_x H$. The claim follows. $\qquad\square$

C. Auxiliary Lemmas

Lemma C.1 (Multidimensional Fekete's lemma [100]). *Let $f:\mathbb{N}_+^d\to\mathbb{R}$ be superadditive with respect to each variable, i.e.,*

$$f(x_1,\,\dots,\,x_i+t,\,\dots,\,x_d) \geq f(x_1,\,\dots,\,x_i,\,\dots,\,x_d) + f(x_1,\,\dots,\,t,\,\dots,\,x_d) \tag{C.1}$$

for any $x_1,\,\dots,\,x_d,\,t\in\mathbb{N}$ and any $i=1,\,\dots,\,d$. Then,

$$\lim_{x\to\infty}\frac{f(x)}{x_1\dots x_d} = \sup_{x\in\mathbb{N}_+^d}\frac{f(x)}{x_1\dots x_d}. \tag{C.2}$$

Lemma C.2. *Let $H\in\mathcal{O}_\Lambda$ be positive semi-definite. Then, for any $x\in\Lambda$,*

$$\partial_x H \leq \|H\|_\infty. \tag{C.3}$$

Proof. We have

$$-\frac{\|H\|_\infty}{2}\, \mathbb{I} \le H - \frac{\|H\|_\infty}{2}\, \mathbb{I} \le \frac{\|H\|_\infty}{2}\, \mathbb{I}, \tag{C.4}$$

therefore

$$\partial_x H \le 2\left\|H - \frac{\|H\|_\infty}{2}\, \mathbb{I}\right\|_\infty \le \|H\|_\infty\,. \tag{C.5}$$

The claim follows. $\qquad\square$

Proposition C.1. *Let* $\Lambda_1,\, \dots,\, \Lambda_k$ *be* k *copies of the finite set* Λ. *Then, for any* $\rho \in \mathcal{S}_{\Lambda_1\dots\Lambda_k}$ *and any* $\sigma \in \mathcal{S}_\Lambda$ *we have*

$$\left\|\rho - \sigma^{\otimes k}\right\|_{W_1}^2 \le 2k\,|\Lambda|^2\, S\left(\rho\,\|\,\sigma^{\otimes k}\right). \tag{C.6}$$

Proof. The proof follows the same lines as the proof of [17, Theorem 2]. We have

$$\left\|\rho - \sigma^{\otimes k}\right\|_{W_1} \le \sum_{i=1}^k \left\|\sigma^{\otimes(i-1)} \otimes \rho_{\Lambda_i\dots\Lambda_k} - \sigma^{\otimes i} \otimes \rho_{\Lambda_{i+1}\dots\Lambda_k}\right\|_{W_1}$$

$$\overset{(a)}{\le} |\Lambda| \sum_{i=1}^k \left\|\rho_{\Lambda_i\dots\Lambda_k} - \sigma \otimes \rho_{\Lambda_{i+1}\dots\Lambda_k}\right\|_1$$

$$\overset{(b)}{\le} |\Lambda| \sum_{i=1}^k \sqrt{2\,S\left(\rho_{\Lambda_i\dots\Lambda_k}\,\|\,\sigma \otimes \rho_{\Lambda_{i+1}\dots\Lambda_k}\right)}$$

$$= |\Lambda| \sum_{i=1}^k \sqrt{2\left(S(\rho_{\Lambda_i}) + S(\rho_{\Lambda_{i+1}\dots\Lambda_k}) - S(\rho_{\Lambda_i\dots\Lambda_k}) + S(\rho_{\Lambda_i}\|\sigma)\right)}$$

$$\overset{(c)}{\le} |\Lambda| \sqrt{2k \sum_{i=1}^k \left(S(\rho_{\Lambda_i}) + S(\rho_{\Lambda_{i+1}\dots\Lambda_k}) - S(\rho_{\Lambda_i\dots\Lambda_k}) + S(\rho_{\Lambda_i}\|\sigma)\right)}$$

$$= |\Lambda|\sqrt{2k} \sqrt{\sum_{i=1}^k \left(S(\rho_{\Lambda_i}) + S(\rho_{\Lambda_i}\|\sigma)\right) - S(\rho)} = |\Lambda|\sqrt{2k\,S\left(\rho\,\|\,\sigma^{\otimes k}\right)}. \tag{C.7}$$

(a) follows from Proposition A.2 observing that

$$\mathrm{Tr}_{\Lambda_i}\left[\sigma^{\otimes(i-1)} \otimes \rho_{\Lambda_i\dots\Lambda_k} - \sigma^{\otimes i} \otimes \rho_{\Lambda_{i+1}\dots\Lambda_k}\right] = 0; \tag{C.8}$$

(b) follows from Pinsker's inequality; (c) follows from the concavity of the square root. The claim follows. $\qquad\square$

Lemma C.3. *We have*

$$\dim \mathcal{W}_k \le D_k \dim \mathcal{V}. \tag{C.9}$$

Proof. Let $A_0,\, \dots,\, A_{q^2-1}$ be a basis of $\mathbb{C}^{q\times q}$ with $A_0 = \mathbb{I}$. For any $x \in \{0,\, \dots,\, q^2 - 1\}^\Lambda$, let

$$A_x = \bigotimes_{i\in\Lambda} A_{x_i}, \tag{C.10}$$

where each A_{x_i} acts on the site i. We have

$$\mathcal{W}_k \subseteq \operatorname{span} \{A_x |\psi\rangle : |\psi\rangle \in \mathcal{V},\, H(x) \leq k\}. \tag{C.11}$$

We also have

$$\mathcal{W}_{k-1} \subseteq \operatorname{span} \{A_x |\psi\rangle : |\psi\rangle \in \mathcal{V},\, H(x) \leq k-1\}, \tag{C.12}$$

and since $\mathcal{W}_k \perp \mathcal{W}_{k-1}$, we have

$$\mathcal{W}_k \subseteq \operatorname{span} \{A_x |\psi\rangle : |\psi\rangle \in \mathcal{V},\, H(x) = k\}. \tag{C.13}$$

Therefore,

$$\dim \mathcal{W}_k \leq \left|H^{-1}(k)\right| \dim \mathcal{V}. \tag{C.14}$$

The claim follows. □

Lemma C.4. *Let* $\Phi \in \mathcal{B}_{\mathbb{Z}^d}^r$. *Then, for any* $\Lambda \in \mathcal{F}_{\mathbb{Z}^d}$ *and any* $x \in \Lambda$ *we have*

$$\partial_x H_\Lambda^\Phi \leq 2 \|\Phi\|_r, \tag{C.15}$$

and

$$\left\|H_\Lambda^\Phi\right\|_L \leq 2 \|\Phi\|_r. \tag{C.16}$$

Proof. We have

$$\partial_x H_\Lambda^\Phi \leq 2 \sum_{x \in X \subseteq \Lambda} \|\Phi(X)\|_\infty \overset{\text{(a)}}{=} 2 \sum_{0 \in X \subseteq \Lambda - x} \|\Phi(X)\|_\infty$$
$$\leq 2 \sum_{0 \in X \in \mathcal{F}_{\mathbb{Z}^d}} \|\Phi(X)\|_\infty \leq 2 \|\Phi\|_r, \tag{C.17}$$

where (a) follows from the translation invariance of Φ. The claim follows. □

References

[1] Monge, G.: Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences de Paris (1781)

[2] Kantorovich, L.V.: On the translocation of masses. In Dokl. Akad. Nauk. USSR (NS) **37**, 199–201 (1942)

[3] Vershik, A.M.: Long history of the Monge–Kantorovich transportation problem. Math. Intell. **35**(4), 1–9 (2013)

[4] Villani, C.: Optimal Transport: old and new, volume 338. Springer Science & Business Media (2008)

[5] Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows: in Metric Spaces and in the Space of Probability Measures. Springer Science & Business Media (2008)

[6] Peyré, G., Cuturi, M.: Computational optimal transport: with applications to data science. Found. Trends® Mach. Learn. **11**(5–6), 355–607 (2019)

[7] Ornstein, D.S.: An application of ergodic theory to probability theory. Ann. Probab. **1**(1), 43–58 (1973)

[8] Ornstein, D.S., Weiss, B.: How sampling reveals a process. Ann. Probab. **18**(3), 905–930 (1990)

[9] Ornstein, D.S., Shields, P.C.: The d-recognition of processes. Adv. Math. **104**(2), 182–224 (1994)

[10] Fernandez, R., Galves, A.: Markov approximations of chains of infinite order. Bull. Braz. Math. Soc. **33**(3), 295–306 (2002)

[11] Csiszár, I., Talata, Z.: On rate of convergence of statistical estimation of stationary ergodic processes. IEEE Trans. Inf. Theory **56**(8), 3637–3641 (2010)

[12] O'Connor, K., McGoff, K., Nobel, A.B: Estimation of stationary optimal transport plans (2021). arXiv preprint arXiv:2107.11858

[13] Gray, R.M., Neuhoff, D.L., Shields, P.C.: A generalization of Ornstein's d distance with applications to information theory. Ann. Probab. 315–328 (1975)

[14] Gray, R.M.: Entropy and Information Theory. Springer Science & Business Media (2011)

[15] Choi, Y., Won, J.-H.: Ornstein auto-encoders. In IJCAI, pp. 2172–2178 (2019)

[16] Choi, Y., Lee, S., Won, J.-H.: Learning from nested data with ornstein auto-encoders. In: International Conference on Machine Learning, pp. 1943–1952. PMLR (2021)

[17] De Palma, G., Marvian, M., Trevisan, D., Lloyd, S.: The quantum wasserstein distance of order 1. IEEE Trans. Inf. Theory **67**(10), 6627–6643 (2021)

[18] Hamazaki, R.: Speed limits for macroscopic transitions. PRX Quantum **3**, 020319 (2022)

[19] De Palma, G., Rouzé, C.: Quantum concentration inequalities. Ann. Henri Poincaré **23**, 3391–3429 (2022)

[20] De Palma, G., Marvian, M., Rouzé, C., França, D.S.: Limitations of variational quantum algorithms: a quantum optimal transport approach. PRX Quantum **4**, 010309 (2023)

[21] Chou, C.-N., Love, P.J., Sandhu, J.S., Shi, J.: Limitations of local quantum algorithms on random MAX-k-XOR and Beyond. In: Bojańczyk, M., Merelli, E., Woodruff, D.P. (eds), 49th International Colloquium on Automata, Languages, and Programming (ICALP 2022), vol. 229 of Leibniz International Proceedings in Informatics (LIPIcs), pp. 41:1–41:20, Dagstuhl, Germany (2022). Schloss Dagstuhl – Leibniz-Zentrum für Informatik

[22] Li, L., Bu, K., Koh, D.E., Jaffe, A., Lloyd, S.: Wasserstein Complexity of Quantum Circuits (2022). arXiv:2208.06306

[23] Rouzé, C., França, D.S.: Learning quantum many-body systems from a few copies (2021). arXiv:2107.03333

[24] Maciejewski, F.B., Puchała, Z., Oszmaniec, M.: Exploring quantum average-case distances: proofs, properties, and examples. IEEE Trans. Inf. Theory 1–1 (2023)

[25] Kiani, B.T., De Palma, G., Marvian, M., Liu, Z.-W., Lloyd, S.: Learning quantum data with the quantum earth mover's distance. Quantum Sci. Technol. **7**(4), 045002 (2022)

[26] Herr, D., Obert, B., Rosenkranz, M.: Anomaly detection with variational quantum generative adversarial networks. Quantum Sci. Technol. **6**(4), 045004 (2021)

[27] Anschuetz, E.R., Kiani, B.T.: Quantum variational algorithms are swamped with traps. Nat. Commun. **13**(1), 7760 (2022)

[28] Coyle, B.: Machine learning applications for noisy intermediate-scale quantum computers. Ph.D. thesis, University of Edinburgh (2022)

[29] Hirche, C., Rouzé, C., França, D.S.: Quantum differential privacy: an information theory perspective (2022). arXiv:2202.10717

[30] Angrisani, A., Doosti, M., Kashefi, E.: Differential privacy amplification in quantum and quantum-inspired algorithms (2022). arXiv:2203.03604

[31] Duvenhage, R., Mapaya, M.: Quantum Wasserstein distance of order 1 between channels (2022). arXiv:2210.03483

[32] Bratteli, O., Robinson, D.W.: Operator Algebras and Quantum Statistical Mechanics I: C*- and W*-Algebras. Symmetry Groups. Decomposition of States. Theoretical and Mathematical Physics. Springer, Berlin (2013)

[33] Bratteli, O., Robinson, D.W.: Operator Algebras and Quantum Statistical Mechanics II: Equilibrium States Models in Quantum Statistical Mechanics. Theoretical and Mathematical Physics. Springer, Berlin (2013)

[34] Naaijkens, P.: Quantum Spin Systems on Infinite Lattices: A Concise Introduction. Lecture Notes in Physics. Springer International Publishing (2017)

[35] Alicki, R., Alicki, I.T.P.A.R., Fannes, M.: Quantum Dynamical Systems. Oxford University Press (2001)

[36] Barnum, H.: Quantum rate-distortion coding. Phys. Rev. A **62**(4), 042309 (2000)

[37] Devetak, I., Berger, T.: Quantum rate-distortion theory for IID sources. In: Proceedings of 2001 IEEE International Symposium on Information Theory (IEEE Cat. No. 01CH37252), pp. 276. IEEE (2001)

[38] Devetak, I., Berger, T.: Quantum rate-distortion theory for memoryless sources. IEEE Trans. Inf. Theory **48**(6), 1580–1589 (2002)

[39] Chen, X.-Y., Wang, W.-M.: Entanglement information rate distortion of a quantum Gaussian source. IEEE Trans. Inf. Theory **54**(2), 743–748 (2008)

[40] Datta, N., Hsieh, M.-H., Wilde, M.M.: Quantum rate distortion, reverse Shannon theorems, and source-channel separation. IEEE Trans. Inf. Theory **59**(1), 615–630 (2013)

[41] Datta, N., Hsieh, M.-H., Wilde, M.M., Winter, A.: Quantum-to-classical rate distortion coding. J. Math. Phys. **54**(4), 042201 (2013)

[42] Wilde, M.M., Datta, N., Hsieh, M.-H., Winter, A.: Quantum rate-distortion coding with auxiliary resources. IEEE Trans. Inf. Theory **59**(10), 6755–6773 (2013)

[43] Salek, S., Cadamuro, D., Kammerlander, P., Wiesner, K.: Quantum rate-distortion coding of relevant information. IEEE Trans. Inf. Theory **65**(4), 2603–2613 (2018)

[44] Carlen, E.A., Maas, J.: An analog of the 2-Wasserstein metric in non-commutative probability under which the Fermionic Fokker-Planck equation is gradient flow for the entropy. Commun. Math. Phys. **331**(3), 887–926 (2014)

[45] Carlen, E.A., Maas, J.: Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance. J. Funct. Anal. **273**(5), 1810–1869 (2017)

[46] Carlen, E.A., Maas, J.: Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems. J. Stat. Phys. **178**(2), 319–378 (2020)

[47] Rouzé, C., Datta, N.: Concentration of quantum states from quantum functional and transportation cost inequalities. J. Math. Phys. **60**(1), 012202 (2019)

[48] Datta, N., Rouzé, C.: Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality. Ann. Henri Poincaré **21**, 2115–2150 (2020)

[49] Van Tan, V., Hasegawa, Y.: Geometrical bounds of the irreversibility in Markovian systems. Phys. Rev. Lett. **126**, 010601 (2021)

[50] Wirth, M.: A dual formula for the noncommutative transport distance. J. Stat. Phys. **187**(2), 1–18 (2022)

[51] Gao, L., Junge, M., LaRacuente, N.: Fisher information and logarithmic Sobolev inequality for matrix-valued functions. Ann. Henri Poincaré **21**(11), 3409–3478 (2020)

[52] Chen, Y., Georgiou, T.T., Ning, L., Tannenbaum, A.: Matricial Wasserstein-1 distance. IEEE Control Syst. Lett. **1**(1), 14–19 (2017)

[53] Ryu, E.K., Chen, Y., Li, W., Osher, S.: Vector and matrix optimal mass transport: theory, algorithm, and applications. SIAM J. Sci. Comput. **40**(5), A3675–A3698 (2018)

[54] Chen, Y., Georgiou, T.T., Tannenbaum, A.: Matrix optimal mass transport: a quantum mechanical approach. IEEE Trans. Autom. Control **63**(8), 2612–2619 (2018)

[55] Chen, Y., Georgiou, T.T., Tannenbaum, A.: Wasserstein geometry of quantum states and optimal transport of matrix-valued measures. In: Emerging Applications of Control and Systems Theory, pp. 139–150. Springer (2018)

[56] Agredo, J.: A Wasserstein-type distance to measure deviation from equilibrium of quantum Markov semigroups. Open Syst. Inf. Dyn. **20**(02), 1350009 (2013)

[57] Agredo, J.: On exponential convergence of generic quantum Markov semigroups in a Wasserstein-type distance. Int. J. Pure Appl. Math. **107**(4), 909–925 (2016)

[58] Ikeda, K.: Foundation of quantum optimal transport and applications. Quantum Inf. Process. **19**(1), 25 (2020)

[59] Golse, F., Mouhot, C., Paul, T.: On the mean field and classical limits of quantum mechanics. Commun. Math. Phys. **343**(1), 165–205 (2016)

[60] Caglioti, E., Golse, F., Paul, T.: Towards optimal transport for quantum densities. Annali Scuola Normale Superiore-Classe di Scienze, pp. 49–4, (2022)

[61] Golse, F.: The quantum N-body problem in the mean-field and semiclassical regime. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **376**(2118), 20170229 (2018)

[62] Golse, F., Paul, T.: The Schrödinger equation in the mean-field and semiclassical regime. Arch. Ration. Mech. Anal. **223**(1), 57–94 (2017)

[63] Golse, F., Paul, T.: Wave packets and the quadratic Monge–Kantorovich distance in quantum mechanics. C.R. Math. **356**(2), 177–197 (2018)

[64] Caglioti, E., Golse, F., Paul, T.: Quantum optimal transport is cheaper. J. Stat. Phys. **181**(1), 149–162 (2020)

[65] Friedland, S., Eckstein, M., Cole, S., Życzkowski, K.: Quantum Monge–Kantorovich problem and transport distance between density matrices. Phys. Rev. Lett. **129**, 110402 (2022)

[66] Cole, S., Eckstein, M., Friedland, S., Życzkowski, K.: Quantum Optimal Transport (2021). arXiv:2105.06922

[67] Duvenhage, R.: Optimal quantum channels. Phys. Rev. A **104**, 032604 (2021)

[68] Bistroń, R., Eckstein, M., Życzkowski, K.: Monotonicity of a quantum 2-Wasserstein distance. J. Phys. A Math. Theor. **56**(9), 095301 (2023)

[69] Van Tan, V., Saito, K.: Thermodynamic unification of optimal transport: thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits. Phys. Rev. X **13**, 011013 (2023)

[70] Duvenhage, R.: Quadratic Wasserstein metrics for von Neumann algebras via transport plans. J. Oper. Theory **88**(2), 289–308 (2022)

[71] Duvenhage, R.: Wasserstein distance between noncommutative dynamical systems (2021). arXiv:2112.12532

[72] Duvenhage, R., Skosana, S., Snyman, M.: Extending quantum detailed balance through optimal transport (2022). arXiv preprint arXiv:2206.15287

[73] De Palma, G., Trevisan, D.: Quantum optimal transport with quantum channels. Ann. Henri Poincaré **22**(10), 3199–3234 (2021)

[74] Duvenhage, R., Snyman, M.: Balance between quantum Markov semigroups. Ann. Henri Poincaré **19**(6), 1747–1786 (2018)

[75] Agredo, J., Fagnola, F.: On quantum versions of the classical Wasserstein distance. Stochastics **89**(6–7), 910–922 (2017)

[76] Życzkowski, K., Slomczynski, W.: The Monge distance between quantum states. J. Phys. A Math. Gen. **31**(45), 9095 (1998)

[77] Życzkowski, K., Slomczynski, W.: The Monge metric on the sphere and geometry of quantum states. J. Phys. A Math. Gen. **34**(34), 6689 (2001)

[78] Bengtsson, I., Życzkowski, K.: Geometry of Quantum States: An Introduction to Quantum Entanglement. Cambridge University Press (2017)

[79] Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press (2010)

[80] Wilde, M.M.: Quantum Information Theory. Cambridge University Press, 2 edition (2017)

[81] Holevo, A.S.: Quantum Systems, Channels, Information: A Mathematical Introduction. Texts and Monographs in Theoretical Physics. De Gruyter (2019)

[82] Jakšić, V., Pillet, C.-A., Tauber, C.: Approach to equilibrium in translation-invariant quantum systems: some structural results. Annales Henri Poincaré (2023). https://doi.org/10.1007/s00023-023-01281-6

[83] Griffiths, R.B., Ruelle, D.: Strict convexity ("continuity") of the pressure in lattice systems. Commun. Math. Phys. **23**(3), 169–175 (1971)

[84] Roos, H.: Strict convexity of the pressure: a note on a paper of R. B. Griffiths and D. Ruelle. Commun. Math. Phys. **36**(4), 263–276 (1974)

[85] Israel, R.B.: Convexity in the Theory of Lattice Gases. Princeton Series in Physics. Princeton University Press (2015)

[86] Ruelle, D.: Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics. Cambridge Mathematical Library. Cambridge University Press, Cambridge (2004)

[87] Jakšić, V., Pillet, C.-A., Tauber, C.: A note on adiabatic time evolution and quasi-static processes in translation-invariant quantum systems. Ann. Henri Poincaré (2023). https://doi.org/10.1007/s00023-023-01282-5

[88] Van Enter, A.C.D., Fernández, R., Sokal, A.D.: Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory. J. Stat. Phys. **72**(5), 879–1167 (1993)

[89] Lieb, E.H.: Convex trace functions and the Wigner–Yanase–Dyson conjecture. Adv. Math. **11**(3), 267–288 (1973)

[90] Polyanskiy, Y., Yihong, W.: Wasserstein continuity of entropy and outer bounds for interference channels. IEEE Trans. Inf. Theory **62**(7), 3992–4002 (2016)

[91] Osborne, T., Winter, A.: A quantum generalisation of Talagrand's inequality. Tobias J. Osborne's Research Notes. https://tjoresearchnotes.wordpress.com/2009/02/13/a-quantum-generalisation-of-talagrands-inequality (2009)

[92] Eldar, L., Harrow, A.W.: Local Hamiltonians whose ground states are hard to approximate. In: 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS), pp. 427–438 (2017)

[93] Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256**(3), 810–864 (2009)

[94] Gao, L., Rouzé, C.: Ricci curvature of quantum channels on non-commutative transportation metric spaces (2021). arXiv:2108.10609

[95] Bardet, I., Capel, Á., Gao, L., Lucia, A., Pérez-García, D., Rouzé, C.: Entropy decay for Davies semigroups of a one dimensional quantum lattice (2021). arXiv:2112.00601

[96] Bardet, I., Capel, Á., Gao, L., Lucia, A., Pérez-García, D., Rouzé, C.: Rapid thermalization of spin chain commuting Hamiltonians. Phys. Rev. Lett. **130**, 060401 (2023)

[97] Capel, Á., Rouzé, C., França, D.S.: The modified logarithmic Sobolev inequality for quantum spin systems: classical and commuting nearest neighbour interactions (2020). arXiv:2009.11817

[98] Fröhlich, J., Ueltschi, D.: Some properties of correlations of quantum lattice systems in thermal equilibrium. J. Math. Phys. **56**(5), 053302 (2015)

[99] Alicki, R., Fannes, M.: Quantum Dynamical Systems. Oxford University Press, Oxford (2001)

[100] Capobianco, S.: Multidimensional cellular automata and generalization of Fekete's lemma. Discrete Math. Theor. Comput. Sci. **10**(3) (2008)

Giacomo De Palma
Department of Mathematics
University of Bologna
Bologna
Italy
e-mail: giacomo.depalma@unibo.it

Dario Trevisan
Department of Mathematics
University of Pisa
Pisa
Italy