Optimal light cone for macroscopic particle transport in long-range systems: A quantum speed limit approach

Tan Van Vu[1,2], Tomotaka Kuwahara[1,3], and Keiji Saito[4]

[1]Analytical Quantum Complexity RIKEN Hakubi Research Team, RIKEN Center for Quantum Computing (RQC), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
[2]Yukawa Institute for Theoretical Physics, Kyoto University, Kitashirakawa Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan
[3]PRESTO, Japan Science and Technology (JST), Kawaguchi, Saitama 332-0012, Japan
[4]Department of Physics, Kyoto University, Kyoto 606-8502, Japan

arXiv:2307.01059v2 [quant-ph] 18 Sep 2024

Understanding the ultimate rate at which information propagates is a pivotal issue in nonequilibrium physics. Nevertheless, the task of elucidating the propagation speed inherent in quantum bosonic systems presents challenges due to the unbounded nature of their interactions. In this study, we tackle the problem of macroscopic particle transport in a long-range generalization of the lattice Bose-Hubbard model through the lens of the quantum speed limit. By developing a unified approach based on optimal transport theory, we rigorously prove that the minimum time required for macroscopic particle transport is always bounded by the distance between the source and target regions, while retaining its significance even in the thermodynamic limit. Furthermore, we derive an upper bound for the probability of observing a specific number of bosons inside the target region, thereby providing additional insights into the dynamics of particle transport. Our results hold true for arbitrary initial states under both long-range hopping and long-range interactions, thus resolving an open problem of particle transport in generic bosonic systems.

1 Introduction

The investigation of the velocity at which information propagates is a central topic in quantum mechanics. On this subject, two primary approaches—the Lieb-Robinson bound [1] and the quantum speed limit [2]—have attracted great attention. The Lieb-Robinson bound shows the existence of an effective light cone, outside of which information propagation undergoes an exponential decay with distance. Hence, it clarifies the *unreachable* region of information propagation for a given time. This bound is not only fundamental for establishing causality in physics but also plays a pivotal role in analyzing universal properties in many-body systems [3–16], such as the clustering theorem in ground states [5, 6], the area law of entanglement [8, 9], and computational complexity [12, 13], to name only a few. On the other hand, the quantum speed limit addresses the minimum time required for the propagation of information to a *reachable* regime. In the celebrated Mandelstam-Tamm speed limit, the minimum time required for the initial density matrix to change to another state is bounded from below by the ratio of the distance between the two states to the energy fluctuation. To date, the quantum speed limit has been generalized and applied to various fields [17–35]. Essentially, these two notions of speed limits, despite having been developed independently in the literature, offer complementary perspectives on the fundamental question of how much time is required to propagate information. It is thus expected that the proper utilization of these two notions in a unified manner will yield physically novel results and have significant methodological implications, especially for open problems that were previously intractable through the isolated use of either method.

Let us consider the primary problems that Lieb and Robinson have addressed. In the early days of studies, the Lieb-Robinson bound was explored in the context of short-range interacting spin systems. It was shown that information propagation is effectively confined within the linear light cone, which grows at a finite speed. In scenarios involving long-range interactions, however, the existence of linear light cones becomes significantly intricate, as such interactions have the capability to instantaneously transmit information to distant places [36]. Here, long-range interaction means that the interaction strength between distinct sites obeys a power-law decay as $1/d^{\alpha}$ with distance d. Intuitively, the shape of the light cone is dependent on the exponent α and may no longer be linear. Given the ubiquity of long-range systems in nature [37, 38] and their fundamental interest, it is of great importance to elucidate the shape of the light cone with respect to the power decay α. Thus far,

Tan Van Vu: tan.vu@yukawa.kyoto-u.ac.jp
Tomotaka Kuwahara: tomotaka.kuwahara@riken.jp
Keiji Saito: keiji.saitoh@scphys.kyoto-u.ac.jp

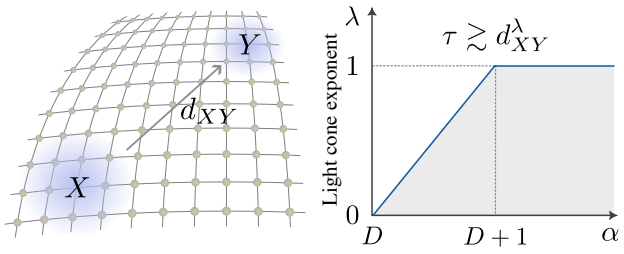

Figure 1: (a) Schematic of long-range bosonic systems, wherein bosons are transported from region X to a separate region Y with a distance of d_{XY}. (b) Diagram of our first main result on the minimum operational time τ for particle transport. The gray area denotes the prohibited region of the exponent.

a comprehensive characterization of the shape of the light cone has been established for long-range interacting quantum spin and fermionic systems [39–48]. In Ref. [42], the condition for the linear light cone was proven to be $\alpha > 2D + 1$, where D denotes the spatial dimension. Furthermore, for the entire regime of $\alpha > 2D$, the optimal form of the light cone has been identified as $\tau \gtrsim d^{\min(1,\alpha-2D)}$ [46].

As another important class, quantum bosonic systems (e.g., the paradigmatic Bose-Hubbard model [49]) hold significant importance in the study of cold atoms in experimental setups. Nevertheless, the theoretical investigation for bosonic systems has persisted due to the unbounded nature of interactions in such systems. Classifying the effective light cone within the Bose-Hubbard model has been substantiated through both experimental observations and numerical simulations using ultracold gases confined in optical lattices [50–54]. While the speeds of bosonic transport and information propagation have been studied for bosonic systems with short-range hoppings [55–59], the exploration of long-range cases remains in an early stage of development [60–62]. As the first step toward the Lieb-Robinson bound for long-range cases, the problem of macroscopic bosonic transport has been studied. Here, macroscopic transport refers to the transfer of a macroscopic number of bosons (i.e., comparable to the total boson number) between separated regions. Concerning this, Faupin and coworkers [60] have recently demonstrated the finite velocity of macroscopic particle transport by showing the existence of the linear light cone for $\alpha > D+2$[1]. However, we are still far from the complete classification of the effective light cone for macroscopic particle transport. In particular, it is a critical problem to identify the optimal form of the light cone for macroscopic particle transport in bosonic systems, as has been studied for quantum spin and fermionic systems.

In this study, we resolve the above problem for macroscopic particle transport with a *quantum speed limit approach*. We consider the problem of particle transport in a generalization of the Bose-Hubbard model with *long-range hopping* and *long-range interactions*, spanning the *entire regime* of $\alpha > D$. We develop a unified speed limit technique based on optimal transport theory, and we demonstrate the finite speed of macroscopic bosonic transport. Specifically, we rigorously prove that the minimum time required for the transport of a macroscopic number of bosons is always bounded by the distance d as $\tau \gtrsim d^{\min(1,\alpha-D)}$ [cf. Eq. (5); see Fig. 1 for illustration]. The exponent in this bound is optimal, as it can be achieved through a constructive state-transfer protocol. Additionally, we derive a limit for the probability of transporting a specific number of bosons into a target region [cf. Eq. (8)]. We emphasize that all our results unconditionally hold for arbitrary initial states including pure and mixed states. Notably, our proofs are significantly simple, necessitating only a few fundamental properties of the Wasserstein distance. These findings highlight the efficacy of our unified approach in overcoming the primary open problems related to speed in quantum many-body physics.

2 Results

We consider a generic model of bosons on an arbitrary D-dimensional lattice, wherein bosons can hop between arbitrary sites and interact with each other. The system Hamiltonian is time dependent (i.e., an external control protocol can be applied to the system) and can be expressed in the following form:

$$H_t := \sum_{i\neq j\in\Lambda} J_{ij}(t)\hat{b}_i^\dagger \hat{b}_j + \sum_{Z\subseteq\Lambda} h_Z(t). \tag{1}$$

Here, Λ denotes the set of all the sites in the lattice, $\hat{b}_i^\dagger$ and $\hat{b}_i$ are the bosonic creation and annihilation operators for site i, respectively, and $h_Z(t)$ is an arbitrary function of $\{\hat{n}_i\}_{i\in Z}$, where $\hat{n}_i := \hat{b}_i^\dagger \hat{b}_i$ is the number operator. Note that $h_Z(t)$ does not need to be local (i.e., Z can be arbitrarily large). In other words, both *long-range hopping* and *long-range interactions* are allowed in our setup. The hopping terms $J_{ij}(t)$ are symmetric [i.e., $J_{ij}(t) = J_{ji}(t)$] and upper bounded by a power law in the Euclidean distance, i.e., $|J_{ij}(t)| \leq J/\|i-j\|^\alpha$ for some positive constant J, where $\|\cdot\|$ denotes the Euclidean norm and the power decay α satisfies $\alpha > D$[2]. Examples include the standard Bose-Hubbard model, specified by $J_{ij}(t) = 0$ for any non-neighboring sites i and j and

[1] Although the range of α considered in Ref. [60] is stated as $\alpha \geq D+3$, it can be improved to $\alpha > D+2$ by setting $p = 2$ instead of $p = \alpha - D - 1$.

[2] Note that in the $\alpha \leq D$ regime, particle transport speed can be effectively infinite. To elaborate further, there exists a transport protocol that transfers a single particle over a distance d within time $\tau = O(\log d)$ when $\alpha = D$, and $\tau = O(1)$ when $\alpha < D$ [63]. Given our focus on scenarios involving finite bosonic transport speeds, the region where transport speed becomes infinite is excluded from our study.

$\sum_{Z\subseteq\Lambda} h_Z(t) = (U/2)\sum_i \hat{n}_i(\hat{n}_i - 1) - \mu\sum_i \hat{n}_i$, where U and μ are real constants. Let ϱ_t be the system's density matrix at time t. Then, its time evolution is described by the von Neumann equation ($\hbar = 1$ for simplicity),

$$\dot{\varrho}_t = -i[H_t, \varrho_t]. \tag{2}$$

It can be easily shown that the total number of particles in this system is conserved, denoted hereafter by $\mathcal{N}$.

We introduce some notations that will be used in this study. For an arbitrary set X of lattice sites, $X_c \coloneqq \Lambda \setminus X$ is the set of sites that are not present in X, and $|X|$ denotes its cardinality. The distance between any two disjoint sets X and Y is defined as the minimum distance between sites belonging to each of them,

$$d_{XY} \coloneqq \min_{i\in X,\, j\in Y} \|i - j\|. \tag{3}$$

The number operator of local bosons occupied inside region X is given by $\hat{n}_X \coloneqq \sum_{i\in X} \hat{n}_i$, and the corresponding concentration operator is $\overline{n}_X \coloneqq \hat{n}_X/\mathcal{N}$. Let $\Pi_{\vec{N}} = |\vec{N}\rangle\langle\vec{N}|$ be the projection onto the Mott state $|\vec{N}\rangle$ (i.e., $\hat{n}_i|\vec{N}\rangle = n_i|\vec{N}\rangle$ for $\vec{N} = [n_i]_{i\in\Lambda}$). Then, the projection $P_{\hat{n}_X\le N_0}$ is defined as

$$P_{\hat{n}_X\le N_0} \coloneqq \sum_{\vec{N}:\langle\vec{N}|\hat{n}_X|\vec{N}\rangle\le N_0} \Pi_{\vec{N}}. \tag{4}$$

The expectation of an observable A with respect to state ϱ is denoted as $\langle A\rangle_\varrho \coloneqq \mathrm{tr}\{A\varrho\}$.

To fully resolve the open problem of bosonic transport, a novel mathematical framework needs to be developed. Unlike the previous approach based on adiabatic spacetime localization observables in Ref. [60], which could only achieve a partial resolution, we here tackle the problem using a quantum speed limit approach. It is important to note that the quantum speed limit primarily focuses on determining the minimum time required for state changes, but it may not provide accurate predictions for many-body systems [64]. On the other hand, the Lieb-Robinson bound offers valuable insights into many-body systems but does not furnish information regarding the speed of transitions inside the light cone. Consequently, it seems reasonable to anticipate that unifying these two concepts could yield a novel bound that takes advantage of their respective strengths and establishes new constraints that neither concept can fully encompass on its own. To achieve this, we develop a unified speed limit based on optimal transport theory [cf. Eq. (20)]. This approach enables us to integrate geometric structures of the underlying dynamics into the speed limit, allowing for meaningful limits to be derived, even when dealing with large-scale systems.

Given the aforementioned setup, we are now ready to explain the results, whose proofs are presented in Appendix A. Our first main result is the following statement, which describes a fundamental limit on the operational time required for a macroscopic transport of bosons in long-range systems.

Theorem 1. *Consider a situation where a fraction $\mu \in (0, 1]$ of all bosons is transported from region X to a distant region Y. Then, the operational time τ required for this macroscopic bosonic transport is lower bounded by the distance between the two regions as*

$$\tau \ge \kappa_1^\epsilon d_{XY}^{\min(1,\alpha-D-\epsilon)}. \tag{5}$$

Here, $0 < \epsilon < \alpha - D$ is an arbitrary number, $\kappa_1^\epsilon \coloneqq [J\gamma\zeta(\alpha-\alpha_\epsilon-D+1)]^{-1}\mu$ is a constant independent of the system size, $\alpha_\epsilon \coloneqq \min(1, \alpha - D - \epsilon)$, γ is a $O(1)$ constant [cf. Eq. (24)], and ζ denotes the Riemann zeta function.

We briefly explain the role of the auxiliary variable ϵ. In simple terms, it is introduced to address the case of $D < \alpha \le D + 1$. For $\alpha > D + 1$, we can choose $\epsilon = \alpha - D - 1$; thus, $\kappa_1^\epsilon = [J\gamma\zeta(\alpha - D)]^{-1}\mu$ and $\min(1, \alpha-D-\epsilon) = 1$. In the case $D < \alpha \le D+1$, the bound (5) explicitly reads

$$\tau \ge \frac{\mu}{J\gamma\zeta(1+\epsilon)} d_{XY}^{\alpha-D-\epsilon}. \tag{6}$$

As seen, ϵ can be arbitrarily small but should not be zero as the function ζ can diverge.

Some remarks on Theorem 1 are in order. (i) First, macroscopic transport here means that an average number $\mu\mathcal{N}$ of bosons is transported from X to Y, where μ is a $O(1)$ constant. Since bosons are indistinguishable, a fraction μ of the total bosons must be effectively transported from region X to region Y if the following inequality is satisfied:

$$x_Y(\tau) \ge x_{X_c}(0) + \mu, \tag{7}$$

where $x_Z(t)$ denotes the boson concentration inside region Z at time t. Inequality (7) can be interpreted as follows: Even if all bosons outside of X are transported to Y, it is still insufficient to achieve the desired concentration; consequently, at least a fraction μ of the bosons must be effectively transported from X to Y. The result holds for arbitrary initial states, including both pure and mixed states, and for the entire range of the power decay $\alpha > D$. Notably, this type of time constraint cannot be deduced by the conventional quantum speed limits [2, 18].

(ii) Second, bound (5) can be physically interpreted as follows. For the case $\alpha > D + 1$, the bound reads $\tau \ge O(d_{XY})$, implying that bosonic transport always takes time at least proportional to the distance. Surprisingly, this implication is the same as in the case of short-range hopping [33], indicating that long-range hopping and long-range interactions do not enhance the speed of macroscopic bosonic transport[3]. On the

[3]Note that it may not be the case when considering the speed of information propagation.

other hand, in the case $D < \alpha \leq D+1$, the minimum operational time is bounded in terms of the power decay α as $\tau \geq O(d_{XY}^{\alpha-D})$, indicating the effect of long-range hopping on speeding up bosonic transport.

(iii) Last, we discuss the optimality of the bound (5) and show that it is optimal. The optimality here is interpreted in the sense that the power coefficient associated with the distance term in the bound is optimal. In other words, we need only show that there exist transport protocols such that bosonic transport can be accomplished within time $\tau = O(d_{XY}^{\min(1,\alpha-D)})$. To this end, we first consider the case of a single particle (i.e., $\mathcal{N} = 1$). In this case, it was shown that one could construct a rapid state-transfer protocol such that the particle can be transferred between two sites of distance d within time $\tau = O(d)$ for $\alpha > D+1$ and $\tau = O(d^{\alpha-D})$ for $D < \alpha \leq D+1$ [63] (see Theorem 11 therein). Therefore, it can be concluded that the bound (5) is optimal. We note that while the form of the optimal light cone is the same as that obtained in Ref. [63], the previous study only considers noninteracting particles, which can be simplified to the problem of single free particle. In contrast, our result applies to interacting systems, where boson-boson interactions can be arbitrary and play a crucial role in facilitating bosonic transport (cf. Lemmas 5 and 7). Next, we show a protocol that saturates the bound for $\alpha > D+1$ in the general case of an extensive boson number (i.e., $\mathcal{N} \gg 1$). We consider a one-dimensional lattice of length L for simplicity and assume that the initial state is a Mott state $|\mathcal{N}, 0, \ldots, 0\rangle$. Here, $|m\rangle \otimes |n\rangle$ is denoted by $|m, n\rangle$ for convenience. Initially, all bosons are concentrated at site 1, and we construct a protocol that transports all bosons to site L within time τ. We can sequentially transport bosons within $L-1$ steps, where all bosons are transferred from site k to site $k+1$ at each step $k = 1, \ldots, L-1$. The transformation in each step can be achieved as $|\mathcal{N}, 0\rangle \to |\mathcal{N}-1, 1\rangle \to |1, \mathcal{N}-1\rangle \to |0, \mathcal{N}\rangle$. The time consumed in each step is equal to $\pi/(2J) + \pi/(J\sqrt{\mathcal{N}}) = O(1)$ (cf. Lemmas 6 and 7). Thus, the total time for transporting $\mathcal{N}$ bosons over a distance of $L-1$ is $O(L)$.

Theorem 1 concerns the minimum time required to transport an average number of bosons between two regions. In the sequel, we focus on the probability of finding a number of bosons occupied inside the target region within a finite time. Our second main result, which is stated below, establishes an upper bound on this probability in terms of the operational time and the distance.

Theorem 2. *Assume that the initial boson number outside of region X does not exceed N_0, i.e., $\langle P_{\hat{n}_{X_c} \leq N_0} \rangle_{\varrho_0} = 1$. Then, the probability of finding bosons inside a disjoint region Y is upper bounded by the operational time τ and the distance between the two regions as*

$$\langle P_{\hat{n}_Y \geq N_0 + \Delta N_0} \rangle_{\varrho_\tau} \leq \frac{\kappa_2^\epsilon \mathcal{N}}{\Delta N_0} \frac{\tau}{d_{XY}^{\min(1,\alpha-D-\epsilon)}}. \tag{8}$$

Here, $\mathcal{N}$ is the total number of bosons, $0 < \epsilon < \alpha - D$ is an arbitrary number, and $\kappa_2^\epsilon := J\gamma\zeta(\alpha - \alpha_\epsilon - D + 1)$ is a constant independent of the system size.

This result holds under a generic setting for both initial states and system dynamics, same as in Theorem 1. It can be interpreted as follows. Initially, there are at most N_0 bosons outside of X; therefore, in order to observe more than N_0 bosons inside Y, a certain number of bosons must be transferred from X to Y. Bound (8) indicates that the probability of finding at least $N_0 + \Delta N_0$ inside Y is at most linear in time τ and inversely proportional to boson number ΔN_0 and distance d_{XY}. While the inequality (8) is valid for any boson number ΔN_0, it is particularly relevant and meaningful to discuss macroscopic transport [i.e., $\Delta N_0/\mathcal{N} =: \mu$ is a $O(1)$ constant]. In this case, we obtain

$$\langle P_{\overline{n}_Y \geq \xi} \rangle_{\varrho_\tau} \leq \kappa_2^\epsilon \mu^{-1} \tau d_{XY}^{-\min(1,\alpha-D-\epsilon)}, \tag{9}$$

where $\xi := (N_0 + \Delta N_0)/\mathcal{N} \geq x_{X_c}(0) + \mu$. In the case $\alpha > D+2$, a similar probability constraint has been derived for pure states in the absence of long-range interactions [60] [see Eq. (7) therein], remaining an open question for the case of $D < \alpha \leq D+2$. Our result thus completely resolves this open problem of bosonic transport for generic long-range systems.

It is worth noting that, at the qualitative level, Theorem 1 can be viewed as a corollary of Theorem 2. For arbitrary $\mu' \in (0, \mu)$, by applying Markov's inequality, we obtain

$$\begin{aligned} \langle P_{\overline{n}_Y \leq x_{X_c}(0) + \mu'} \rangle_{\varrho_\tau} &= \langle P_{\mathbb{1} - \overline{n}_Y \geq 1 - x_{X_c}(0) - \mu'} \rangle_{\varrho_\tau} \\ &\leq \frac{\langle \mathbb{1} - \overline{n}_Y \rangle_{\varrho_\tau}}{1 - x_{X_c}(0) - \mu'} \\ &= \frac{1 - x_Y(\tau)}{1 - x_{X_c}(0) - \mu'}. \end{aligned} \tag{10}$$

Thus, the probability of finding bosons inside region Y can be lower bounded as

$$\langle P_{\overline{n}_Y \geq x_{X_c}(0) + \mu'} \rangle_{\varrho_\tau} \geq \frac{x_Y(\tau) - x_{X_c}(0) - \mu'}{1 - x_{X_c}(0) - \mu'}. \tag{11}$$

When $x_Y(\tau) \geq x_{X_c}(0) + \mu$, using Eq. (11), Theorem 2 can deduce a statement that is qualitatively equivalent to Theorem 1 as

$$\tau \geq \frac{\mu'(\mu - \mu')}{\mu[1 - x_{X_c}(0) - \mu']} \kappa_1^\epsilon d_{XY}^{\min(1,\alpha-D-\epsilon)}, \tag{12}$$

which can be optimized with $\mu' = 1 - x_{X_c}(0) - \sqrt{[1 - x_{X_c}(0)][1 - x_{X_c}(0) - \mu]}$. While the bound (12) is looser than the bound (5) in Theorem 1, it captures the same fundamental limit regarding the speed of bosonic transport.

It is worthwhile to investigate the extent to which our results can be extended. We demonstrate that our main results cannot hold in more general classes of quantum many-boson systems. This implies that, in these cases, particle transport at *supersonic* speeds is possible. Eisert and Gross [65] have previously established a well-known result that showcases an *exponential* acceleration in quantum communication. However, the model used in their study was highly artificial, and there may be more naturally occurring categories of many-boson systems that could potentially exhibit finite-speed quantum transport. Our objective here is to illustrate that even by making slight modifications to the Bose-Hubbard model, we can observe an *infinite* acceleration in quantum propagation.

In the sequel, we show that the speed of macroscopic bosonic transport can be supersonic in systems with interaction-induced tunneling terms [66]. These systems have undergone intensive theoretical investigation [67] and have been observed in experiments [68].

Theorem 3. *The speed of particle transport can be infinite in a one-dimensional bosonic system of size L with the time-dependent Hamiltonian given by*

$$H_t = \sum_{\langle i,j\rangle} J_{ij}(t)\hat{b}_i^\dagger \hat{b}_j + \sum_{\langle i,j\rangle} T_{ij}(t)\hat{b}_i^\dagger(\hat{n}_i+\hat{n}_j)\hat{b}_j + \frac{1}{2}\sum_i U_i(t)\hat{n}_i(\hat{n}_i-1) + \sum_i \mu_i(t)\hat{n}_i. \tag{13}$$

Here, $\langle i,j\rangle$ denotes an ordered pair of neighboring sites (i.e., $|i-j|=1$).

We provide a proof outline of Theorem 3. We prove by systematically constructing a protocol that transfers a macroscopic number of bosons to a distant region within a constant of time. Let the initial state be a Mott state of the form $|\psi_0\rangle = |L,0,\ldots,0\rangle$; that is, there are L bosons at site 1 and no boson at any other site $s \in [2,L]$. We consider a control protocol that transfers the initial state to the final state $|0,\ldots,0,L\rangle$ (i.e., all bosons are transported to site L within time τ). The protocol is constructed as follows.

At each step $k = 1,\ldots,L-1$, all the hopping and interaction terms are set to zero, except for sites k and $k+1$. In other words, the bosonic system effectively becomes a two-site system consisting of sites k and $k+1$. By performing a subprotocol described below, the state will be transformed as $|L,0\rangle \to |0,L\rangle$. Consequently, after the final step, the system state $|\psi_\tau\rangle$ will become $|0,\ldots,0,L\rangle$ as desired.

The subprotocol that performs the state transformation $|L,0\rangle \to |0,L\rangle$ is composed of three stages, where the first performs $|L,0\rangle \to |L-1,1\rangle$, the second transforms $|L-1,1\rangle \to |1,L-1\rangle$, and the last performs $|1,L-1\rangle \to |0,L\rangle$[4]. The Hamiltonian in the first stage is given in the following form:

$$H_1 = J(\hat{b}_k^\dagger\hat{b}_{k+1} + \hat{b}_{k+1}^\dagger\hat{b}_k) - U\hat{n}_k^2 + U(2L-1)\hat{n}_k + J\Big[\hat{b}_{k+1}^\dagger(\hat{n}_{k+1}+\hat{n}_k)\hat{b}_k + \hat{b}_k^\dagger(\hat{n}_{k+1}+\hat{n}_k)\hat{b}_{k+1}\Big]. \tag{14}$$

As proved in Lemma 5, the transformation $|L,0\rangle \to |L-1,1\rangle$ can be accomplished within time $\pi/(2JL\sqrt{L})$ by the above Hamiltonian with the $U \to \infty$ limit. In the second stage, the Hamiltonian reads

$$H_2 = J(\hat{b}_k^\dagger\hat{b}_{k+1} + \hat{b}_{k+1}^\dagger\hat{b}_k) + J\Big[\hat{b}_{k+1}^\dagger(\hat{n}_{k+1}+\hat{n}_k)\hat{b}_k + \hat{b}_k^\dagger(\hat{n}_{k+1}+\hat{n}_k)\hat{b}_{k+1}\Big]. \tag{15}$$

As shown in Lemma 4, the transformation $|L-1,1\rangle \to |1,L-1\rangle$ can be attained within time $\pi/(2JL)$ by this Hamiltonian. In the last stage, the Hamiltonian is given by

$$H_3 = J(\hat{b}_k^\dagger\hat{b}_{k+1} + \hat{b}_{k+1}^\dagger\hat{b}_k) - U\hat{n}_{k+1}^2 + U(2L-1)\hat{n}_{k+1} + J\Big[\hat{b}_{k+1}^\dagger(\hat{n}_{k+1}+\hat{n}_k)\hat{b}_k + \hat{b}_k^\dagger(\hat{n}_{k+1}+\hat{n}_k)\hat{b}_{k+1}\Big]. \tag{16}$$

As proved in Lemma 5, the transformation $|1,L-1\rangle \to |0,L\rangle$ can be achieved within time $\pi/(2JL\sqrt{L})$ by the Hamiltonian H_3 in the $U \to \infty$ limit.

For the protocol constructed above, the total time consumed for the bosonic transport is

$$\tau = \frac{\pi}{2J}\sum_{k=1}^{L-1}\left(\frac{1}{L} + \frac{2}{L\sqrt{L}}\right) < \frac{\pi}{J} \tag{17}$$

for any $L \geq 3$. As can be observed, macroscopic transport can always be accomplished within a constant of time, independent of the distance between the regions. Therefore, the speed of macroscopic bosonic transport is not finite in bosonic systems with interaction-induced tunneling terms.

3 Conclusion

Using optimal transport theory, we elucidated fundamental limits on the speed of macroscopic particle transport in long-range bosonic systems from the perspective of the quantum speed limit. The results hold for the entire range $\alpha > D$ of power decay and arbitrary initial states under generic long-range characteristics. Our findings completely resolve the critical open problem of the speed of macroscopic particle transport. It is worth noting that the speed of bosonic

[4] There exists another simple protocol that sequentially transports bosons as $|L,0\rangle \to |L-1,1\rangle \to |L-2,2\rangle \to \cdots \to |0,L\rangle$. In this case, the total time required to transport L bosons over a distance of $L-1$ in the thermodynamic limit $L \to \infty$ is given by $\pi^2/(2J)$ (cf. Lemma 5).

transport may not be finite in the presence of higher-order hopping terms (i.e., terms like $\hat{b}_i\hat{b}_j\hat{b}_k^\dagger\hat{b}_l^\dagger$), such as interaction-induced tunneling terms in Eq. (13). We also note that our study does not delve into the exploration of the Lieb-Robinson bound concerning the speed of information propagation. This crucial aspect remains unaddressed and is reserved for future research.

Acknowledgements

The authors thank Marius Lemm for the helpful communication. T.V.V. was supported by JSPS KAKENHI Grant No. JP23K13032. T.K. acknowledges Hakubi projects of RIKEN and was supported by JST, PRESTO Grant No. JPMJPR2116. K.S. was supported by JSPS KAKENHI Grant No. JP23K25796.

References

[1] Elliott H. Lieb and Derek W. Robinson. "The finite group velocity of quantum spin systems". Commun. Math. Phys. **28**, 251–257 (1972).

[2] L. Mandelstam and Ig. Tamm. "The uncertainty relation between energy and time in non-relativistic quantum mechanics". J. Phys. USSR **9**, 249–254 (1945).

[3] M. B. Hastings and Xiao-Gang Wen. "Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance". Phys. Rev. B **72**, 045141 (2005).

[4] S. Bravyi, M. B. Hastings, and F. Verstraete. "Lieb-Robinson bounds and the generation of correlations and topological quantum order". Phys. Rev. Lett. **97**, 050401 (2006).

[5] Matthew B. Hastings and Tohru Koma. "Spectral gap and exponential decay of correlations". Commun. Math. Phys. **265**, 781–804 (2006).

[6] Bruno Nachtergaele and Robert Sims. "Lieb-Robinson bounds and the exponential clustering theorem". Commun. Math. Phys. **265**, 119–130 (2006).

[7] Tobias J. Osborne. "Efficient approximation of the dynamics of one-dimensional quantum spin systems". Phys. Rev. Lett. **97**, 157202 (2006).

[8] M B Hastings. "An area law for one-dimensional quantum systems". J. Stat. Mech.: Theory Exp. **2007**, P08024 (2007).

[9] Karel Van Acoleyen, Michaël Mariën, and Frank Verstraete. "Entanglement rates and area laws". Phys. Rev. Lett. **111**, 170501 (2013).

[10] Daniel A. Roberts and Brian Swingle. "Lieb-Robinson bound and the butterfly effect in quantum field theories". Phys. Rev. Lett. **117**, 091602 (2016).

[11] Eiki Iyoda, Kazuya Kaneko, and Takahiro Sagawa. "Fluctuation theorem for many-body pure quantum states". Phys. Rev. Lett. **119**, 100601 (2017).

[12] Jeongwan Haah, Matthew B. Hastings, Robin Kothari, and Guang Hao Low. "Quantum algorithm for simulating real time evolution of lattice Hamiltonians". SIAM J. Comput. **52**, FOCS18–250 (2018).

[13] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. "Sample-efficient learning of interacting quantum systems". Nat. Phys. **17**, 931–935 (2021).

[14] Álvaro M. Alhambra and J. Ignacio Cirac. "Locally accurate tensor networks for thermal states and time evolution". PRX Quantum **2**, 040331 (2021).

[15] Tomotaka Kuwahara and Keiji Saito. "Exponential clustering of bipartite quantum entanglement at arbitrary temperatures". Phys. Rev. X **12**, 021022 (2022).

[16] Chi-Fang Chen, Andrew Lucas, and Chao Yin. "Speed limits and locality in many-body quantum dynamics". Rep. Prog. Phys. **86**, 116001 (2023).

[17] Armin Uhlmann. "An energy dispersion estimate". Phys. Lett. A **161**, 329–331 (1992).

[18] Norman Margolus and Lev B. Levitin. "The maximum speed of dynamical evolution". Physica D **120**, 188–195 (1998).

[19] A. del Campo, I. L. Egusquiza, M. B. Plenio, and S. F. Huelga. "Quantum speed limits in open system dynamics". Phys. Rev. Lett. **110**, 050403 (2013).

[20] Sebastian Deffner and Eric Lutz. "Quantum speed limit for non-Markovian dynamics". Phys. Rev. Lett. **111**, 010402 (2013).

[21] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho. "Quantum speed limit for physical processes". Phys. Rev. Lett. **110**, 050402 (2013).

[22] Diego Paiva Pires, Marco Cianciaruso, Lucas C. Céleri, Gerardo Adesso, and Diogo O. Soares-Pinto. "Generalized geometric quantum speed limits". Phys. Rev. X **6**, 021031 (2016).

[23] Debasis Mondal, Chandan Datta, and Sk Sazim. "Quantum coherence sets the quantum speed limit for mixed states". Phys. Lett. A **380**, 689–695 (2016).

[24] Sebastian Deffner. "Geometric quantum speed limits: a case for Wigner phase space". New J. Phys. **19**, 103018 (2017).

[25] B. Shanahan, A. Chenu, N. Margolus, and A. del Campo. "Quantum speed limits across the quantum-to-classical transition". Phys. Rev. Lett. **120**, 070401 (2018).

[26] Francesco Campaioli, Felix A. Pollock, Felix C. Binder, and Kavan Modi. "Tightening quantum

speed limits for almost all states". Phys. Rev. Lett. **120**, 060409 (2018).

[27] Ken Funo, Naoto Shiraishi, and Keiji Saito. "Speed limit for open quantum systems". New J. Phys. **21**, 013006 (2019).

[28] Luis Pedro García-Pintos and Adolfo del Campo. "Quantum speed limits under continuous quantum measurements". New J. Phys. **21**, 033012 (2019).

[29] Shuning Sun, Yonggang Peng, Xianghong Hu, and Yujun Zheng. "Quantum speed limit quantified by the changing rate of phase". Phys. Rev. Lett. **127**, 100404 (2021).

[30] Tan Van Vu and Yoshihiko Hasegawa. "Geometrical bounds of the irreversibility in Markovian systems". Phys. Rev. Lett. **126**, 010601 (2021).

[31] Naoto Shiraishi and Keiji Saito. "Speed limit for open systems coupled to general environments". Phys. Rev. Res. **3**, 023074 (2021).

[32] Ryusuke Hamazaki. "Speed limits for macroscopic transitions". PRX Quantum **3**, 020319 (2022).

[33] Tan Van Vu and Keiji Saito. "Topological speed limit". Phys. Rev. Lett. **130**, 010402 (2023).

[34] Yoshihiko Hasegawa. "Unifying speed limit, thermodynamic uncertainty relation and Heisenberg principle via bulk-boundary correspondence". Nat. Commun. **14**, 2828 (2023).

[35] Sebastian Deffner and Steve Campbell. "Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control". J. Phys. A **50**, 453001 (2017).

[36] Jens Eisert, Mauritz van den Worm, Salvatore R. Manmana, and Michael Kastner. "Breakdown of quasilocality in long-range quantum lattice models". Phys. Rev. Lett. **111**, 260401 (2013).

[37] M. Saffman, T. G. Walker, and K. Mølmer. "Quantum information with Rydberg atoms". Rev. Mod. Phys. **82**, 2313–2363 (2010).

[38] Nicolò Defenu, Tobias Donner, Tommaso Macrì, Guido Pagano, Stefano Ruffo, and Andrea Trombettoni. "Long-range interacting quantum systems". Rev. Mod. Phys. **95**, 035002 (2023).

[39] Michael Foss-Feig, Zhe-Xuan Gong, Charles W. Clark, and Alexey V. Gorshkov. "Nearly linear light cones in long-range interacting quantum systems". Phys. Rev. Lett. **114**, 157201 (2015).

[40] Takuro Matsuta, Tohru Koma, and Shu Nakamura. "Improving the Lieb-Robinson bound for long-range interactions". Ann. Henri Poincaré **18**, 519–528 (2016).

[41] Chi-Fang Chen and Andrew Lucas. "Finite speed of quantum scrambling with long range interactions". Phys. Rev. Lett. **123**, 250605 (2019).

[42] Tomotaka Kuwahara and Keiji Saito. "Strictly linear light cones in long-range interacting systems of arbitrary dimensions". Phys. Rev. X **10**, 031010 (2020).

[43] Dominic V. Else, Francisco Machado, Chetan Nayak, and Norman Y. Yao. "Improved Lieb-Robinson bound for many-body Hamiltonians with power-law interactions". Phys. Rev. A **101**, 022333 (2020).

[44] Minh C. Tran, Andrew Y. Guo, Abhinav Deshpande, Andrew Lucas, and Alexey V. Gorshkov. "Optimal state transfer and entanglement generation in power-law interacting systems". Phys. Rev. X **11**, 031016 (2021).

[45] Tomotaka Kuwahara and Keiji Saito. "Absence of fast scrambling in thermodynamically stable long-range interacting systems". Phys. Rev. Lett. **126**, 030604 (2021).

[46] Minh C. Tran, Andrew Y. Guo, Christopher L. Baldwin, Adam Ehrenberg, Alexey V. Gorshkov, and Andrew Lucas. "Lieb-Robinson light cone for power-law interactions". Phys. Rev. Lett. **127**, 160401 (2021).

[47] Chi-Fang Chen and Andrew Lucas. "Optimal Frobenius light cone in spin chains with power-law interactions". Phys. Rev. A **104**, 062420 (2021).

[48] Zongping Gong, Tommaso Guaita, and J. Ignacio Cirac. "Long-range free fermions: Lieb-Robinson bound, clustering properties, and topological phases". Phys. Rev. Lett. **130**, 070401 (2023).

[49] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. "Many-body physics with ultracold gases". Rev. Mod. Phys. **80**, 885–964 (2008).

[50] Andreas M Läuchli and Corinna Kollath. "Spreading of correlations and entanglement after a quench in the one-dimensional Bose-Hubbard model". J. Stat. Mech.: Theory Exp. **2008**, P05018 (2008).

[51] Giuseppe Carleo, Federico Becca, Laurent Sanchez-Palencia, Sandro Sorella, and Michele Fabrizio. "Light-cone effect and supersonic correlations in one- and two-dimensional bosonic superfluids". Phys. Rev. A **89**, 031602 (2014).

[52] Lorenzo Cevolani, Giuseppe Carleo, and Laurent Sanchez-Palencia. "Protected quasilocality in quantum systems with long-range interactions". Phys. Rev. A **92**, 041603 (2015).

[53] Marc Cheneau, Peter Barmettler, Dario Poletti, Manuel Endres, Peter Schauß, Takeshi Fukuhara, Christian Gross, Immanuel Bloch, Corinna Kollath, and Stefan Kuhr. "Light-cone-like spreading of correlations in a quantum many-body system". Nature **481**, 484–487 (2012).

[54] Yosuke Takasu, Tomoya Yagami, Hiroto Asaka, Yoshiaki Fukushima, Kazuma Nagao, Shimpei Goto, Ippei Danshita, and Yoshiro Takahashi. "Energy redistribution and spatiotemporal evolution of correlations after a sudden quench of the Bose-Hubbard model". Sci. Adv. **6**, eaba9255 (2020).

[55] Norbert Schuch, Sarah K. Harrison, Tobias J. Osborne, and Jens Eisert. "Information propagation for interacting-particle systems". Phys. Rev. A **84**, 032309 (2011).

[56] Zhiyuan Wang and Kaden R.A. Hazzard. "Tightening the Lieb-Robinson bound in locally interacting systems". PRX Quantum **1**, 010303 (2020).

[57] Tomotaka Kuwahara and Keiji Saito. "Lieb-Robinson bound and almost-linear light cone in interacting boson systems". Phys. Rev. Lett. **127**, 070403 (2021).

[58] Chao Yin and Andrew Lucas. "Finite speed of quantum information in models of interacting bosons at finite density". Phys. Rev. X **12**, 021039 (2022).

[59] Tomotaka Kuwahara, Tan Van Vu, and Keiji Saito. "Effective light cone and digital quantum simulation of interacting bosons". Nat. Commun. **15**, 2520 (2024).

[60] Jérémy Faupin, Marius Lemm, and Israel Michael Sigal. "Maximal speed for macroscopic particle transport in the Bose-Hubbard model". Phys. Rev. Lett. **128**, 150602 (2022).

[61] Jérémy Faupin, Marius Lemm, and Israel Michael Sigal. "On Lieb-Robinson bounds for the Bose-Hubbard model". Commun. Math. Phys. **394**, 1011–1037 (2022).

[62] Marius Lemm, Carla Rubiliani, Israel Michael Sigal, and Jingxuan Zhang. "Information propagation in long-range quantum many-body systems". Phys. Rev. A **108**, L060401 (2023).

[63] Minh C. Tran, Chi-Fang Chen, Adam Ehrenberg, Andrew Y. Guo, Abhinav Deshpande, Yifan Hong, Zhe-Xuan Gong, Alexey V. Gorshkov, and Andrew Lucas. "Hierarchy of linear light cones with long-range interactions". Phys. Rev. X **10**, 031009 (2020).

[64] Marin Bukov, Dries Sels, and Anatoli Polkovnikov. "Geometric speed limit of accessible many-body state preparation". Phys. Rev. X **9**, 011034 (2019).

[65] J. Eisert and D. Gross. "Supersonic quantum communication". Phys. Rev. Lett. **102**, 240501 (2009).

[66] Tomasz Sowiński, Omjyoti Dutta, Philipp Hauke, Luca Tagliacozzo, and Maciej Lewenstein. "Dipolar molecules in optical lattices". Phys. Rev. Lett. **108**, 115301 (2012).

[67] Omjyoti Dutta, Mariusz Gajda, Philipp Hauke, Maciej Lewenstein, Dirk-Sören Lühmann, Boris A Malomed, Tomasz Sowiński, and Jakub Zakrzewski. "Non-standard Hubbard models in optical lattices: a review". Rep. Prog. Phys. **78**, 066001 (2015).

[68] Ole Jürgensen, Florian Meinert, Manfred J. Mark, Hanns-Christoph Nägerl, and Dirk-Sören Lühmann. "Observation of density-induced tunneling". Phys. Rev. Lett. **113**, 193003 (2014).

[69] Soheil Kolouri, Se Rim Park, Matthew Thorpe, Dejan Slepcev, and Gustavo K. Rohde. "Optimal mass transport: Signal processing and machine-learning applications". IEEE Signal Process. Mag. **34**, 43–59 (2017).

[70] Steven Haker, Lei Zhu, Allen Tannenbaum, and Sigurd Angenent. "Optimal mass transport for registration and warping". Int. J. Comput. Vision **60**, 225–240 (2004).

[71] Gao Huang, Chuan Guo, Matt J Kusner, Yu Sun, Fei Sha, and Kilian Q Weinberger. "Supervised word mover's distance". In Advances in Neural Information Processing Systems. Volume 29. (2016).

[72] Geoffrey Schiebinger, Jian Shu, Marcin Tabaka, Brian Cleary, Vidya Subramanian, Aryeh Solomon, Joshua Gould, Siyan Liu, Stacie Lin, Peter Berube, et al. "Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming". Cell **176**, 928–943 (2019).

[73] Erik Aurell, Carlos Mejía-Monasterio, and Paolo Muratore-Ginanneschi. "Optimal protocols and optimal transport in stochastic thermodynamics". Phys. Rev. Lett. **106**, 250601 (2011).

[74] Muka Nakazato and Sosuke Ito. "Geometrical aspects of entropy production in stochastic thermodynamics based on wasserstein distance". Phys. Rev. Res. **3**, 043093 (2021).

[75] Andreas Dechant. "Minimum entropy production, detailed balance and Wasserstein distance for continuous-time Markov processes". J. Phys. A **55**, 094001 (2022).

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

[77] Cédric Villani. "Optimal Transport: Old and New". Springer. Berlin, Heidelberg (2008).

[78] Mark Kac. "Random walk and the theory of Brownian motion". Am. Math. Mon. **54**, 369 (1947).

A Methods

A.1 Unified speed limit based on optimal transport theory

The key ingredient in the proofs of our results is a novel unified speed limit. This speed limit is derived using optimal transport theory, which plays crucial roles in various disciplines, such as statistics and machine learning [69], computer vision [70], linguistics [71], molecular biology [72], and stochastic thermodynamics [73–76]. Here we describe the unified speed limit and its derivation.

First, we begin with a very general setup, which is applicable to both classical and quantum systems. We consider a time evolution of a vector state $\boldsymbol{x}_t \coloneqq [x_1(t), \dots, x_N(t)]^\top \in \mathbb{R}^N_{\geq 0}$ over an undirected graph $\mathcal{G}(V, E)$ with the vertex set $V = \{1, \dots, N\}$ and edge set $E = \{(i, j)\}$. For each vertex i, let $B_i \coloneqq \{j \,|\, (i, j) \in E\}$ denote the set of neighboring vertices of i. The dynamics of $\boldsymbol{x}_t$ is described by the following deterministic equation:

$$\dot{x}_i(t) = \sum_{j \in B_i} f_{ij}(t), \tag{18}$$

where $f_{ij}(t) = -f_{ji}(t)$ denotes the flow exchange between vertices i and j for $i \neq j$. Note that the sum $\sum_{i=1}^N x_i(t)$ is time-invariant.

Next, we introduce the discrete L^1-Wasserstein distance between two distributions $\boldsymbol{x}$ and $\boldsymbol{y}$. Suppose that we have a transport plan that redistributes $\boldsymbol{x}$ to $\boldsymbol{y}$ by sending an amount of $\pi_{ij} \geq 0$ from x_j to y_i with a cost of $c_{ij} \geq 0$ per unit mass for any i and j. Here, $\pi = [\pi_{ij}] \in \mathbb{R}^{N \times N}_{\geq 0}$ is a joint probability distribution such that its marginal distributions reduce to $\sum_j \pi_{ij} = y_i$ and $\sum_j \pi_{ji} = x_i$, defining an admissible transport plan. The L^1-Wasserstein distance is then defined as the minimum transport cost over all feasible plans, given by

$$\mathcal{W}(\boldsymbol{x}, \boldsymbol{y}) \coloneqq \min_{\pi \in \mathcal{C}(\boldsymbol{x}, \boldsymbol{y})} \sum_{i,j} c_{ij} \pi_{ij}, \tag{19}$$

where $\mathcal{C}(\boldsymbol{x}, \boldsymbol{y})$ denotes the set of couplings π. The cost matrix $[c_{ij}]$ can be specified arbitrarily, as long as the following two requirements are fulfilled: (i) symmetry ($c_{ij} = c_{ji}\ \forall i, j$) and (ii) the triangle inequality ($c_{ij} + c_{jk} \geq c_{ik} \forall i, j, k$). The freedom to set up the cost matrix is one of the powerful aspects of the Wasserstein distance. This freedom enables us to include the geometric structure of the underlying dynamics, leading to the derivation of novel physical implications.

Given the setup above, by generalizing the speed limit obtained in Ref. [33], we prove that the minimum time required to transform state $\boldsymbol{x}_0$ into $\boldsymbol{x}_\tau$ is always lower bounded as

$$\tau \geq \frac{\mathcal{W}(\boldsymbol{x}_0, \boldsymbol{x}_\tau)}{\left\langle \sum_{(i,j) \in E} c_{ij} |f_{ij}(t)| \right\rangle_\tau}, \tag{20}$$

where $\langle z_t \rangle_\tau \coloneqq \tau^{-1} \int_0^\tau dt\, z_t$ is the time-average quantity of z_t. The unified speed limit (20) includes all the essence to extend the use of the conventional quantum speed limits to a wide range of quantum many-body problems with various geometric structures. The crucial step in applying this bound effectively is the thoughtful selection of the cost matrix $[c_{ij}]$. With appropriate choices, our new bound (20) can play roles similar to the Lieb-Robinson bound. In the specific case when $c_{ij} = 1$ for any $(i, j) \in E$ and c_{ij} represents the shortest-path distance between vertices i and j otherwise, the unified speed limit recovers the topological speed limit [33].

The unified speed limit can be derived with the help of the Kantorovich-Rubinstein duality [77], which expresses the Wasserstein distance in an alternative variational form as (see Lemma 8 for the proof)

$$\mathcal{W}(\boldsymbol{x}, \boldsymbol{y}) = \max_{\boldsymbol{\phi}} \boldsymbol{\phi}^\top (\boldsymbol{x} - \boldsymbol{y}), \tag{21}$$

where the maximum is over all vectors $\boldsymbol{\phi} = [\phi_1, \dots, \phi_N]^\top$ satisfying $|\phi_i - \phi_j| \leq c_{ij}\ \forall i, j$. Notice that $x_i(\tau) = x_i(0) + \int_0^\tau dt \sum_{j \in B_i} f_{ij}(t)$. Then, by applying the Kantorovich-Rubinstein duality, we can prove Eq. (20) as

follows:

$$\begin{aligned}\mathcal{W}(\boldsymbol{x}_0,\boldsymbol{x}_\tau) &= \max_{\boldsymbol{\phi}} \boldsymbol{\phi}^\top(\boldsymbol{x}_\tau-\boldsymbol{x}_0)\\ &= \max_{\boldsymbol{\phi}} \sum_i \phi_i \int_0^\tau dt \sum_{j\in B_i} f_{ij}(t)\\ &= \max_{\boldsymbol{\phi}} \sum_{(i,j)\in E} (\phi_i-\phi_j)\int_0^\tau dt\, f_{ij}(t)\\ &\le \max_{\boldsymbol{\phi}} \sum_{(i,j)\in E} |\phi_i-\phi_j|\int_0^\tau dt\, |f_{ij}(t)|\\ &\le \int_0^\tau dt \sum_{(i,j)\in E} c_{ij}|f_{ij}(t)|\\ &= \tau\Big\langle \sum_{(i,j)\in E} c_{ij}|f_{ij}(t)|\Big\rangle_\tau. \end{aligned} \tag{22}$$

In what follows, we apply the unified speed limit to prove our main results, Theorems 1 and 2.

A.2 Proof of Theorem 1

First, we explain some useful notations that will be used later. For any site i, we denote by $i[r]$ the ball of radius $r\ge 0$, which contains all sites j such that $\|i-j\|<r$. Using this definition, the set $\Lambda\setminus\{i\}$ can be decomposed as

$$\Lambda\setminus\{i\} = \cup_{\ell=1}^{\infty}(i[\ell+1]\setminus i[\ell]). \tag{23}$$

In addition, we define γ as a constant such that the following inequality holds for any $r\ge 0$:

$$|i[r+1]\setminus i[r]| \le \gamma r^{D-1}. \tag{24}$$

We consider a $|\Lambda|$-dimensional vector $\boldsymbol{x}_t$, where $x_i(t)$ denotes the boson concentration occupied at site i [i.e., $x_i(t) \coloneqq \mathrm{tr}\{\overline{n}_i\varrho_t\}$]. Evidently, $\sum_i x_i(t)=1$ for any t. To define the Wasserstein distance between these vectors, the cost matrix $[c_{ij}]$ needs to be specified. We consider the following cost matrix:

$$c_{ij} = \|i-j\|^{\alpha_\epsilon}, \tag{25}$$

where $0<\epsilon<\alpha-D$ is an arbitrary number and $\alpha_\epsilon \coloneqq \min(1,\alpha-D-\epsilon)$ is defined for convenience. Since $\|i-j\|+\|j-k\|\ge\|i-k\|$ and $a^z+b^z\ge(a+b)^z$ for any $a,b\ge 0$ and $0\le z\le 1$, we can verify that

$$c_{ij}+c_{jk}\ge c_{ik} \tag{26}$$

for any $i,j,k\in\Lambda$. As can be observed, this distance includes spatial information of the lattice and can be the order of the system size.

We are now ready to prove Theorem 1. Using the relation $[\hat{b}_i,\hat{n}_i]=\hat{b}_i$, we can show that the time evolution of $x_i(t)$ is governed by the following equation:

$$\dot{x}_i(t) = \sum_{j(\neq i)} 2J_{ij}(t)\mathcal{N}^{-1}\,\mathrm{Im}\left[\mathrm{tr}\left\{\hat{b}_i^\dagger\hat{b}_j\varrho_t\right\}\right] =: \sum_{j(\neq i)}\phi_{ij}(t). \tag{27}$$

It can be easily verified that $\phi_{ij}(t)=-\phi_{ji}(t)$. Therefore, the dynamics of $\boldsymbol{x}_t$ can be mapped to Eq. (18), where the graph $\mathcal{G}$ is determined by $V=\Lambda$ and $B_i=\Lambda\setminus\{i\}$. Applying the unified speed limit (20) yields

$$\tau \ge \frac{\mathcal{W}(\boldsymbol{x}_0,\boldsymbol{x}_\tau)}{\langle\Phi_t\rangle_\tau}, \tag{28}$$

where we define $\Phi_t \coloneqq (1/2)\sum_{i\neq j} c_{ij}|\phi_{ij}(t)|$. Since a fraction μ of bosons must be transported from X to Y [i.e., the inequality (7) is satisfied], we have

$$\begin{aligned}\mu \le x_Y(\tau)-x_{X_c}(0) &= \sum_{i\in Y,j\in\Lambda}\pi_{ij} - \sum_{i\in\Lambda,j\in X_c}\pi_{ij}\\ &= \sum_{i\in Y,j\in X}\pi_{ij} - \sum_{i\in Y_c,j\in X_c}\pi_{ij}\\ &\le \sum_{i\in Y,j\in X}\pi_{ij}. \end{aligned} \tag{29}$$

Therefore,

$$\mathcal{W}(\boldsymbol{x}_0, \boldsymbol{x}_\tau) \geq \left(\min_{i\in Y, j\in X} c_{ij} \right) \sum_{i\in Y, j\in X} \pi_{ij} \geq \mu d_{XY}^{\alpha_\epsilon}. \tag{30}$$

On the other hand, by applying the Cauchy-Schwarz inequality, $\phi_{ij}(t)$ can be upper bounded as

$$|\phi_{ij}(t)| \leq |J_{ij}(t)|[x_i(t) + x_j(t)]. \tag{31}$$

Consequently, the velocity term Φ_t is bounded from above as follows:

$$\begin{aligned}
\Phi_t &\leq \sum_i x_i(t) \sum_{j(\neq i)} |J_{ij}(t)| \|i-j\|^{\alpha_\epsilon} \\
&\leq \sum_i x_i(t) \sum_{\ell=1}^{\infty} \sum_{j\in(i[\ell+1]\setminus i[\ell])} \frac{J}{\|i-j\|^{\alpha-\alpha_\epsilon}} \\
&\leq J\gamma \sum_i x_i(t) \sum_{\ell=1}^{\infty} \frac{1}{\ell^{\alpha-\alpha_\epsilon-D+1}} \\
&= J\gamma\zeta(\alpha-\alpha_\epsilon-D+1).
\end{aligned} \tag{32}$$

Here, we use the facts that $\|i-j\| \geq \ell$ for any $j \in (i[\ell+1]\setminus i[\ell])$ and $|i[\ell+1]\setminus i[\ell]| \leq \gamma\ell^{D-1}$ in the third line, and $\zeta(\cdot)$ denotes the Riemann zeta function. Define $\kappa_1^\epsilon := [J\gamma\zeta(\alpha-\alpha_\epsilon-D+1)]^{-1}\mu$. Note that κ_1^ϵ is independent of the system size and finite because $\alpha-\alpha_\epsilon-D+1 \geq 1+\epsilon$. Combining Eqs. (28), (30), and (32) yields

$$\tau \geq \kappa_1^\epsilon d_{XY}^{\alpha_\epsilon}, \tag{33}$$

which completes the proof.

A.3 Proof of Theorem 2

The proof strategy is similar to that of Theorem 1. The different point is that we consider the probability distribution of boson numbers at all sites instead of the vector of average boson concentrations. Define the probability $p_{\vec{N}}(t) := \mathrm{tr}\{\Pi_{\vec{N}}\varrho_t\}$. Since the total boson number is invariant, we need only consider feasible states $|\vec{N}\rangle$ that satisfy $\mathrm{tr}\{\hat{n}_\Lambda\Pi_{\vec{N}}\} = \mathcal{N}$. The time evolution of the probability distribution $\boldsymbol{p}_t = [p_{\vec{N}}(t)]^\top$ can be derived from the von Neumann equation as

$$\begin{aligned}
\dot{p}_{\vec{N}}(t) &= -i\,\mathrm{tr}\{\Pi_{\vec{N}}[H_t, \varrho_t]\} \\
&= \sum_{i\neq j} iJ_{ij}(t)\sqrt{n_i n_j'}\left(\left\langle \vec{N}\middle|\varrho_t\middle|\vec{N}'\right\rangle - \left\langle \vec{N}'\middle|\varrho_t\middle|\vec{N}\right\rangle \right) \\
&=: \sum_{i\neq j} \varphi_{\vec{N}\vec{N}'}(t),
\end{aligned} \tag{34}$$

where $\vec{N}'$ is obtained from $\vec{N}$ by setting $n_i' = n_i - 1$, $n_j' = n_j + 1$, and $n_k' = n_k$ for all $k \neq i, j$. Note that $\varphi_{\vec{N}\vec{N}'}(t) = -\varphi_{\vec{N}'\vec{N}}(t) \in \mathbb{R}$. Two states $\vec{N}$ and $\vec{M}$ are said to be neighboring if there exist $i \neq j$ such that $|n_i - m_i| = |n_j - m_j| = 1$ and $n_k = m_k$ for any $k \neq i, j$. For such neighboring states, the transport cost is defined as $c_{\vec{N}\vec{N}'} = \|i-j\|^{\alpha_\epsilon}$. For arbitrary states $\vec{N}$ and $\vec{M}$, the cost is defined as the shortest-path cost over all possible paths connecting the two states,

$$c_{\vec{M}\vec{N}} = \min \sum_{k=1}^{K} c_{\vec{N}_k\vec{N}_{k-1}}, \tag{35}$$

where $\vec{N}_0 = \vec{N}$, $\vec{N}_K = \vec{M}$, and $\vec{N}_{k-1}$ and $\vec{N}_k$ are neighboring states for all $1 \leq k \leq K$. Obviously, $c_{\vec{N}\vec{N}} = 0$. Using the defined shortest-path costs, we consider the following Wasserstein distance:

$$\mathcal{W}(\boldsymbol{p}, \boldsymbol{q}) := \min_{\pi\in\mathcal{C}(\boldsymbol{p},\boldsymbol{q})} \sum_{\vec{M},\vec{N}} c_{\vec{M}\vec{N}}\pi_{\vec{M}\vec{N}}. \tag{36}$$

Utilizing the unified speed limit (20) for this Wasserstein distance, we can arrive at the following inequality:

$$\frac{1}{2}\int_0^\tau dt \sum_{\vec{N}} \sum_{i\neq j} c_{\vec{N}\vec{N}'} |\varphi_{\vec{N}\vec{N}'}(t)| \geq \mathcal{W}(\boldsymbol{p}_0, \boldsymbol{p}_\tau). \tag{37}$$

Furthermore, applying the Cauchy-Schwarz inequality, we obtain

$$|\varphi_{\vec{N}\vec{N}'}(t)| \leq |J_{ij}(t)| \left[n_i p_{\vec{N}}(t) + n'_j p_{\vec{N}'}(t) \right]. \tag{38}$$

Consequently, the left-hand side of Eq. (37) can be upper bounded as

$$\begin{aligned}
\frac{1}{2}\sum_{\vec{N}} \sum_{i\neq j} c_{\vec{N}\vec{N}'} |\varphi_{\vec{N}\vec{N}'}(t)| &\leq \sum_{\vec{N}} p_{\vec{N}}(t) \sum_i n_i \sum_{j(\neq i)} \frac{J}{\|i-j\|^{\alpha-\alpha_\epsilon}} \\
&\leq \sum_{\vec{N}} p_{\vec{N}}(t) \mathcal{N} J \gamma \zeta (\alpha - \alpha_\epsilon - D + 1) \\
&= \mathcal{N} J \gamma \zeta (\alpha - \alpha_\epsilon - D + 1).
\end{aligned} \tag{39}$$

Define $\mathcal{S}_0 = \{\vec{N} \mid \sum_{i\in X_c} n_i \leq N_0\}$ and $\mathcal{S}_\tau = \{\vec{N} \mid \sum_{i\in Y} n_i \geq N_0 + \Delta N_0\}$. To reach state $\vec{M} \in \mathcal{S}_\tau$ from state $\vec{N} \in \mathcal{S}_0$ by hopping bosons between sites, an amount of ΔN_0 bosons must be transferred from sites of X to sites of Y. Let $\vec{N}_{i_0} \to \vec{N}_{i_1} \to \cdots \to \vec{N}_{i_L}$ be a sequence of states that transfers a boson from site $i_0 \in X$ to site $i_L \in Y$ by iteratively hopping bosons between sites i_{l-1} and i_l for $1 \leq l \leq L$. Then, by applying the triangle inequality, we obtain that the cost of transferring one boson is lower bounded as

$$\sum_{l=1}^{L} c_{\vec{N}_{i_l}\vec{N}_{i_{l-1}}} = \sum_{l=1}^{L} \|i_l - i_{l-1}\|^{\alpha_\epsilon} \geq \|i_L - i_0\|^{\alpha_\epsilon} \geq d_{XY}^{\alpha_\epsilon}. \tag{40}$$

Therefore, we immediately obtain $c_{\vec{M}\vec{N}} \geq \Delta N_0 d_{XY}^{\alpha_\epsilon}$ for any $\vec{M} \in \mathcal{S}_\tau$ and $\vec{N} \in \mathcal{S}_0$. Since $\langle P_{\hat{n}_{X_c} > N_0} \rangle_{\varrho_0} = 0$, $p_{\vec{N}}(0) = 0$ for any $\vec{N} \notin \mathcal{S}_0$. Combining this with $p_{\vec{N}}(0) = \sum_{\vec{M}} \pi_{\vec{M}\vec{N}}$ immediately derives $\pi_{\vec{M}\vec{N}} = 0$ for any $\vec{N} \notin \mathcal{S}_0$ and $\vec{M}$. Using these facts, the Wasserstein distance can be lower bounded as follows:

$$\begin{aligned}
\mathcal{W}(\boldsymbol{p}_0, \boldsymbol{p}_\tau) &= \min_\pi \sum_{\vec{M},\vec{N}} c_{\vec{M}\vec{N}} \pi_{\vec{M}\vec{N}} \\
&\geq \min_\pi \sum_{\vec{M}\in\mathcal{S}_\tau, \vec{N}\in\mathcal{S}_0} c_{\vec{M}\vec{N}} \pi_{\vec{M}\vec{N}} \\
&\geq \Delta N_0 d_{XY}^{\alpha_\epsilon} \min_\pi \sum_{\vec{M}\in\mathcal{S}_\tau} \sum_{\vec{N}\in\mathcal{S}_0} \pi_{\vec{M}\vec{N}} \\
&= \Delta N_0 d_{XY}^{\alpha_\epsilon} \min_\pi \sum_{\vec{M}\in\mathcal{S}_\tau} p_{\vec{M}}(\tau) \\
&= \Delta N_0 d_{XY}^{\alpha_\epsilon} \langle P_{\hat{n}_Y \geq N_0 + \Delta N_0} \rangle_{\varrho_\tau}.
\end{aligned} \tag{41}$$

Here, we use the relation $\sum_{\vec{N}\in\mathcal{S}_0} \pi_{\vec{M}\vec{N}} = \sum_{\vec{N}} \pi_{\vec{M}\vec{N}} = p_{\vec{M}}(\tau)$ in the fourth line. Combining Eqs. (37), (39), and (41) yields

$$\langle P_{\hat{n}_Y \geq N_0 + \Delta N_0} \rangle_{\varrho_\tau} \leq \frac{\mathcal{N} J \gamma \zeta (\alpha - \alpha_\epsilon - D + 1)\tau}{\Delta N_0 d_{XY}^{\alpha_\epsilon}}, \tag{42}$$

which completes the proof by setting $\kappa_2^\epsilon := J\gamma\zeta(\alpha - \alpha_\epsilon - D + 1)$.

A.4 Useful lemmas

We formally present useful lemmas about state transformation and their proofs in the following.

Lemma 4. *For any $M \geq 3$, the state transformation $|M-1, 1\rangle \leftrightarrow |1, M-1\rangle$ can be accomplished within time $t = \pi/(2JM)$ by the following Hamiltonian:*

$$H = J(\hat{b}_1^\dagger \hat{b}_2 + \hat{b}_2^\dagger \hat{b}_1) + J\left[\hat{b}_2^\dagger (\hat{n}_2 + \hat{n}_1)\hat{b}_1 + \hat{b}_1^\dagger (\hat{n}_2 + \hat{n}_1)\hat{b}_2\right]. \tag{43}$$

Proof. Let $|\psi_t\rangle = \sum_{m=0}^{M} a_m(t)\,|m, M-m\rangle$ be the expression of the time-evolved pure state. Then, the time evolution of $a_m(t)$ can be obtained from the Schrödinger equation $|\dot{\psi}_t\rangle = -iH\,|\psi_t\rangle$ as

$$\begin{aligned}\dot{a}_m(t) &= \langle m, M-m|\dot{\psi}_t\rangle \\ &= -i\,\langle m, M-m|H|\psi_t\rangle \\ &= -iJM\sqrt{m(M-m+1)}a_{m-1}(t) - iJM\sqrt{(m+1)(M-m)}a_{m+1}(t).\end{aligned} \tag{44}$$

Defining $\boldsymbol{a}_t \coloneqq [a_0(t), \ldots, a_M(t)]^\top$, we can rewrite Eq. (44) as

$$\dot{\boldsymbol{a}}_t = -iJMG\boldsymbol{a}_t, \tag{45}$$

where $\boldsymbol{a}_0 = [0, \ldots, 0, 1, 0]^\top$ and the matrix $G = [g_{mn}]_{0\le m,n\le M}$ is given by

$$g_{mn} = \begin{cases} \sqrt{m(M-m+1)} & \text{if } n = m-1, \\ \sqrt{(m+1)(M-m)} & \text{if } n = m+1, \\ 0 & \text{otherwise.} \end{cases} \tag{46}$$

Since $\boldsymbol{a}_t = e^{-iJMGt}\boldsymbol{a}_0$, it needs only show that $|[e^{-i\pi G/2}]_{1(M-1)}| = |[e^{-i\pi G/2}]_{(M-1)1}| = 1$. We divide into two cases: $M = 2\bar{M}-1$ and $M = 2\bar{M}$. We prove the former, and the latter can be proved analogously. Define $\tilde{G} = PGP^{-1} = [\tilde{g}_{mn}]_{0\le m,n\le M}$, where $P = \mathrm{diag}(\gamma_0, \ldots, \gamma_M)$ is a diagonal matrix that will be determined latter. We choose $\{\gamma_m\}$ such that

$$\tilde{g}_{mn} = \begin{cases} M-m+1 & \text{if } n = m-1, \\ m+1 & \text{if } n = m+1, \\ 0 & \text{otherwise.} \end{cases} \tag{47}$$

This can be attained by setting

$$\frac{\gamma_m}{\gamma_{m+1}} = \sqrt{\frac{m+1}{M-m}}. \tag{48}$$

The matrix $\tilde{G}$ is nothing but the Sylvester-Kac matrix. Since $e^{-i\pi G/2} = P^{-1}e^{-i\pi\tilde{G}/2}P$ and $\gamma_1 = \gamma_{M-1}$, $[e^{-i\pi G/2}]_{1(M-1)} = [e^{-i\pi\tilde{G}/2}]_{1(M-1)}$; thus, we need only prove that $|[e^{-i\pi\tilde{G}/2}]_{1(M-1)}| = 1$. It was shown that $\tilde{G}$ has distinct eigenvalues $\{\lambda_k = M-2k\}_{k=0}^{M}$ [78], thus being diagonalizable. Therefore, its left eigenvectors $\{|v_k^l\rangle\}$ and right eigenvectors $\{|v_k^r\rangle\}$ are orthogonal. That is, $\langle v_i^l|v_j^r\rangle = \delta_{ij}$ for any i and j. We note that $\langle v_k^r|v_k^r\rangle$ and $\langle v_k^l|v_k^l\rangle$ are not necessarily equal to 1. From the representation $\tilde{G} = \sum_k \lambda_k\,|v_k^r\rangle\langle v_k^l|$, we obtain

$$e^{-i\pi\tilde{G}/2} = \sum_{k=0}^{M} e^{-i\pi\lambda_k/2}\,|v_k^r\rangle\langle v_k^l| = i\sum_{k=0}^{M}(-1)^{\bar{M}+k+1}\,|v_k^r\rangle\langle v_k^l|\,. \tag{49}$$

Next, we analytically calculate the right and left eigenvectors. For each $\lambda_k = M-2k$, from the relation $\tilde{G}\,|v_k^r\rangle = \lambda_k\,|v_k^r\rangle$, we obtain

$$\begin{aligned} v_{k1}^r &= \lambda_k v_{k0}^r, \\ Mv_{k0}^r + 2v_{k2}^r &= \lambda_k v_{k1}^r, \\ (M-1)v_{k1}^r + 3v_{k3}^r &= \lambda_k v_{k2}^r, \\ &\cdots \\ v_{k(M-1)}^r &= \lambda_k v_{kM}^r. \end{aligned} \tag{50}$$

We consider an auxiliary infinite system,

$$\begin{aligned} v_{k1}^r &= \lambda_k v_{k0}^r, \\ Mv_{k0}^r + 2v_{k2}^r &= \lambda_k v_{k1}^r, \\ (M-1)v_{k1}^r + 3v_{k3}^r &= \lambda_k v_{k2}^r, \\ &\cdots \\ v_{k(M-1)}^r + (M+1)v_{k(M+1)}^r &= \lambda_k v_{kM}^r, \\ (M+2)v_{k(M+2)}^r &= \lambda_k v_{k(M+1)}^r, \\ &\cdots \end{aligned} \tag{51}$$

If there exists a nontrivial solution of Eq. (51) for which $v^r_{k(M+1)} = 0$, then we immediately obtain the solution for Eq. (50). Multiplying all equations of Eq. (51) by $1, z, z^2, \dots$ and adding them, we readily obtain

$$\sum_{m=0}^{\infty} v^r_{km}(M-m)z^{m+1} + \sum_{m=1}^{\infty} v^r_{km} m z^{m-1} = \lambda_k \sum_{m=0}^{\infty} v^r_{km} z^m. \tag{52}$$

Defining the generating function $f_k(z) \coloneqq \sum_{m=0}^{\infty} v^r_{km} z^m$, Eq. (52) can be translated to the following form:

$$Mz f_k(z) - z^2 f'_k(z) + f'_k(z) = \lambda_k f_k(z), \tag{53}$$

which yields the following differential equation:

$$f'_k(z) = \frac{\lambda_k - Mz}{1-z^2} f_k(z) \tag{54}$$

with the initial condition $f_k(0) = v^r_{k0}$. Solving this differential equation, $f_k(z)$ can be explicitly obtained as

$$f_k(z) = v^r_{k0}(1+z)^{M-k}(1-z)^k. \tag{55}$$

Consequently, the components of the right eigenvectors can be calculated as

$$v^r_{km} = v^r_{k0} \sum_{i=m-k}^{M-k} (-1)^{m-i} \binom{k}{m-i} \binom{M-k}{i}. \tag{56}$$

We now calculate the left eigenvectors. From the relation $\sum_k v^r_{ki} v^l_{km} = \delta_{im}$, we can proceed as follows:

$$\begin{aligned} z^m &= \sum_{i=0}^{M} z^i \delta_{im} \\ &= \sum_{i=0}^{M} z^i \sum_{k=0}^{M} v^r_{ki} v^l_{km} \\ &= \sum_{k=0}^{M} v^l_{km} \sum_{i=0}^{M} v^r_{ki} z^i \\ &= \sum_{k=0}^{M} v^l_{km} v^r_{k0} (1+z)^{M-k}(1-z)^k, \end{aligned} \tag{57}$$

from which the following relation can be obtained:

$$\frac{z^m}{(1+z)^M} = \sum_{k=0}^{M} v^l_{km} v^r_{k0} \left(\frac{1-z}{1+z}\right)^k. \tag{58}$$

Let $\xi = (1-z)/(1+z)$, Eq. (58) can be rewritten as

$$\sum_{k=0}^{M} v^l_{km} v^r_{k0} \xi^k = 2^{-M}(1-\xi)^m (1+\xi)^{M-m}. \tag{59}$$

By comparing the coefficients of the terms ξ^k, the components of the left eigenvectors can be derived as

$$v^l_{km} = \frac{2^{-M}}{v^r_{k0}} \sum_{i=k-m}^{M-m} (-1)^{k-i} \binom{m}{k-i} \binom{M-m}{i}. \tag{60}$$

Inserting the expressions obtained in Eqs. (56) and (60) to Eq. (49), we can calculate $[e^{-i\pi\tilde{G}/2}]_{1(M-1)}$ as follows:

$$\begin{aligned} [e^{-i\pi\tilde{G}/2}]_{1(M-1)} &= i \sum_{k=0}^{M} (-1)^{\bar{M}+k+1} \left[|v^r_k\rangle\langle v^l_k| \right]_{1(M-1)} \\ &= i \sum_{k=0}^{M} (-1)^{\bar{M}+k+1} v^r_{k1} v^l_{k(M-1)} \\ &= 2^{-M}(-1)^{\bar{M}+1} i \left[M + \sum_{k=0}^{M-1} \frac{(M-2k)^2}{M-k} \binom{M-1}{k} \right] \\ &= (-1)^{\bar{M}+1} i. \end{aligned} \tag{61}$$

Here, we use the following facts to obtain the last line:

$$\sum_{k=0}^{M-1}\binom{M}{k}=2^M-1, \tag{62}$$

$$\sum_{k=0}^{M-1}k\binom{M}{k}=M2^{M-1}-M, \tag{63}$$

$$\sum_{k=0}^{M-1}k(k-1)\binom{M}{k}=M(M-1)2^{M-2}-M(M-1). \tag{64}$$

Therefore, $|[e^{-i\pi\tilde{G}/2}]_{1(M-1)}|=1$, which completes the proof. □

Lemma 5. *For any $M\geq 3$ and $1\leq k\leq M$, the state transformation $|k,M-k\rangle\leftrightarrow|k-1,M-k+1\rangle$ can be accomplished within time $t=\pi/[2JM\sqrt{k(M-k+1)}]$ in the $U\to+\infty$ limit by the following Hamiltonian:*

$$\begin{aligned}H=&J(\hat{b}_1^\dagger\hat{b}_2+\hat{b}_2^\dagger\hat{b}_1)+J\left[\hat{b}_2^\dagger(\hat{n}_2+\hat{n}_1)\hat{b}_1+\hat{b}_1^\dagger(\hat{n}_2+\hat{n}_1)\hat{b}_2\right]\\&-U\hat{n}_2^2+U(2M-2k+1)\hat{n}_2.\end{aligned} \tag{65}$$

Proof. Following the same approach in Lemma 4, we can derive the following differential equation:

$$\dot{\boldsymbol{a}}_t=-iG\boldsymbol{a}_t, \tag{66}$$

where the matrix $G=[g_{mn}]_{0\leq m,n\leq M}$ is given by

$$g_{mn}=\begin{cases}JM\sqrt{m(M-m+1)} & \text{if } n=m-1,\\ JM\sqrt{(m+1)(M-m)} & \text{if } n=m+1,\\ U(M-m)(M+m-2k+1) & \text{if } n=m,\\ 0 & \text{otherwise.}\end{cases} \tag{67}$$

The real symmetric matrix G has distinct real eigenvalues $\lambda_0,\lambda_1,\ldots,\lambda_M$, and their corresponding eigenvectors are denoted by $\{|v_m^r\rangle\}$. In the $U\to+\infty$ limit, we can easily show that

$$\begin{aligned}\lim_{U\to+\infty}[\lambda_{k-1}-U(M-k+1)(M-k)]&=JM\sqrt{k(M-k+1)},\\ \lim_{U\to+\infty}[\lambda_k-U(M-k)(M-k+1)]&=-JM\sqrt{k(M-k+1)},\\ \lim_{U\to+\infty}[\lambda_m-U(M-m)(M+m-2k+1)]&=0\ \forall m\neq k-1,k,\\ \lim_{U\to+\infty}|v_{k-1}^r\rangle&=[\underbrace{0,\ldots,0}_{k-1},1/\sqrt{2},1/\sqrt{2},0,\ldots,0]^\top,\\ \lim_{U\to+\infty}|v_k^r\rangle&=[\underbrace{0,\ldots,0}_{k-1},1/\sqrt{2},-1/\sqrt{2},0,\ldots,0]^\top,\\ \lim_{U\to+\infty}|v_m^r\rangle&=[\delta_{mi}]^\top\ \forall m\neq k-1,k.\end{aligned}$$

Defining $P=(|v_0^r\rangle,\ldots,|v_M^r\rangle)$ and noting that $P^\top=P$ and $P^2=\mathbb{1}$, we obtain $G=P\,\mathrm{diag}\,(\lambda_0,\ldots,\lambda_M)P$. Consequently, the matrix e^{-iGt} can be calculated in the $U\to+\infty$ limit as $e^{-iGt}=P\,\mathrm{diag}\left(e^{-i\lambda_0 t},\ldots,e^{-i\lambda_M t}\right)P$. We need only show that $|[e^{-iGt}]_{k(k-1)}|=|[e^{-iGt}]_{(k-1)k}|=1$ for $t=\pi/[2JM\sqrt{k(M-k+1)}]$. Indeed, we can prove as follows:

$$\begin{aligned}|[e^{-iGt}]_{(k-1)k}|&=\left|e^{-i\lambda_{k-1}t}-e^{-i\lambda_k t}\right|/2\\&=\left|1-e^{i(\lambda_{k-1}-\lambda_k)t}\right|/2\\&=\left|1-e^{2iJM\sqrt{k(M-k+1)}t}\right|/2\\&=\left|1-e^{i\pi}\right|/2\\&=1.\end{aligned} \tag{68}$$

This completes the proof. □

Lemma 6. *For any $M \geq 3$, the state transformation $|M-1,1\rangle \leftrightarrow |1,M-1\rangle$ can be accomplished within time $t = \pi/(2J)$ by the following Hamiltonian:*

$$H = J(\hat{b}_1^\dagger \hat{b}_2 + \hat{b}_2^\dagger \hat{b}_1). \tag{69}$$

Proof. The proof is analogous to that of Lemma 4. □

Lemma 7. *For any $M \geq 3$ and $1 \leq k \leq M$, the state transformation $|k, M-k\rangle \leftrightarrow |k-1, M-k+1\rangle$ can be accomplished within time $t = \pi/[2J\sqrt{k(M-k+1)}]$ in the $U \to +\infty$ limit by the following Hamiltonian:*

$$H = J(\hat{b}_1^\dagger \hat{b}_2 + \hat{b}_2^\dagger \hat{b}_1) - U\hat{n}_2^2 + U(2M-2k+1)\hat{n}_2. \tag{70}$$

Proof. The proof is analogous to that of Lemma 5. □

Lemma 8. *The Wasserstein distance $\mathcal{W}(\boldsymbol{x},\boldsymbol{y}) \coloneqq \min_{\pi\in\mathcal{C}(\boldsymbol{x},\boldsymbol{y})} \sum_{i,j} \pi_{ij} c_{ij}$ can be expressed in the following form:*

$$\mathcal{W}(\boldsymbol{x},\boldsymbol{y}) = \max_{\boldsymbol{\phi}} \boldsymbol{\phi}^\top (\boldsymbol{x}-\boldsymbol{y}), \tag{71}$$

where the maximum is over all vectors $\boldsymbol{\phi} = [\phi_1,\ldots,\phi_N]^\top$ satisfying $|\phi_i - \phi_j| \leq c_{ij}\ \forall i,j$.

Proof. The constraint $\pi \in \mathcal{C}(\boldsymbol{x},\boldsymbol{y})$ on π in the definition of $\mathcal{W}(\boldsymbol{x},\boldsymbol{y})$ can be relaxed to $\pi \in \mathbb{R}_{\geq 0}^{N\times N}$ by adding a term $R(\pi)$ to the variational expression such that $R(\pi) = 0$ for $\pi \in \mathcal{C}(\boldsymbol{x},\boldsymbol{y})$ and $R(\pi) = +\infty$ otherwise. Although there are infinitely many ways to define $R(\pi)$, we conveniently choose

$$R(\pi) = \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N} \left[\boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y} - \sum_{i,j} (g_i + f_j)\pi_{ij} \right]. \tag{72}$$

Consequently, the Wasserstein distance can be expressed as

$$\begin{aligned}
\mathcal{W}(\boldsymbol{x},\boldsymbol{y}) &= \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \left[\sum_{i,j} \pi_{ij} c_{ij} + R(\pi) \right] \\
&= \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N} \left[\sum_{i,j} \pi_{ij}(c_{ij} - g_i - f_j) + \boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y} \right] \\
&= \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N} L(\pi,\boldsymbol{f},\boldsymbol{g}).
\end{aligned} \tag{73}$$

Here, we define $L(\pi,\boldsymbol{f},\boldsymbol{g}) \coloneqq \sum_{i,j} \pi_{ij}(c_{ij} - g_i - f_j) + \boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y}$. Since $L(\pi,\boldsymbol{f},\boldsymbol{g})$ is a linear function in its arguments, the max-min inequality is also an equality, according to von Neumann's minimax theorem. Therefore, we can proceed further from Eq. (73) as

$$\begin{aligned}
\mathcal{W}(\boldsymbol{x},\boldsymbol{y}) &= \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N} L(\pi,\boldsymbol{f},\boldsymbol{g}) \\
&= \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N} \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} L(\pi,\boldsymbol{f},\boldsymbol{g}) \\
&= \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N} \left[\boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y} + \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \sum_{i,j} \pi_{ij}(c_{ij} - g_i - f_j) \right].
\end{aligned} \tag{74}$$

If there exist i and j such that $c_{ij} - g_i - f_j < 0$, then $\min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \sum_{i,j} \pi_{ij}(c_{ij} - g_i - f_j) = -\infty$. Therefore, we need only consider $\boldsymbol{f},\boldsymbol{g} \in \mathbb{R}^N$ such that $c_{ij} - g_i - f_j \geq 0$, or equivalently $f_j + g_i \leq c_{ij}\ \forall i,j$. Therefore, Eq. (74) can be expressed as

$$\mathcal{W}(\boldsymbol{x},\boldsymbol{y}) = \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N : f_j + g_i \leq c_{ij}} \left[\boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y} + \min_{\pi\in\mathbb{R}_{\geq 0}^{N\times N}} \sum_{i,j} \pi_{ij}(c_{ij} - g_i - f_j) \right]. \tag{75}$$

Now, noticing that $\min_{\pi \in \mathbb{R}^{N\times N}_{\geq 0}} \sum_{i,j} \pi_{ij}(c_{ij} - g_i - f_j) = 0$ because $c_{ij} - g_i - f_j \geq 0$, we immediately obtain the following equality from Eq. (75):

$$\mathcal{W}(\boldsymbol{x}, \boldsymbol{y}) = \max_{\boldsymbol{f},\boldsymbol{g}\in\mathbb{R}^N : f_j + g_i \leq c_{ij}} (\boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y}). \tag{76}$$

For any fixed $\boldsymbol{g}$, the optimal $\tilde{\boldsymbol{f}}$ that maximizes $(\boldsymbol{f}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y})$ under the constraints is uniquely determined as

$$\tilde{f}_i = \min_j (c_{ij} - g_j) \ \forall i. \tag{77}$$

Since $c_{ij} \leq c_{kj} + c_{ik} \ \forall k$, we obtain

$$\tilde{f}_i \leq \min_j (c_{kj} - g_j + c_{ik}) = \tilde{f}_k + c_{ik}. \tag{78}$$

This implies $|\tilde{f}_i - \tilde{f}_k| \leq c_{ik}$ as $c_{ik} = c_{ki}$, which also means that the vector $\tilde{\boldsymbol{f}}$ satisfies the 1-Lipschitz condition. Moreover, since $c_{ii} = 0$, we also have

$$\tilde{f}_i \leq -g_i \text{ or } g_i \leq -\tilde{f}_i. \tag{79}$$

Combining these inequalities, we can upper bound the Wasserstein distance as

$$\begin{aligned}
\mathcal{W}(\boldsymbol{x}, \boldsymbol{y}) &= \max_{\boldsymbol{g}\in\mathbb{R}^N} (\tilde{\boldsymbol{f}}^\top \boldsymbol{x} + \boldsymbol{g}^\top \boldsymbol{y}) \\
&\leq \max_{\boldsymbol{g}\in\mathbb{R}^N} (\tilde{\boldsymbol{f}}^\top \boldsymbol{x} - \tilde{\boldsymbol{f}}^\top \boldsymbol{y}) \\
&\leq \max_{\boldsymbol{\phi}\in\mathbb{R}^N : |\phi_i - \phi_j| \leq c_{ij}} \boldsymbol{\phi}^\top (\boldsymbol{x} - \boldsymbol{y}).
\end{aligned} \tag{80}$$

Let $\tilde{\pi}$ be the optimal coupling that attains the Wasserstein distance, that is, $\mathcal{W}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i,j} \tilde{\pi}_{ij} c_{ij}$. Then, we can also prove that

$$\begin{aligned}
\max_{\boldsymbol{\phi}\in\mathbb{R}^N : |\phi_i - \phi_j| \leq c_{ij}} \boldsymbol{\phi}^\top (\boldsymbol{x} - \boldsymbol{y}) &= \max_{\boldsymbol{\phi}\in\mathbb{R}^N : |\phi_i - \phi_j| \leq c_{ij}} \sum_{i,j} (\phi_i \tilde{\pi}_{ji} - \phi_i \tilde{\pi}_{ij}), \\
&= \max_{\boldsymbol{\phi}\in\mathbb{R}^N : |\phi_i - \phi_j| \leq c_{ij}} \sum_{i,j} \tilde{\pi}_{ij} (\phi_j - \phi_i), \\
&\leq \sum_{i,j} \tilde{\pi}_{ij} c_{ij} \\
&= \mathcal{W}(\boldsymbol{x}, \boldsymbol{y}).
\end{aligned} \tag{81}$$

From Eqs. (80) and (81), we immediately obtain the desired result (71). □