Mean-Field Dynamics of the Bose–Hubbard Model in High Dimension

Shahnaz Farhat*, Denis Périce†, Sören Petrat‡

School of Science, Constructor University Bremen,
Campus Ring 1, 28759 Bremen, Germany

January 10, 2025

Abstract

The Bose–Hubbard model effectively describes bosons on a lattice with on-site interactions and nearest-neighbour hopping, serving as a foundational framework for understanding strong particle interactions and the superfluid to Mott insulator transition. This paper aims to rigorously establish the validity of a mean-field approximation for the dynamics of quantum systems in high dimension, using the Bose–Hubbard model on a square lattice as a case study. We prove a trace norm estimate between the one-lattice-site reduced density of the Schrödinger dynamics and the mean-field dynamics in the limit of large dimension. Here, the mean-field approximation is in the hopping amplitude and not in the interaction, leading to a very rich and non-trivial mean-field equation. This mean-field equation does not only describe the condensate, as is the case when the mean-field description comes from a large particle number limit averaging out the interaction, but it allows for a phase transition to a Mott insulator since it contains the full non-trivial interaction. Our work is a rigorous justification of a simple case of the highly successful dynamical mean-field theory (DMFT) for bosons, which somewhat surprisingly yields many qualitatively correct results in three dimensions.

1 Introduction

One of the big aspirations of mathematical physics is to advance our rigorous understanding of phase transitions. Within this research area lots of recent attention has been paid to the phenomenon of Bose–Einstein–Condensation (BEC), a phase of matter of cold Bose gases that has been predicted in 1924 by Bose [6] and Einstein [14, 15]. Since then, BEC has been studied extensively by theoretical physicists, and at least since the 1980's also by mathematical physicists with more rigorous methods. After the first experimental realizations in the labs of Cornell/Wieman [2] and Ketterle [12] in 1995 the study of BEC has received a new wave of attention throughout experimental, theoretical, and mathematical physics. As a recent highlight in mathematical physics, let us mention the rigorous derivation of the Lee–Huang–Yang formula by Fournais and Solovej [18, 19]. The motivation of our work comes from different perspectives:

1. Lots of recent effort has been put into understanding BEC at zero temperature, e.g., in the Gross–Pitaevskii limit [35] and thermodynamic limit [18, 19]. This yields insight into the behavior of the condensate and its excitations, e.g., a rigorous proof of Bogoliubov theory in the Gross–Pitaevskii regime [5], but is very far away from understanding the (thermodynamic) phase transition to BEC.

*Email: `sfarhat@constructor.university`
†Email: `dperice@constructor.university`, corresponding author.
‡Email: `spetrat@constructor.university`

2. A very useful and simple method for studying many-body systems is the mean-field approximation. For bosonic cold atoms, one scales down the interaction potential with the inverse of the particle number N [16] (or density [13]), thus considering weak interaction. Then the inverse particle number can be regarded as a small parameter, and the interaction can be effectively replaced by its mean-field. In this mean-field limit, many physical effects can be rigorously established regarding the dynamics and the low-energy properties. In particular, one can prove the validity of Bogoliubov theory [28, 29], and perturbative expansions beyond Bogoliubov [7, 8]. However, these types of mean-field models do not describe phase transitions.

3. For Bose gases in the continuum, one would ultimately like to prove a thermodynamic phase transition. However, Bose gases on the lattice offer a different possibility for a phase transition, namely a quantum phase transition between a BEC and a localized state usually called a Mott insulator; see, e.g., [37] for a review. There are only few mathematical rigorous works on this topic, e.g., [1].

Our work addresses these points in the following way. 1: We study a limit that may describe a phase transition. 2: Our limit is a mean-field limit, not for large particle number but for large dimension. We hope and indeed show that some of the methods of large N mean-field limits are still relevant for this case. Since in our model the averaging is done over the hopping terms, and the interaction is treated non-perturbatively, our mean-field model is strongly interacting. 3: Our microscopic model is the Bose–Hubbard model, which is a lattice model that has been successfully used to describe the BEC-Mott transition.

Our main result is convergence of the reduced one-lattice-site density matrix of the many-body Schrödinger dynamics to the mean-field dynamics, with an error bound that goes to zero as the dimension d goes to infinity. Other parameters such as the density and the coupling remain fixed. Thus, we rigorously justify the validity of the mean-field approximation for a quantum system in large dimension. We choose the Bose–Hubbard model to illustrate this statement both for its remarkable usefulness in physics and for the technical simplifications it offers as a lattice model. The Bose–Hubbard model is a popular model used to describe bosons on a lattice with on-site interactions, allowing hopping between nearest-neighbor lattice sites. It is well-known for capturing strong interactions between particles [4] and providing one of the simplest descriptions of the Mott transition to date, see [17] and later [26], see also [22, 23].

A common technique to study models such as Bose–Hubbard is Dynamical Mean-field Theory (DMFT). This theory is well-known for its description of the Mott insulator/superfluid phase transition [25, 17]. It is usually formulated via a self-consistency condition for a Green's function. DMFT is typically justified in the physics literature by stating that mean-field theories become exact in the limit of infinite dimensions [31]; see also [30] for fermions. A remarkable fact is that DMFT tends to provide accurate results already in three dimensions [20]. In the literature, the effective equation we are deriving here is often called the mean-field model of Fisher et al. [10], referring to [17]. Our equation can be considered a simple case of DMFT. A more involved mean-field type equation is obtained in [10] by scaling different parts of the hopping term in different ways. In the paper [17] the authors consider the Bose–Hubbard model on a complete graph (the hopping term is of equal strength between all vertices). In comparison, our model has only nearest-neighbor hopping. Rigorous justifications of the effective thermodynamic behavior of the Bose–Hubbard model on a complete graph were obtained in [9]. As mentioned above, in the mathematical literature, mean-field limits are typically considered as many-particle limits for the Bose–Hubbard model [32] or, more generally, for continuous models where the Hartree equation is obtained as effective dynamics (see, e.g., [3, 21] for reviews). This approach requires dividing the interaction term by the number of particles to ensure that the kinetic energy and the interaction energy of the ground state remain of the same order.

Our goal is to provide a rigorous justification, in the $d \to \infty$ limit, that DMFT is a good approximation of the Schrödinger equation in the context of the Bose–Hubbard model. It is interesting to note that, in the large d limit, the roles of the kinetic energy and the interaction between particles are inverted compared to the usual mean-field limit $N \to \infty$. The terms we aim to average in our regime are the hopping terms between nearest-neighbor sites which come from the kinetic energy. Since we only consider on-site interactions, the interaction between particles acts as a one-site operator and therefore does not contribute to correlations between two different lattice sites. For our setting, the basic idea behind the mean-field approximation is that the coordination number of the lattice (the number of nearest neighbors) increases with the dimension. This means we have a mean-field picture locally around every site, which allows us to control the correlations between sites. Note that our main estimates are for the reduced one-lattice-site density matrix, and not for the one-particle reduced density matrix that is usually used to describe convergence in the large N limit.

2 The Model and Main Results

2.1 Model

We consider the d-dimensional square lattice with periodic boundary conditions $\Lambda \coloneqq (\mathbb{Z}/L\mathbb{Z})^d$ of volume $|\Lambda| \coloneqq L^d$, with $L \in \mathbb{N}, L \geqslant 2$. We write $x \sim y$ if $x, y \in \Lambda$ are nearest neighbors. The one-site Hilbert space is $\ell^2(\mathbb{C})$ and its canonical Hilbert basis is denoted by $(|n\rangle)_{n\in\mathbb{N}}$. We define the standard creation and annihilation operators a^*, a satisfying the CCR $[a, a^*] = 1$, $[a, a] = 0 = [a^*, a^*]$; explicitly,

$$
\begin{aligned}
&a\,|0\rangle \coloneqq 0, a\,|n\rangle \coloneqq \sqrt{n}\,|n-1\rangle \;\forall n \in \mathbb{N}^*,\\
&a^*\,|n\rangle \coloneqq \sqrt{n+1}\,|n+1\rangle \;\forall n \in \mathbb{N}.
\end{aligned}
$$

The number operator is given as $\mathcal{N} \coloneqq a^*a$. To simplify our notation in some later proofs, we introduce an order on Λ such that $\forall x \in \Lambda$

$$
\#\{y \in \Lambda | y > x \text{ and } x \sim y\} = \#\{y \in \Lambda | x > y \text{ and } x \sim y\} = d.
$$

For example, the lexicographic order does the job. The Fock space is

$$
\mathcal{F} \coloneqq \ell^2(\mathbb{C})^{\otimes|\Lambda|} \cong \mathcal{F}_+\left(L^2(\Lambda, \mathbb{C})\right) \coloneqq \bigoplus_{k\in\mathbb{N}} L^2(\Lambda, \mathbb{C})^{\otimes_+ k},
$$

where $\otimes_+$ denotes the symmetric tensor product. Given a one-site operator A and $x \in \Lambda$, we define

$$
A_x \coloneqq \left(\bigotimes_{y<x} \mathbb{1}\right) A \left(\bigotimes_{y>x} \mathbb{1}\right).
$$

In the following we define $\langle x, y\rangle$ to mean that $x, y \in \Lambda, x \sim y$ and $x < y$. The kinetic energy is given by the negative second quantized discrete Laplacian

$$
-\,\mathrm{d}\Gamma(\Delta_d) \coloneqq \sum_{\langle x,y\rangle} (a_x^* - a_y^*)(a_x - a_y) = -\sum_{\langle x,y\rangle} \left(a_x^* a_y + a_y^* a_x\right) + 2d \sum_{x\in\Lambda} \mathcal{N}_x.
$$

Furthermore, we denote by $\mathcal{N}_{\mathcal{F}}$ the number operator on Fock space, i.e.,

$$
\mathcal{N}_{\mathcal{F}} \coloneqq d\Gamma(1) := \sum_{x\in\Lambda} \mathcal{N}_x.
$$

Given hopping amplitude $J \in \mathbb{R}$, chemical potential $\mu \in \mathbb{R}$, and coupling constant $U \in \mathbb{R}$, we define the Bose–Hubbard Hamiltonian

$$
\begin{aligned}
H_d &\coloneqq -\frac{J}{2d}\,\mathrm{d}\Gamma(\Delta_d) - \mu\mathcal{N}_{\mathcal{F}} + \frac{U}{2}\sum_{x\in\Lambda} \mathcal{N}_x(\mathcal{N}_x - 1)\\
&= -\frac{J}{2d}\sum_{\langle x,y\rangle} \left(a_x^* a_y + a_y^* a_x\right) + (J-\mu)\sum_{x\in\Lambda}\mathcal{N}_x + \frac{U}{2}\sum_{x\in\Lambda}\mathcal{N}_x(\mathcal{N}_x - 1). \qquad (1)
\end{aligned}
$$

Here, we have scaled down the hopping term with the inverse of the dimension d. The time-dependent Schrödinger equation for $\Psi_d \in \mathcal{F}$ is

$$i\frac{d}{dt}\Psi_d(t) = H_d\Psi_d(t). \tag{2}$$

The idea of the mean-field approximation is the lattice site product state ansatz $\Psi_d \approx \prod_{x\in\Lambda}\varphi_x$ where $\varphi \in \ell^2(\mathbb{C})$ is a one-lattice-site wave function. Such a state $\prod_{x\in\Lambda}\varphi_x$ is sometimes called Gutzwiller product state [24, 34]. Our main results state that if $\Psi_d(0) \approx \prod_{x\in\Lambda}\varphi_x(0)$, then also $\Psi_d(t) \approx \prod_{x\in\Lambda}\varphi_x(t)$ for all times $t > 0$, where $\approx$ is meant in an appropriate reduced sense. We can guess the right mean-field equation for $\varphi(t)$ by writing $a_x = \langle\varphi(t), a\varphi(t)\rangle + \widetilde{a}_x(t)$, and neglect in the hopping term of (1) all terms that are quadratic in $\widetilde{a}(t)$. Then the corresponding mean-field equation is

$$i\partial_t\varphi(t) = h^{\varphi(t)}\varphi(t), \tag{3}$$

where the nonlinear mean-field operator is

$$h^{\varphi} := -J\big(\alpha_\varphi a^* + \overline{\alpha_\varphi}a - |\alpha_\varphi|^2\big) + (J-\mu)\mathcal{N} + \frac{U}{2}\mathcal{N}(\mathcal{N}-1), \tag{4}$$

with order parameter $\alpha_\varphi := \langle\varphi, a\varphi\rangle$. Roughly speaking, $\alpha_\varphi = 0$ indicates a Mott insulator state, whereas $\alpha_\varphi \neq 0$ indicates a superfluid state. The well-posedness of (3) is discussed in Section 3.

The approximation $\Psi_d \approx \prod_{x\in\Lambda}\varphi_x$ is not expected to hold in $\mathcal{F}$, but in the sense of reduced lattice site density matrices. Given Ψ_d, let us first define the corresponding positive trace one operator $\gamma_d \in \mathcal{L}^1(\mathcal{F})$, which satisfies the von Neumann equation

$$i\partial_t\gamma_d(t) = [H_d, \gamma_d(t)]\,. \tag{5}$$

We define its first reduced one-lattice-site density matrix as

$$\gamma_{\mathrm{d}}^{(1)} := \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\mathrm{Tr}_{\Lambda\setminus\{x\}}\left(\gamma_d\right).$$

The operator $\gamma_{\mathrm{d}}^{(1)} : \ell^2(\mathbb{C}) \to \ell^2(\mathbb{C})$ is not to be confused with the reduced one-particle density matrix $\gamma^{(1)}_{\mathrm{particle}} : L^2(\Lambda) \to L^2(\Lambda)$ defined via its integral kernel $\gamma^{(1)}_{\mathrm{particle}}(x,y) = \langle\Psi_d, a_y^*a_x\Psi_d\rangle$ with $x, y \in \Lambda$. Given $\varphi \in \ell^2(\mathbb{C})$, let us also introduce the corresponding orthogonal projections

$$p = p_\varphi := |\varphi\rangle\langle\varphi|, \quad \text{and} \quad q = q_\varphi := 1 - p_\varphi. \tag{6}$$

In our main results we prove convergence of $\gamma_d^{(1)}(t)$ to $p_\varphi(t)$.

2.2 Main Results

Our main results are estimates for the trace-norm difference of $\gamma_d^{(1)}(t)$ and $p_\varphi(t)$, which we denote by $\|\gamma_d^{(1)}(t) - p_\varphi(t)\|_{\mathcal{L}^1}$. We prove two similar estimates. Theorem 1 proves an error bound that holds for any value of the parameters J, μ, U of the Bose–Hubbard model (1). The convergence rate is slightly worse than $\frac{1}{\sqrt{d}}$. However, we need to assume stronger conditions on the initial data, and the bound contains a double exponential growth in time. On the other hand, Theorem 2 holds only for repulsive interaction, i.e., $U > 0$. However, it holds for a larger class of initial data, and the error bound only grows exponentially in time. Our first main result is the following.

Theorem 1. *Let γ_d be the solution to (5) with initial data $\gamma_d(0) \in \mathcal{L}^1(\mathcal{F})$, and let φ be the solution to (3) with initial data $\varphi(0) \in \ell^2(\mathbb{C})$ such that $\|\varphi\|_{\ell^2} = 1$. Let p_φ be defined as in (6). We assume that there exist $a, c > 0$ such that*

$$\forall n \in \mathbb{N}, \mathrm{Tr}\left(p_\varphi(0)\mathbb{1}_{\mathcal{N}=n}\right) \leqslant ce^{-\frac{n}{a}} \text{ and } \mathrm{Tr}\left(\gamma_d^{(1)}(0)\mathbb{1}_{\mathcal{N}=n}\right) \leqslant ce^{-\frac{n}{a}}. \tag{7}$$

Then for all $t \in \mathbb{R}_+$ we have

$$\left\|\gamma_d^{(1)}(t) - p_\varphi(t)\right\|_{\mathcal{L}^1} \leqslant \sqrt{2}\left(\left\|\gamma_{\mathrm{d}}^{(1)}(0) - p_\varphi(0)\right\|_{\mathcal{L}^1} + \frac{C_2 e^{C_1 t} + \mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}}}{d\left(C_4 + 2\left(\sqrt{2(a+e)}e^{\frac{C_1}{2}t}+1\right)\sqrt{\ln(d+1)}\right)}\right)^{\frac{1}{2}}$$
$$e^{JC_3\left(C_4+2\left(\sqrt{2(a+e)}e^{\frac{C_1}{2}t}+1\right)\sqrt{\ln(d+1)}\right)t}, \tag{8}$$

with the following constants independent of d and t:

$$\begin{aligned}
C_1 &:= 2eJ\max\left(\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right), 1\right),\\
C_2 &:= 4\left(c(1+a)+e^{-1}\right)\left(2+4(a+e)\right),\\
C_3 &:= \left(\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)+1\right)^{\frac{1}{2}},\\
C_4 &:= 4\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}}+2.
\end{aligned}$$

Note that the d-dependent terms on the right-hand side of (8) are small when $d \to \infty$, since

$$\frac{1}{\left(d\sqrt{\ln(d+1)}\right)^{\frac{1}{2}}} e^{C\sqrt{\ln(d+1)}t} = e^{C\sqrt{\ln(d+1)}t - \frac{1}{2}\ln(d) - \frac{1}{4}\ln(\ln(d+1))} \underset{d\to\infty}{\longrightarrow} 0$$

for any $C, t > 0$. Our second main result is the following.

Theorem 2. *Let γ_d be the solution to (5) with initial data $\gamma_d(0) \in \mathcal{L}^1(\mathcal{F})$, and let φ be the solution to (3) with initial data $\varphi(0) \in \ell^2(\mathbb{C})$ such that $\|\varphi\|_{\ell^2} = 1$. Let p_φ, q_φ be defined as in (6). We assume that there is $C > 0$ such that*

$$\mathrm{Tr}(p_\varphi(0)\mathcal{N}^4) \leqslant C,$$

and that $U > 0$. Then there exists $C(J,\mu,U) > 0$ such that for all $t \in \mathbb{R}_+$ we have

$$\left\|\gamma_d^{(1)}(t) - p_\varphi(t)\right\|_{\mathcal{L}^1} \leqslant C(J,\mu,U)e^{C(J,\mu,U)(1+t^7)}\left(\mathrm{Tr}\left(\gamma_d^{(1)}(0)\left(q_\varphi(0)\mathcal{N}^2 q_\varphi(0) + q_\varphi(0)\right)\right) + \frac{1}{d}\right)^{1/2}. \tag{9}$$

Note that for initial Gutzwiller product states $\Psi_d(0) = \prod_{x\in\Lambda}\varphi(0)$, we have

$$\mathrm{Tr}\left(\gamma_d^{(1)}(0)\left(q_\varphi(0)\mathcal{N}^2 q_\varphi(0) + q_\varphi(0)\right)\right) = 0. \tag{10}$$

More generally, assuming $\mathrm{Tr}\left(\gamma_d^{(1)}(0)q_\varphi(0)\mathcal{N}^2 q_\varphi(0)\right) \leqslant d^{-1}$ and $\mathrm{Tr}\left(\gamma_d^{(1)}(0)q_\varphi(0)\right) \leqslant d^{-1}$, the estimate (9) becomes

$$\left\|\gamma_d^{(1)}(t) - p_\varphi(t)\right\|_{\mathcal{L}^1} \leqslant C(J,\mu,U)e^{C(J,\mu,U)(1+t^7)}\frac{1}{\sqrt{d}}.$$

From the perspective of the law of large numbers, this is the expected optimal convergence rate. (However, this convergence rate obviously does not explain why our approximation is so successful even for $d = 3$.) Note also that the bound (9) can be written in more detail as

$$\left\|\gamma_d^{(1)}(t) - p_\varphi(t)\right\|_{\mathcal{L}^1}$$
$$\leqslant \left(\frac{1}{d}\frac{1}{U} + \widetilde{C}(J,\mu,U)\left(1+\frac{1}{U^2}\right)e^{\widetilde{C}(J,\mu,U)\sum_{j=1}^{7}t^j}\left(\mathrm{Tr}\left(\gamma_d^{(1)}(0)\left(q_\varphi(0)\mathcal{N}^2 q_\varphi(0) + q_\varphi(0)\right)\right) + \frac{1}{d}\right)\right)^{1/2}, \tag{11}$$

where $\widetilde{C}(J,\mu,U)$ depends polynomially on the parameters J, μ, U of the Bose–Hubbard model. The divergence for small U comes from our use of an energy estimate, as outlined below (around Equation (13)).

The trace norm convergence of Theorems 1 and 2 in particular implies convergence of the order parameter α, meaning

$$\alpha_{\text{micro}}(t) := \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d(t), a_x\Psi_d(t)\rangle \to \alpha_{\varphi(t)} := \langle\varphi(t), a\varphi(t)\rangle \text{ as } d\to\infty. \tag{12}$$

Note that for initial data $\Psi_d(0)$ with a fixed particle number the left-hand side of (12) is zero (since H_d is particle number conserving), but in general our initial data live on Fock space where the particle number is not fixed.

Both theorems are proven using a Gronwall estimate for $\operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)} q_\varphi)$. This quantity heuristically counts the average number of lattice sites that do not follow the product state ansatz. It is inspired by the corresponding quantity for the weak coupling limit introduced by Pickl [33]. The main technical challenge is then caused by the unboundedness of the creation and annihilation operators in the hopping term, a bit analogous to the technical problems that arise when considering the weak coupling limit with singular interactions; see, e.g., [27]. More concretely, we need to bound $\operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)q_\varphi)$ in terms of $\operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)} q_\varphi)$ or terms that go to zero as $d\to\infty$. This we do in two different ways, leading to the two main theorems. For Theorem 1, we introduce a new moment method. For this, we first separate

$$\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)q_\varphi\right) = \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)\mathbb{1}_{\mathcal{N}<M}q_\varphi\right) + \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)\mathbb{1}_{\mathcal{N}\geqslant M}q_\varphi\right).$$

Then the first term can simply be bounded by M, whereas we use moment estimates to bound the second term by $e^{MC(t)}d^{-1}$. Then it turns out to be possible to optimize in M to close the Gronwall argument. For Theorem 2, we proceed using an energy estimate inspired by [28] (which deals with proving Bogoliubov theory for the dynamics of the weakly interacting Bose gas). The idea is to write the Hamiltonian (1) as

$$H_d = \sum_{x\in\Lambda} h_x^{\varphi(t)} + \widetilde{H}(t).$$

Here, $\widetilde{H}(t)$ describes the excitations around our product state ansatz. A similar splitting was used in [28], where, after a unitary transformation, $\widetilde{H}(t)$ converges to a Bogoliubov Hamiltonian. The energy of the excitations should now be defined as

$$E^{\text{exc}}(t) := \left\langle\Psi_d, \left(\widetilde{H}(t) + \sum_{x\in\Lambda} q_x h_x^{\varphi(t)} q_x\right)\Psi_d\right\rangle, \tag{13}$$

where the first term corresponds to the kinetic energy, and the second term to the mean-field energy of the excitations (including the interaction energy). The energy $E^{\text{exc}}(t)$ is not conserved since excitations from the product state can be created and annihilated. However, it can be bounded using a Gronwall argument. The crucial point here is that the interaction term only as

$$\sum_{x\in\Lambda} q_x\mathcal{N}_x(\mathcal{N}_x-1)q_x$$

and hence the Gronwall argument does not produce higher powers than $q_x\mathcal{N}_x^2 q_x$. This ultimately allows us to control terms involving $q_x\mathcal{N}_x^2 q_x$ or $q_x\mathcal{N}_x q_x$, in particular $\operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)q_\varphi)$.

The rest of the paper is organized as follows. We first prove global well-posedness of the mean-field equation (3) in Section 3. Then, we discuss some preliminary estimates in Section 4. In particular, we prove properties of a two-lattice-site reduced density matrix, preliminary bounds for the mean-field and Bose–Hubbard energies, and conservation laws and propagation estimates for the mean-field equation and the Bose–Hubbard dynamics. Furthermore, we compute $\partial_t \operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)}(t)q_\varphi(t))$, and prove bounds on all the terms in this time derivative except for $\operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)q_\varphi)$. Then, Theorem 1 is proven in Section 5, and Theorem 2 is proven in Section 6, each using a different method to control $\operatorname{Tr}(\gamma_{\mathrm{d}}^{(1)} q_\varphi(\mathcal{N}+1)q_\varphi)$.

3 Global Well-posedness of the Mean-Field Dynamics

The goal of this section is to prove well-posedness of the mean field dynamics. The mean-field dynamics can be written as

$$\begin{cases} i\partial_t \varphi = h^{\alpha_\varphi}\varphi = A\varphi + F(\varphi), \\ \varphi(0) = \varphi_0, \end{cases} \tag{14}$$

where the linear operator A and the nonlinear operator F are defined as

$$A := (J-\mu)\mathcal{N} + \frac{U}{2}\mathcal{N}(\mathcal{N}-1), \tag{15}$$

$$F(\varphi) := -J\left(\alpha_\varphi a^* + \overline{\alpha_\varphi} a - |\alpha_\varphi|^2\right)\varphi. \tag{16}$$

When examining the semilinear equation (14) above, we cannot directly apply fixed-point arguments to study global well-posedness because both A and F are unbounded operators, and the nonlinear operator F is not Lipschitz continuous. Therefore, a different approach is required. Our strategy is to approximate the nonlinear term so that it becomes Lipschitz continuous. This allows us to establish the existence of a unique solution to the approximated problem by standard methods. Then, we show that the obtained solution converges to the solution of the untruncated mean-field equation.

3.1 Approximating the Mean-field Dynamics

Let $M > 0$ and consider the truncated creation and annihilation operators

$$a_M := a\mathbb{1}_{\mathcal{N}\leqslant M}, \qquad a_M^* := \mathbb{1}_{\mathcal{N}\leqslant M}a^* \tag{17}$$

and

$$\alpha_M := \langle \varphi_M, a_M\varphi_M\rangle, \tag{18}$$

where φ_M is the solution to the approximated problem

$$\begin{cases} i\partial_t \varphi_M = h_M^{\alpha_M}\varphi_M = A\varphi_M + F_M(\varphi_M), \\ \varphi_M(0) = \varphi_0 \in \mathcal{D}(\mathcal{N}^2), \end{cases} \tag{19}$$

where we have introduced the approximated nonlinear operator F_M as

$$F_M(\varphi_M) := -J\left(\alpha_M a_M^* + \overline{\alpha_M} a_M - |\alpha_M|^2\right)\varphi_M. \tag{20}$$

The solution to (19) solves the weak form of the preceding nonlinear equation (19) which is usually known as Duhamel formula

$$\begin{cases} \varphi_M(t) = \tilde{\varphi}_0(t) - i\int_0^t e^{-i(t-s)A}F_M(\varphi_M(s))\,ds, \\ \tilde{\varphi}_0(t) := e^{-itA}\varphi_0, \qquad \varphi_0 \in \mathcal{D}(\mathcal{N}^2). \end{cases} \tag{21}$$

Remark 3. *Note the following:*

1. *For the weak formulation* (21) *of the approximated nonlinear equation* (19)*, the existence of a unique local solution can be established using fixed-point arguments for a broader class of initial data, specifically $\varphi_0 \in \ell^2(\mathbb{C})$, which leads to the existence of a unique local solution $\varphi_M \in C([0,T],\ell^2(\mathbb{C}))$. This follows from the fact that F_M is a nonlinear bounded operator satisfying $F_M(\varphi_M) \in C([0,T],\ell^2(\mathbb{C}))$. However, to extend the solution to global times, we rely on the conservation laws, which require the initial data $\varphi_0 \in \mathcal{D}(\mathcal{N}^2)$, as A remains an unbounded operator.*

2. *Since $F_M(\varphi_M) \in C([0,T], \ell^2(\mathbb{C}))$, to ensure the equivalence between* (19) *and* (21)*, it is enough to restrict our analysis to initial data $\varphi_0 \in \mathcal{D}(\mathcal{N}^2)$. For further details, see [11, Lemma 4.1.1, Proposition 4.1.6 and Corollary 4.1.8].*

3. *One could in addition truncate the unbounded linear term A. On the one hand, this approach ensures equivalence between* (19) *and* (21) *for all initial data $\varphi_0 \in \ell^2(\mathbb{C})$. On the other hand, it allows us to obtain a unique global strong solution $\varphi_M \in C(\mathbb{R}, \ell^2(\mathbb{C}))$. However, ensuring convergence to the solution of the mean-field dynamics becomes more complicated.*

3.2 Properties of the Approximate Solution

In this subsection, we state some conservation laws for the approximated problem.

Lemma 4. *Assume that φ_M is a solution to* (19) *with $\|\varphi_0\|_{\ell^2} = 1$. Then, the following holds:*

(i) $\|\varphi_M(t)\|_{\ell^2} = \|\varphi_0\|_{\ell^2} = 1$.

(ii) $\langle \varphi_M(t), \mathcal{N}\varphi_M(t)\rangle = \langle \varphi_0, \mathcal{N}\varphi_0\rangle$.

(iii) $\langle \varphi_M(t), h_M^{\alpha_M}\varphi_M(t)\rangle = \langle \varphi_0, h_M^{\alpha_M}\varphi_0\rangle$.

(iv) $|\alpha_M| \leqslant \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2}$.

(v) Assume that $\varphi_0 \in \mathcal{D}(\mathcal{N}^k)$. Then, there exists a constant $C > 0$ such that

$\langle \varphi_M(t), \mathcal{N}^k\varphi_M(t)\rangle \leqslant \sum_{j=0}^{2k-2} \left(CJk\|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2}\right)^j \left\langle \varphi_0, (\mathcal{N}+j)^{k-\frac{j}{2}}\varphi_0 \right\rangle \frac{t^j}{j!}$.

Proof. Statement (i) is true by definition of the truncation. For (ii), note that

$$[h_M^{\alpha_M}, \mathcal{N}] = -J(-\alpha_M a_M^* + \overline{\alpha_M} a_M). \tag{22}$$

This gives

$$\frac{d}{dt}\langle \varphi_M, \mathcal{N}\varphi_M\rangle = i\langle \varphi_M, [h_M^{\alpha_M}, \mathcal{N}]\varphi_M\rangle = -iJ\langle \varphi_M, (-\alpha_M a_M^* + \overline{\alpha_M} a_M)\varphi_M\rangle = 0.$$

For (iii), we have

$$\begin{aligned}\frac{d}{dt}\langle \varphi_M, h_M^{\alpha_M}\varphi_M\rangle &= i\langle \varphi_M, [h_M^{\alpha_M}, h_M^{\alpha_M}]\varphi_M\rangle + \langle \varphi_M, \partial_t h_M^{\alpha_M}\varphi_M\rangle \\ &= -J\left\langle \varphi_M, \left(-\partial_t\alpha_M a_M^* + \partial_t\overline{\alpha_M} a_M - \partial_t\alpha_M\overline{\alpha_M} - \partial_t\overline{\alpha_M}\alpha_M\right)\varphi_M\right\rangle = 0.\end{aligned}$$

Statement (iv) follows directly from Cauchy–Schwarz and (i)-(ii). For (v), note that

$$\begin{aligned}\frac{d}{dt}\langle \varphi_M, \mathcal{N}^k\varphi_M\rangle &= -iJ\alpha_M\langle \varphi_M, [a_M^*, \mathcal{N}^k]\varphi_M\rangle - iJ\overline{\alpha_M}\langle \varphi_M, [a_M, \mathcal{N}^k]\varphi_M\rangle \\ &= -iJ\alpha_M\langle \varphi_M, \mathbb{1}_{\mathcal{N}\leqslant M}[a^*, \mathcal{N}^k]\varphi_M\rangle - iJ\overline{\alpha_M}\langle \varphi_M, [a, \mathcal{N}^k]\mathbb{1}_{\mathcal{N}\leqslant M}\varphi_M\rangle \\ &= 2J\mathrm{Im}\left(\alpha_M\langle \varphi_M, \mathbb{1}_{\mathcal{N}\leqslant M}[a^*, \mathcal{N}^k]\varphi_M\rangle\right) \\ &\leqslant 2k|J||\alpha_M||\langle \varphi_M, \mathbb{1}_{\mathcal{N}\leqslant M}a^*\mathcal{N}^{k-1}\varphi_M\rangle| \\ &\leqslant 2k|J|\|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2}\langle \varphi_M, (\mathcal{N}+1)^{k-\frac{1}{2}}\varphi_M\rangle.\end{aligned} \tag{23}$$

Iterating this $(2k-2)$ times leads to

$$\langle \varphi_M, \mathcal{N}^k\varphi_M\rangle \leqslant \sum_{j=0}^{2k-2} \left(C|J|k\|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2}\right)^j \left\langle \varphi_0, (\mathcal{N}+j)^{k-\frac{j}{2}}\varphi_0 \right\rangle \frac{t^j}{j!}. \tag{24}$$

□

3.3 Global Well-posedness of the Approximated Problem

For the approximated problem, proving global well-posedness is straightforward due to the use of standard techniques such as fixed-point arguments, particularly because the nonlinearity in this case is Lipschitz. We have the following results.

Lemma 5. *For any fixed $M > 0$, we have the following statements:*

(i) There exists a unique global strong solution $\varphi_M(\cdot) \in \mathcal{C}(\mathbb{R}, \mathcal{D}(\mathcal{N}^2))$ of the Duhamel formula (21).

(ii) There exists a unique global strong solution $\varphi_M(\cdot) \in \mathcal{C}(\mathbb{R}, \mathcal{D}(\mathcal{N}^2)) \cap \mathcal{C}^1(\mathbb{R}, \ell^2(\mathbb{C}))$ of the approximated problem (19).

Proof. Let $X = \mathcal{C}([0,T], \mathcal{D}(\mathcal{N}^2))$ denote the space of continuous functions from $[0,T]$ to $\mathcal{D}(\mathcal{N}^2)$, equipped with the norm

$$|||\varphi||| := \sup_{t\in[0,T]} \|\varphi(t)\|_{\mathcal{D}(\mathcal{N}^2)}, \qquad \|\varphi(t)\|^2_{\mathcal{D}(\mathcal{N}^2)} = \|\varphi(t)\|^2_{\ell^2} + \|\mathcal{N}^2\varphi(t)\|^2_{\ell^2}.$$

Note that $\big(\mathcal{D}(\mathcal{N}^2), \|\cdot\|_{\mathcal{D}(\mathcal{N}^2)}\big)$ is a Banach space. For a fixed $M > 0$, we define the map $\Gamma_M : X \to X$ by

$$\Gamma_M(\varphi)(t) := \tilde{\varphi}_0(t) - i\int_0^t e^{-i(t-s)A} F_M(\varphi(s))ds,$$

with $\tilde{\varphi}_0(t) := e^{-itA}\varphi_0$ and where A and F_M are defined in (15)and (20). We can check that, for any $T > 0$, the map Γ_M is Lipschitz-continuous. More precisely, we claim that for all $\varphi_1, \varphi_2 \in X$,

$$|||\Gamma_M(\varphi_1) - \Gamma_M(\varphi_2)||| \leqslant C(M,J,T)|||\varphi_1 - \varphi_2|||,$$

where $C(M,J,T) > 0$ is defined as

$$C(M,J,T) := MT|J|(6c^2 + 6c^2M^2 + 10c^4),$$

and $c > 0$ as

$$c := \max_{i=1,2} \sup_{[0,T]} \|\varphi_i(t)\|_{\mathcal{D}(\mathcal{N}^2)} < \infty.$$

To prove the claim, we need first to establish some useful estimates. To this end, we denote $\alpha_{M,i}(s) := \langle \varphi_i(s), a_M\varphi_i(s)\rangle$. We have, for $k \geqslant 0$ and for $i = 1, 2$,

$$\|\mathcal{N}^k a^\sharp_M \varphi_i(s)\|_{\ell^2} \leqslant M^{k+\frac{1}{2}}\|\varphi_i(s)\|_{\ell^2} \leqslant M^{k+\frac{1}{2}}c, \quad \sharp \in \{\ \ , *\}, \tag{25}$$

$$|\alpha_{M,i}(s)| \leqslant \sqrt{M}\|\varphi_i(s)\|^2_{\ell^2} \leqslant \sqrt{M}c^2, \tag{26}$$

$$|\alpha_{M,1}(s) - \alpha_{M,2}(s)| \leqslant 2\sqrt{M}c\|\varphi_1(s) - \varphi_2(s)\|_{\ell^2}, \tag{27}$$

$$\big||\alpha_{M,1}(s)|^2 - |\alpha_{M,2}(s)|^2\big| \leqslant 4Mc^3\|\varphi_1(s) - \varphi_2(s)\|_{\ell^2}. \tag{28}$$

We have then for all $t \in [0,T]$,

$$\begin{aligned}
\|\Gamma_M(\varphi_1)(t) - \Gamma_M(\varphi_2)(t)\|_{\ell^2} &= \left\|\int_0^t e^{-i(t-s)A}\left(F_M(\varphi_1(s)) - F_M(\varphi_2(s))\right) ds\right\|_{\ell^2} \\
&\leqslant |J|\int_0^t \Bigg(\left\|\overline{\alpha_{M,1}(s)}a_M\varphi_1(s) - \overline{\alpha_{M,2}(s)}a_M\varphi_2(s)\right\|_{\ell^2} \\
&\qquad + \|\alpha_{M,1}(s)a^*_M\varphi_1(s) - \alpha_{M,2}(s)a^*_M\varphi_2(s)\|_{\ell^2} \\
&\qquad + \left\||\alpha_{M,1}(s)|^2\varphi_1(s) - |\alpha_{M,2}(s)|^2\varphi_2(s)\right\|_{\ell^2}\Bigg) ds.
\end{aligned}$$

We begin by considering the first term. Using (25) with $k = 0$, along with (26) and (27), we get

$$\begin{aligned}
&\left\|\overline{\alpha_{M,1}(s)}a_M\varphi_1(s) - \overline{\alpha_{M,2}(s)}a_M\varphi_2(s)\right\|_{\ell^2} \\
&\quad \leqslant |\alpha_{M,1}(s) - \alpha_{M,2}(s)| \|a_M\varphi_1(s)\|_{\ell^2} + |\alpha_{M,2}(s)| \|a_M(\varphi_1(s) - \varphi_2(s))\|_{\ell^2} \\
&\quad \leqslant 3Mc^2 \|\varphi_1(s) - \varphi_2(s)\|_{\ell^2}.
\end{aligned}$$

Similarly, for the second term, applying (25) with $k = 0$, along with (26) and (27), we obtain

$$\|\alpha_{M,1}(s)a_M^*\varphi_1(s) - \alpha_{M,2}(s)a_M^*\varphi_2(s)\|_{\ell^2} \leqslant 3Mc^2\|\varphi_1(s) - \varphi_2(s)\|_{\ell^2}.$$

Finally, for the last term, by using (25) with $k = 0$, along with (26) and (28), we obtain

$$\begin{aligned}
&\left\||\alpha_{M,1}(s)|^2\varphi_1(s) - |\alpha_{M,2}(s)|^2\varphi_2(s)\right\|_{\ell^2} \\
&\quad \leqslant \left||\alpha_{M,1}(s)|^2 - |\alpha_{M,2}(s)|^2\right| \|\varphi_1(s)\|_{\ell^2} + |\alpha_{M,2}(s)|^2 \|\varphi_1(s) - \varphi_2(s)\|_{\ell^2} \\
&\quad \leqslant 5Mc^4 \|\varphi_1(s) - \varphi_2(s)\|_{\ell^2}.
\end{aligned}$$

To summarize, we obtain

$$\|\Gamma_M(\varphi_1)(t) - \Gamma_M(\varphi_2)(t)\|_{\ell^2} \leqslant MT|J|(6c^2 + 5c^4)|||\varphi_1 - \varphi_2|||. \tag{29}$$

More generally, for any $k \geqslant 0$, we have

$$\|\mathcal{N}^k\left(\Gamma_M(\varphi_1)(t) - \Gamma_M(\varphi_2)(t)\right)\|_{\ell^2} \leqslant MT|J|(6M^kc^2 + 5c^4) \sup_{[0,T]} \|\varphi_1(s) - \varphi_2(s)\|_{\mathcal{D}(\mathcal{N}^k)}. \tag{30}$$

Specifically, for the other component of the norm, we have

$$\|\mathcal{N}^2\left(\Gamma_M(\varphi_1)(t) - \Gamma_M(\varphi_2)(t)\right)\|_{\ell^2} \leqslant MT|J|(6M^2c^2 + 5c^4)|||\varphi_1 - \varphi_2|||. \tag{31}$$

By combining the two estimates (29) and (31) above, we obtain

$$|||\Gamma_M(\varphi_1) - \Gamma_M(\varphi_2)||| \leqslant MT|J|(6c^2 + 6c^2M^2 + 10c^4)|||\varphi_1 - \varphi_2|||.$$

Considering the semilinear equation of the form (19), and noting that our nonlinearity is Lipschitz continuous (or can be made a contraction by choosing T sufficiently small), we can approach the problem in two ways. First, we can apply the local well-posedness results from [11], specifically [11, Lemma 4.3.2 and Proposition 4.3.3], to obtain a unique local solution. Then, we can extend this solution globally using [11, Theorem 4.3.4] by employing conservation laws, including the norm and the moment bounds (i)-(v) in Lemma 4. On the other hand, we can establish global well-posedness by directly applying the Banach fixed point arguments. To this end, we consider the closed ball on the Banach space X defined by

$$B_X(\tilde{\varphi}_0, R] := \{\varphi \in X; \quad |||\varphi - \tilde{\varphi}_0||| \leqslant R\}.$$

Then, we check that for $T > 0$ small enough and for $R > 0$ large enough, the map Γ_M satisfies the condition of the Banach fixed point theorem, namely

- Γ_M maps $B_X(\tilde{\varphi}_0, R]$ into itself,
- $\Gamma_M : (X, |||\cdot|||) \to (X, |||\cdot|||)$ is a contraction map,

guaranteeing the existence of a fixed point ($\varphi_M = \Gamma_M(\varphi_M) \in X$). The solution can then be extended globally using the same conservation laws employed for the first approach.

Remark 6. *In the above theorem, we can also apply the fixed point theorem in the Banach space* $X = C([0,T], \ell^2(\mathbb{C}))$*, which guarantees the existence of a unique local solution* $\varphi_M \in C([0,T], \ell^2(\mathbb{C}))$*. Subsequently, we can globalize the solution (which is equivalent to obtaining an estimate of the norm* $\|\varphi_M(t)\|_{\ell^2}$ *on* $[0,T]$*) for the set of initial data* $\varphi_0 \in \mathcal{D}(\mathcal{N}^2)$*, ensuring that*

$$\|\varphi_M(t)\|_{\ell^2} = \|\varphi_0\|_{\ell^2}.$$

The above estimate implies that

$$\lim_{t\uparrow T} \|\varphi_M(t)\|_{\ell^2} = \|\varphi_0\|_{\ell^2} < \infty,$$

which guarantees that the solution does not blow up in finite time and thus $T = \infty$*.*

□

3.4 Convergence

By (i) and (v) from Lemma 4, we have that $(\varphi_M)_{M\in\mathbb{N}}$ and $(\mathcal{N}^k\varphi_M)_{M\in\mathbb{N}}$ are bounded sequences in the Hilbert space $\ell^2(\mathbb{C})$. Then there exist a convergent subsequence still denoted by $(\varphi_M)_{M\in\mathbb{N}}$ such that

- φ_M converges weakly to φ and the limit is unique,
- $\mathcal{N}^k\varphi_M$ converges weakly to $\mathcal{N}^k\varphi$ for all $k \in \mathbb{R}^+$.

As a consequence of this convergence, we have

$$\|\varphi\|_{\ell^2} \leqslant \liminf_{M\to\infty} \|\varphi_M\|_{\ell^2},$$
$$\|\mathcal{N}^k\varphi\|_{\ell^2} \leqslant \liminf_{M\to\infty} \|\mathcal{N}^k\varphi_M\|_{\ell^2}.$$

This in fact implies strong convergence.

Lemma 7. *Let* $(\varphi_M)_M$ *be a sequence of solutions to* (19) *with* $\|\varphi_0\|_{\ell^2} = 1$ *and* φ *its associated weak limit. Then we have for all* $k \geqslant 0$*,*

$$\|\mathcal{N}^k(\varphi_M(t) - \varphi(t))\|_{\ell^2} \underset{M\to\infty}{\longrightarrow} 0. \tag{32}$$

Proof. This follows from the weak* convergence in the Banach space $\mathcal{L}^1(\ell^2(\mathbb{C})) = (\mathcal{K}(\ell^2(\mathbb{C})))^*$, where $\mathcal{L}^1(\ell^2(\mathbb{C}))$ and $\mathcal{K}(\ell^2(\mathbb{C}))$ denote the space of trace-class and compact operators, respectively. Let $p_{\varphi_M} = |\varphi_M\rangle\langle\varphi_M|$ be the projection onto the state φ_M, so taht in particular $p_{\varphi_M}^2 = p_{\varphi_M}$. We also have

$$\mathrm{Tr}(p_{\varphi_M}) = \|\varphi_M\|_{\ell^2}^2 = 1, \qquad \mathrm{Tr}(\mathcal{N}^k p_{\varphi_M}) = \langle\varphi_M, \mathcal{N}^k\varphi_M\rangle < \infty.$$

The second bound is a consequence of part (v) of Lemma 4. This ensures the existence of a subsequence, still denoted by $(p_{\varphi_M})_M$, such that

$$p_{\varphi_M} \overset{*}{\rightharpoonup} \nu \quad \text{as } M \to \infty \quad \text{weakly * in } \mathcal{L}^1(\ell^2(\mathbb{C})),$$

$$\mathcal{N}^k p_{\varphi_M} \overset{*}{\rightharpoonup} \mathcal{N}^k\nu \quad \text{as } M \to \infty \quad \text{weakly * in } \mathcal{L}^1(\ell^2(\mathbb{C})).$$

For any compact operator $B \in \mathcal{K}(\ell^2(\mathbb{C}))$, this implies

$$\mathrm{Tr}(p_{\varphi_M}B) \to \mathrm{Tr}(\nu B) \quad \text{and} \quad \mathrm{Tr}(\mathcal{N}^k p_{\varphi_M}B) \to \mathrm{Tr}(\mathcal{N}^k\nu B) \quad \text{as } M \to \infty.$$

For $k \geqslant 0$, this leads to

$$\mathrm{Tr}(\mathcal{N}^k p_{\varphi_M}) = \mathrm{Tr}(\mathcal{N}^{-1}\mathcal{N}^{k+1}p_{\varphi_M}) \to \mathrm{Tr}(\mathcal{N}^{-1}\mathcal{N}^{k+1}\nu) = \mathrm{Tr}(\mathcal{N}^k\nu). \tag{33}$$

Specifically, for $k=0$ we have

$$\mathrm{Tr}(p_{\varphi_M}) = \|\varphi_M\|_{\ell^2}^2 \to \mathrm{Tr}(\nu),$$

which implies $\mathrm{Tr}(\nu)=1$. Now, using results from [36], we obtain strong convergence for all $k \geqslant 0$,

$$\|\mathcal{N}^k p_{\varphi_M} - \mathcal{N}^k \nu\|_{\mathcal{L}^1} = \mathrm{Tr}\left(\left|\mathcal{N}^k p_{\varphi_M} - \mathcal{N}^k \nu\,\right|\right) \to 0 \quad \text{as } M \to \infty.$$

Since both p_{φ_M} and ν are bounded in norm, we conclude that $p_{\varphi_M}^2 = p_{\varphi_M}$ converges strongly to $\nu = \nu^2$. Therefore, ν is a projection, and $\nu = P_\chi = |\chi\rangle\langle\chi|$. Additionally, we have

$$\mathrm{Tr}(p_{\varphi_M}\nu) = |\langle\varphi_M, \chi\rangle|^2 \to \mathrm{Tr}(\nu^2) = \mathrm{Tr}(\nu) = 1.$$

On the other hand, we also know

$$\langle\varphi_M, \chi\rangle \to \langle\varphi, \chi\rangle,$$

which implies $|\langle\varphi, \chi\rangle|^2 = 1$, leading to $P_\varphi = P_\chi$. Therefore, by exploiting (33), we have

$$\mathcal{N}^k\varphi_M \rightharpoonup \mathcal{N}^k\varphi \quad \text{and} \quad \|\mathcal{N}^k\varphi_M\|_{\ell^2} \to \|\mathcal{N}^k\varphi\|_{\ell^2}.$$

Since ℓ^2 is a Hilbert space, these results imply strong convergence

$$\|\mathcal{N}^k(\varphi_M - \varphi)\|_{\ell^2} \to 0 \quad \text{as } M \to \infty.$$

□

Next, we show that the limit indeed satisfies the corresponding mean-field equation.

Lemma 8. *Let $(\varphi_M)_M$ be a sequence of solutions to* (19) *and assume $\|\varphi_0\|_{\ell^2} = 1$. Then the limit φ satisfies the Duhamel version of the mean-field dynamics* (14),

$$\varphi(t) = \tilde{\varphi}_0(t) - i\int_0^t e^{-i(t-s)A} F(\varphi(s))\, ds, \tag{34}$$

with $\tilde{\varphi}_0(t) := e^{-itA}\varphi_0$ and where F is defined in (16).

Proof. Let us start by establishing some useful estimates. Since both φ_M and $\mathcal{N}^k\varphi_M$ converge weakly to φ and $\mathcal{N}^k\varphi$, respectively, we have

$$\|\varphi\|_{\ell^2} \leqslant \liminf_{M\to\infty} \|\varphi_M\|_{\ell^2} = \|\varphi_0\|_{\ell^2} = 1, \tag{35}$$

$$\|\mathcal{N}^{1/2}\varphi\|_{\ell^2} \leqslant \liminf_{M\to\infty} \|\mathcal{N}^{1/2}\varphi_M\|_{\ell^2} = \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2}, \tag{36}$$

$$|\alpha_\varphi| = |\langle\varphi, a\varphi\rangle| \leqslant \|\varphi\|_{\ell^2}\|a\varphi\|_{\ell^2} \leqslant \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2}. \tag{37}$$

Moreover, we also have

$$\begin{aligned}
|\alpha_M - \alpha_\varphi| &= |\langle\varphi_M, a_M\varphi_M\rangle - \langle\varphi, a\varphi\rangle| \\
&\leqslant |\langle\varphi_M, (a_M - a)\varphi_M\rangle| + |\langle\varphi_M, a(\varphi_M - \varphi)\rangle| + |\langle(\varphi_M - \varphi), a\varphi\rangle| \\
&\leqslant \|\varphi_M\|_{\ell^2}\|a\mathbb{1}_{\mathcal{N}>M}\varphi_M\|_{\ell^2} + \|a^*\varphi_M\|_{\ell^2}\|\varphi_M - \varphi\|_{\ell^2} + \|a\varphi\|_{\ell^2}\|\varphi_M - \varphi\|_{\ell^2} \\
&\leqslant 2\|(\mathcal{N}+1)^{1/2}\varphi_0\|_{\ell^2}\|\varphi_M - \varphi\|_{\ell^2} + \mathcal{E}_1(M),
\end{aligned}$$

where we have introduced $\mathcal{E}_1(M)$ as

$$\mathcal{E}_1(M) := \|a\mathbb{1}_{\mathcal{N}>M}\varphi_M\|_{\ell^2}. \tag{38}$$

Then we estimate

$$\left\|\varphi(t) - \tilde{\varphi}_0(t) - i\int_0^t e^{-i(t-s)A} F(\varphi(s))\, ds\right\|_{\ell^2} \tag{39}$$

$$
\begin{aligned}
&\leqslant \|\varphi_M(t) - \varphi(t)\|_{\ell^2} && (40)\\
&\quad + \left\| \varphi_M(t) - \tilde{\varphi}_0(t) - i\int_0^t e^{-i(t-s)A} F_M(\varphi_M(s))\,ds \right\|_{\ell^2} && (41)\\
&\quad + \left\| \int_0^t e^{-i(t-s)A} F_M(\varphi_M(s))\,ds - \int_0^t e^{-i(t-s)A} F(\varphi(s))\,ds \right\|_{\ell^2} && (42)
\end{aligned}
$$

The first term (40) converges to zero by Lemma 7. The second term (41) is zero because φ_M is a solution to the approximated problem (21). It remains to estimate the difference between the nonlinear parts,

$$
\begin{aligned}
(42) \leqslant |J| \int_0^t \Big(&\left\| \overline{\alpha_M(s)} a_M \varphi_M(s) - \overline{\alpha_\varphi(s)} a \varphi(s) \right\|_{\ell^2} && (43a)\\
&+ \left\| \alpha_M(s) a_M^* \varphi_M(s) - \alpha_\varphi(s) a^* \varphi(s) \right\|_{\ell^2} && (43b)\\
&+ \left\| |\alpha_M(s)|^2 \varphi_M(s) - |\alpha_\varphi(s)|^2 \varphi(s) \right\|_{\ell^2} \Big) ds. && (43c)
\end{aligned}
$$

For (43a), we find

$$
\begin{aligned}
(43\text{a}) &\leqslant |\alpha_M - \alpha| \|a_M \varphi_M\|_{\ell^2} + |\alpha| \|(a_M - a)\varphi_M\|_{\ell^2} + |\alpha| \|a(\varphi_M - \varphi)\|_{\ell^2}\\
&\leqslant \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2} \left(2\|(\mathcal{N}+1)^{1/2}\varphi_0\|_{\ell^2} \|\varphi_M - \varphi\|_{\ell^2} + \|\mathcal{N}^{1/2}(\varphi_M - \varphi)\|_{\ell^2} + 2\mathcal{E}_1(M) \right),
\end{aligned}
$$

with $\mathcal{E}_1(M)$ from (38). The first and the second term go to zero as $M \to \infty$. It remains to check $\mathcal{E}_1(M) \to 0$ as $M \to \infty$. Indeed by (24), we have for some $C > 0$ that

$$
\begin{aligned}
\mathcal{E}_1(M) &= \|a \mathbb{1}_{\mathcal{N}>M} \varphi_M\|_{\ell^2}\\
&\leqslant \|\mathbb{1}_{\mathcal{N}+1>M} (\mathcal{N}+1)^{-1/2}\|_{\mathcal{L}} \|(\mathcal{N}+1)^{1/2} a \varphi_M\|_{\ell^2}\\
&\leqslant \frac{1}{\sqrt{M}} \underbrace{\left(\sum_{j=0}^{2} \left(2C|J| \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2} \right)^j \left\langle \varphi_0, (\mathcal{N}+j)^{2-\frac{j}{2}} \varphi_0 \right\rangle \frac{t^j}{j!} \right)^{1/2}}_{<\infty \text{ since } \varphi_0 \in \mathcal{D}(\mathcal{N}^2)}.
\end{aligned}
$$

So, the term $\mathcal{E}_1(M)$ goes to zero as $M \to \infty$. Similarly, for (43b) we find that

$$
\begin{aligned}
(43\text{b}) &\leqslant |\alpha_M - \alpha| \|a_M^* \varphi_M\|_{\ell^2} + |\alpha| \|(a_M^* - a^*)\varphi_M\|_{\ell^2} + |\alpha| \|a^*(\varphi_M - \varphi)\|_{\ell^2}\\
&\leqslant 2\|(\mathcal{N}+1)^{1/2}\varphi_0\|_{\ell^2}^2 \|\varphi_M - \varphi\|_{\ell^2} + \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2} \|(\mathcal{N}+1)^{1/2}(\varphi_M - \varphi)\|_{\ell^2}\\
&\quad + \|(\mathcal{N}+1)^{1/2}\varphi_0\|_{\ell^2} \mathcal{E}_1(M) + \|\mathcal{N}^{1/2}\varphi_0\|_{\ell^2} \mathcal{E}_2(M),
\end{aligned}
$$

where we have introduced

$$
\mathcal{E}_2(M) := \|\mathbb{1}_{\mathcal{N}>M} a^* \varphi_M\|_{\ell^2}. \tag{44}
$$

By the same arguments as for (43a), the term (43b) goes to zero as $M \to \infty$. It remains to estimate the last term (43c). We have

$$
\begin{aligned}
(43\text{c}) &= \left\| |\alpha_M|^2 \varphi_M - |\alpha_\varphi|^2 \varphi \right\|_{\ell^2}\\
&\leqslant \left| |\alpha_M|^2 - |\alpha_\varphi|^2 \right| \|\varphi_M\|_{\ell^2} + |\alpha_\varphi|^2 \|\varphi_M - \varphi\|_{\ell^2}\\
&\leqslant |\alpha_M(\overline{\alpha_M} - \overline{\alpha_\varphi}) + \overline{\alpha_\varphi}(\alpha_M - \alpha_\varphi)| + |\alpha_\varphi|^2 \|\varphi_M - \varphi\|_{\ell^2}\\
&\leqslant (|\alpha_M| + |\alpha_\varphi|) |\alpha_M - \alpha_\varphi| + |\alpha_\varphi|^2 \|\varphi_M - \varphi\|_{\ell^2},
\end{aligned}
$$

which also converges to zero by the same arguments as for the previous two terms. □

4 Preliminaries

4.1 Reduced Densities Matrices

Given a density matrix $\gamma_d \in \mathcal{L}^1(\mathcal{F})$ we define a two-lattice-site reduced density matrix

$$\gamma_d^{(2)} := \frac{1}{d|\Lambda|}\sum_{\langle x,y\rangle} \mathrm{Tr}_{\Lambda\setminus\{x,y\}}(\gamma_d) = \frac{1}{2d|\Lambda|}\sum_{\substack{x,y\in\Lambda\\ x\sim y}} \mathrm{Tr}_{\Lambda\setminus\{x,y\}}(\gamma_d).$$

Note that this two-lattice-site reduced density matrix is symmetrized over all interacting pairs of sites and not over all pairs of sites. The normalization factor $2d|\Lambda|$ is indeed the number of interacting pairs of sites. Note that $\gamma_d^{(2)}$ is symmetric, i.e.,

$$\forall A, B \in \mathcal{L}(\ell^2(\mathbb{C})), \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, A\otimes B\right) = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, B\otimes A\right),$$

and reduces to $\gamma_d^{(1)}$, i.e.,

$$\mathrm{Tr}_1\left(\gamma_{\mathrm{d}}^{(2)}\right) = \mathrm{Tr}_2\left(\gamma_{\mathrm{d}}^{(2)}\right) = \gamma_{\mathrm{d}}^{(1)}.$$

Moreover, if $C\in\mathcal{L}(\ell^2(\mathbb{C}))$ and $D\in\mathcal{L}(\ell^2(\mathbb{C})^{\otimes 2})$, then it follows directly from its definition that

$$\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\mathrm{Tr}(\gamma_d C_x) = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\, C\right),$$
$$\frac{1}{2d|\Lambda|}\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\mathrm{Tr}(\gamma_d D_{x,y}) = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, D\right).$$

Furthermore, the following standard results hold. If $A\in\mathcal{L}(\ell^2(\mathbb{C}))$ is self adjoint such that $A\geqslant 0$ or $\gamma_{\mathrm{d}}^{(1)}\, A\in\mathcal{L}^1(\ell^2(\mathbb{C}))$, then

$$\left[\sum_{x\in\Lambda} A_x, \gamma_d\right]^{(1)} = \left[A, \gamma_{\mathrm{d}}^{(1)}\right]. \tag{45}$$

If $B\in\mathcal{L}(\ell^2(\mathbb{C})^{\otimes 2})$ is self adjoint such that $B\geqslant 0$ or $\gamma_{\mathrm{d}}^{(2)}\, B\in\mathcal{L}^1(\ell^2(\mathbb{C})^{\otimes 2})$, then

$$\frac{1}{2d}\left[\sum_{\substack{x,y\in\Lambda\\ x\sim y}} B_{x,y}, \gamma_d\right]^{(1)} = \mathrm{Tr}_1\left(\left[B, \gamma_{\mathrm{d}}^{(2)}\right]\right) + \mathrm{Tr}_2\left(\left[B, \gamma_{\mathrm{d}}^{(2)}\right]\right). \tag{46}$$

4.2 Energy Bounds

With the definitions of one and two-lattice-site density matrices we can rewrite the energy per lattice site as

$$\frac{\mathrm{Tr}(\gamma_d H_d)}{|\Lambda|} = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\left((J-\mu)\mathcal{N} + \frac{U}{2}\mathcal{N}(\mathcal{N}-1)\right)\right) - J\,\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, a^*\otimes a\right). \tag{47}$$

Note that the mean-field energy can be written as

$$\langle\varphi, h^\varphi\varphi\rangle = J\left(\langle\varphi,\mathcal{N}\varphi\rangle - |\alpha_\varphi|^2\right) - \mu\langle\varphi,\mathcal{N}\varphi\rangle + \frac{U}{2}\langle\varphi,\mathcal{N}(\mathcal{N}-1)\varphi\rangle. \tag{48}$$

The following bounds allow us to control the Bose–Hubbard energy and the mean-field energy in terms of moments of the number operator.

Lemma 9. *Let $\gamma_d \in \mathcal{L}^1(\mathcal{F})$ and $\varphi \in \ell^2(\mathbb{C})$. Then there exists $C > 0$ such that, for $U = 0$,*

$$|\langle \varphi, h^\varphi \varphi \rangle| \leqslant C \langle \varphi, \mathcal{N} \varphi \rangle, \tag{49}$$

$$\frac{|\mathrm{Tr}(\gamma_d H_d)|}{|\Lambda|} \leqslant C\left(1 + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)} \mathcal{N}\right)\right), \tag{50}$$

and, for $U \neq 0$,

$$|\langle \varphi, h^\varphi \varphi \rangle| \leqslant C\left(1 + \langle \varphi, \mathcal{N}^2 \varphi \rangle\right), \tag{51}$$

$$\frac{|\mathrm{Tr}(\gamma_d H_d)|}{|\Lambda|} \leqslant C\left(1 + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)} \mathcal{N}^2\right)\right). \tag{52}$$

Proof. Using Cauchy–Schwarz's inequality we have

$$|\alpha_\varphi|^2 = |\langle \varphi, a\varphi \rangle|^2 \leqslant \|\varphi\|^2 \|a\varphi\|^2 = \langle a\varphi, a\varphi \rangle = \langle \varphi, a^* a \varphi \rangle = \langle \varphi, \mathcal{N} \varphi \rangle. \tag{53}$$

Recalling (48), this immediately yields (49) and (51). In order to obtain (50) and (52), we estimate the two-site term in (47) with Cauchy–Schwarz to obtain

$$\left|\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)} a^* \otimes a\right)\right| \leqslant \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)} \mathcal{N}\right)^{\frac{1}{2}} \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)} (\mathcal{N}+1)\right)^{\frac{1}{2}} \leqslant \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)} \mathcal{N}\right) + 1.$$

□

4.3 Conservation Laws

For both the Bose–Hubbard model (1) and the mean-field model (4) the total particle number and the total energy are conserved. Furthermore, one can control higher powers of the total particle number. Let us show this first for the mean-field equation. The total particle number is conserved since

$$\begin{aligned} i\partial_t \langle \varphi, \mathcal{N}\varphi \rangle = \langle \varphi, [\mathcal{N}, h^\varphi] \varphi \rangle &= -J\left(\alpha_\varphi \langle \varphi, [\mathcal{N}, a^*] \varphi \rangle + \overline{\alpha_\varphi} \langle \varphi, [\mathcal{N}, a] \varphi \rangle\right) \\ &= -J\left(\alpha_\varphi \langle \varphi, a^* \varphi \rangle - \overline{\alpha_\varphi} \langle \varphi, a\varphi \rangle\right) \\ &= -J\left(|\alpha_\varphi|^2 - |\alpha_\varphi|^2\right) \\ &= 0. \end{aligned} \tag{54}$$

The energy is conserved since

$$\begin{aligned} i\partial_t \langle \varphi, h^\varphi \varphi \rangle = \langle \varphi, \partial_t h^\varphi \varphi \rangle &= -J \langle \varphi, (\partial_t \alpha_\varphi a^* + \partial_t \overline{\alpha_\varphi} a - \overline{\alpha_\varphi} \partial_t \alpha_\varphi - \alpha_\varphi \partial_t \overline{\alpha_\varphi}) \varphi \rangle \\ &= -J\left(\overline{\alpha_\varphi} \partial_t \alpha_\varphi + \alpha_\varphi \partial_t \overline{\alpha_\varphi} - \overline{\alpha_\varphi} \partial_t \alpha_\varphi - \alpha_\varphi \partial_t \overline{\alpha_\varphi}\right) \\ &= 0. \end{aligned}$$

Moreover, we can prove two different bounds for controlling powers of the number operator, which we will use for our two main theorems.

Proposition 10. *Let φ solve* (3) *with $\varphi(0) \in \ell^2(\mathbb{C})$. Let $k \in \mathbb{N}/2, k \geqslant 1$ and $t \in \mathbb{R}_+$. Then*

$$\mathrm{Tr}\left(p_\varphi(t)\mathcal{N}^k\right) \leqslant \left(\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}^k\right) + e^{-1}k^k\right) e^{2eJk\mathrm{Tr}(p_\varphi(0)\mathcal{N})^{\frac{1}{2}} t}, \tag{55}$$

$$\mathrm{Tr}\left(p_\varphi(t)\mathcal{N}^k\right) \leqslant \sum_{l=0}^{2(k-1)} \binom{2k}{l} \left(J\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}} t\right)^l \mathrm{Tr}\left(p_\varphi(0)\left(\mathcal{N}+l\right)^{k-\frac{l}{2}}\right). \tag{56}$$

Proof. Let $n \in \mathbb{N}$. Recalling the mean-field dynamics (3), we find

$$i\partial_t \left\langle \varphi, (\mathcal{N}+n)^k \varphi \right\rangle = \left\langle \varphi, \left[(\mathcal{N}+n)^k, h^\varphi\right] \varphi \right\rangle$$

$$
\begin{aligned}
&= -J\left\langle\varphi,\left[(\mathcal{N}+n)^k,\alpha_\varphi a^*+\overline{\alpha_\varphi}a\right]\varphi\right\rangle\\
&= J\alpha_\varphi\left\langle\varphi,\left[a^*,(\mathcal{N}+n)^k\right]\varphi\right\rangle+J\overline{\alpha_\varphi}\left\langle\varphi,\left[a,(\mathcal{N}+n)^k\right]\varphi\right\rangle\\
&= 2iJ\ \mathrm{Im}\left[\overline{\alpha_\varphi}\left\langle\varphi,\left[a,(\mathcal{N}+n)^k\right]\varphi\right\rangle\right].
\end{aligned}\tag{57}
$$

Now, let $\mathcal{A}\in\mathcal{L}^1\left(\ell^2(\mathbb{C})\right)$ be the positive operator defined as

$$
\mathcal{A}^{k-1}:=(\mathcal{N}+n+1)^k-(\mathcal{N}+n)^k\leqslant k\,(\mathcal{N}+n+1)^{k-1}.\tag{58}
$$

Since $a\,(\mathcal{N}+n)^k=(\mathcal{N}+n+1)^k\,a$ we find

$$
\left\langle\varphi,\left[a,(\mathcal{N}+n)^k\right]\varphi\right\rangle=\left\langle\varphi,\mathcal{A}^{k-1}a\varphi\right\rangle=\left\langle\mathcal{A}^{\frac{k}{2}-\frac{1}{4}}\varphi,\mathcal{A}^{\frac{k}{2}-\frac{3}{4}}a\varphi\right\rangle
$$

so with Cauchy–Schwarz's inequality,

$$
\begin{aligned}
\left|\left\langle\varphi,\left[a,(\mathcal{N}+n)^k\right]\varphi\right\rangle\right| &\leqslant \left\langle\varphi,\mathcal{A}^{k-\frac{1}{2}}\varphi\right\rangle^{\frac{1}{2}}\left\langle\varphi,a^*\mathcal{A}^{k-\frac{3}{2}}a\varphi\right\rangle^{\frac{1}{2}}\\
&\leqslant k\left\langle\varphi,(\mathcal{N}+n+1)^{k-\frac{1}{2}}\varphi\right\rangle^{\frac{1}{2}}\left\langle\varphi,a^*\,(\mathcal{N}+n+1)^{k-\frac{3}{2}}\,a\varphi\right\rangle^{\frac{1}{2}}\\
&= k\left\langle\varphi,(\mathcal{N}+n+1)^{k-\frac{1}{2}}\varphi\right\rangle^{\frac{1}{2}}\left\langle\varphi,(\mathcal{N}+n)^{k-\frac{3}{2}}\,\mathcal{N}\varphi\right\rangle^{\frac{1}{2}}\\
&\leqslant k\left\langle\varphi,(\mathcal{N}+n+1)^{k-\frac{1}{2}}\varphi\right\rangle.
\end{aligned}\tag{59}
$$

Combining (57) with (59) and also (53) we conclude

$$
\left|\partial_t\left\langle\varphi,(\mathcal{N}+n)^k\varphi\right\rangle\right|\leqslant 2Jk\,\langle\varphi,\mathcal{N}\varphi\rangle^{\frac{1}{2}}\left\langle\varphi,(\mathcal{N}+n+1)^{k-\frac{1}{2}}\varphi\right\rangle.\tag{60}
$$

Proof of (56)**.** By induction on k, we prove that, $\left\langle\varphi(0),\mathcal{N}^k\varphi(0)\right\rangle<\infty$ implies that for all $n\in\mathbb{N}$,

$$
\left\langle\varphi(t),(\mathcal{N}+n)^k\varphi(t)\right\rangle\leqslant\sum_{l=0}^{2(k-1)}\binom{2k}{l}\left(J\,\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^{\frac{1}{2}}\,t\right)^l\left\langle\varphi(0),(\mathcal{N}+n+l)^{k-\frac{l}{2}}\varphi(0)\right\rangle.\tag{61}
$$

Then (56) follows for $n=0$. The inequality is indeed true for $k=1$ since $\langle\varphi,(\mathcal{N}+n)\,\varphi\rangle$ is conserved, see (54). For the induction step, we assume (61) holds for some k and that

$$
\left\langle\varphi(0),\mathcal{N}^{k+\frac{1}{2}}\varphi(0)\right\rangle<\infty,
$$

and we now prove (61) for $k+\frac{1}{2}$ instead of k. Using (60) with $k+\frac{1}{2}$ instead of k, and using the conservation of $\langle\varphi,\mathcal{N}\varphi\rangle$ we find

$$
\left|\partial_t\left\langle\varphi,(\mathcal{N}+n)^{k+\frac{1}{2}}\varphi\right\rangle\right|\leqslant J(2k+1)\,\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^{\frac{1}{2}}\left\langle\varphi,(\mathcal{N}+n+1)^k\varphi\right\rangle.
$$

Integrating over time and inserting (61) we conclude

$$
\begin{aligned}
&\left\langle\varphi(t),(\mathcal{N}+n)^{k+\frac{1}{2}}\varphi(t)\right\rangle\\
&\quad\leqslant\left\langle\varphi(0),(\mathcal{N}+n)^{k+\frac{1}{2}}\varphi(0)\right\rangle+J(2k+1)\,\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^{\frac{1}{2}}\int_0^t\left\langle\varphi(\tau),(\mathcal{N}+n+1)^k\varphi(\tau)\right\rangle d\tau\\
&\quad=\left\langle\varphi(0),(\mathcal{N}+n)^{k+\frac{1}{2}}\varphi(0)\right\rangle+J(2k+1)\,\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^{\frac{1}{2}}
\end{aligned}
$$

$$\sum_{l=0}^{2(k-1)} \binom{2k}{l} \left(J\left\langle\varphi(0),\mathcal{N}\varphi(0)\right\rangle^{\frac{1}{2}}\right)^{l} \left\langle\varphi(0),(\mathcal{N}+n+l+1)^{k-\frac{l}{2}}\varphi(0)\right\rangle \int_0^t \tau^l d\tau$$
$$= \left\langle\varphi(0),(\mathcal{N}+n)^{k+\frac{1}{2}}\varphi(0)\right\rangle$$
$$+ \sum_{l=0}^{2(k-1)} \binom{2k+1}{l+1} \left(J\left\langle\varphi(0),\mathcal{N}\varphi(0)\right\rangle^{\frac{1}{2}} t\right)^{l+1} \left\langle\varphi(0),(\mathcal{N}+n+l+1)^{k-\frac{l}{2}}\varphi(0)\right\rangle$$
$$= \left\langle\varphi(0),(\mathcal{N}+n)^{k+\frac{1}{2}}\varphi(0)\right\rangle$$
$$+ \sum_{l=1}^{2(k-1)+1} \binom{2k+1}{l} \left(J\left\langle\varphi(0),\mathcal{N}\varphi(0)\right\rangle^{\frac{1}{2}} \tau\right)^{l} \left\langle\varphi(0),(\mathcal{N}+n+l)^{k+\frac{1}{2}-\frac{l}{2}}\varphi(0)\right\rangle$$
$$\leqslant \sum_{l=0}^{2\left(k+\frac{1}{2}-1\right)} \binom{2\left(k+\frac{1}{2}\right)}{l} \left(J\left\langle\varphi(0),\mathcal{N}\varphi(0)\right\rangle^{\frac{1}{2}} t\right)^{l} \left\langle\varphi(0),(\mathcal{N}+n+l)^{k+\frac{1}{2}-\frac{l}{2}}\varphi(0)\right\rangle .$$

which concludes the induction.

Proof of (55)**.** Since

$$N \geqslant 1 \implies (N+1)^k = N^k e^{k\ln\left(1+\frac{1}{N}\right)} \leqslant N^k e^{\frac{k}{N}}$$

we notice that

$$N \geqslant k \implies (N+1)^k \leqslant eN^k.$$

Next, we continue from (60) for $n=0$, and introduce a cutoff, to obtain

$$\begin{aligned}
\left|\partial_t \mathrm{Tr}\left(p_\varphi(t)\mathcal{N}^k\right)\right| &\leqslant 2Jk\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}} \mathrm{Tr}\left(p_\varphi(t)\left(\mathcal{N}+1\right)^k\right)\\
&= 2Jk\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}} \mathrm{Tr}\left(p_\varphi(t)\left(\mathcal{N}+1\right)^k\left(\mathbb{1}_{\mathcal{N}<k}+\mathbb{1}_{\mathcal{N}\geqslant k}\right)\right)\\
&\leqslant 2Jk\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}} \left(\mathrm{Tr}\left(p_\varphi(t)k^k\right)+e\mathrm{Tr}\left(p_\varphi(t)\mathcal{N}^k\right)\right)\\
&= 2Jk\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}} \left(k^k+e\mathrm{Tr}\left(p_\varphi(t)\mathcal{N}^k\right)\right). \qquad (62)
\end{aligned}$$

With Gronwall's lemma we conclude that

$$\mathrm{Tr}\left(p_\varphi(t)\mathcal{N}^k\right) \leqslant \left(\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}^k\right)+e^{-1}k^k\right)e^{2eJk\mathrm{Tr}\left(p_\varphi(0)\mathcal{N}\right)^{\frac{1}{2}}t}.$$

□

For the Bose–Hubbard model (1) the total energy $\mathrm{Tr}\left(\gamma_d H_d\right)$ and the total particle number $\mathrm{Tr}(\gamma_{\mathrm{d}}^{(1)}\mathcal{N})$ are conserved as well. Moreover, we can prove bounds analogous to the mean-field dynamics for powers of the total number of particles. Note first that we can rewrite the Hamiltonian H_d as

$$H_d = \sum_{x\in\Lambda} h_x^\varphi - \frac{J}{2d}\sum_{\substack{x,y\in\Lambda\\ x\sim y}} \left(a_x^*-\overline{\alpha}_\varphi\right)\left(a_y-\alpha_\varphi\right), \qquad (63)$$

with $\alpha_\varphi = \langle\varphi, a\varphi\rangle$. Then, by using (45) and (46) we find that the one-lattice-site reduced density matrix satisfies

$$\begin{aligned}
i\partial_t\,\gamma_{\mathrm{d}}^{(1)} = [H,\gamma_d]^{(1)} &= \left[\sum_{x\in\Lambda} h_x^\varphi,\gamma_d\right]^{(1)} - \frac{J}{2d}\left[\sum_{\substack{x,y\in\Lambda\\ x\sim y}} \left(a_x^*-\overline{\alpha}\right)\left(a_y-\alpha\right),\gamma_d\right]^{(1)}\\
&= \left[h^\varphi,\gamma_{\mathrm{d}}^{(1)}\right] - J\mathrm{Tr}_2\left(\left[\left(a^*-\overline{\alpha}\right)\otimes\left(a-\alpha\right)+\left(a-\alpha\right)\otimes\left(a^*-\overline{\alpha}\right),\gamma_{\mathrm{d}}^{(2)}\right]\right). \qquad (64)
\end{aligned}$$

We have the following propagation bounds.

Proposition 11. *Let $\gamma_{\mathrm{d}}^{(1)}$ solve (64), let $k \in \mathbb{N}/2, k \geqslant 1$, $t \in \mathbb{R}_+$, and $\gamma_{\mathrm{d}}^{(1)}(0)\mathcal{N}^k \in \mathcal{L}^1(\ell^2(\mathbb{C}))$. Then*

$$\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)\mathcal{N}^k\right) \leqslant \left(\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(0)\mathcal{N}^k\right) + e^{-1}k^k\right) e^{2eJkt}. \tag{65}$$

Proof. Similarly to (58), let us define

$$\mathcal{A}^{k-1} := (\mathcal{N}+1)^k - \mathcal{N}^k \leqslant k\,(\mathcal{N}+1)^{k-1}.$$

Then Cauchy–Schwarz yields

$$\begin{aligned}
\left|\partial_t \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)\right| &\leqslant 2J\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\left[a_1, \mathcal{N}_1^k\right] a_2\right)\right| \\
&= 2J\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, a_2\mathcal{A}_1^{k-1}a_1\right)\right| \\
&\leqslant 2J\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, a_2\mathcal{A}_1^{k-1}a_2^*\right)^{\frac{1}{2}} \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, a_1^*\mathcal{A}_1^{k-1}a_1\right)^{\frac{1}{2}} \\
&\leqslant 2Jk\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,(\mathcal{N}_1+1)^{k-1}(\mathcal{N}_2+1)\right)^{\frac{1}{2}} \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)^{\frac{1}{2}}
\end{aligned}$$

Since $\left[(\mathcal{N}_1+1)^{k-1}, (\mathcal{N}_2+1)\right] = 0$, by Young's inequality,

$$(\mathcal{N}_1+1)^{k-1}(\mathcal{N}_2+1) \leqslant \left(1-\frac{1}{k}\right)(\mathcal{N}_1+1)^k + \frac{1}{k}(\mathcal{N}_2+1)^k.$$

Introducing a cutoff similarly to (62), we conclude that

$$\begin{aligned}
\left|\partial_t \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)\right| &\leqslant 2Jk\left(\left(1-\frac{1}{k}\right)\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}(\mathcal{N}_1+1)^k\right) + \frac{1}{k}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}(\mathcal{N}_2+1)^k\right)\right)^{\frac{1}{2}} \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)^{\frac{1}{2}} \\
&= 2Jk\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(\mathcal{N}+1)^k\right)^{\frac{1}{2}} \operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)^{\frac{1}{2}} \\
&\leqslant 2Jk\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)^{\frac{1}{2}}\left(k^k + e\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)\right)^{\frac{1}{2}} \\
&\leqslant 2Jk\left(k^k + e\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\mathcal{N}^k\right)\right).
\end{aligned}$$

With Gronwall's lemma we conclude that

$$\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)\mathcal{N}^k\right) \leqslant \left(\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(0)\mathcal{N}^k\right) + e^{-1}k^k\right) e^{2eJkt}.$$

□

4.4 Gronwall Estimate

Both Theorems 1 and 2 are proven via a Gronwall estimate for the quantity $\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\, q\right)$. This is directly related to the trace norm difference of reduced density matrices, analogous to the case of the weak coupling limit [27], as the following Lemma shows.

Lemma 12. *Let p be a rank one projection and γ a positive trace 1 operator on $\ell^2(\mathbb{C})$ and $q := 1-p$. Then*

$$2\operatorname{Tr}(\gamma q) \leqslant \|\gamma - p\|_{\mathcal{L}^1} \leqslant 2\sqrt{2}\sqrt{\operatorname{Tr}(\gamma q)}. \tag{66}$$

Proof. In order to get the upper bound in (66), we first notice that since $\gamma \leqslant 1$ and $\operatorname{Tr}(\gamma) = \operatorname{Tr}(p) = 1$,

$$\|p\gamma p - p\|_{\mathcal{L}^1} = \operatorname{Tr}\left((1-\gamma)\,p\right) = 1 - \operatorname{Tr}(\gamma p) = \operatorname{Tr}(\gamma(1-p)) = \operatorname{Tr}(\gamma q),$$

so

$$
\begin{aligned}
\|\gamma - p\|_{\mathcal{L}^1} &= \|(p+q)\gamma(p+q) - p\|_{\mathcal{L}^1} \\
&\leqslant 2\mathrm{Tr}\,(\gamma q) + 2\,\|q\gamma p\|_{\mathcal{L}^1} \\
&\leqslant 2\mathrm{Tr}\,(\gamma q) + 2\sqrt{\mathrm{Tr}\,(\gamma q)}\sqrt{\mathrm{Tr}\,(\gamma p)} \\
&= 2\sqrt{\mathrm{Tr}\,(\gamma q)}\left(\sqrt{\mathrm{Tr}\,(\gamma q)} + \sqrt{1 - \mathrm{Tr}\,(\gamma q)}\right) \\
&\leqslant 2\sqrt{2}\sqrt{\mathrm{Tr}\,(\gamma q)},
\end{aligned}
$$

where we used $\sqrt{x} + \sqrt{1-x} \leqslant \sqrt{2}$ for $0 \leqslant x \leqslant 1$. The lower bound follows directly from

$$
\mathrm{Tr}\,(\gamma q) = \mathrm{Tr}\,((p-\gamma)\,p) \leqslant \|\gamma - p\|_{\mathcal{L}^1}\,.
$$

□

Next, we compute the time derivative of $\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)$ and estimate some of the appearing terms. This is analogous to the estimates in the weak coupling limit, see, e.g., [33, Lemma 3.2]. The only term that causes technical difficulties is $\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\,(\mathcal{N}+1)\,q\right)$, and Sections 5 and 6 are devoted to controlling this term in different ways, leading to our two main theorems.

Proposition 13. *Let γ_d solve* (5) *with normalized initial data $\gamma_d(0) \in \mathcal{L}^1(\mathcal{F})$ and φ solve* (3) *with normalized initial data $\varphi(0) \in \ell^2(\mathbb{C})$. We define $p := |\varphi\rangle\langle\varphi|$ and $q := 1 - p$. Then*

$$
\begin{aligned}
&\left|\partial_t \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)\right| \\
&\leqslant J\left(\mathrm{Tr}\,(p\mathcal{N}) + 1\right)^{\frac{1}{2}}\left(8\mathrm{Tr}\,(p\mathcal{N})^{\frac{1}{2}}\,\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right) + 4\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)^{\frac{1}{2}}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\,(\mathcal{N}+1)\,q\right)^{\frac{1}{2}} + \frac{\mathrm{Tr}\,(p\mathcal{N})^{\frac{1}{2}}}{d}\right).
\end{aligned} \tag{67}
$$

Proof. **Computation of the time derivative.** We introduce the self-adjoint operator

$$
A := (a^* - \overline{\alpha_\varphi}) \otimes (a - \alpha_\varphi) + (a - \alpha_\varphi) \otimes (a^* - \overline{\alpha_\varphi})\,.
$$

With (64), we start by computing

$$
\begin{aligned}
i\partial_t \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right) &= \mathrm{Tr}\left(\left[h^\varphi, \gamma_{\mathrm{d}}^{(1)}\right] q\right) - J\mathrm{Tr}\left(\left[A, \gamma_{\mathrm{d}}^{(2)}\right] q_1\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,[h^\varphi, q]\right) \\
&= J\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,[A, q_1]\right) \\
&= 2iJ\mathrm{Im}\left[\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,Aq_1\right)\right].
\end{aligned} \tag{68}
$$

Inserting resolution of identities $1 = p + q$, we get

$$
\begin{aligned}
&\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,Aq_1\right) \\
&= \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,p_1p_2Aq_1p_2\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,p_1p_2Aq_1q_2\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,p_1q_2Aq_1p_2\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,p_1q_2Aq_1q_2\right) \\
&\quad + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,q_1p_2Aq_1p_2\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,q_1p_2Aq_1q_2\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,q_1q_2Aq_1p_2\right) + \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,q_1q_2Aq_1q_2\right).
\end{aligned}
$$

Note that $q_1p_2Aq_1p_2$ and $q_1q_2Aq_1q_2$ are self adjoint and hence do not contribute to 68. This is also the case for $q_1p_2Aq_1q_2$ and $q_1q_2Aq_1p_2$ which are each others complex conjugate. Furthermore, $p_1p_2Aq_1p_2 = 0$ by definition of A. Then, by symmetry, we see that $p_1q_2Aq_1p_2$ is also not contributing, since

$$
\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,p_1q_2Aq_1p_2\right) = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,q_1p_2Ap_1q_2\right) = \overline{\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\,p_1q_2Aq_1p_2\right)}.
$$

Thus, we are left with

$$i\partial_t \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\, q\right) = 2iJ\mathrm{Im}\left[\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1p_2Aq_1q_2\right)\right] + 2iJ\mathrm{Im}\left[\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1q_2Aq_1q_2\right)\right]. \tag{69}$$

Estimation of the $p_1p_2Aq_1q_2$ term. Since $pq=0$,

$$\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1p_2Aq_1q_2\right) = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1p_2(a_1^*a_2+a_1a_2^*)q_1q_2\right),$$

and by symmetry of $\gamma_{\mathrm{d}}^{(2)}$,

$$\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1p_2a_1^*a_2q_1q_2\right) = \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1p_2a_1a_2^*q_1q_2\right).$$

Then, we use Cauchy–Schwarz to estimate, for any $\epsilon>0$,

$$\begin{aligned}
\left|\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(2)}\, p_1p_2Aq_1q_2\right)\right| &= \frac{1}{d\,|\Lambda|}\left|\sum_{\substack{x,y\in\Lambda\\ x\sim y}} \mathrm{Tr}\left(\gamma_d p_xp_ya_x^*a_yq_xq_y\right)\right| \\
&\leqslant \frac{1}{d\,|\Lambda|}\sum_{x\in\Lambda}\left|\mathrm{Tr}\left(q_x\gamma_d^{\frac12}\cdot\gamma_d^{\frac12}\sum_{y\in\Lambda,x\sim y} p_xp_ya_x^*a_yq_y\right)\right| \\
&\leqslant \frac{1}{2d\epsilon\,|\Lambda|}\sum_{x\in\Lambda}\mathrm{Tr}\left(q_x\gamma_d\right) + \frac{\epsilon}{2d\,|\Lambda|}\sum_{x\in\Lambda}\mathrm{Tr}\left(\gamma_d\sum_{\substack{y\in\Lambda,x\sim y\\ z\in\Lambda,x\sim z}} p_xp_ya_x^*a_yq_yq_za_z^*a_xp_zp_x\right) \\
&= \frac{1}{2d\epsilon}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\, q\right) + \epsilon\frac{\mathrm{Tr}\left(p\mathcal{N}\right)}{2d\,|\Lambda|}\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\sum_{z\in\Lambda,x\sim z}\mathrm{Tr}\left(\gamma_dp_xp_ya_yq_yq_za_z^*p_z\right) \\
&= \frac{1}{2d\epsilon}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\, q\right) + \epsilon\frac{\mathrm{Tr}\left(p\mathcal{N}\right)}{2d\,|\Lambda|}\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\mathrm{Tr}\left(\gamma_dp_xp_ya_yq_ya_y^*p_y\right) \\
&\quad + \epsilon\frac{\mathrm{Tr}\left(p\mathcal{N}\right)}{2d\,|\Lambda|}\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\sum_{\substack{z\in\Lambda,x\sim z\\ z\neq y}}\mathrm{Tr}\left(q_y\gamma_dq_zp_xp_ya_ya_z^*p_z\right).
\end{aligned}$$

The last two summands can be estimated as

$$\begin{aligned}
\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\mathrm{Tr}\left(\gamma_dp_xp_ya_yq_ya_y^*p_y\right) \leqslant \sum_{\substack{x,y\in\Lambda\\ x\sim y}}\mathrm{Tr}\left(\gamma_dp_xp_ya_ya_y^*p_y\right) &= \mathrm{Tr}\left(p(\mathcal{N}+1)\right)\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\mathrm{Tr}\left(\gamma_dp_xp_y\right) \\
&\leqslant 2d\,|\Lambda|\left(\mathrm{Tr}\left(p\mathcal{N}\right)+1\right),
\end{aligned}$$

and

$$\begin{aligned}
\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\sum_{\substack{z\in\Lambda,x\sim z\\ z\neq y}}\mathrm{Tr}\left(q_y\gamma_dq_zp_xp_ya_ya_z^*p_z\right) &\leqslant \sum_{\substack{x,y\in\Lambda\\ x\sim y}}\sum_{\substack{z\in\Lambda,x\sim z\\ z\neq y}}\mathrm{Tr}\left(\gamma_dq_zp_xp_ya_ya_y^*p_y\right)^{\frac12}\mathrm{Tr}\left(\gamma_dq_yp_za_za_z^*p_z\right)^{\frac12} \\
&= \mathrm{Tr}\left(p(\mathcal{N}+1)\right)\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\sum_{\substack{z\in\Lambda,x\sim z\\ z\neq y}}\mathrm{Tr}\left(\gamma_dq_zp_xp_y\right)^{\frac12}\mathrm{Tr}\left(\gamma_dq_yp_z\right)^{\frac12} \\
&\leqslant \left(\mathrm{Tr}\left(p\mathcal{N}\right)+1\right)\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\sum_{z\in\Lambda,x\sim z}\mathrm{Tr}\left(\gamma_dq_z\right)^{\frac12}\mathrm{Tr}\left(\gamma_dq_y\right)^{\frac12} \\
&= \left(\mathrm{Tr}\left(p\mathcal{N}\right)+1\right)\sum_{x\in\Lambda}\left(\sum_{y\in\Lambda,x\sim y}\mathrm{Tr}\left(\gamma_dq_y\right)^{\frac12}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&\leqslant 2d\left(\operatorname{Tr}(p\mathcal{N})+1\right)\sum_{\substack{x,y\in\Lambda\\ x\sim y}}\operatorname{Tr}\left(\gamma_d q_y\right)\\
&=4d^2\left|\Lambda\right|\left(\operatorname{Tr}(p\mathcal{N})+1\right)\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right).
\end{aligned}
$$

Using these two estimates and then choosing $\epsilon^{-1} \coloneqq 2d\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}$, we obtain

$$
\begin{aligned}
\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1p_2Aq_1q_2\right)\right| &\leqslant \frac{1}{2d\epsilon}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+\epsilon\operatorname{Tr}(p\mathcal{N})\left(\operatorname{Tr}(p\mathcal{N})+1\right)\\
&\quad+2d\epsilon\operatorname{Tr}(p\mathcal{N})\left(\operatorname{Tr}(p\mathcal{N})+1\right)\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)\\
&=\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\left(2\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+\frac{1}{2d}\right).
\end{aligned}
\tag{70}
$$

Estimation of the $p_1q_2Aq_1q_2$ term. Since $pq=0$,

$$
\begin{aligned}
\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2Aq_1q_2\right)&=\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1^*a_2q_1q_2\right)+\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1a_2^*q_1q_2\right)-\alpha_\varphi\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1^*q_1q_2\right)\\
&\quad-\overline{\alpha_\varphi}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1q_1q_2\right).
\end{aligned}
$$

We estimate

$$
\begin{aligned}
\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1^*a_2q_1q_2\right)\right|&\leqslant\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2\mathcal{N}_1p_1\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}q_2q_1a_2a_2^*q_2\right)^{\frac{1}{2}}\\
&=\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}q_1q_2\left(\mathcal{N}_2+1\right)q_2\right)^{\frac{1}{2}}\\
&\leqslant\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\left(\mathcal{N}+1\right)q\right)^{\frac{1}{2}},\\
\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1^*q_1q_2\right)\right|&\leqslant\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2\mathcal{N}_1p_1\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}q_1q_2\right)^{\frac{1}{2}}\\
&=\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}q_1q_2\right)^{\frac{1}{2}}\leqslant\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right),
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1a_2^*q_1q_2\right)\right|&\leqslant\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\mathcal{N}q\right)^{\frac{1}{2}},\\
\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2a_1q_1q_2\right)\right|&\leqslant\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right).
\end{aligned}
$$

Inserting these estimates yields

$$
\begin{aligned}
&\left|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(2)}p_1q_2Aq_1q_2\right)\right|\\
&\quad\leqslant 2\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\left(\left|\alpha_\varphi\right|\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\left(\mathcal{N}+1\right)q\right)^{\frac{1}{2}}\right)\\
&\quad\leqslant 2\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\left(\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\left(\mathcal{N}+1\right)q\right)^{\frac{1}{2}}\right).
\end{aligned}
\tag{71}
$$

Conclusion. Inserting (70) and (71) into (69) we obtain

$$
\begin{aligned}
&\left|\partial_t\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)\right|\\
&\quad\leqslant 2J\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\left(2\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+\frac{1}{2d}\right)\\
&\qquad+4J\left(\operatorname{Tr}(p\mathcal{N})+1\right)^{\frac{1}{2}}\left(\operatorname{Tr}(p\mathcal{N})^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)^{\frac{1}{2}}\operatorname{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\left(\mathcal{N}+1\right)q\right)^{\frac{1}{2}}\right)
\end{aligned}
$$

$$= J\left(\mathrm{Tr}\left(p\mathcal{N}\right)+1\right)^{\frac{1}{2}}\left(8\mathrm{Tr}\left(p\mathcal{N}\right)^{\frac{1}{2}}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)+4\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\right)^{\frac{1}{2}}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\left(\mathcal{N}+1\right)q\right)^{\frac{1}{2}}+\frac{\mathrm{Tr}\left(p\mathcal{N}\right)^{\frac{1}{2}}}{d}\right).$$

□

5 Proof of Theorem 1

In this Section we prove Theorem 1 by estimating the term $\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}q\left(\mathcal{N}+1\right)q\right)$ from the Gronwall estimate in Proposition 13 using a moment method. The main idea is to use the following estimates obtained by iterating the Cauchy–Schwarz inequality, along with the moment bounds we obtained in Section 4.3.

Lemma 14. *Let $k\in\mathbb{N}$ and $\gamma,p\in\mathcal{L}^1(\ell^2(\mathbb{C}))$. We assume that $0\leqslant\gamma\leqslant 1$, that p is a rank one projection and $p\mathcal{N}^k,\gamma\mathcal{N}^k\in\mathcal{L}^1(\ell^2(\mathbb{C}))$. Then*

$$\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)q\right)\leqslant\mathrm{Tr}\left(\gamma q\right)^{1-2^{-k}}\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)^{2^k}q\right)^{2^{-k}}, \tag{72}$$

$$\mathrm{Tr}\left(\gamma q\mathcal{N}^k q\right)\leqslant 2\mathrm{Tr}\left(\gamma\mathcal{N}^k\right)+2\mathrm{Tr}\left(p\mathcal{N}^k\right). \tag{73}$$

Proof. **Proof of (72).** We proceed by induction on k. The inequality is trivial for $k=0$. With the Cauchy–Schwarz inequality,

$$\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)^{2^k}q\right)\leqslant\mathrm{Tr}\left(\gamma q\right)^{\frac{1}{2}}\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)^{2^{k+1}}q\right)^{\frac{1}{2}},$$

so assuming the result holds for given $k\in\mathbb{N}$ we can bound

$$\begin{aligned}\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)q\right)&\leqslant\mathrm{Tr}\left(\gamma q\right)^{1-2^{-k}}\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)^{2^k}q\right)^{2^{-k}}\\&\leqslant\mathrm{Tr}\left(\gamma q\right)^{1-2^{-k}+2^{-(k+1)}}\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)^{2^{k+1}}q\right)^{2^{-(k+1)}}\\&=\mathrm{Tr}\left(\gamma q\right)^{1-2^{-(k+1)}}\mathrm{Tr}\left(\gamma q\left(\mathcal{N}+1\right)^{2^{k+1}}q\right)^{2^{-(k+1)}},\end{aligned}$$

i.e., the result holds for $k+1$, closing the induction argument.
Proof of (73). With the Cauchy–Schwarz inequality,

$$\begin{aligned}\mathrm{Tr}\left(\gamma q\mathcal{N}^k q\right)&=\mathrm{Tr}\left(\gamma\mathcal{N}^k\right)-\mathrm{Tr}\left(\gamma p\mathcal{N}^k p\right)-\mathrm{Tr}\left(\gamma p\mathcal{N}^k q\right)-\mathrm{Tr}\left(\gamma q\mathcal{N}^k p\right)\\&=\mathrm{Tr}\left(\gamma\mathcal{N}^k\right)-\mathrm{Tr}\left(\gamma p\mathcal{N}^k p\right)+2\sqrt{\mathrm{Tr}\left(\gamma p\mathcal{N}^k p\right)}\sqrt{\mathrm{Tr}\left(\gamma q\mathcal{N}^k q\right)}\\&\leqslant\mathrm{Tr}\left(\gamma\mathcal{N}^k\right)+\mathrm{Tr}\left(\gamma p\mathcal{N}^k p\right)+\frac{1}{2}\mathrm{Tr}\left(\gamma q\mathcal{N}^k q\right),\end{aligned}$$

so

$$\mathrm{Tr}\left(\gamma q\mathcal{N}^k q\right)\leqslant 2\mathrm{Tr}\left(\gamma\mathcal{N}^k\right)+2\mathrm{Tr}\left(\gamma p\mathcal{N}^k p\right)\leqslant 2\mathrm{Tr}\left(\gamma\mathcal{N}^k\right)+2\mathrm{Tr}\left(p\mathcal{N}^k\right).$$

□

5.1 The Moment Method

We will prove Theorem 1 by showing that the probability of having a large lattice site occupation outside the product state structure is small. We use the following basic Calculus estimates.

Lemma 15. *Let* $(u_n)_{n\in\mathbb{N}} \subset \mathbb{R}_+$*. Then*

$$\exists a > 0 \ s.t. \ \forall n \in \mathbb{N}, u_n \leqslant e^{-\frac{n}{a}} \implies \forall k \in \mathbb{N}, \sum_{n\in\mathbb{N}} n^k u_n \leqslant (1+a)a^k k!, \tag{74}$$

and conversely,

$$\exists b > 0 \ s.t. \ \forall k \in \mathbb{N}, \sum_{n\in\mathbb{N}} n^k u_n \leqslant b^k k! \implies \forall M \in \mathbb{N}, \sum_{n=M}^{\infty} (n+1)u_n \leqslant (2+4b)e^{-\frac{M}{2b}}. \tag{75}$$

Proof. **Proof of** (74)**.** The function

$$f_a : \mathbb{R}_+ \quad \to \quad \mathbb{R}_+, x \quad \mapsto \quad x^k e^{-\frac{x}{a}}$$

is increasing up to ak and decreasing afterwards. Thus, by series-integral comparison,

$$\begin{aligned}\sum_{n\in\mathbb{N}} f_a(n) &\leqslant \int\limits_{\mathbb{R}_+} f_a(x)dx + f_a(\lfloor ak\rfloor) + f_a(\lceil ak\rceil) = a^k \left(ak! + f_1\left(a^{-1}\lfloor ak\rfloor\right) + f_1\left(a^{-1}\lceil ak\rceil\right)\right) \\ &\leqslant a^k\left(ak! + 2f_1(k)\right) = a^k\left(ak! + 2\left(\frac{k}{e}\right)^k\right).\end{aligned}$$

If $k \geqslant 1$, inserting the Stirling lower approximation,

$$\sqrt{2\pi k}\left(\frac{k}{e}\right)^k \leqslant k! \tag{76}$$

yields

$$\sum_{n\in\mathbb{N}} n^k e^{-\frac{n}{a}} \leqslant a^k k!\left(a + \sqrt{\frac{2}{\pi k}}\right) \leqslant (1+a)\,a^k k!.$$

The statement also holds for $k = 0$ since

$$\sum_{n\in\mathbb{N}} e^{-\frac{n}{a}} \leqslant 1 + \int\limits_{\mathbb{R}_+} e^{-\frac{x}{a}}dx = 1 + a.$$

Proof of (75)**.** If $0 < a < \frac{1}{b}$ and $M \in \mathbb{N}$ we find

$$\begin{aligned}\sum_{n=M}^{\infty}(n+1)u_n e^{aM} &\leqslant \sum_{n\in\mathbb{N}}(n+1)u_n e^{an} = \sum_{n,k\in\mathbb{N}}(n+1)\frac{(an)^k}{k!}u_n = \sum_{k\in\mathbb{N}}\frac{a^k}{k!}\left(\sum_{n\in\mathbb{N}} n^{k+1}u_n + \sum_{n\in\mathbb{N}} n^k u_n\right) \\ &\leqslant \sum_{k\in\mathbb{N}}\frac{a^k}{k!}\left(b^{k+1}(k+1)! + b^k k!\right) = \sum_{k\in\mathbb{N}}\left(b(k+1)(ab)^k + (ab)^k\right) \\ &= \frac{b}{(1-ab)^2} + \frac{1}{1-ab}.\end{aligned}$$

Choosing $a = \frac{1}{2b}$ yields

$$\sum_{n=M}^{\infty}(n+1)u_n \leqslant \frac{1-ab+b}{(1-ab)^2}e^{-aM} = (2+4b)e^{-\frac{M}{2b}}.$$

$\square$

With this we can prove our first main theorem.

Proof of Theorem 1. **Controlling** $\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\,(\mathcal{N}+1)\,\mathbb{1}_{\mathcal{N}\geqslant M}q\right)$ **with moments.** Let $k\in\mathbb{N}$. Applying (74) from Lemma 15 first to $u_n \coloneqq \mathrm{Tr}\,(p(0)\mathbb{1}_{\mathcal{N}=n})$ and then to $u_n \coloneqq \mathrm{Tr}\left(\gamma_d^{(1)}(0)\mathbb{1}_{\mathcal{N}=n}\right)$ while using the assumption (7) from Theorem 1, we obtain directly

$$\mathrm{Tr}\left(p(0)\mathcal{N}^k\right)\leqslant c(1+a)a^k k!, \tag{77}$$

$$\mathrm{Tr}\left(\gamma_d^{(1)}(0)\mathcal{N}^k\right)\leqslant c(1+a)a^k k!. \tag{78}$$

For $k\geqslant 1$, we use first (73) from Lemma 14, then the moment bounds (55) from Proposition 10 and (65) from Proposition 11, then (77) and (78), and Stirling's approximation (76), and find

$$\begin{aligned}
\sum_{n\in\mathbb{N}} n^k\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)q\mathbb{1}_{\mathcal{N}=n}q\right) &= \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)q(t)\mathcal{N}^k q(t)\right)\\
&\leqslant 2\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)\mathcal{N}^k\right)+2\mathrm{Tr}\left(p(t)\mathcal{N}^k\right)\\
&\leqslant 2\left(\mathrm{Tr}\left(p(0)\mathcal{N}^k\right)+e^{-1}k^k\right)e^{2eJk\mathrm{Tr}(p(0)\mathcal{N})^{\frac12}t}\\
&\quad+2\left(\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(0)\mathcal{N}^k\right)+e^{-1}k^k\right)e^{2eJkt}\\
&\leqslant 2\left(\mathrm{Tr}\left(p(0)\mathcal{N}^k\right)+\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(0)\mathcal{N}^k\right)+2e^{-1}k^k\right)e^{C_1kt}\\
&\leqslant 4\left(c(1+a)a^k k!+e^{-1}k^k\right)e^{C_1kt}\\
&\leqslant 4\left(c(1+a)a^k+\frac{e^{k-1}}{\sqrt{2\pi k}}\right)k!e^{C_1kt}\\
&\leqslant 4\left(c(1+a)+e^{-1}\right)\left(a^k+e^k\right)k!e^{C_1kt}\\
&\leqslant 4\left(c(1+a)+e^{-1}\right)\left((a+e)e^{C_1t}\right)^k k!.
\end{aligned}$$

This is also valid for $k=0$, so (75) from Lemma 15 implies

$$\begin{aligned}
\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)q(t)\left(\mathcal{N}+1\right)\mathbb{1}_{\mathcal{N}\geqslant M}q(t)\right) &= \sum_{n=M}^{\infty}(n+1)\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t)q(t)\mathbb{1}_{\mathcal{N}=n}q(t)\right)\\
&\leqslant 4\left(c(1+a)+e^{-1}\right)\left(2+4(a+e)e^{C_1t}\right)e^{-\frac{M}{2(a+e)}e^{-C_1t}}\\
&\leqslant C_2 e^{C_1t-\frac{M}{2(a+e)}e^{-C_1t}}.
\end{aligned}\tag{79}$$

Conclusion of the proof. Let $M\in\mathbb{N}^*$. We use the beginning of the Gronwall estimate from Proposition 13 while introducing a cutoff on $\mathcal{N}$, and then Proposition 10 to find, for any $\epsilon>0$,

$$\begin{aligned}
&\left|\partial_t\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)\right|\\
&\leqslant JC_3\Bigg(8\mathrm{Tr}\,(p\mathcal{N})^{\frac12}\,\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)+4\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)^{\frac12}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\,(\mathcal{N}+1)\left(\mathbb{1}_{\mathcal{N}<M}+\mathbb{1}_{\mathcal{N}\geqslant M}\right)q\right)^{\frac12}\\
&\qquad+\frac{\mathrm{Tr}\,(p\mathcal{N})^{\frac12}}{d}\Bigg)\\
&\leqslant JC_3\left(\left(8\mathrm{Tr}\,(p\mathcal{N})^{\frac12}+4\sqrt{M}\right)\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)+4\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)^{\frac12}\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\,(\mathcal{N}+1)\,\mathbb{1}_{\mathcal{N}\geqslant M}q\right)^{\frac12}+\frac{\mathrm{Tr}\,(p\mathcal{N})^{\frac12}}{d}\right)\\
&\leqslant JC_3\left(\left(8\mathrm{Tr}\,(p\mathcal{N})^{\frac12}+4\sqrt{M}+4\epsilon^{-1}\right)\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\right)+\epsilon\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}\,q\,(\mathcal{N}+1)\,\mathbb{1}_{\mathcal{N}\geqslant M}q\right)+\frac{\mathrm{Tr}\,(p\mathcal{N})^{\frac12}}{d}\right).
\end{aligned}$$

Next, we insert (79) and use the conservation of the mean-field number of particles (see (54)). Then the choice $\epsilon \coloneqq d^{-1}\, e^{\frac{M}{2(a+e)}e^{-C_1 t}}$ yields

$$\begin{aligned}
&\left|\partial_t \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t) q(t)\right)\right| \\
&\quad\leqslant JC_3 \left(\left(8\mathrm{Tr}\left(p(0)\mathcal{N}\right)^{\frac{1}{2}} + 4\sqrt{M} + 4\epsilon^{-1} \right) \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t) q(t)\right) + \epsilon C_2 e^{C_1 t - \frac{M}{2(a+e)}e^{-C_1 t}} + \frac{\mathrm{Tr}\left(p(0)\mathcal{N}\right)^{\frac{1}{2}}}{d} \right) \\
&\quad\leqslant JC_3 \left(\left(8\mathrm{Tr}\left(p(0)\mathcal{N}\right)^{\frac{1}{2}} + 4\sqrt{M} + 4d e^{-\frac{M}{2(a+e)}e^{-C_1 t}} \right) \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t) q(t)\right) + \frac{C_2 e^{C_1 t} + \mathrm{Tr}\left(p(0)\mathcal{N}\right)^{\frac{1}{2}}}{d} \right).
\end{aligned}$$

Then we choose[1]

$$M \coloneqq \left\lceil 2(a+e)e^{C_1 t} \ln\left(\frac{d}{\sqrt{\ln(d+1)}} \right) \right\rceil.$$

Observing that for $d \geqslant 1$ we have

$$\ln\left(\frac{d}{\sqrt{\ln(d+1)}} \right) \leqslant \ln(d+1),$$

this choice implies

$$\begin{aligned}
\sqrt{M} + d e^{-\frac{M}{2(a+e)}e^{-C_1 t}} &\leqslant \left(2(a+e)e^{C_1 t} \ln\left(\frac{d}{\sqrt{\ln(d+1)}} \right) + 1 \right)^{\frac{1}{2}} + \sqrt{\ln(d+1)} \\
&\leqslant \left(\sqrt{2(a+e)} e^{\frac{C_1}{2} t} + 1 \right) \sqrt{\ln(d+1)} + 1.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\left|\partial_t \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t) q(t)\right)\right| \leqslant JC_3 \Bigg(&\left(2C_4 + 4\left(\sqrt{2(a+e)} e^{\frac{C_1}{2} t} + 1 \right) \sqrt{\ln(d+1)} \right) \mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t) q(t)\right) \\
&+ \frac{C_2 e^{C_1 t} + \mathrm{Tr}\left(p(0)\mathcal{N}\right)^{\frac{1}{2}}}{d} \Bigg).
\end{aligned}$$

Noticing that the time dependent coefficients in the above expression are non-decreasing in time, we can use Gronwall's lemma to obtain

$$\begin{aligned}
\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(t) q(t)\right) \leqslant &\left(\mathrm{Tr}\left(\gamma_{\mathrm{d}}^{(1)}(0) q(0)\right) + \frac{C_2 e^{C_1 t} + \mathrm{Tr}\left(p(0)\mathcal{N}\right)^{\frac{1}{2}}}{d\left(2C_4 + 4\left(\sqrt{2(a+e)} e^{\frac{C_1}{2} t} + 1 \right) \sqrt{\ln(d+1)} \right)} \right) \\
&e^{JC_3\left(2C_4 + 4\left(\sqrt{2(a+e)} e^{\frac{C_1}{2} t} + 1 \right) \sqrt{\ln(d+1)} \right) t}.
\end{aligned}$$

Finally, using (66) from Lemma 12 proves Theorem 1. □

[1]Let us comment on the choice of the cutoff parameter. Optimizing in M requires to solve, for $x \geqslant 0$,

$$\begin{aligned}
\sqrt{x} = d e^{-\frac{x}{2(a+e)}e^{-C_1 t}} &\iff e^{\frac{x}{a+e}e^{-C_1 t}} x = d^2 \iff e^{\frac{x}{a+e}e^{-C_1 t}} \frac{x}{a+e} e^{-C_1 t} = \frac{d^2}{a+e} e^{-C_1 t} \\
&\iff \frac{x}{a+e} e^{-C_1 t} = W_0\left(\frac{d^2}{a+e} e^{-C_1 t} \right) \iff x = (a+e) e^{C_1 t} W_0\left(\frac{d^2}{a+e} e^{-C_1 t} \right),
\end{aligned}$$

where W_0 is the principal branch of the Lambert W function. Our choice of M comes from the fact that

$$W_0(x) \underset{x\to\infty}{=} \ln\left(\frac{x}{\ln(x)} \right) + o(1).$$

6 Proof of Theorem 2

In this section we prove Theorem 2 using an energy estimate. Recall that the Bose–Hubbard Hamiltonian H_d can be written as a sum of two time-dependent quantities,

$$H_d = \sum_{x\in\Lambda} h_x^{\alpha_\varphi}(t) + \tilde{H}(t),$$

where $\alpha_\varphi(t) := \langle \varphi(t), a\varphi(t)\rangle$. Here, $h_x^{\alpha_\varphi}(t)$ is the mean-field operator from (4), i.e.,

$$h_x^{\alpha_\varphi}(t) := -J\Big[\alpha_\varphi(t)a_x^* + \overline{\alpha_\varphi(t)}a_x - |\alpha_\varphi(t)|^2\Big] + (J-\mu)\mathcal{N}_x + \frac{U}{2}\mathcal{N}_x(\mathcal{N}_x - 1), \tag{80}$$

and $\tilde{H}(t)$ can be computed as

$$\begin{aligned}\tilde{H}(t) : = &-\frac{J}{2d}\sum_{<x,y>}\Big(p_x(t)p_y(t)K_{x,y}^{(2)}q_x(t)q_y(t) + p_x(t)q_y(t)K_{x,y}^{(2)}q_x(t)p_y(t)\Big) + h.c.\\ &-\frac{J}{d}\sum_{<x,y>} p_x(t)q_y(t)K_{x,y}^{(3)}(t)q_x(t)q_y(t) + h.c.\\ &-\frac{J}{2d}\sum_{<x,y>} q_x(t)q_y(t)K_{x,y}^{(4)}(t)q_x(t)q_y(t),\end{aligned} \tag{81}$$

where

$$K_{x,y}^{(2)} := a_x^* a_y + a_y^* a_x, \tag{82}$$
$$K_{x,y}^{(3)}(t) := K_{x,y}^{(2)} - \alpha_\varphi(t)a_x^* - \overline{\alpha_\varphi(t)}a_x, \tag{83}$$
$$K_{x,y}^{(4)}(t) := K_{x,y}^{(3)}(t) - \alpha_\varphi(t)a_y^* - \overline{\alpha_\varphi(t)}a_y + 2|\alpha_\varphi(t)|^2. \tag{84}$$

Here, the superscript i in the expression $K_{x,y}^{(i)}$ refers to the number of q's that accompany it in the expression of $\tilde{H}$ in (81). Note that $K_{x,y}^{(2)}$ does not depend on t whereas the other terms $K_{x,y}^{(3)}$ and $K_{x,y}^{(4)}$ do through the term $\alpha_\varphi(t)$.

For our proof we define the quantities

$$f(t) := \frac{1}{|\Lambda|}\left\langle \Psi_d(t), \left(H_d + \sum_{x\in\Lambda}\big(q_x(t)h_x^{\alpha_\varphi}(t)q_x(t) - h_x^{\alpha_\varphi}(t) + cq_x(t)\big)\right)\Psi_d(t)\right\rangle \tag{85}$$

with $c > 0$, and

$$g(t) := \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\big\langle \Psi_d(t), \big(q_x(t)\mathcal{N}_x^2 q_x(t) + q_x(t)\big)\Psi_d(t)\big\rangle. \tag{86}$$

The idea of the proof is the following. In the Gronwall estimate from Proposition 13, the problematic term was $\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d(t), q_x(t)\mathcal{N}_x^2 q_x(t)\Psi_d(t)\rangle$. Hence, one might want to attempt to do a joint Gronwall argument for this and the original quantity $\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d(t), q_x(t)\Psi_d(t)\rangle$ that we want to estimate, i.e., a Gronwall argument for g. However, if one computes the time derivative of g, one finds higher and higher powers $q\mathcal{N}^k q$ that need to be controlled, so the Gronwall argument cannot be closed. The trick is to do instead a Gronwall argument for f. Except for the cq term, f represents the energy of deviations from the lattice product state structure. The technical advantage for the Gronwall argument is that $\langle \Psi_d(t), H_d\Psi_d(t)\rangle$ is conserved, and

$$qh^{\alpha_\varphi}q - h^{\alpha_\varphi} = -h^{\alpha_\varphi}p - ph^{\alpha_\varphi} + ph^{\alpha_\varphi}p,$$

so the $\mathcal{N}^2$ term from the interaction appears always together with at least one p projection. And all powers of $\mathcal{N}$ can be controlled when traced out against p due to Proposition 10. Hence, we can close a Gronwall estimate for f. Finally, one can prove that $Cg - d^{-1} \leqslant f \leqslant Cg + d^{-1}$. Hence, g can be estimated in terms of its initial data and an error d^{-1}, which, together with Lemma 12, implies Theorem 2.

In the following, we start by proving the equivalence of f and g up to an error d^{-1} in Section 6.1. Then, in Section 6.2, we prove the Gronwall estimate for f. We conclude with the proof of Theorem 2 in Section 6.3.

Notation. In the following estimates, we use the quantities $C > 0$, $C(J, \mu, U) > 0$, and $\tilde{C}(t) > 0$ with the following definitions:

- C is a positive constant that is independent of all parameters of the model.
- $C(J, \mu, U)$ is a positive constant that depends on the parameters J, μ, U only polynomially, and is independent of the initial conditions and time t.
- $\tilde{C}(t)$ is a positive quantity that may depend on $C(J, \mu, U)$, the initial data $\langle \varphi(0), \mathcal{N}^j \varphi(0) \rangle$ for $j \leqslant 4$, and polynomially on time t.

For convenience, these quantities may change from one line to the next in the subsequent estimates.

6.1 Equivalence of f and g

We start by presenting an estimate for a slightly modified f.

Proposition 16. *There exist $C > 0$ such that for all $\epsilon > 0$ we have*

$$\begin{aligned}
&\frac{1}{|\Lambda|} \left| \left\langle \Psi_d, \left(\tilde{H} + \sum_{x \in \Lambda} q_x \left(h_x^{\alpha_\varphi} - \frac{U}{2} \mathcal{N}_x^2 \right) q_x \right) \Psi_d \right\rangle \right| \\
&\quad \leqslant C \left(1 + J^2 + \left(J - \mu - \frac{U}{2} \right)^2 \right) \left(1 + \frac{1}{\epsilon} + \langle \varphi(0), \mathcal{N} \varphi(0) \rangle^2 \right) \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \langle \Psi_d, q_x \Psi_d \rangle \\
&\qquad + \epsilon \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \langle \Psi_d, q_x \mathcal{N}_x^2 q_x \Psi_d \rangle + \frac{1}{d},
\end{aligned} \tag{87}$$

Proof. Recalling the definition of $\tilde{H}$ in (81), we find

$$\frac{1}{|\Lambda|} \langle \Psi_d, \tilde{H} \Psi_d \rangle = -\frac{J}{2d} \frac{1}{|\Lambda|} \sum_{<x,y>} \langle \Psi_d, p_x p_y K_{x,y}^{(2)} q_x q_y \Psi_d \rangle + h.c. \tag{88}$$

$$- \frac{J}{2d} \frac{1}{|\Lambda|} \sum_{<x,y>} \langle \Psi_d, p_x q_y K_{x,y}^{(2)} q_x p_y \Psi_d \rangle + h.c. \tag{89}$$

$$- \frac{J}{d} \frac{1}{|\Lambda|} \sum_{<x,y>} \langle \Psi_d, p_x q_y K_{x,y}^{(3)} q_x q_y \Psi_d \rangle + h.c. \tag{90}$$

$$- \frac{J}{2d} \frac{1}{|\Lambda|} \sum_{<x,y>} \langle \Psi_d, q_x q_y K_{x,y}^{(4)} q_x q_y \Psi_d \rangle, \tag{91}$$

where the terms $K_{x,y}^{(2)}$, $K_{x,y}^{(3)}$ and $K_{x,y}^{(4)}$ are defined in (82), (83) and (84). Let us estimate the above equation term by term. The *pp*-*qq* term of (88) has already been estimated in the proof of Proposition 13. Here, we find it slightly more convenient to choose $\epsilon^{-1} \coloneqq 2d \mathrm{Tr}\,(p\mathcal{N}) (\mathrm{Tr}\,(p\mathcal{N}) + 1)$, so instead of (70) we arrive at

$$|(88)| \leqslant \left(1 + J^2 \langle \varphi(0), (\mathcal{N} + 1) \varphi(0) \rangle^2 \right) \frac{1}{|\Lambda|} \sum_{x \in \Lambda} \langle \Psi_d, q_x \Psi_d \rangle + \frac{1}{d}. \tag{92}$$

For the pq-qp term of (89) we find, using Cauchy–Schwarz,

$$\begin{aligned}
|(89)| &\leqslant 2\left|-\frac{J}{2d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\langle a_x p_x q_y \Psi_d, a_y p_y q_x \Psi_d\rangle\right| \\
&\leqslant \frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\langle \varphi(0), \mathcal{N}\varphi(0)\rangle \|q_x\Psi_d\|\|q_y\Psi_d\| \\
&\leqslant 2J\langle \varphi(0), \mathcal{N}\varphi(0)\rangle \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\Psi_d\rangle.
\end{aligned}$$

The pq-qq terms of (90) have already been estimated in the proof of Proposition 13. We find it convenient to introduce $\epsilon > 0$, so continuing from (71) and using Cauchy–Schwarz, we find

$$|(90)| \leqslant 2\left(3J + \frac{J^2}{\epsilon} + 4J\langle \varphi(0), \mathcal{N}\varphi(0)\rangle\right)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\Psi_d\rangle + 2\epsilon\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\mathcal{N}_x^2 q_x\Psi_d\rangle. \tag{93}$$

Finally, the terms involving a^*a in (91) can be estimated directly with Cauchy–Schwarz. We find

$$\begin{aligned}
&\left|-\frac{J}{2d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\langle \Psi_d, q_x q_y a_x^* a_y q_x q_y \Psi_d\rangle\right| \\
&\leqslant \frac{J}{2d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\sqrt{\langle \Psi_d, q_y q_x \mathcal{N}_x q_x \Psi_d\rangle}\sqrt{\langle \Psi_d, q_x q_y \mathcal{N}_y q_y \Psi_d\rangle} \\
&\leqslant \frac{1}{2d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\sqrt{J\|q_x\Psi_d\|\|\mathcal{N}_x q_x\Psi_d\|}\sqrt{J\|q_y\Psi_d\|\|\mathcal{N}_y q_y\Psi_d\|} \\
&\leqslant \frac{1}{2d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\left(\frac{J^2}{2\epsilon}\|q_x\Psi_d\|^2 + \frac{\epsilon}{2}\|\mathcal{N}_x q_x\Psi_d\|^2\right)^{1/2}\left(\frac{J^2}{2\epsilon}\|q_y\Psi_d\|^2 + \frac{\epsilon}{2}\|\mathcal{N}_y q_y\Psi_d\|^2\right)^{1/2} \\
&\leqslant \left(\frac{J^2}{2\epsilon}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\|q_x\Psi_d\|^2 + \frac{\epsilon}{2}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\|\mathcal{N}_x q_x\Psi_d\|^2\right)^{1/2}\left(\frac{J^2}{2\epsilon}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\|q_x\Psi_d\|^2 + \frac{\epsilon}{2}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\|\mathcal{N}_x q_x\Psi_d\|^2\right)^{1/2} \\
&\leqslant \frac{J^2}{2\epsilon}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\Psi_d\rangle + \frac{\epsilon}{2}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\mathcal{N}_x^2 q_x\Psi_d\rangle.
\end{aligned}$$

The terms involving α_φ can be estimated in the same way, using additionally that

$$|\alpha_\varphi(t)| = |\langle \varphi(t), a\varphi(t)\rangle| \leqslant \|a\varphi(t)\| = \sqrt{\langle \varphi(0), \mathcal{N}\varphi(0)\rangle}. \tag{94}$$

Combining these bounds yields

$$|(91)| \leqslant C\left(\frac{1}{\epsilon} + \frac{J^2}{\epsilon} + J^2\langle \varphi(0), \mathcal{N}\varphi(0)\rangle\right)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\Psi_d\rangle + C\epsilon\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\mathcal{N}_x^2 q_x\Psi_d\rangle. \tag{95}$$

Thus, altogether, we get for some $C > 0$ that

$$\begin{aligned}
\left|\frac{1}{|\Lambda|}\langle \Psi_d, \tilde{H}\Psi_d\rangle\right| &\leqslant C\epsilon\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\mathcal{N}_x^2 q_x\Psi_d\rangle + \frac{1}{d} \\
&\quad + C\left(1+J^2\right)\left(1 + \frac{1}{\epsilon} + \langle \varphi(0), \mathcal{N}\varphi(0)\rangle + \langle \varphi(0), \mathcal{N}\varphi(0)\rangle^2\right)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle \Psi_d, q_x\Psi_d\rangle.
\end{aligned} \tag{96}$$

Similarly, we can use Cauchy–Schwarz and again (94) to show that for some $C > 0$,

$$\begin{aligned}
&\left| \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \left(h_x^{\alpha_\varphi} - \frac{U}{2}\mathcal{N}_x^2 \right) q_x \Psi_d \rangle \right| \\
&\quad = \left| \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \left\langle \Psi_d, q_x \left(-J\alpha_\varphi a_x^* - J\overline{\alpha_\varphi} a_x + J|\alpha_\varphi|^2 + \left(J - \mu - \frac{U}{2} \right) \mathcal{N}_x \right) q_x \Psi_d \right\rangle \right| \\
&\quad \leqslant C \left(1 + J^2 + \left(J - \mu - \frac{U}{2} \right)^2 \right) \left(\frac{1}{\epsilon} + \langle \varphi(0), \mathcal{N}\varphi(0) \rangle \right) \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \Psi_d \rangle \\
&\qquad + \epsilon \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \mathcal{N}_x^2 q_x \Psi_d \rangle.
\end{aligned} \tag{97}$$

Combining both bounds yields (87). □

This proposition allows us to prove the equivalence of f and g up to an error d^{-1}.

Proposition 17. *Let f and g be defined as in* (85) *and* (86)*. Then, for $U > 0$ and for some $C > 0$, we have the equivalence*

$$\frac{U}{4} g - \frac{1}{d} \leqslant f \leqslant C \left(1 + J^2 + U + \left(J - \mu - \frac{U}{2} \right)^2 \right) \left(1 + \frac{1}{U} + \langle \varphi(0), \mathcal{N}\varphi(0) \rangle^2 \right) g + \frac{1}{d}. \tag{98}$$

Proof. We start by proving the lower bound on f from (98). From Proposition 16 we know that

$$\begin{aligned}
&\frac{1}{|\Lambda|} \left\langle \Psi_d, \left(\tilde{H} + \sum_{x\in\Lambda} q_x \left(h_x^{\alpha_\varphi} - \frac{U}{2}\mathcal{N}_x^2 \right) q_x \right) \Psi_d \right\rangle \\
&\quad \geqslant -C \left(1 + J^2 + \left(J - \mu - \frac{U}{2} \right)^2 \right) \left(1 + \frac{1}{\epsilon} + \langle \varphi(0), \mathcal{N}\varphi(0) \rangle^2 \right) \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \Psi_d \rangle - \frac{1}{d} \\
&\qquad - \epsilon \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \mathcal{N}_x^2 q_x \Psi_d \rangle.
\end{aligned} \tag{99}$$

Hence,

$$\begin{aligned}
f &= \frac{1}{|\Lambda|} \left\langle \Psi_d, \left(\tilde{H} + \sum_{x\in\Lambda} q_x \left(h_x^{\alpha_\varphi} - \frac{U}{2}\mathcal{N}_x^2 \right) q_x \right) \Psi_d \right\rangle \\
&\quad + \frac{U}{2} \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \mathcal{N}_x^2 q_x \Psi_d \rangle + \frac{c}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \Psi_d \rangle \\
&\geqslant \left(\frac{U}{2} - \epsilon \right) \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \mathcal{N}_x^2 q_x \Psi_d \rangle \\
&\quad + \left(c - C \left(1 + J^2 + \left(J - \mu - \frac{U}{2} \right)^2 \right) \left(1 + \frac{1}{\epsilon} + \langle \varphi(0), \mathcal{N}\varphi(0) \rangle^2 \right) \right) \frac{1}{|\Lambda|} \sum_{x\in\Lambda} \langle \Psi_d, q_x \Psi_d \rangle - \frac{1}{d}.
\end{aligned}$$

Then the lower bound on f from (98) follows by choosing

$$c = C \left(1 + J^2 + \left(J - \mu - \frac{U}{2} \right)^2 \right) \left(1 + \frac{1}{\epsilon} + \langle \varphi(0), \mathcal{N}\varphi(0) \rangle^2 \right) + \frac{U}{4}, \qquad \epsilon = \frac{U}{4}. \tag{100}$$

For the upper bound on f from (98), note that

$$\begin{aligned}
&\left|\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d, q_x h_x^{\alpha_\varphi} q_x \Psi_d\rangle\right| \\
&= \left|\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left\langle\Psi_d, q_x\left(-J\alpha_\varphi a_x^* - J\overline{\alpha_\varphi}a_x + J|\alpha_\varphi|^2 + (J-\mu)\mathcal{N}_x + \frac{U}{2}\mathcal{N}_x(\mathcal{N}_x - 1)\right)q_x\Psi_d\right\rangle\right| \qquad (101)\\
&\leqslant C\left(1+|J|+U+\left|J-\mu-\frac{U}{2}\right|\right)(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x^2 q_x + q_x)\Psi_d\rangle.
\end{aligned}$$

Using this and the bound (96) from the proof of Proposition 16 for $\epsilon = 1$, the choice (100) for the constant c yields

$$\begin{aligned}
|f| &= \left|\frac{1}{|\Lambda|}\langle\Psi_d,\tilde{H}\Psi_d\rangle + \frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d, q_x h_x^{\alpha_\varphi} q_x\Psi_d\rangle + \frac{c}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d, q_x\Psi_d\rangle\right| \\
&\leqslant C\left(1+J^2+U+\left(J-\mu-\frac{U}{2}\right)^2\right)\left(1+\frac{1}{U}+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^2\right)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x^2 q_x + q_x)\Psi_d\rangle \\
&\quad + \frac{1}{d}.
\end{aligned}$$

□

6.2 Proof of Gronwall Estimate for f

In the computation of the time derivative of f we need to control in particular $\dot{\tilde{H}}$, the time derivative of $\tilde{H}$. Its computation in straightforward but a bit lengthy. The key point is to write this time derivative in such a way that it contains the commutator $[\tilde{H}, q_x h_x^{\alpha_\varphi} q_x - h_x^{\alpha_\varphi}]$, which we will later use for cancellations.

Proposition 18. *The expectation of $\dot{\tilde{H}}$ can be written as*

$$\begin{aligned}
\frac{1}{|\Lambda|}\langle\Psi_d,\dot{\tilde{H}}\Psi_d\rangle &= -\frac{i}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,[\tilde{H}, q_x h_x^{\alpha_\varphi} q_x - h_x^{\alpha_\varphi}]\Psi_d\rangle + \mathcal{R} \\
&= \frac{J}{2d}\frac{i}{|\Lambda|}\sum_{<x,y>}\langle\Psi_d,[\tilde{H}_{x,y}, q_x h_x^{\alpha_\varphi} q_x - h_x^{\alpha_\varphi} + q_y h_y^{\alpha_\varphi} q_y - h_y^{\alpha_\varphi}]\Psi_d\rangle + \mathcal{R},
\end{aligned} \qquad (102)$$

with $\tilde{H}_{x,y}$ refers to the terms in (81) *such that $\tilde{H} = \sum_{<x,y>}\tilde{H}_{x,y}$ and where the rest term $\mathcal{R}\equiv\mathcal{R}(t)$ is given by*

$$\mathcal{R} := -\frac{J}{d}\frac{i}{|\Lambda|}\sum_{<x,y>}\left\langle\Psi_d,\left(p_x h_x^{\alpha_\varphi} p_y K_{x,y}^{(2)} q_x q_y + q_x h_x^{\alpha_\varphi} p_x q_y K_{x,y}^{(2)} p_x p_y\right)\Psi_d\right\rangle + h.c. \qquad (103)$$

$$-\frac{J}{d}\frac{i}{|\Lambda|}\sum_{<x,y>}\left\langle\Psi_d, q_y p_x\left(h_x^{\alpha_\varphi} + h_y^{\alpha_\varphi} p_y\right)K_{x,y}^{(2)} q_x p_y\Psi_d\right\rangle + h.c. \qquad (104)$$

$$\begin{aligned}
&+\frac{J}{d}\frac{i}{|\Lambda|}\sum_{<x,y>}\Big\langle\Psi_d, q_y\Big((h_x^{\alpha_\varphi} p_x - p_x h_x^{\alpha_\varphi} - p_x h_y^{\alpha_\varphi} p_y)K_{x,y}^{(3)} \\
&\qquad + p_x K_{x,y}^{(3)}(p_x h_x^{\alpha_\varphi} + p_y h_y^{\alpha_\varphi})\Big)q_x q_y\Psi_d\Big\rangle + h.c.
\end{aligned} \qquad (105)$$

$$+\frac{J}{d}\frac{i}{|\Lambda|}\sum_{<x,y>}\langle\Psi_d, q_x q_y K_{x,y}^{(4)} p_x h_x^{\alpha_\varphi} q_x q_y\Psi_d\rangle + h.c. \qquad (106)$$

$$-\frac{J}{2d}\frac{1}{|\Lambda|}\sum_{<x,y>}\left\langle\Psi_d,\left(p_x q_y\dot{K}_{x,y}^{(3)} q_x q_y + \frac{1}{2}q_x q_y\dot{K}_{x,y}^{(4)} q_x q_y\right)\Psi_d\right\rangle + h.c.. \qquad (107)$$

Proof. We start by gathering some useful computations,

$$\dot{\alpha}_\varphi = i\mu\alpha_\varphi - iU\langle\varphi, \mathcal{N}a\varphi\rangle, \tag{108}$$
$$\overline{\alpha_\varphi}\dot{\alpha}_\varphi + \alpha_\varphi\dot{\overline{\alpha_\varphi}} = 2U\mathrm{Im}\left(\langle\varphi, \mathcal{N}a\varphi\rangle\overline{\alpha_\varphi}\right), \tag{109}$$
$$\dot{h}_x^{\alpha_\varphi} = -J\dot{\alpha}_\varphi a_x^* - J\dot{\overline{\alpha_\varphi}}a_x + 2JU\mathrm{Im}\left(\langle\varphi, \mathcal{N}a\varphi\rangle\overline{\alpha_\varphi}\right), \tag{110}$$
$$\dot{K}_{x,y}^{(2)} = 0, \tag{111}$$
$$\dot{K}_{x,y}^{(3)} = -\dot{\alpha}_\varphi a_x^* - \dot{\overline{\alpha_\varphi}}a_x, \tag{112}$$
$$\dot{K}_{x,y}^{(4)} = -\dot{\alpha}_\varphi(a_x^* + a_y^*) - \dot{\overline{\alpha_\varphi}}(a_x + a_y) + 4U\mathrm{Im}\left(\langle\varphi, \mathcal{N}a\varphi\rangle\overline{\alpha_\varphi}\right). \tag{113}$$

Starting from the definition (81) of $\tilde{H}$, using these relations and $i\dot{p}_x = [h_x^{\alpha_\varphi}, p_x]$, $i\dot{q}_x = [h_x^{\alpha_\varphi}, q_x]$, we arrive at

$$\begin{aligned}
\dot{\tilde{H}} = \frac{iJ}{2d}\sum_{<x,y>}\Bigg(&[h_x^{\alpha_\varphi}, p_x]p_yK_{x,y}^{(2)}q_xq_y + p_x[h_y^{\alpha_\varphi}, p_y]K_{x,y}^{(2)}q_xq_y + p_xp_yK_{x,y}^{(2)}[h_x^{\alpha_\varphi}, q_x]q_y + p_xp_yK_{x,y}^{(2)}q_x[h_y^{\alpha_\varphi}, q_y]\\
&+ [h_x^{\alpha_\varphi}, p_x]q_yK_{x,y}^{(2)}q_xp_y + p_x[h_y^{\alpha_\varphi}, q_y]K_{x,y}^{(2)}q_xp_y + p_xq_yK_{x,y}^{(2)}[h_x^{\alpha_\varphi}, q_x]p_y + p_xq_yK_{x,y}^{(2)}q_x[h_y^{\alpha_\varphi}, p_y]\\
&+ [h_x^{\alpha_\varphi}, q_x]q_yK_{x,y}^{(2)}p_xp_y + q_x[h_y^{\alpha_\varphi}, q_y]K_{x,y}^{(2)}p_xp_y + q_xq_yK_{x,y}^{(2)}[h_x^{\alpha_\varphi}, p_x]p_y + q_xq_yK_{x,y}^{(2)}p_x[h_y^{\alpha_\varphi}, p_y]\\
&+ [h_x^{\alpha_\varphi}, q_x]p_yK_{x,y}^{(2)}p_xq_y + q_x[h_y^{\alpha_\varphi}, p_y]K_{x,y}^{(2)}p_xq_y + q_xp_yK_{x,y}^{(2)}[h_x^{\alpha_\varphi}, p_x]q_y + q_xp_yK_{x,y}^{(2)}p_x[h_y^{\alpha_\varphi}, q_y]\Bigg)\\
+ \frac{iJ}{d}\sum_{<x,y>}\Bigg(&[h_x^{\alpha_\varphi}, p_x]q_yK_{x,y}^{(3)}q_xq_y + p_x[h_y^{\alpha_\varphi}, q_y]K_{x,y}^{(3)}q_xq_y + p_xq_yK_{x,y}^{(3)}[h_x^{\alpha_\varphi}, q_x]q_y + p_xq_yK_{x,y}^{(3)}q_x[h_y^{\alpha_\varphi}, q_y]\\
&+ [h_x^{\alpha_\varphi}, q_x]q_yK_{x,y}^{(3)}p_xq_y + q_x[h_y^{\alpha_\varphi}, q_y]K_{x,y}^{(3)}p_xq_y + q_xq_yK_{x,y}^{(3)}[h_x^{\alpha_\varphi}, p_x]q_y + q_xq_yK_{x,y}^{(3)}p_x[h_y^{\alpha_\varphi}, q_y]\Bigg)\\
+ \frac{iJ}{2d}\sum_{<x,y>}\Bigg(&[h_x^{\alpha_\varphi}, q_x]q_yK_{x,y}^{(4)}q_xq_y + q_x[h_y^{\alpha_\varphi}, q_y]K_{x,y}^{(4)}q_xq_y + q_xq_yK_{x,y}^{(4)}[h_x^{\alpha_\varphi}, q_x]q_y + q_xq_yK_{x,y}^{(4)}q_x[h_y^{\alpha_\varphi}, q_y]\Bigg)\\
+ \frac{J}{d}\sum_{<x,y>}\Bigg[&p_xq_y\big(\dot{\alpha}_\varphi a_x^* - \dot{\overline{\alpha_\varphi}}a_x\big)q_xq_y + q_xq_y\big(\dot{\alpha}_\varphi a_x^* - \dot{\overline{\alpha_\varphi}}a_x\big)p_xq_y\Bigg]\\
+ \frac{J}{2d}\sum_{<x,y>}&q_xq_y\Bigg(\dot{\alpha}_\varphi(a_x^* + a_y^*) + \dot{\overline{\alpha_\varphi}}(a_x + a_y) - 4U\mathrm{Im}(\langle\varphi, \mathcal{N}a\varphi\rangle\overline{\alpha_\varphi})\Bigg)q_xq_y.
\end{aligned}$$

To obtain (102), we isolate the first part on the right-hand side of (102) and define the rest as the remainder term $\mathcal{R}$. □

Next, we estimate the rest term in Proposition 18.

Proposition 19. *The rest term $\mathcal{R}(t)$ in Proposition 18 satisfies the bound*

$$|\mathcal{R}(t)| \leqslant \tilde{C}(t)\left(\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d(t), (q_x(t)\mathcal{N}_xq_x(t) + q_x(t))\Psi_d(t)\rangle + \frac{1}{d}\right),$$

where

$$\begin{aligned}\tilde{C}(t) &= C(J,\mu,U)\left(1 + \langle\varphi(0), \mathcal{N}\varphi(0)\rangle^2\right)\\ &\quad\left(1 + \sum_{j=0}^{6}\left(8J\langle\varphi(0), \mathcal{N}\varphi(0)\rangle^{1/2}\right)^j\left\langle\varphi(0), (\mathcal{N}+j)^{4-\frac{j}{2}}\varphi(0)\right\rangle\frac{t^j}{j!}\right),\end{aligned} \tag{114}$$

with $C(J,\mu,U) > 0$ depending polynomially on the model parameters J, μ and U.

Proof. We need to estimate each term in $\mathcal{R}$. We start by explaining in detail how to estimate one of the terms in (103). By Cauchy–Schwarz and Hölder's inequality we have

$$
\begin{aligned}
&\left|\frac{J}{d|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\ x\sim y}}\langle\Psi_d, p_x h_x^{\alpha_\varphi} p_y a_x^* a_y q_x q_y \Psi_d\rangle\right| \\
&\leqslant \frac{J}{d|\Lambda|}\sum_{x\in\Lambda}\|a_x^* q_x\Psi_d\|\left\|h_x^{\alpha_\varphi} p_x\left(\sum_{\substack{y\in\Lambda\\ x\sim y}} q_y a_y^* p_y\right)\Psi_d\right\| \\
&\leqslant \frac{2J}{|\Lambda|}\sum_{x\in\Lambda}\sqrt{\langle\varphi(0),(\mathcal{N}+1)\varphi(0)\rangle}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\|a_x^* q_x\Psi_d\|\left(\frac{1}{2d}+\frac{1}{(2d)^2}\sum_{\substack{y\in\Lambda\\ y\neq z, x\sim y}}\sum_{\substack{z\in\Lambda\\ x\sim z}}\|q_z\Psi_d\|\|q_y\Psi_d\|\right)^{\frac{1}{2}} \\
&\leqslant \frac{2J}{|\Lambda|}\sum_{x\in\Lambda}\sqrt{\langle\varphi(0),(\mathcal{N}+1)\varphi(0)\rangle}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\|a_x^* q_x\Psi_d\|\left(\frac{1}{2d}+\frac{1}{4d}\sum_{\substack{y\in\Lambda\\ y\neq z, x\sim y}}\|q_y\Psi_d\|^2+\frac{1}{4d}\sum_{\substack{z\in\Lambda\\ x\sim z}}\|q_z\Psi_d\|^2\right)^{\frac{1}{2}} \\
&\leqslant CJ\sqrt{\langle\varphi(0),(\mathcal{N}+1)\varphi(0)\rangle}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\left(\frac{1}{|\Lambda|}\sum_x\|a_x^* q_x\Psi_d\|^2+\frac{1}{d}\right. \\
&\qquad\qquad\left.+\frac{1}{d}\frac{1}{|\Lambda|}\underbrace{\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\ x\sim y}}\|q_y\Psi_d\|^2}_{=2d\sum_{x\in\Lambda}\|q_x\Psi_d\|^2}+\frac{1}{d}\frac{1}{|\Lambda|}\underbrace{\sum_{x\in\Lambda}\sum_{\substack{z\in\Lambda\\ x\sim z}}\|q_z\Psi_d\|^2}_{=2d\sum_{x\in\Lambda}\|q_x\Psi_d\|^2}\right) \\
&\leqslant CJ\sqrt{\langle\varphi(0),(\mathcal{N}+1)\varphi(0)\rangle}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\left(\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x+q_x)\Psi_d\rangle+\frac{1}{d}\right).
\end{aligned}
$$

The other terms of (103) can be estimated analogously, so we arrive at

$$
|(103)|\leqslant CJ\left(1+\langle\varphi(0),(\mathcal{N})\varphi(0)\rangle\right)\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\left(\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x+q_x)\Psi_d\rangle+\frac{1}{d}\right).
$$

In order to bound (104), we use Cauchy–Schwarz to find

$$
\begin{aligned}
&\left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\ x\sim y}}\langle\Psi_d, q_y p_x h_x^{\alpha_\varphi} a_x^* a_y q_x p_y\Psi_d\rangle\right| \\
&\quad=\left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\ x\sim y}}\langle h_x^{\alpha_\varphi} p_x a_y^* q_y\Psi_d, a_x^* q_x p_y\Psi_d\rangle\right| \\
&\quad\leqslant\frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\ x\sim y}}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\left(\|q_x\Psi_d\|^2+\|\mathcal{N}_x^{1/2}q_x\Psi_d\|^2\right)^{1/2}\left(\|q_y\Psi_d\|^2+\|\mathcal{N}_y^{1/2}q_y\Psi_d\|^2\right)^{1/2} \\
&\quad\leqslant CJ\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x+q_x)\Psi_d\rangle.
\end{aligned}
$$

Estimating the other terms of (104) in an analogous way, we get

$$
|(104)|\leqslant CJ\left(1+\sqrt{\langle\varphi(0),(\mathcal{N})\varphi(0)\rangle}\right)\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x+q_x)\Psi_d\rangle.
$$

To bound (105), we use Cauchy–Schwarz to estimate

$$\begin{aligned}
&\left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\langle\Psi_d, q_y h_x^{\alpha_\varphi} p_x a_x^* a_y q_x q_y \Psi_d\rangle\right| \\
&\quad= \left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\langle a_y^* q_y\Psi_d, h_x^{\alpha_\varphi} p_x a_x^* q_x q_y \Psi_d\rangle\right| \\
&\quad\leqslant \frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\|a_y^* q_y\Psi_d\|\Big(\langle\Psi_d, q_x a_x \underbrace{p_x (h_x^{\alpha_\varphi})^2 p_x}_{=\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle p_x} a_x^* q_x\Psi_d\rangle\Big)^{\frac{1}{2}} \\
&\quad\leqslant \frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\Big(\langle\Psi_d, q_y(\mathcal{N}_y+1)q_y\Psi_d\rangle\Big)^{\frac{1}{2}}\Big(\langle\Psi_d, q_x(\mathcal{N}_x+1)q_x\Psi_d\rangle\Big)^{\frac{1}{2}} \\
&\quad\leqslant CJ\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x + q_x)\Psi_d\rangle,
\end{aligned}$$

and similarly

$$\begin{aligned}
&\left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\langle\Psi_d, q_y p_x h_y^{\alpha_\varphi} p_y a_x^* a_y q_x q_y \Psi_d\rangle\right| \\
&\quad\leqslant \frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle}\|q_y\Psi_d\|\left(\left\langle\Psi_d, q_x q_y a_y^* \underbrace{p_y (h_y^{\alpha_\varphi})^2 p_y}_{=\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle p_y} a_y q_y\Psi_d\right\rangle\right)^{1/2} \\
&\quad= \frac{J}{d}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\sum_{\substack{y\in\Lambda\\x\sim y}}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle}\|q_y\Psi_d\|\left(\langle\Psi_d, q_x q_y a_y^* p_y a_y q_y\Psi_d\rangle\right)^{1/2} \\
&\quad\leqslant CJ\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x + q_x)\Psi_d\rangle.
\end{aligned}$$

For (106), we directly find

$$\begin{aligned}
\left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{<x,y>}\langle\Psi_d, q_x q_y a_x^* a_y p_x h_x^{\alpha_\varphi} q_x q_y\Psi_d\rangle\right| &= \left|\frac{J}{d}\frac{1}{|\Lambda|}\sum_{<x,y>}\langle h_x^{\alpha_\varphi} p_x a_x q_x q_y\Psi_d, a_y q_x q_y\Psi_d\rangle\right| \\
&\leqslant CJ\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d, q_x\mathcal{N}_x q_x\Psi_d\rangle.
\end{aligned}$$

Estimating the other terms in an analogous way, we obtain

$$|(104)+(105)+(106)| \leqslant CJ\left(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle\right)\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_x q_x + q_x)\Psi_d\rangle.$$

Using

$$\begin{aligned}
|\dot{\alpha}| &\leqslant C\left(|\mu|\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle} + U\langle\varphi,(\mathcal{N}+1)^{3/2}\varphi\rangle\right) \\
|\dot{\overline{\alpha_\varphi}}\alpha_\varphi + \overline{\alpha_\varphi}\dot{\alpha}_\varphi| &\leqslant CU\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle}\langle\varphi,(\mathcal{N}+1)^{3/2}\varphi\rangle,
\end{aligned}$$

we furthermore find

$$\begin{aligned}
|(107)| \leqslant& CJ\left(|\mu+U|\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle} + U\langle\varphi,\mathcal{N}^{3/2}\varphi\rangle + U\sqrt{\langle\varphi(0),\mathcal{N}\varphi(0)\rangle}\sqrt{\langle\varphi,(\mathcal{N}+1)^{3/2}\varphi\rangle}\right) \\
&\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,\big(q_x\mathcal{N}_x q_x + q_x\big)\Psi_d\rangle,
\end{aligned}$$

Combining all estimates, we arrive at the bound

$$|\mathcal{R}| \leqslant C(J,\mu,U)\,(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle)\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle}\left(\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\mathcal{N}_xq_x+q_x)\Psi_d\rangle+\frac{1}{d}\right),$$

where $C(J,\mu,U)$ is a polynomial in J, μ and U. We also have by Cauchy–Schwarz

$$\begin{aligned}\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle &= \left\langle\varphi,\left(-J\big(\alpha_\varphi a^*+\overline{\alpha_\varphi}a-|\alpha_\varphi|^2\big)+(J-\mu)\mathcal{N}+\frac{U}{2}\mathcal{N}(\mathcal{N}-1)\right)^2\varphi\right\rangle\\ &\leqslant C(J,\mu,U)\left(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^2\right)\left(1+\langle\varphi,\mathcal{N}^4\varphi\rangle\right).\end{aligned}\tag{115}$$

The proposition is proven by using the propagation bound (56) from Proposition (10) for $k=4$, since then

$$\begin{aligned}\sqrt{\langle\varphi,(h^{\alpha_\varphi})^2\varphi\rangle} &\leqslant C(J,\mu,U)\,(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle)\left(1+\sqrt{\langle\varphi,\mathcal{N}^4\varphi\rangle}\right)\\ &\leqslant C(J,\mu,U)\,(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle)\\ &\quad\left(1+\sum_{j=0}^{6}\frac{\left(8J\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^{1/2}t\right)^j}{j!}\left\langle\varphi(0),(\mathcal{N}+j)^{4-\frac{j}{2}}\varphi(0)\right\rangle\right).\end{aligned}\tag{116}$$

□

With this proposition we can now prove a Gronwall estimate for f.

Proposition 20. *For f as defined in (85), we have for all $t\in\mathbb{R}$,*

$$f(t)\leqslant e^{\int_0^t\tilde{C}(s)ds}f(0)+\frac{1}{d}\int_0^t\left(1+\tilde{C}(s)\right)e^{\int_s^t\tilde{C}(r)dr}ds,\tag{117}$$

with

$$\begin{aligned}\tilde{C}(t) &= \frac{C(J,\mu,U)}{U}\left(1+\frac{1}{U}+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^2\right)\\ &\quad\left(1+\sum_{j=0}^{6}\left(8J\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^{1/2}\right)^j\left\langle\varphi(0),(\mathcal{N}+j)^{4-\frac{j}{2}}\varphi(0)\right\rangle\frac{t^j}{j!}\right),\end{aligned}\tag{118}$$

where $C(J,\mu,U)>0$ is a polynomial in J, μ and U.

Proof. Using $\dot{H}_d=0=\sum_{x\in\Lambda}\dot{h}_x^{\alpha_\varphi}+\dot{\tilde{H}}$, the time derivative of f can be computed as

$$\begin{aligned}\dot{f} &= \frac{i}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,[H_d,H_d+q_xh_x^{\alpha_\varphi}q_x-h_x^{\alpha_\varphi}+cq_x]\Psi_d\rangle+\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\dot{h}_x^{\alpha_\varphi}q_x-\dot{h}_x^{\alpha_\varphi})\Psi_d\rangle\\ &\quad-\frac{i}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,[h_x^{\alpha_\varphi},q_xh_x^{\alpha_\varphi}q_x+cq_x]\Psi_d\rangle\\ &= \frac{i}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,[\tilde{H},q_xh_x^{\alpha_\varphi}q_x-h_x^{\alpha_\varphi}+cq_x]\Psi_d\rangle+\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,(q_x\dot{h}_x^{\alpha_\varphi}q_x+\dot{\tilde{H}})\Psi_d\rangle.\end{aligned}$$

Using Proposition 18 we get

$$\dot{f}=\frac{i}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,[\tilde{H},cq_x]\Psi_d\rangle+\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,q_x\dot{h}_x^{\alpha_\varphi}q_x\Psi_d\rangle+\mathcal{R}.$$

For the first two terms of this expression, we find

$$
\begin{aligned}
&\left|\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d, q_x\dot{h}_x^{\alpha_\varphi}q_x\Psi_d\rangle\right| \\
&\quad=\left|\frac{1}{|\Lambda|}\left\langle\Psi_d, q_x\left(-J\dot{\alpha}_\varphi a_x^* - J\dot{\overline{\alpha_\varphi}}a_x + 2JU\mathrm{Im}(\langle\varphi,\mathcal{N}a\varphi\rangle\overline{\alpha_\varphi})\right)q_x\Psi_d\right\rangle\right| \\
&\quad\leqslant C(J,\mu,U)\left(1+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle\right)\left(1+\sqrt{\langle\varphi,\mathcal{N}^2\varphi\rangle}\right)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left\langle\Psi_d,\left(q_x\mathcal{N}_x^2q_x+q_x\right)\Psi_d\right\rangle,
\end{aligned}
$$

and

$$
\begin{aligned}
&\left|\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,[\tilde{H},cq_x]\Psi_d\rangle\right| \\
&\quad\leqslant C(J,\mu,U)\left(1+\frac{1}{U}+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^2\right)\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left\langle\Psi_d,(q_x\mathcal{N}_x^2q_x+q_x)\Psi_d\right\rangle+\frac{1}{d}.
\end{aligned}
$$

These two estimates, together with the estimate on $\mathcal{R}$ from Proposition 19 and the equivalence of f and g up to an error d^{-1} from Proposition 17 imply

$$
\left|\frac{d}{dt}f(t)\right|\leqslant\tilde{C}(t)\left(f(t)+\frac{1}{d}\right)+\frac{1}{d},
$$

where $\tilde{C}(t)$ depends on the initial data and on the other parameters of our model as defined in (118). With Gronwall's lemma we arrive at (117). $\square$

6.3 Conclusion of the Proof

We combine the above results to prove our second main result.

Proof of Theorem 2. We use the equivalence of f and g up to an error d^{-1} from Proposition 17, and the Gronwall estimate for f from Proposition 20 to find

$$
\begin{aligned}
&\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,q_x\Psi_d\rangle \\
&\quad\leqslant\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d,\left(q_x\mathcal{N}_x^2q_x+q_x\right)\Psi_d\rangle \\
&\quad\leqslant\frac{4}{U}\left(f+\frac{1}{d}\right) \\
&\quad\leqslant\frac{4}{U}e^{\int_0^t\tilde{C}(s)ds}f(0)+\frac{1}{d}\left(\frac{4}{U}+\frac{4}{U}\int_0^t\left(1+\tilde{C}(s)\right)e^{\int_s^t\tilde{C}(r)dr}ds\right) \\
&\quad\leqslant\frac{C}{U}\left(1+J^2+U+\left(J-\mu-\frac{U}{2}\right)^2\right)\left(1+\frac{1}{U}+\langle\varphi(0),\mathcal{N}\varphi(0)\rangle^2\right)e^{\int_0^t\tilde{C}(s)ds} \\
&\qquad\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left\langle\Psi_d(0),(q_x(0)\mathcal{N}_x^2q_x(0)+q_x(0))\Psi_d(0)\right\rangle+\frac{1}{d}\frac{4}{U}\left(1+e^{\int_0^t\tilde{C}(s)ds}+\int_0^t(1+\tilde{C}(s))e^{\int_s^t\tilde{C}(r)dr}ds\right),
\end{aligned}
\tag{119}
$$

where $\tilde{C}(t)$ is defined in (118). Now note that since $\mathrm{Tr}(p(0)\mathcal{N}^4)\leqslant C$, we get that $\tilde{C}(t)$ satisfies

$$
\tilde{C}(t)\leqslant C(J,\mu,U)\left(1+\sum_{j=1}^{6}t^j\right)
$$

where $C(J,\mu,U)>0$ depends polynomially on the parameters of our model J, μ and U. Thus, (119) can be estimated as

$$\begin{aligned}
&\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\langle\Psi_d(t),q_x(t)\Psi_d(t)\rangle \\
&\quad\leqslant \frac{1}{d}\frac{1}{U}+C(J,\mu,U)e^{C(J,\mu,U)\sum_1^7|t|^j}\left(1+\frac{1}{U^2}\right)\left(\frac{1}{|\Lambda|}\sum_{x\in\Lambda}\left\langle\Psi_d(0),\left(q_x(0)\mathcal{N}_x^2q_x(0)+q_x(0)\right)\Psi_d(0)\right\rangle+\frac{1}{d}\right),
\end{aligned} \tag{120}$$

and the Theorem follows from using (66) from Lemma 12. $\square$

References

[1] M. Aizenman, E. Lieb, R. Seiringer, J. Solovej, and J. Yngvason. Bose–Einstein quantum phase transition in an optical lattice model. *Phys. Rev. A*, 70:023612, 2004.

[2] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell. Observation of Bose–Einstein condensation in a dilute atomic vapor. *Science*, 269(5221):198–201, 1995.

[3] N. Benedikter, M. Porta, and B. Schlein. *Effective Evolution Equations from Quantum Dynamics.* SpringerBriefs in Mathematical Physics. Springer, 2016.

[4] I. Bloch, J. Dalibard, and W. Zwerger. Many-body physics with ultracold gases. *Rev. Mod. Phys.*, 80:885–964, 2007.

[5] C. Boccato, C. Brennecke, S. Cenatiempo, and B. Schlein. Bogoliubov theory in the Gross–Pitaevskii limit. *Acta Math.*, 222(2):219–335, 2019.

[6] Bose. Plancks Gesetz und Lichtquantenhypothese. *Z. Phys.*, 26:178–181, 1924.

[7] L. Boßmann, S. Petrat, P. Pickl, and A. Soffer. Beyond Bogoliubov dynamics. *Pure Appl. Anal.*, 3(4):677–726, 2021.

[8] L. Boßmann, S. Petrat, and R. Seiringer. Asymptotic expansion of low-energy excitations for weakly interacting bosons. *Forum Math. Sigma*, 9:e28, 2021.

[9] J.-B. Bru and T. C. Dorlas. Exact solution of the infinite-range-hopping Bose–Hubbard model. *J. Stat. Phys.*, 113, 2003.

[10] K. Byczuk and D. Vollhardt. Correlated bosons on a lattice: Dynamical mean-field theory for Bose–Einstein condensed and normal phases. *Phys. Rev. B*, 77(235106), 2008.

[11] T. Cazenave and A. Haraux. *An Introduction to Semilinear Evolution Equations.* Oxford lecture series in mathematics and its applications, 1998.

[12] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle. Bose–Einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.*, 75(22):3969–3973, 1995.

[13] J. Dereziński and M. Napiórkowski. Excitation spectrum of interacting bosons in the mean-field infinite-volume limit. *Ann. Henri Poincaré*, 15(12):2409–2439, 2014.

[14] A. Einstein. Quantentheorie des einatomigen idealen Gases. *Sitzber. Kgl. Preuss. Akad. Wiss.*, pages 261–267, 1924.

[15] A. Einstein. Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. *Sitzber. Kgl. Preuss. Akad. Wiss.*, pages 3–14, 1925.

[16] L. Erdős and H.-T. Yau. Derivation of the nonlinear Schrödinger equation from a many-body Coulomb system. *Adv. Theor. and Math. Phys.*, 5(6):1169–1205, 2001.

[17] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher. Boson localization and the superfluid-insulator transition. *Phys. Rev. B*, 40:546–570, Jul 1989.

[18] S. Fournais and J. P. Solovej. The energy of dilute Bose gases. *Ann. Math.*, 192:893–976, 2020.

[19] S. Fournais and J. P. Solovej. The energy of dilute Bose gases II: the general case. *Invent. Math.*, 232:863–994, 2022.

[20] A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68:13–125, 01 1996.

[21] F. Golse. *On the Dynamics of Large Particle Systems in the Mean Field Limit*, chapter I, page 1–144. Springer International Publishing, 2016.

[22] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature*, 415:39–44, 2002.

[23] M. Greiner, O. Mandel, T. Rom, A. Altmeyer, A. Widera, T. Hänsch, and I. Bloch. Quantum phase transition from a superfluid to a Mott insulator in an ultracold gas of atoms. *Physica B: Condensed Matter*, 329-333:11–12, 2003. Proceedings of the 23rd International Conference on Low Temperature Physics.

[24] M. C. Gutzwiller. Effect of correlation on the ferromagnetism of transition metals. *Phys. Rev. Lett.*, 10, 1963.

[25] W.-J. Hu and N.-H. Tong. Dynamical mean-field theory for the Bose–Hubbard model. *Phys. Rev. B*, 80:245110, Dec 2009.

[26] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller. Cold bosonic atoms in optical lattices. *Phys. Rev. Lett.*, 81:3108–3111, Oct 1998.

[27] A. Knowles and P. Pickl. Mean-field dynamics: singular potentials and rate of convergence. *Commun. Math. Phys.*, 298(1):101–138, 2010.

[28] M. Lewin, P. T. Nam, and B. Schlein. Fluctuations around Hartree states in the mean field regime. *Amer. J. Math.*, 137(6):1613–1650, 2015.

[29] M. Lewin, P. T. Nam, S. Serfaty, and J. P. Solovej. Bogoliubov spectrum of interacting Bose gases. *Commun. Pure Appl. Math.*, 68(3):413–471, 2015.

[30] W. Metzner and D. Vollhardt. Correlated lattice fermions in $d = \infty$ dimensions. *Phys. Rev. Lett.*, 62:324–327, 1989.

[31] E. Pavarini, E. Koch, A. Lichtenstein, and D. Vollhardt. *Dynamical Mean-Field Theory of Correlated Electrons*. Institute for Advanced Simulation, 2022.

[32] E. Picari, A. Ponno, and L. Zanelli. Mean field derivation of DNLS from the Bose-–Hubbard model. *Ann. Henri Poincaré*, 23, 10 2021.

[33] P. Pickl. A simple derivation of mean field limits for quantum systems. *Lett. Math. Phys.*, 97(2):151–164, 2011.

[34] D. S. Rokhsar and B. G. Kotliar. Gutzwiller projection for bosons. *Phys. Rev. B*, 44, 1991.

[35] B. Schlein. Bose gases in the Gross–Pitaevskii limit: A survey of some rigorous results. In R. Frank, A. Laptev, M. Lewin, and R. Seiringer, editors, *The Physics and Mathematics of Elliott Lieb*, volume II, pages 277–305. EMS Press, 2022.

[36] B. Simon. *Trace Ideals and Their Applications.* Mathematical surveys and monographs, 2005.

[37] W. Zwerger. Mott–Hubbard transition of cold atoms in optical lattices. *J. Opt. B*, 5:9–16, 2003.