Forum of Mathematics, Pi (2023), **Vol. 11**:e32 1–131
doi:10.1017/fmp.2023.31

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

The Random Phase Approximation for Interacting Fermi Gases in the Mean-Field Regime

Martin Ravn Christiansen, Christian Hainzl and Phan Thành Nam

Department of Mathematics, Ludwig Maximilian University of Munich, Germany;
E-mails: christiansen@math.lmu.de, hainzl@math.lmu.de, nam@math.lmu.de.

Received: 4 May 2022; **Revised:** 25 May 2023; **Accepted:** 6 November 2023

Abstract

We present a general approach to justify the random phase approximation for the homogeneous Fermi gas in three dimensions in the mean-field scaling regime. We consider a system of N fermions on a torus, interacting via a two-body repulsive potential proportional to $N^{-\frac{1}{3}}$. In the limit $N \to \infty$, we derive the exact leading order of the correlation energy and the bosonic elementary excitations of the system, which are consistent with the prediction of the random phase approximation in the physics literature.

Contents

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

1. Introduction

In the 1940s, experiments on the cohesive energy and specific heat of alkali atoms[1] showed a large discrepancy with theoretical calculations based solely on the Hartree–Fock approximation [3], further complicated by the fact that second-order perturbation theory failed because it yielded infinities. Motivated by this unfortunate situation, Bohm and Pines in four seminal papers [11, 12, 13, 32] introduced the random phase approximation (RPA) as a useful tool for studying the properties of a high-density electron gas moving in a background of uniform positive charge, called jellium. In the Bohm–Pines RPA approach, the electron gas could be decoupled into collective plasmon excitations and quasi-electrons that interacted via a screened Coulomb interaction. The latter fact justified the independent particle approach commonly used for many-body fermion systems. Their work was also in good agreement with experimental data, the culmination of which was the experimental detection of plasmons [42, 17].

The microscopic derivation of the RPA has led to notable work by theoretical physicists since the 1950s. In 1957, Gell-Mann and Brueckner [20] derived the correlation energy of the electron gas in the high density limit by using a formal summation of a particular class of Feynman diagrams. Although each diagram is divergent in itself, it turned out that the sum is finite. This diagrammatic picture further suggested that the main contribution to the ground-state energy came from the interaction of pairs of fermions, one from inside and one from outside the Fermi ball. Shortly thereafter, Sawada [36] and Sawada–Brueckner–Fukuda–Brout [37] interpreted these pairs as bosons and obtained the correlation energy by diagonalizing an effective Hamiltonian which is quadratic with respect to the bosonic particle pairs. Since then, the random phase approximation has become a cornerstone in the physics of condensed matter and nuclear physics [34], also playing a significant role in bosonic field theory [26], in the quark-gluon plasma [41] and especially in computational chemistry and materials science. Although originally proposed for an electron gas, it is applicable to a wide variety of fermionic systems.

[1]When calculated in the Hartree–Fock approximation, the cohesive energy of metals is off by an order of magnitude compared to experiments on alkali metals, as described in [33 , p. 80]. The same is true for the specific heat, as theoretically calculated in [3].

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

The complete derivation of the RPA from first principles, namely from the microscopic Schrödinger equation, has, however, long been a major open problem in mathematical physics. Recently, some rigorous results on the correlation energy have been derived in the mean-field regime for small interaction potentials by Hainzl–Porta–Rexze [24] (perturbative results) and by Benedikter–Nam–Porta–Schlein–Seiringer [4, 5, 6] (non-perturbative results).

The aim of the present paper is to justify the RPA for a large class of interaction potentials in the mean-field regime, addressing not only the ground state energy but also the excitation spectrum. As we will explain below, the correlation structure of Fermi gases can indeed be described correctly by treating appropriate pairs of fermions as bosons. The corresponding bosonic Hamiltonian can be handled by Bogolubov's diagonalization method, thus putting the description in the physics literature [20, 36, 37] on a firm mathematical footing. Although this general point of view has been employed in [24, 5, 6], we will provide a new bosonization approach to fermionic systems which enables us to not only extend the study on the ground state energy initiated in [24, 5, 6] but also obtain all bosonic elementary excitations predicted in the physics literature, thus justifying the RPA in the mean-field regime. In the long run, we expect that the tools developed in our work will pave the way towards the Coulomb gas in the thermodynamic limit.

1.1. Model

We consider a system of N (spinless) fermions on the torus $\mathbb{T}^3 = [0, 2\pi]^3$ (with periodic boundary conditions), interacting via a bounded potential $V : \mathbb{T}^3 \to \mathbb{R}$. The system is described by the Hamiltonian

$$H_N = H_{\mathrm{kin}} + k_F^{-1} H_{\mathrm{int}} = \sum_{i=1}^{N} (-\Delta_i) + k_F^{-1} \sum_{1 \le i < j \le N} V\left(x_i - x_j\right), \tag{1.1}$$

which acts on the fermionic space

$$\mathcal{H}_N = \bigwedge^N \mathfrak{h}, \quad \mathfrak{h} = L^2(\mathbb{T}^3). \tag{1.2}$$

Here, the coupling constant $k_F^{-1} > 0$ corresponds to the interaction strength. We will focus on the mean-field regime $k_F^{-1} \sim N^{-\frac{1}{3}}$, where the kinetic and interaction energies are comparable. More precisely, we assume that

$$N = |B_F| = \frac{4\pi}{3} k_F^3 (1 + o(1)_{k_F \to \infty}), \quad B_F = \overline{B}(0, k_F) \cap \mathbb{Z}^3, \tag{1.3}$$

namely, the Fermi ball B_F is completely filled by N integer points. In this case, the kinetic operator H_{kin} has a unique, non-degenerate ground state which is the Fermi state

$$\psi_{\mathrm{FS}} = \bigwedge_{p \in B_F} u_p, \quad u_p(x) = (2\pi)^{-\frac{3}{2}} e^{i p \cdot x}. \tag{1.4}$$

More generally, the eigenstates of H_{kin} can be written explicitly in terms of the plane waves $\left(u_p\right)_{p \in \mathbb{Z}^3}$. However, the spectrum of the interacting operator H_N is highly nontrivial, and its computation often requires suitable approximations.

We assume that V is of positive type, namely, its Fourier transform satisfies $\hat{V} \ge 0$ with

$$V(x) = \frac{1}{(2\pi)^3} \sum_{k \in \mathbb{Z}^3} \hat{V}_k e^{i k \cdot x} \quad \text{with} \quad \hat{V}_k = \int_{\mathbb{T}^3} V(x) e^{-i k \cdot x} \, dx. \tag{1.5}$$

Under our assumption, H_N is a self-adjoint operator on $\mathcal{H}_N$ with domain $D(H_N) = D(H_{\text{kin}}) = \bigwedge^N H^2(\mathbb{T}^3)$. Moreover, H_N is bounded from below and has compact resolvent. We are interested in the asymptotic behavior of the low-lying spectrum of H_N when $N \to \infty$ and $k_F \to \infty$.

One of the most famous approximations for fermions is the Hartree–Fock theory, where one restricts the states under consideration to the set of all Slater determinants $g_1 \wedge g_2 \cdots \wedge g_N$ with $\{g_i\}_{i=1}^N$ orthonormal in $L^2(\mathbb{T}^3)$. The precision of the Hartree–Fock energy is an interesting subject, which has been studied for Coulomb systems by Bach [1] and Graf–Solovej [22]. In general, the Hartree–Fock minimizer could be different from the Fermi state ψ_{FS}; see [21] for an estimate for Coulomb systems. However, in the mean-field model that we are considering here, the Hartree–Fock minimizer coincides with ψ_{FS}; see [6, Theorem A.1] for a precise statement. Thus, to obtain the correction to the ansatz of plane waves, we have to understand the correlation structure of the system.[2]

To go beyond the ansatz of plane waves, the first step is the extraction of the energy of the Fermi state. For computational purposes, it is convenient to use the second quantization language. For every $p \in \mathbb{Z}^3$, we denote by $c_p^* = c^*(u_p)$, $c_p = c(u_p)$ the fermionic creation and annihilation operators associated to the plane-wave state u_p. These operators act on the fermionic Fock space

$$\mathcal{F}^-(\mathfrak{h}) = \bigoplus_{N=0}^{\infty} \bigwedge^N \mathfrak{h} \tag{1.6}$$

and obey the canonical anticommutation relations (CAR)

$$\{c_p, c_q\} = \{c_p^*, c_q^*\} = 0, \quad \{c_p, c_q^*\} = \delta_{p,q}, \quad p, q \in \mathbb{Z}^3, \tag{1.7}$$

where $\{A, B\} = AB + BA$. The Hamiltonian operator H_N in (1.1) can be expressed as

$$H_N = H_{\text{kin}} + k_F^{-1} H_{\text{int}} = \sum_{p \in \mathbb{Z}^3} |p|^2 c_p^* c_p + \frac{k_F^{-1}}{2(2\pi)^3} \sum_{k \in \mathbb{Z}^3} \sum_{p,q \in \mathbb{Z}^3} \hat{V}_k c_{p+k}^* c_{q-k}^* c_q c_p. \tag{1.8}$$

Thanks to the CAR (1.7), it is straightforward to see that the Fermi state obeys, for all $p \in \mathbb{Z}^3$,

$$c_p^* c_p \psi_{\text{FS}} = 1_{B_F}(p) \psi_{\text{FS}} = \begin{cases} \psi_{\text{FS}} & p \in B_F \\ 0 & p \in B_F^c, \end{cases} \tag{1.9}$$

where $1_{B_F}(\cdot)$ denotes the indicator function of the Fermi ball B_F. Thus, the kinetic energy of the Fermi state is

$$\langle \psi_{\text{FS}}, H_{\text{kin}} \psi_{\text{FS}} \rangle = \sum_{p \in \mathbb{Z}^3} |p|^2 \langle \psi_{\text{FS}}, c_p^* c_p \psi_{\text{FS}} \rangle = \sum_{p \in \mathbb{Z}^3} 1_{B_F}(p) |p|^2 \|\psi_{\text{FS}}\|^2 = \sum_{p \in B_F} |p|^2. \tag{1.10}$$

Hence, we can define the *localized kinetic operator* $H'_{\text{kin}} : D(H_{\text{kin}}) \subset \mathcal{H}_N \to \mathcal{H}_N$ by

$$H'_{\text{kin}} = H_{\text{kin}} - \langle \psi_{\text{FS}}, H_{\text{kin}} \psi_{\text{FS}} \rangle = \sum_{p \in B_F^c} |p|^2 c_p^* c_p - \sum_{p \in B_F} |p|^2 c_p c_p^*. \tag{1.11}$$

We refer to this operator as being ‘localized’ since extracting $\langle \psi_{\text{FS}}, H_{\text{kin}} \psi_{\text{FS}} \rangle$ in this manner can be seen as changing the point of reference from the vacuum state Ω to the Fermi state ψ_{FS}, so H'_{kin} can be seen as a kind of expansion of H_{kin} around ψ_{FS}.

Note that it is clear from the first identity in (1.11) that H'_{kin} is nonnegative since ψ_{FS} is the ground state of H_{kin}. However, the positivity of H'_{kin} is unclear from the second identity in (1.11) since the

[2]The Slater determinants are the least correlated states among all fermionic wave functions (they are eigenfunctions of non-interacting Hamiltonians).

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

difference of two operators which are nonnegative may not have a sign. The resolution of this apparent paradox lies in the underlying Hilbert space: in the N-body space $\mathcal{H}_N$, we always have

$$N = \sum_{p\in\mathbb{Z}^3} c_p^* c_p = \sum_{p\in B_F} (1 - c_p c_p^*) + \sum_{p\in B_F^c} c_p^* c_p = |B_F| - \sum_{p\in B_F} c_p c_p^* + \sum_{p\in B_F^c} c_p^* c_p. \tag{1.12}$$

Therefore, the assumption $|B_F| = N$ implies the *particle-hole symmetry*

$$\mathcal{N}_E = \sum_{p\in B_F^c} c_p^* c_p = \sum_{p\in B_F} c_p c_p^* \quad \text{on } \mathcal{H}_N, \tag{1.13}$$

namely, the *excitation number operator* (which counts the number of particles outside the Fermi state) coincides with the *hole number operator* (which counts the number of holes inside the Fermi state). Consequently, the kinetic operator in (1.11) can be rewritten as

$$H'_{\text{kin}} = \sum_{p\in B_F^c} ||p|^2 - \zeta| \, c_p^* c_p + \sum_{p\in B_F} ||p|^2 - \zeta| \, c_p c_p^* \tag{1.14}$$

for any $\zeta \in [\sup_{p\in B_F} |p|^2, \inf_{p\in B_F^c} |p|^2]$, which is clearly nonnegative.

For the interaction operator, it is convenient to use the factorized form

$$\begin{aligned} H_{\text{int}} &= \frac{1}{2(2\pi)^3} \sum_{k\in\mathbb{Z}^3} \sum_{p,q\in\mathbb{Z}^3} \hat{V}_k c_{p+k}^* c_{q-k}^* c_q c_p \\ &= \frac{1}{2(2\pi)^3} \sum_{k\in\mathbb{Z}^3} \hat{V}_k (\mathrm{d}\Gamma(e^{-ik\cdot x})^* \mathrm{d}\Gamma(e^{-ik\cdot x}) - N), \end{aligned} \tag{1.15}$$

where

$$\mathrm{d}\Gamma(e^{-ik\cdot x}) = \sum_{p,q\in\mathbb{Z}^3} \left\langle u_p, e^{-ik\cdot x} u_q \right\rangle c_p^* c_q = \sum_{p,q\in\mathbb{Z}^3} \delta_{p,q-k} c_p^* c_q = \sum_{p\in\mathbb{Z}^3} c_p^* c_{p+k}. \tag{1.16}$$

Note that for any $k \in \mathbb{Z}^3_* = \mathbb{Z}^3 \setminus \{0\}$, we have

$$\mathrm{d}\Gamma(e^{-ik\cdot x})\psi_{\text{FS}} = \sum_{p\in\mathbb{Z}^3} c_p^* c_{p+k} \psi_{\text{FS}} = \sum_{p\in L_{-k}} c_p^* c_{p+k} \psi_{\text{FS}} \tag{1.17}$$

since the summand $c_p^* c_{p+k}\psi_{\text{FS}}$ in (1.17) does not vanish if and only if $p \in L_{-k}$, where the *lune*

$$L_k = B_F^c \cap (B_F + k) = \left\{ p \in \mathbb{Z}^3 \mid |p - k| \le k_F < |p| \right\} \tag{1.18}$$

will play an important role in our analysis. In particular, using (1.9) and the CAR again, we find that for all $k \in \mathbb{Z}^3_*$,

$$\left\| \mathrm{d}\Gamma(e^{-ik\cdot x})\psi_{\text{FS}} \right\|^2 = \sum_{p\in L_{-k}} \left\| c_p^* c_{p+k} \psi_{\text{FS}} \right\|^2 = \sum_{p\in L_{-k}} 1 = |L_{-k}| = |L_k| \,. \tag{1.19}$$

Thus, the interaction energy of the Fermi state is given by

$$\langle \psi_{\text{FS}}, H_{\text{int}} \psi_{\text{FS}} \rangle = \frac{N\,(N-1)}{2(2\pi)^3} \hat{V}_0 + \frac{1}{2(2\pi)^3} \sum_{k\in\mathbb{Z}^3_*} \hat{V}_k \left(|L_k| - N\right), \tag{1.20}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where we see the direct and exchange energies (involving $\hat{V}_0$ and $\{\hat{V}_k\}_{k\neq 0}$, respectively). We can define the *localized interaction operator*

$$H'_{\text{int}} = H_{\text{int}} - \langle \psi_{\text{FS}}, H_{\text{int}} \psi_{\text{FS}} \rangle = \frac{1}{2(2\pi)^3} \sum_{k\in\mathbb{Z}^3_*} \hat{V}_k \big(\mathrm{d}\Gamma(e^{-ik\cdot x})^* \mathrm{d}\Gamma(e^{-ik\cdot x}) - |L_k| \big). \tag{1.21}$$

In summary, with H'_{kin} and H'_{int} defined in (1.11) and (1.21), we can write

$$H_N = E_{\text{FS}} + H'_{\text{kin}} + k_F^{-1} H'_{\text{int}}, \quad E_{\text{FS}} = \langle \psi_{\text{FS}}, H_N \psi_{\text{FS}} \rangle\,. \tag{1.22}$$

Note that in the prior works [24, 5, 6], the localization procedure was carried out by employing what is known as the particle-hole transformation, which maps the Fermi state ψ_{FS} to the vacuum; see, for example, [6, Eq. (1.20)] for an analogue of (1.22). However, in the present paper we do not follow this approach since we prefer to work on the N-body Hilbert space.

1.2. Random phase approximation

In this subsection, we explain the ideas of the bosonization approach to the random phase approximation. On the one hand, in the original approach [11, 12, 13, 32], Bohm and Pines considered fluctuations of density in the momentum representation where the plasma momenta and the effective particle momenta of different wavelengths k, l are coupled by phases $e^{i(k-l)\cdot x_j}$, summing over the 'random' particle positions x_j. The assumption that the phases average toward zero for a large number of particles is originally called the 'random phase approximation'. On the other hand, after the work of Sawada [36] and Sawada–Brueckner–Fukuda–Brout [37], the term RPA has been widely used in the physics literature in the context of a quasi-bosonic Hamiltonian, where a quasi-boson consists of a particle-hole pair. The quasi-bosonic approach is used not only for Coulomb gases but also in a much broader context, especially in nuclear matter (for a standard textbook, see [18, p. 156] for Coulomb gases and [18, pp. 540-543] for nuclear matter).

In the present paper, we will focus on building a mathematical formulation of the quasi-bosonic approach for general potentials and eventually apply this theory to regular potentials. In the long run, we hope that this general theory will also be helpful for singular potentials, in particular for Coulomb gases where the next-order correction to the bosonization picture matters (in [15], we used the formulation provided in the present paper to find the analogue of the Gell-Mann–Brueckner formula for the mean-field Coulomb gas, which shows how important it is to carry the non-bosonic part in the calculation at least to the leading order).

Now let us explain the bosonization argument in detail. Roughly speaking, the RPA suggests that the fermionic correlation can be described by a Hamiltonian which is quadratic in suitable *bosonic* creation and annihilation operators. To explain the heuristic bosonization argument, let us decompose further the interaction terms in (1.21) by defining, for every $k \in \mathbb{Z}^3_*$,

$$\mathrm{d}\Gamma\big(e^{-ik\cdot x}\big) = \mathrm{d}\Gamma\big((P_{B_F} + P_{B_F^c}) e^{-ik\cdot x} (P_{B_F} + P_{B_F^c})\big) = \tilde{B}_k + \tilde{B}^*_{-k} + D_k, \tag{1.23}$$

where P_{B_F} and $P_{B_F^c}$ are projections in the one-fermion Hilbert space and

$$\tilde{B}_k = \mathrm{d}\Gamma\big(P_{B_F} e^{-ik\cdot x} P_{B_F^c}\big) = \sum_{p,q\in\mathbb{Z}^3} \Big\langle u_p, P_{B_F} e^{-ik\cdot x} P_{B_F^c} u_q \Big\rangle c^*_p c_q = \sum_{p\in L_k} c^*_{p-k} c_p, \tag{1.24}$$

$$D_k = \mathrm{d}\Gamma\big(P_{B_F} e^{-ik\cdot x} P_{B_F}\big) + \mathrm{d}\Gamma\big(P_{B_F^c} e^{-ik\cdot x} P_{B_F^c}\big) = \sum_{p\in B_F\cap(B_F+k)} c^*_{p-k} c_p + \sum_{p\in B_F^c\cap(B_F^c+k)} c^*_{p-k} c_p.$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Note that for all $k \in \mathbb{Z}^3_*$, we have $D^*_k = D_{-k}$ and

$$[\tilde{B}_k, \tilde{B}_{-k}] = [\tilde{B}_{-k}, D_k] = [\tilde{B}^*_k, D_k] = 0, \tag{1.25}$$

which can be seen from the identity $[\mathrm{d}\Gamma(X), \mathrm{d}\Gamma(Y)] = \mathrm{d}\Gamma([X, Y])$ and (1.24). Due to the symmetry between k and $-k$, it is convenient to introduce the set[3]

$$\mathbb{Z}^3_+ = \left(\left\{x_1 > 0\right\} \cup \{x_1 = 0, x_2 > 0\} \cup \{x_1 = x_2 = 0, x_3 > 0\}\right) \cap \mathbb{Z}^3_* \tag{1.26}$$

such that

$$\mathbb{Z}^3_+ \cup \left(-\mathbb{Z}^3_+\right) = \mathbb{Z}^3_*, \quad \mathbb{Z}^3_+ \cap \left(-\mathbb{Z}^3_+\right) = \emptyset. \tag{1.27}$$

Using this notation and the assumption $\hat{V}_k = \hat{V}_{-k}$, we can rewrite the interaction operator in (1.21) as

$$\begin{aligned} k_F^{-1} H'_{\mathrm{int}} &= \frac{k_F^{-1}}{2(2\pi)^3} \sum_{k \in \mathbb{Z}^3_*} \hat{V}_k \left(\left(\tilde{B}_k + \tilde{B}^*_{-k} + D_k\right)^* \left(\tilde{B}_k + \tilde{B}^*_{-k} + D_k\right) - |L_k| \right) \\ &= \sum_{k \in \mathbb{Z}^3_+} \left(H^k_{\mathrm{int}} - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3} |L_k| \right) + \frac{k_F^{-1}}{(2\pi)^3} \sum_{k \in \mathbb{Z}^3_*} \hat{V}_k \left(\tilde{B}^*_k D_k + D^*_k \tilde{B}_k + \frac{1}{2} D^*_k D_k \right), \end{aligned} \tag{1.28}$$

where for each $k \in \mathbb{Z}^3_+$, we denote

$$\begin{aligned} H^k_{\mathrm{int}} &= \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(\left(\tilde{B}_k + \tilde{B}^*_{-k}\right)^* \left(\tilde{B}_k + \tilde{B}^*_{-k}\right) + \left(\tilde{B}_{-k} + \tilde{B}^*_k\right)^* \left(\tilde{B}_{-k} + \tilde{B}^*_k\right) \right) \\ &= \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(\{\tilde{B}^*_k, \tilde{B}_k\} + \{\tilde{B}^*_{-k}, \tilde{B}_{-k}\} + 2\tilde{B}^*_k \tilde{B}^*_{-k} + 2\tilde{B}_{-k}\tilde{B}_k \right). \end{aligned} \tag{1.29}$$

Now let us introduce the quasi-bosonicity. From the CAR (1.7), it is straightforward to see that

$$\left[\tilde{B}_k, \tilde{B}_l\right] = \left[\tilde{B}^*_k, \tilde{B}^*_l\right] = 0, \quad \left[\tilde{B}_k, \tilde{B}^*_l\right] = |L_k|\, \delta_{k,l} - \sum_{p \in L_k \cap L_l} c_{p-l} c^*_{p-k} - \sum_{p \in L_k \cap (L_l - l + k)} c^*_{p-k+l} c_p \tag{1.30}$$

for all $k, l \in \mathbb{Z}^3_*$, where $[A, B] = AB - BA$. Hence, on states with few excitations (e.g., the expectation value of $\mathcal{N}_E$ is much smaller than $|L_k| \sim \min\{|k| k_F^2, k_F^3\}$), the rescaled operators $\tilde{B}'_k = |L_k|^{-\frac{1}{2}} \tilde{B}_k$ obey the commutation relations

$$\left[\tilde{B}'_k, \tilde{B}'_l\right] = \left[\left(\tilde{B}'_k\right)^*, \left(\tilde{B}'_l\right)^*\right] = 0, \quad \left[\tilde{B}'_k, \left(\tilde{B}'_l\right)^*\right] \approx \delta_{k,l} \tag{1.31}$$

for all $k, l \in \mathbb{Z}^3_*$, in direct analogy with the canonical commutation relations (CCR) obeyed by a set of *bosonic* creation and annihilation operators a^*_k, a_k indexed by $\mathbb{Z}^3_*$,

$$[a_k, a_l] = \left[a^*_k, a^*_l\right] = 0, \quad \left[a_k, a^*_l\right] = \delta_{k,l}. \tag{1.32}$$

Since the relation $[\tilde{B}'_k, (\tilde{B}'_l)^*] \approx \delta_{k,l}$ is only approximate, we call these operators *quasi*-bosonic.

In view of the quasi-bosonicity of these operators, in the form (1.28) of H'_{int}, we call the first sum on the right-hand side of this equation the *bosonizable terms*, while the second sum constitutes the *non-bosonizable terms* which are regarded as error terms. The bosonizable part H^k_{int} can be viewed as a quadratic Hamiltonian in the bosonic setting, which can be diagonalized by Bogolubov transformations.

[3] The exact definition of $\mathbb{Z}^3_+$ is not important, only that it satisfies $\mathbb{Z}^3_+ \cup \left(-\mathbb{Z}^3_+\right) = \mathbb{Z}^3_*$ and $\mathbb{Z}^3_+ \cap \left(-\mathbb{Z}^3_+\right) = \emptyset$.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

This is the spirit of what we will do, but there is a catch: the kinetic operator H'_{kin} cannot be written in terms of $\tilde{B}_k$. The solution is to further decompose the operators $\tilde{B}_k$ by defining the *excitation operators*

$$b_{k,p} = c^*_{p-k} c_p, \quad b^*_{k,p} = c^*_p c_{p-k}, \quad k \in \mathbb{Z}^3_*,\ p \in L_k. \tag{1.33}$$

The name is due to the fact that the action of $b^*_{k,p}$ is to create a state at momentum $p \in B_F^c$ and annihilate a state at momentum $p - k \in B_F$.

Since H^k_{int} is quadratic in terms of $\tilde{B}_k$, it is also quadratic in terms of $b^*_{k,p}$, namely,

$$\begin{aligned} H^k_{\text{int}} = & \sum_{p,q \in L_k} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{k,p} b_{k,q} + b_{k,q} b^*_{k,p} \right) + \sum_{p,q \in L_{-k}} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{-k,p} b_{-k,q} + b_{-k,q} b^*_{-k,p} \right) \\ & + \sum_{p \in L_k} \sum_{q \in L_{-k}} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{k,p} b^*_{-k,q} + b_{-k,q} b_{k,p} \right) + \sum_{p \in L_{-k}} \sum_{q \in L_k} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{-k,p} b^*_{k,q} + b_{k,q} b_{-k,p} \right). \end{aligned} \tag{1.34}$$

The reason that the operators $b_{k,p}$ are preferable to the operators $\tilde{B}_k$ is that they satisfy the following commutation relation with the kinetic operator (see (1.74) below)

$$\left[H'_{\text{kin}}, b^*_{k,p} \right] = 2\lambda_{k,p} b^*_{k,p}, \quad \lambda_{k,p} = \frac{1}{2}(|p|^2 - |p-k|^2). \tag{1.35}$$

Note that $\lambda_{k,p} \geq \frac{1}{2}$ (first, $\lambda_{k,p} > 0$ since $p \in L_k$; moreover, $|p|^2 - |p-k|^2$ is an integer as $p, k \in \mathbb{Z}^3$). This is to be compared with the bosonic setting: if the operators a_k obey the CCR (1.32), then

$$\left[\sum_l \varepsilon_l a^*_l a_l, a^*_k \right] = \varepsilon_k a^*_k. \tag{1.36}$$

Therefore, viewing $b^*_{k,p}$ as being analogous to a bosonic creation operator, we get

$$H'_{\text{kin}} \approx \sum_{k \in \mathbb{Z}^3_*} \sum_{p \in L_k} 2\lambda_{k,p} b^*_{k,p} b_{k,p} = \sum_{k \in \mathbb{Z}^3_+} \left(\sum_{p \in L_k} 2\lambda_{k,p} b^*_{k,p} b_{k,p} + \sum_{p \in L_{-k}} 2\lambda_{-k,p} b^*_{-k,p} b_{-k,p} \right). \tag{1.37}$$

Combining (1.34) and (1.37), we arrive at a Hamiltonian quadratic in terms of the operators $b_{k,p}$, which could be treated in the bosonic interpretation. Note that $b_{k,p} \psi_{\text{FS}} = 0$ for all $k \in \mathbb{Z}^3_*$, $p \in L_k$, and hence, the Fermi state plays the role of the bosonic vacuum.

Overview of the heuristic assumptions behind the random phase approximation

In the physics literature [36, 37], the RPA entails two assumptions:

1. That the excitation operators $b^*_{k,p}, b_{k,p}$ in (1.33) can be treated as *bosonic* creation and annihilation operators, and that the operators $b_{k,p}$ and $b_{l,q}$ with $k \neq l$ can be considered as acting on independent Fock spaces. Mathematically, we thus expect that the approximate canonical commutation relations (CCR)

$$\left[b_{k,p}, b_{l,q} \right] = \left[b^*_{k,p}, b^*_{l,q} \right] = 0, \quad \left[b_{k,p}, b^*_{l,q} \right] \approx \delta_{k,l} \delta_{p,q} \tag{1.38}$$

should hold in an appropriate sense.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

2. That the operator in (1.22) can be approximated by an effective Hamiltonian which is quadratic in terms of $b^*_{k,p}$ and $b_{k,p}$. This is already true for the interaction part $\sum_{k\in\mathbb{Z}^3_+} H^k_{\mathrm{int}}$ in (1.34), and in the RPA, the *non-bosonizable terms*

$$\frac{k_F^{-1}}{(2\pi)^3}\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\left(\tilde{B}^*_k D_k + D^*_k\tilde{B}_k + \frac{1}{2}D^*_k D_k\right) \tag{1.39}$$

are simply dropped. Moreover, the kinetic operator H'_{kin} is not exactly of the desired form, but it can be replaced by the right side of (1.37). All this leads to the effective Hamiltonian

$$\sum_{k\in\mathbb{Z}^3_+} H_{\mathrm{Bog},k} = \sum_{k\in\mathbb{Z}^3_+}\left(2\sum_{p\in L_k}\lambda_{k,p}b^*_{k,p}b_{k,p} + 2\sum_{p\in L_{-k}}\lambda_{-k,p}b^*_{-k,p}b_{-k,p} + H^k_{\mathrm{int}} - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|\right) \tag{1.40}$$

acting on the bosonic Fock space $\bigoplus_{k\in\mathbb{Z}^3_+}\mathcal{F}^+\left(\ell^2\left(L_k\cup L_{-k}\right)\right)$.

Consequently, since the operators $b_{k,p}$ and $b_{l,q}$ with $k\neq l$ are considered as acting independently, we can diagonalize separately each quadratic bosonic Hamiltonian $H_{\mathrm{Bog},k}$ by a Bogolubov transformation $\mathcal{U}_k$ on $\mathcal{F}^+\left(\ell^2\left(L_k\cup L_{-k}\right)\right)$ such that

$$\mathcal{U}_k H_{\mathrm{Bog},k}\mathcal{U}^*_k = 2\,\mathrm{tr}\left(\widetilde{E}_k - h_k\right) - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k| + 2\sum_{p\in L_k\cup L_{-k}}\langle e_p, \widetilde{E}_k e_q\rangle b^*_{k,p}b_{k,q}, \tag{1.41}$$

where for every $k\in\mathbb{Z}^3_*$, we denote the following quantities on $\ell^2(L_k)$:

$$\widetilde{E}_k = (h_k^{\frac{1}{2}}\left(h_k + 2P_{v_k}\right)h_k^{\frac{1}{2}})^{\frac{1}{2}},\quad h_k e_p = \lambda_{k,p}e_p,\quad P_{v_k} = |v_k\rangle\langle v_k|,\quad v_k = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}}\sum_{p\in L_k} e_p \tag{1.42}$$

with $(e_p)_{p\in L_k}$ the standard orthonormal basis of $\ell^2(L_k)$.

Summing over k, we obtain the *correlation energy* (see Proposition 7.1)

$$E_{\mathrm{corr}} = \sum_{k\in\mathbb{Z}^3_+}\left(\mathrm{tr}\left(\widetilde{E}_k - h_k\right) - \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}|L_k|\right) = \sum_{k\in\mathbb{Z}^3_*}\frac{1}{\pi}\int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda^2_{k,p}+t^2}\right)dt, \tag{1.43}$$

where $F(x) = \log(1+x) - x$. All in all, the RPA thus suggests that up to a unitary transformation, we expect that

$$H_N \approx E_{\mathrm{FS}} + E_{\mathrm{corr}} + 2\sum_{k\in\mathbb{Z}^3_*}\sum_{p,q\in L_k}\langle e_p, \widetilde{E}_k e_q\rangle b^*_{k,p}b_{k,p}, \tag{1.44}$$

at least on states with few excitations.

Prediction of the correlation energy and the excitation spectrum

Equation (1.44) leads immediately to the following approximation for the ground state energy

$$\inf\sigma\left(H_N\right) \approx E_{\mathrm{FS}} + E_{\mathrm{corr}}, \tag{1.45}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

which coincides with [37, Eq. (34)],[4] where the authors derived it from the effective operator of equation (1.40) and also explained the connection to the original work of Gell-Mann–Brueckner [20]. See also [35, Eq. (9.54)] and [18, Eq. (12.53)] for this expression of the ground state energy.

More importantly, (1.44) also suggests that the excitation spectrum of H_N could be described in terms of the eigenvalues of $2\widetilde{E}_k$, which correspond to the *bosonic elementary excitations* and can be explicitly computed.

Indeed, for every eigenvalue ϵ of $\widetilde{E}_k$, we may find an eigenvector $w \in \ell^2(L_k)$ such that

$$\epsilon^2 w = \widetilde{E}_k^2 w = h_k^{\frac{1}{2}} \left(h_k + 2P_{v_k}\right) h_k^{\frac{1}{2}} w = h_k^2 w + 2\langle h_k^{\frac{1}{2}} v_k, w\rangle h_k^{\frac{1}{2}} v_k. \tag{1.46}$$

But either ϵ is also an eigenvalue of h_k or $\epsilon^2 - h_k^2$ is invertible. In the latter case, we can write

$$w = 2\langle h_k^{\frac{1}{2}} v_k, w\rangle (\epsilon^2 - h_k^2)^{-1} h_k^{\frac{1}{2}} v_k, \tag{1.47}$$

and taking the inner product with $h_k^{\frac{1}{2}} v_k$ and cancelling the factors of $\langle h_k^{\frac{1}{2}} v_k, w\rangle$ yields

$$1 = 2\langle v_k, h_k(\epsilon^2 - h_k^2)^{-1} v_k\rangle = \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3} \sum_{p\in L_k} \frac{\lambda_{k,p}}{\epsilon^2 - \lambda_{k,p}^2}, \tag{1.48}$$

which appears in [37, Eq. (6)]. The sum can be rewritten as

$$1 = \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \sum_{p\in B_F} \frac{|k|^2}{(\epsilon - k\cdot p)^2 - \left(\frac{1}{2}|k|^2\right)^2}. \tag{1.49}$$

The formula (1.49) allows to compute all eigenvalues of $\widetilde{E}_k$ outside the spectrum of h_k.

In the physically relevant case of the Coulomb potential where $\hat{V}_k k_F^{-1}$ is replaced by $4\pi e^2|k|^{-2}$, one can immediately derive the famous plasmon frequency from (1.49): for $|k| \ll k_F^{1/2}$, the largest eigenvalue ϵ is proportional to $k_F^{3/2}$ (see [14, Eq. (2.27)–(2.54)] for a detailed explanation), and its leading order behavior can be computed easily in the thermodynamic limit (including also a factor of 2 for the electron spin states)

$$\epsilon^2 = \frac{4\pi e^2}{(2\pi)^3} \int_{\overline{B}(0,k_F)} \frac{\epsilon^2}{(\epsilon - k\cdot p)^2 - \left(\frac{1}{2}|k|^2\right)^2}\, dp \approx \frac{2e^2}{(2\pi)^2} \mathrm{Vol}(\overline{B}(0,k_F)) = \frac{2e^2}{3\pi} k_F^3 = 2\pi n e^2, \tag{1.50}$$

where $n = \frac{N}{\mathcal{V}} = \frac{1}{3\pi^2} k_F^3$ is the number density of the system. Recalling that the relevant operator is $2\widetilde{E}_k$ rather than $\widetilde{E}_k$ and that $\frac{\hbar^2}{2m} = 1$, this yields an excitation energy of

$$2\epsilon \approx 2\sqrt{2\pi n e^2} = \hbar\sqrt{\frac{4\pi n e^2}{m}} = \hbar\omega_{\mathrm{plasmon}}, \tag{1.51}$$

where $\omega_{\mathrm{plasmon}} = \sqrt{4\pi n e^2 m^{-1}}$ is called the plasmon frequency in [33, Eq. (3-90)] and [18, Eq. (15.16) - (15.18)]. Note that the Coulomb potential is special as it makes the right-hand side of (1.51) independent of k. See also [4, 14] where (1.51) was discussed.

Establishing the above heuristic computation is a longstanding problem in mathematical physics. In the present paper, we will give a rigorous formulation for the operator approximation (1.44) and then

[4]Provided one replaces $(2\pi)^3$ with the volume Ω of the box, includes a spin factor and inserts the Coulomb potential, $\hat{V}_k k_F^{-1} = 4\pi e^2|k|^{-2}$.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

use this to justify the prediction of the correlation energy and the bosonic elementary excitations for a wide class of bounded potentials in the mean-field regime.

1.3. Main results

Our first result is the following rigorous formulation of the operator approximation (1.44).

Theorem 1.1 (Operator formulation of the RPA). *Let $V : \mathbb{T}^3 \to \mathbb{R}$ obey $\hat{V}_k \geq 0$ and $\hat{V}_{-k} = \hat{V}_k$ for all $k \in \mathbb{Z}^3$, and assume furthermore that $\sum_{k\in\mathbb{Z}^3} \hat{V}_k |k| < \infty$. Consider the Hamiltonian H_N given in (1.1) with $N = |B_F|$. Let the operators H'_{kin}, $\mathcal{N}_E$, $\widetilde{E}_k - h_k$ be defined in (1.11), (1.13), (1.42). Let the energies E_{FS}, E_{corr} be defined in (1.22), (1.43). Then there exists a unitary transformation $\mathcal{U} : \mathcal{H}_N \to \mathcal{H}_N$ such that*

$$\mathcal{U} H_N \mathcal{U}^* = E_{\text{FS}} + E_{\text{corr}} + H_{\text{eff}} + \mathcal{E}_{\mathcal{U}}, \tag{1.52}$$

where the effective operator $H_{\text{eff}} : \mathcal{H}_N \to \mathcal{H}_N$ is

$$H_{\text{eff}} = H'_{\text{kin}} + 2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p,q\in L_k} \left\langle e_p, (\widetilde{E}_k - h_k) e_q \right\rangle b^*_{k,p} b_{k,q} \tag{1.53}$$

and the error operator $\mathcal{E}_{\mathcal{U}} : \mathcal{H}_N \to \mathcal{H}_N$ obeys the operator inequality: for every constant $\epsilon > 0$,

$$\pm \mathcal{E}_{\mathcal{U}} \leq C k_F^{-\frac{1}{94}+\epsilon} \left(k_F^{-1} \mathcal{N}_E H'_{\text{kin}} + H'_{\text{kin}} + k_F \right), \quad k_F \to \infty. \tag{1.54}$$

The unitary operator in Theorem 1.1 is given explicitly as $\mathcal{U} = e^{\mathcal{J}} e^{\mathcal{K}}$, where $\mathcal{K}$ and $\mathcal{J}$ are given in (1.78) and (1.85), respectively (the transformations $e^{\mathcal{K}}$ and $e^{\mathcal{J}}$ are studied in detail in Sections 5 and 9).

Remark 1.1. The operator $\mathcal{N}_E H'_{\text{kin}}$ on the right-hand side of (1.54) is nothing but the 'bosonic kinetic operator', due to the following remarkable identity (see Proposition 10.1):

$$2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p\in L_k} \lambda_{k,p} b^*_{k,p} b_{k,p} = \mathcal{N}_E H'_{\text{kin}}. \tag{1.55}$$

Thus, in Theorem 1.1, we control the error in the random phase approximation using only the fermionic and bosonic kinetic operators, which is very natural.

Remark 1.2. In the expansion (1.52), E_{FS} is of order k_F^5, and E_{corr} is of order k_F. As we will argue below, when we apply this to the low-lying eigenstates with energy $E_{\text{FS}} + O(k_F)$, the expectation of the effective Hamiltonian H_{eff} in (1.53) is of order k_F, while the error term $\mathcal{E}_{\mathcal{U}}$ in (1.54) is of order $O(k_F^{1-\frac{1}{94}+\epsilon}) = o(k_F)$.

In order to put Theorem 1.1 to good use, we need some a priori estimate on the low-lying eigenstates of the Hamiltonian H_N. We have the following:

Theorem 1.2 (A priori estimate for eigenstates). *Let V and $\mathcal{U}$ be as in Theorem 1.1. Let $\Psi \in D\left(H'_{\text{kin}}\right)$ be a normalized eigenstate of H_N with energy $\langle \Psi, H_N \Psi \rangle \leq E_{\text{FS}} + \kappa k_F$ for some constant $\kappa > 0$ independent of k_F. Then,*

$$\left\langle \Psi, \left(H'_{\text{kin}} + k_F^{-1} \mathcal{N}_E H'_{\text{kin}} \right) \Psi \right\rangle \leq C(\kappa+1)^2 k_F$$

for a constant $C > 0$ depending only on V. The same bound holds with Ψ replaced by $\mathcal{U}\Psi$.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Remark 1.3. Thanks to the inequality $\mathcal{N}_E \leq H'_{\text{kin}}$ (see [6, Lemma 2.4] and also Proposition 2.1 below), Theorem 1.2 implies that for an eigenstate Ψ of H_N with energy $\langle \Psi, H_N \Psi \rangle \leq E_{\text{FS}} + O(k_F)$, we have

$$\langle \Psi, \mathcal{N}_E \Psi \rangle \leq \left\langle \Psi, H'_{\text{kin}} \Psi \right\rangle = O(k_F). \tag{1.56}$$

Thus, the number of excitations is much smaller than the total number of particles ($k_F \sim N^{1/3} \ll N$). While (1.56) has been derived in [24, 6] for every state with energy $\langle \Psi, H_N \Psi \rangle \leq E_{\text{FS}} + O(k_F)$ (at least for a class of potentials V), the improved bound in Theorem 1.2 is deeper, and the eigenstate assumption plays a crucial role in the proof.

From Theorems 1.1 and 1.2, we can deduce immediately the asymptotic formula (1.45) on the ground state energy up to an error $o(k_F)$. Indeed, the energy upper bound is given by the trial state $\mathcal{U}^* \psi_{\text{FS}}$, while the energy lower bound follows from the obvious operator inequality $\widetilde{E}_k \geq h_k$. Moreover, our approach is quantitative, and we can derive (1.45) with explicit error estimates.

Theorem 1.3 (Ground state energy). *Let V be as in Theorem 1.1. Then for all $\epsilon > 0$,*

$$\inf \sigma(H_N) = E_{\text{FS}} + E_{\text{corr}} + O(k_F^{1-\frac{1}{94}+\epsilon}), \quad k_F \to \infty.$$

Here are some remarks concerning Theorem 1.3.

Remark 1.4. The method of our proof can be adapted to give the upper bound under the weaker condition $\sum_{k \in \mathbb{Z}^3} \hat{V}_k^2 |k| < \infty$ (see [8, Appendix A] for a derivation of the upper bound under this weaker condition). Additionally, under this condition it can be shown that

$$\frac{1}{\pi} \sum_{k \in \mathbb{Z}^3_*} \int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3} \sum_{p \in L_k} \frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2} \right) dt = \frac{k_F}{\pi} \sum_{k \in \mathbb{Z}^3_*} |k| \int_0^\infty F\left(\frac{\hat{V}_k}{(2\pi)^2} I(t) \right) dt + o(k_F), \tag{1.57}$$

where $F(x) = \log(1 + x) - x$ and $I(t) = 1 - t \tan^{-1}\left(t^{-1}\right)$ (this essentially amounts to replacing the Riemann sum by the integral and can be done by following either the proof of [5, Eq. (5.15)] or the analysis in Appendix A; the condition $\sum \hat{V}_k^2 |k| < \infty$ ensures that the main contribution comes from $|k| \sim O(1)$). Hence, Theorem 1.3 implies that

$$\inf \sigma(H_N) = E_{\text{FS}} + \frac{k_F}{\pi} \sum_{k \in \mathbb{Z}^3_*} |k| \int_0^\infty F\left(\frac{\hat{V}_k}{(2\pi)^2} I(t) \right) dt + o(k_F). \tag{1.58}$$

A result similar to ours, namely, the bound (1.58) for all potentials satisfying $\sum_k \hat{V}_k |k| < \infty$, has been independently obtained in [8], based on a refinement of the method in [5, 6].[5] The bound (1.58) was proved earlier in [5, 6], under the additional assumption that the Fourier coefficients $\hat{V}_k$ be finitely supported and that $\|\hat{V}\|_{\ell^1}$ be sufficiently small. For small $\hat{V}_k$, the logarithm of equation (1.58) can be expanded for

$$\sigma(H_N) = E_{\text{FS}} - \frac{1 - \log(2)}{6(2\pi)^4} k_F \sum_{k \in \mathbb{Z}^3_*} \hat{V}_k^2 |k| \left(1 + O\left(\hat{V}_k\right)\right) + o(k_F), \tag{1.59}$$

which was first proved in [24].

Remark 1.5. A further refinement of our method allows a derivation of a rigorous energy upper bound for all potentials satisfying $\sum_{k \neq 0} \hat{V}_k^2 < \infty$; see [15]. This covers the case of the Coulomb potential $\hat{V}_k = 4\pi e^2 |k|^{-2}$, where the correlation energy is given by the left-hand side of (1.57) which is of order $k_F \log k_F$ plus a correlation exchange correction of order k_F (the correlation exchange contribution

[5] Note that the conventions of the Fourier transform and scaling of H_N in [5, 6, 8] differ from ours.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

comes from the fact that the purely bosonic picture is not exact; it is different from the exchange energy which is part of E_{FS}). In particular, for the Coulomb potential, the right-hand side of (1.57) diverges, whereas the left-hand side does not, and hence the discrete form in (1.57) is arguably more fundamental than the continuous form. It is interesting that in our method the discrete version of the correlation energy always appears naturally.

Besides containing the information of the ground state energy, another decisive consequence of the operator statement in Theorem 1.1 is that it allows us to obtain *all bosonic elementary excitations* predicted in the physics literature. We have the following:

Theorem 1.4 (Bosonic elementary excitations). *Let V and $\mathcal{U}$ be as in Theorem 1.1. Let $\Psi \in \mathcal{H}_N$ be a normalized wave function such that $\mathcal{N}_E \Psi = \Psi$ and $\langle \Psi, H'_{\mathrm{kin}} \Psi \rangle = O(k_F)$. Then for all $\epsilon > 0$, we have*

$$\langle \Psi, \mathcal{U} H_N \mathcal{U}^* \Psi \rangle = E_{\mathrm{FS}} + E_{\mathrm{corr}} + \langle \Psi, H_{\mathrm{eff}}|_{\mathcal{N}_E = 1} \Psi \rangle + O(k_F^{1-\frac{1}{94}+\epsilon}),$$

where

$$H_{\mathrm{eff}}|_{\mathcal{N}_E=1} = 2 \sum_{k \in \mathbb{Z}^3_*} \sum_{p,q \in L_k} \langle e_p, \widetilde{E}_k e_q \rangle b^*_{k,p} b_{k,q} = \tilde{U} \left(\bigoplus_{k \in \mathbb{Z}^3_*} 2 \widetilde{E}_k \right) \tilde{U}^* \tag{1.60}$$

on the space $\{\Psi \in \mathcal{H}_N \mid \mathcal{N}_E \Psi = \Psi\}$, and

$$\tilde{U} : \bigoplus_{k \in \mathbb{Z}^3_*} L^2(L_k) \to \{\Psi \in \mathcal{H}_N \mid \mathcal{N}_E \Psi = \Psi\} \tag{1.61}$$

is a unitary isomorphism defined by

$$\tilde{U} \bigoplus_{k \in \mathbb{Z}^3_*} \varphi_k = \sum_{k \in \mathbb{Z}^3_*} b^*_k(\varphi_k)\, \psi_{\mathrm{FS}} = \sum_{k \in \mathbb{Z}^3_*} \sum_{p \in L_k} \langle e_p, \varphi_k \rangle b^*_{k,p} \psi_{\mathrm{FS}}. \tag{1.62}$$

Recall that all eigenvalues of $\widetilde{E}_k$ can be computed explicitly from the spectrum of h_k and (1.49). From Theorem 1.1 and Theorem 1.4, we may say that up to the unitary transformation $\mathcal{U}$, the RPA is exact for the $\{\mathcal{N}_E = 1\}$ eigenspace of the effective Hamiltonian H_{eff}. To our knowledge, this is the first rigorous derivation of the bosonic elementary excitations from first principles.

Remark 1.6. For every fixed $k \in \mathbb{Z}^3_*$, in the limit $k_F \to \infty$, most eigenvalues of $\widetilde{E}_k$ are of order k_F, but the lowest eigenvalue of $\widetilde{E}_k$ is of order $o(k_F)$. This absence of a one-body spectral gap corresponds to the expected fact that the excitation spectrum of $k_F^{-1} H_N$ becomes continuous in the limit $k_F \to \infty$. Therefore, in principle, it is very difficult to extract useful information by analyzing the full spectrum of H_N. The significance of Theorem 1.4 is to offer a nontrivial statement on the bosonic excitations by analyzing exactly the spectrum of the effective Hamiltonian instead of looking directly at the spectrum of H_N.

Remark 1.7. In Theorem 1.4, the restriction to the $\mathcal{N}_E = 1$ eigenspace is important. Obviously, the effective Hamiltonian (1.53) does not coincide with that in the heuristic formula (1.44). Hence, it is natural to ask what to make of the assumption of the RPA that the effective Hamiltonian should behave like a diagonalized bosonic Hamiltonian. To approach this question, we note that using (1.55), we can rewrite the effective Hamiltonian in (1.53) as

$$H_{\mathrm{eff}} = 2 \sum_{k \in \mathbb{Z}^3_*} \sum_{p,q \in L_k} \langle e_p, \widetilde{E}_k e_q \rangle b^*_{k,p} b_{k,q} - (\mathcal{N}_E - 1)\, H'_{\mathrm{kin}}. \tag{1.63}$$

Since this operator commutes with $\mathcal{N}_E$, we can restrict H_{eff} to the eigenspaces of $\mathcal{N}_E$. Doing so, we see that the trivial eigenspace $\{\mathcal{N}_E = 0\} = \mathrm{span}\,(\psi_{\mathrm{FS}})$ exactly corresponds to the ground state energy

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

which is already addressed in Theorem 1.3. For the first nontrivial eigenspace $\{\mathcal{N}_E = 1\}$, we do indeed obtain the expected operator

$$H_{\text{eff}}|_{\mathcal{N}_E=1} = 2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p,q\in L_k} \langle e_p, \widetilde{E}_k e_q\rangle b^*_{k,p} b_{k,q}, \tag{1.64}$$

as in the heuristic formula (1.44). Moreover, the second identity in (1.60) tells us that $H_{\text{eff}}|_{\mathcal{N}_E=1}$ can be diagonalized explicitly on $\{\mathcal{N}_E = 1\}$, which is important for applications.

More generally, we can also consider the higher excitation sectors $\{\mathcal{N}_E = M\}$ for $M \in \mathbb{N}$.

Theorem 1.5 (Higher excitations). *Let V and $\mathcal{U}$ be as in Theorem 1.1. Let $1 \le M \le O(k_F)$. Let $\Psi \in \mathcal{H}_N$ be a normalized wave function such that $\mathcal{N}_E\Psi = M\Psi$ and $\langle \Psi, H'_{\text{kin}}\Psi\rangle \le O(k_F)$. Then for all $\epsilon > 0$, we have*

$$\langle \Psi, \mathcal{U} H_N \mathcal{U}^* \Psi\rangle = E_{\text{FS}} + E_{\text{corr}} + \langle \Psi, H_{\text{eff}}|_{\mathcal{N}_E=M}\, \Psi\rangle + O(k_F^{1-\frac{1}{94}+\epsilon}),$$

where

$$H_{\text{eff}}|_{\mathcal{N}_E=M} = 2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p,q\in L_k} \langle e_p, (\widetilde{E}_k - (1 - M^{-1})h_k) e_q\rangle b^*_{k,p} b_{k,q}.$$

Remark 1.8. For $M \ge 2$, the operator $H_{\text{eff}}|_{\mathcal{N}_E=M}$ in Theorem 1.5 cannot be diagonalized explicitly as in (1.60). The quasi-bosonic property is insufficient to guarantee that it is diagonalizable, even approximately. Understanding the behaviour of H_{eff} on higher eigenspaces and reconciling the RPA thus appears to be an interesting but nontrivial task. Some progress in this direction was done in [14] where the norm $\|(H_{\text{eff}} - M\epsilon)\Psi\|$ was estimated for suitable trial states.

1.4. Proof strategy

Now let us explain some key ingredients of the proof. Following [37], our approach consists of studying pair-excitations $b^*_{k,p} = c^*_p c_{p-k}$, where c_{p-k} annihilates a particle with momentum $p - k$ (i.e., creates a hole in the Fermi ball), and c^*_p creates a particle outside the Fermi ball. These operators $b_{k,p}$, $b^*_{k,p}$ satisfy the bosonic commutation relations in an appropriate sense. This enables the use of a quasi-bosonic Bogolubov transformation to diagonalize the original fermionic operator. A main achievement of the present work is the analytical elaboration of this bosonic picture.

In [5, 6], a different, collective bosonization approach was developed by averaging the pair-excitations $b^*_{k,p}$ on 'patches' near the surface of the Fermi ball, thus realizing strengthened versions of the bosonic commutation relations which make the comparison with the purely bosonic computation significantly easier. In the present paper, we show that the bosonization idea can be implemented directly for pairs of fermions without such an averaging procedure. In our opinion, this new approach is conceptually closer to the physics of the problem and more transparent for applications. In particular, it allows us to obtain all bosonic elementary excitations as in Theorem 1.4. Moreover, the new method is potentially applicable to Coulomb systems, where the correlation exchange correction to the purely bosonic computation plays an important role; see [15] for a rigorous ground state energy upper bound.

In the context of interacting Bose gases, Bogolubov transformations based on another approximate CCR have been used to study the excitation spectrum; see, for example, [38, 23, 9, 25]. However, for the fermionic problem considered in the present paper, the approximate CCR holds in a very different setting and requires distinct estimation techniques.

Now let us provide further details.

Bosonization method

The driving concept of the random phase approximation is the bosonization of fermionic pairs. We must therefore argue why the excitation operators

$$b_{k,p} = c^*_{p-k} c_p, \quad b^*_{k,p} = c^*_p c_{p-k}, \quad p \in L_k = (B_F + k) \setminus B_F \tag{1.65}$$

obey an approximate CCR. Consider for simplicity the case $k = l$: then computation shows that for any $p, q \in L_k$, $[b_{k,p}, b_{k,q}] = [b^*_{k,p}, b^*_{k,q}] = 0$, but

$$[b_{k,p}, b^*_{k,q}] = \delta_{p,q} - \delta_{p,q}(c^*_p c_p + c_{p-k} c^*_{p-k}). \tag{1.66}$$

In general, thanks to Pauli's exclusion principle ($c^*_p c_p,\ c_p c^*_p \le 1$), the error term in (1.66) satisfies the simple bound $\delta_{p,q}(c^*_p c_p + c_{p-k} c^*_{p-k}) \le 2\delta_{p,q}$, but this is even bigger than the leading term $\delta_{p,q}$. The key observation is that although these errors terms can not be considered to be small individually, they are so on average. For instance,

$$\sum_{p,q \in L_k} \delta_{p,q}(c^*_p c_p + c_{p-k} c^*_{p-k}) = \sum_{p \in L_k} c^*_p c_p + \sum_{p \in L_k} c_{p-k} c^*_{p-k} \le 2\mathcal{N}_E, \tag{1.67}$$

where $\mathcal{N}_E$ is the 'excitation number operator' defined in (1.13). Thus, for states where the expectation value of $\mathcal{N}_E$ is much smaller than $\sum_{p,q \in L_k} \delta_{p,q} = |L_k| \sim \min\{|k| k_F^2, k_F^3\}$, one may expect that the contribution of the non-bosonic error terms are also smaller than the leading bosonic behaviour. Justifying this idea rigorously is one of the main results of this paper.

Note that unlike the works [24, 5, 6], we do not employ the 'particle-hole transformation' R, which maps ψ_{FS} to the vacuum, so that we always work directly on the space $\mathcal{H}_N$.

A priori estimates

As explained above, to apply the bosonization method, we need to show that the expectation of $\mathcal{N}_E$ against low-lying eigenstates of H_N is much smaller than $|L_k| \sim \min\left\{k_F^2 |k|, k_F^3\right\}$.

Using the condition $\sum_{k \in \mathbb{Z}^3} \hat{V}_k |k| < \infty$ and a variant of Onsager's lemma, we can prove that

$$H_N \ge E_{\mathrm{FS}} + H'_{\mathrm{kin}} - C k_F. \tag{1.68}$$

Consequently, if Ψ is any eigenstate for H_N satisfying $\langle \Psi, H_N \Psi \rangle \le E_{\mathrm{FS}} + C k_F$, then

$$\left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle \le C k_F. \tag{1.69}$$

Since $H'_{\mathrm{kin}} \ge \mathcal{N}_E$, which was already explained in [6], this implies that $\langle \Psi, \mathcal{N}_E \Psi \rangle \le C k_F \ll |L_k|$. For V sufficiently small, this bound was first proved in [24] (by a different method), and it was also used in [6]. In practice, we will also need a stronger a priori estimate, namely,

$$\left\langle \Psi, k_F^{-1} \mathcal{N}_E H'_{\mathrm{kin}} \Psi \right\rangle \le C k_F \tag{1.70}$$

as stated in Theorem 1.2. This we will obtain by employing a bootstrapping argument for eigenstates, inspired by the 'improved condensation' in the context of Bose gases in [38, 23, 29, 30]. In [6], an analogue of equation (1.70) was proved for a modified ground state by using a 'localization in Fock space' technique. In comparison, our estimate of equation (1.70) is obtained in a far more direct fashion and yields a uniform bound for all low-lying eigenstates. In particular, thanks to (1.69) and (1.70),

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

the operator estimate in Theorem 1.1 leads to direct consequences on the ground state energy and the excitation spectrum of H_N.

Removing the non-bosonizable terms

An important ingredient of the RPA is that the non-bosonizable terms

$$\frac{k_F^{-1}}{2(2\pi)^3}\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\left(2\,\mathrm{Re}\left(\tilde{B}_k^*+\tilde{B}_{-k}\right)D_k+D_k^*D_k\right) \tag{1.71}$$

are negligible to the leading order of the correlation energy. Here, we offer a direct estimate for these terms, which is simpler than the strategy proposed in [6] and does not require a smallness condition on V. More precisely, in Theorem 2.4, we will prove that the non-bosonizable terms are bounded by $o(1)(k_F^{-1}\mathcal{N}_E H'_{\mathrm{kin}}+H'_{\mathrm{kin}}+k_F)$, and hence, the expectation against the low-lying eigenstates of H_N is of order $o(k_F)$ due to the a priori estimates mentioned before.

Bosonization of the kinetic operator and the excitation number operator

Concerning the bosonizable terms, while the interaction terms can be interpreted directly as a quadratic Hamiltonian in the quasi-bosonic picture as in (1.34), the treatment of the kinetic operator is more subtle. In fact, (1.37) does not hold as a direct operator approximation. Instead, we will justify it by appealing to the commutator relation

$$\left[H'_{\mathrm{kin}},b_{l,q}^*\right]\approx\left[2\sum_{k\in\mathbb{Z}^3_*}\sum_{p\in L_k}\lambda_{k,p}b_{k,p}^*b_{k,p},b_{l,q}^*\right]. \tag{1.72}$$

This commutator relation ensures that the difference

$$H'_{\mathrm{kin}}-2\sum_{k\in\mathbb{Z}^3_*}\sum_{p\in L_k}\lambda_{k,p}b_{k,p}^*b_{k,p} \tag{1.73}$$

is essentially invariant under the Bogolubov transformations introduced later, which is sufficient for our purpose. The approximation (1.72) is a consequence of the exact commutation relation (1.35): For every $p\in L_k=B_F^c\cap(B_F+k)$, by the CAR, we have

$$\begin{aligned}\left[H'_{\mathrm{kin}},b_{k,p}^*\right]&=\sum_{q\in B_F^c}|q|^2\left[c_q^*c_q,c_p^*c_{p-k}\right]-\sum_{q\in B_F}|q|^2\left[c_qc_q^*,c_p^*c_{p-k}\right]\\&=\sum_{q\in B_F^c}|q|^2\left[c_q^*c_q,c_p^*\right]c_{p-k}-\sum_{q\in B_F}|q|^2\,c_p^*\left[c_qc_q^*,c_{p-k}\right]\\&=\sum_{q\in B_F^c}|q|^2\,\delta_{q,p}c_q^*c_{p-k}-\sum_{q\in B_F}|q|^2\,\delta_{q,p-k}c_p^*c_{p-k}\\&=|p|^2c_p^*c_{p-k}-|p-k|^2c_p^*c_{p-k}=\left(|p|^2-|p-k|^2\right)b_{k,p}^*.\end{aligned} \tag{1.74}$$

A similar strategy was used in [6], although the analysis there is more complicated due to the averaging technique of the 'patches'. In particular, the operators on 'patches' in [6] do not obey the exact commutator relation $[H'_{\mathrm{kin}},b_{k,p}^*]=2\lambda_{k,p}b_{k,p}^*$, and so the kinetic operator has to be handled by an additional linearization argument.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Note that in the same manner of the dispersion relation in (1.74), we also have

$$\begin{aligned}[\mathcal{N}_E, b_{k,p}] &= \sum_{q\in B_F^c}[c_q^* c_q, c_{p-k}^* c_p] = \sum_{q\in B_F^c}\left(c_q^*[c_q, c_{p-k}^* c_p] + [c_q^*, c_{p-k}^* c_p]c_q\right)\\ &= \sum_{q\in B_F^c}\left(c_q^*\left(-c_{p-k}^*\{c_q,c_p\} + \{c_q, c_{p-k}^*\}c_p\right) + \left(-c_{p-k}^*\{c_q^*, c_p\} + \{c_q^*, c_{p-k}^*\}c_p\right)c_q\right)\\ &= -\sum_{q\in B_F^c}\left(\delta_{q,p}c_{p-k}^*\right)c_q = -c_{p-k}^* c_p = -b_{k,p} \end{aligned} \tag{1.75}$$

for all $k \in \mathbb{Z}_*^3$ and $p \in L_k$. This means that $\mathcal{N}_E$ plays the same role as the number operator in the bosonic picture.

Bogolubov transformation I

We will estimate the contribution of high momenta separately and only diagonalize the effective operator in (1.40) for low momenta. For this reason, we define a cutoff set

$$S_C = \overline{B}\left(0, k_F^\gamma\right) \cap \mathbb{Z}_+^3, \tag{1.76}$$

where $\gamma \in (0,1]$ will be optimized later. For a given k_F, we then diagonalize only

$$H'_{\text{eff}} = \sum_{k\in S_C}\left(2\sum_{p\in L_k}\lambda_{k,p}b_{k,p}^* b_{k,p} + 2\sum_{p\in L_{-k}}\lambda_{-k,p}b_{-k,p}^* b_{-k,p} + H_{\text{int}}^k\right) \tag{1.77}$$

and treat the remaining terms with $k \in \mathbb{Z}_+^3 \backslash S_C$ as an error term. As $\overline{B}\left(0, k_F^\gamma\right) \cap \mathbb{Z}_+^3$ forms an exhaustion of $\mathbb{Z}_+^3$, all terms are thus nonetheless diagonalized in the limit $k_F \to \infty$.

Inspired by the exact bosonic diagonalization (see Theorem 3.1 for details), we take the diagonalizing Bogolubov transformation to be of the form $e^{\mathcal{K}}$ for a generator $\mathcal{K} : \mathcal{H}_N \to \mathcal{H}_N$ defined by

$$\mathcal{K} = \sum_{k\in S_C}\left(\sum_{p\in L_k}\sum_{q\in L_{-k}}\langle e_p, K_k e_{-q}\rangle\left(b_{k,p}b_{-k,q} - b_{-k,q}^* b_{k,p}^*\right)\right), \tag{1.78}$$

where the transformation kernels $K_k : \ell^2(L_k) \to \ell^2(L_k)$, $k \in \mathbb{Z}_+^3$, are defined by

$$K_k = -\frac{1}{2}\log\left(h_k^{-\frac{1}{2}}\left(h_k^{\frac{1}{2}}\left(h_k + 2P_{v_k}\right)h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} h_k^{-\frac{1}{2}}\right) \tag{1.79}$$

with h_k, P_{v_k} as defined in equation (1.42). With this choice, we find that

$$e^{\mathcal{K}}H'_{\text{eff}}e^{-\mathcal{K}} \approx \sum_{k\in S_C\cup(-S_C)}\left(\operatorname{tr}\left(E_k - h_k\right) + 2\sum_{p,q\in L_k}\langle e_p, E_k e_q\rangle b_{k,p}^* b_{k,q}\right) \tag{1.80}$$

for

$$E_k = e^{-K_k}h_k e^{-K_k} \tag{1.81}$$

and by the commutation relation of equation (1.72), that

$$e^{\mathcal{K}}\left(H'_{\text{kin}} - 2\sum_{k\in S_C\cup(-S_C)}\sum_{p\in L_k}\lambda_{k,p}b_{k,p}^* b_{k,p}\right)e^{-\mathcal{K}} \approx H'_{\text{kin}} - 2\sum_{k\in S_C\cup(-S_C)}\sum_{p\in L_k}\lambda_{k,p}b_{k,p}^* b_{k,p} \tag{1.82}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

so by the equations (1.77), (1.80) and (1.82), noting also that $\langle e_p, h_k e_q \rangle = \delta_{p,q}\lambda_{k,p}$,

$$e^{\mathcal{K}}\left(H'_{\text{kin}} + \sum_{k\in S_C\cup(-S_C)} H^k_{\text{int}}\right)e^{-\mathcal{K}} \approx H'_{\text{kin}} + \sum_{k\in S_C\cup(-S_C)}\left(\operatorname{tr}(E_k - h_k) + 2\sum_{p,q\in L_k}\langle e_p, (E_k - h_k)e_q\rangle b^*_{k,p}b_{k,q}\right). \quad (1.83)$$

On the right side of (1.83), the constant $\sum_{k\in S_C\cup(-S_C)}\operatorname{tr}(E_k - h_k)$ captures correctly the leading order of the correlation energy E_{corr}. However, although E_k is isospectral to

$$\widetilde{E}_k = h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}} = \left(h_k^{\frac{1}{2}}(h_k + 2P_{v_k})h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} \geq h_k, \quad (1.84)$$

the operator $E_k - h_k$ is not non-negative. Thus the term $2\sum_{p,q\in L_k}\langle e_p, (E_k - h_k)e_q\rangle b^*_{k,p}b_{k,q}$ – a kind of second quantization of $E_k - h_k$ – cannot be ignored for the lower bound.

The Bogolubov transformation used in this part is analogous to that of [6]. It was proved in [6] that if V is small, then the quantization of $E_k - h_k$ can be controlled by H'_{kin}, leading to the desired lower bound on the ground state energy. In order to treat an arbitrary potential, we will instead utilize a second Bogolubov transformation which effectively replaces E_k by $\widetilde{E}_k$ in (1.83).

Bogolubov transformation II

We define the second Bogolubov transformation $e^{\mathcal{J}}$ for a generator $\mathcal{J} : \mathcal{H}_N \to \mathcal{H}_N$ defined by

$$\mathcal{J} = \sum_{k\in S_C\cup(-S_C)}\sum_{p,q\in L_k}\langle e_p, J_k e_q\rangle b^*_{k,p}b_{k,q}, \quad (1.85)$$

where $J_k = \log(U_k)$ denotes the (principal) logarithm of the unitary transformation $U_k : \ell^2(L_k) \to \ell^2(L_k)$ defined by

$$U_k = \left(h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} h_k^{-\frac{1}{2}} e^{K_k}. \quad (1.86)$$

This is precisely the unitary transformation which satisfies

$$U_k E_k U_k^* = h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}} = \left(h_k^{\frac{1}{2}}(h_k + 2P_{v_k})h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} = \widetilde{E}_k, \quad (1.87)$$

as is easily verified. This transformation acts such that

$$e^{\mathcal{J}}\left(\sum_{p,q\in L_k}\langle e_p, E_k e_q\rangle b^*_{k,p}b_{k,q}\right)e^{-\mathcal{J}} \approx \sum_{p,q\in L_k}\langle e_p, \widetilde{E}_k e_q\rangle b^*_{k,p}b_{k,q}, \quad (1.88)$$

and thanks to the relation of equation (1.72), also

$$e^{\mathcal{J}}\left(H'_{\text{kin}} - 2\sum_{k\in S_C\cup(-S_C)}\sum_{p\in L_k}\lambda_{k,p}b^*_{k,p}b_{k,p}\right)e^{-\mathcal{J}} \approx H'_{\text{kin}} - 2\sum_{k\in S_C\cup(-S_C)}\sum_{p\in L_k}\lambda_{k,p}b^*_{k,p}b_{k,p}, \quad (1.89)$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

so all in all,

$$e^{\mathcal{J}}e^{\mathcal{K}}\left(H'_{\text{kin}}+\sum_{k\in S_C\cup(-S_C)}H^k_{\text{int}}\right)e^{-\mathcal{K}}e^{-\mathcal{J}}$$
$$\approx\sum_{k\in S_C\cup(-S_C)}\operatorname{tr}(E_k-h_k)+H'_{\text{kin}}+\sum_{k\in S_C\cup(-S_C)}\sum_{p,q\in L_k}\left\langle e_p,(\widetilde{E}_k-h_k)e_q\right\rangle b^*_{k,p}b_{k,q}. \tag{1.90}$$

As $\widetilde{E}_k-h_k\geq 0$, the last term can now be dropped and the energy lower bound concluded. The cutoff S_C can be removed at the end without serious difficulties. On the technical level, the second Bogolubov transformation is an important new tool to remove the smallness condition of [6], thus enabling us to work with a significantly larger class of interaction potentials. In the independent work [8], the idea of using the second Bogolubov transformation has also been introduced to refine the method in [5, 6].

Elementary excitations

The key ingredient to obtain all bosonic elementary excitations is the formula (1.60) in Theorem 1.4. To prove this, note that $H_{\text{eff}}|_{\mathcal{N}_E=1}$ commutes with $\mathcal{N}_E$ and the total momentum $P=\sum_{p\in\mathbb{Z}^3_*}pc^*_pc_p$, so we may restrict H_{eff} to the simultanous eigenspaces of $\mathcal{N}_E$ and P, which are

$$\{\Psi\in\mathcal{H}_N\mid\mathcal{N}_E\Psi=\Psi,\ P\Psi=k\Psi\}=\operatorname{span}(b^*_{k,p}\psi_{\text{FS}})_{p\in L_k}=\left\{b^*_k(\varphi)\psi_{\text{FS}}\mid\varphi\in L^2(L_k)\right\}. \tag{1.91}$$

It turns out that the mapping $U_k:L^2(L_k)\to\{\Psi\in\mathcal{H}_N\mid\mathcal{N}_E\Psi=\Psi,\ P\Psi=k\Psi\}$ defined by

$$U_k\varphi=b^*_k(\varphi)\psi_{\text{FS}},\quad \varphi\in L^2(L_k) \tag{1.92}$$

is a unitary isomorphism with the property that

$$H_{\text{eff}}|_{\mathcal{N}_E=1}=U_k(2\widetilde{E}_k)U^*_k. \tag{1.93}$$

Summing over different momenta k's, we obtain the transformation $\tilde{U}$ introduced in (1.62).

In summary, our approach is different from the previous works [24, 5, 6] in many aspects. On the conceptual level, our direct bosonization method (i.e., working directly with the operators $b_{k,p}$ instead of averaging them on ‘patches’) allows us to stick closely to the heuristic argument of the physics literature and to obtain not only the ground state energy but also all bosonic elementary excitations, thus leading to the first complete justification of the RPA in the mean-field regime.

Although our general ideas are very transparent, to realize the whole procedure on a rigorous basis, we will need to develop several new estimates to justify all of the approximations made. In the rest of the paper, we will show how to implement the proof strategy rigorously.

Outline of the paper. In Section 2, we prove some general estimates involving the kinetic operator H_{kin} and bound the non-bosonizable terms. In Section 3, we review the theory of bosonic Bogolubov transformations; in particular, we review how one may explicitly define a Bogolubov transformation which diagonalizes a given positive-definite quadratic Hamiltonian. We then apply the bosonic theory to our study of the Fermi gas where we implement the diagonalization procedure in the quasi-bosonic framework. This is done by introducing the quasi-bosonic quadratic Hamiltonian in Section 4 and the quasi-bosonic Bogolubov transformation $e^{\mathcal{K}}$ in Section 5 (these notations mirror the exact bosonic ones as closely as possible such that the bosonic theory is easily transferred to the quasi-bosonic setting). In this way, the quasi-bosonic analysis reduces to that of a collection of exact bosonic quadratic Hamiltonians plus correlation exchange terms – error terms which arise due to the deviation from the exact CCR. In Section 6, we estimate the exchange terms, reducing the analysis of these to the associated one-body operators of the bosonic problem. The one-body operators are studied separately in Section 7. In this part, we will need several estimates of Riemann sums, which are collected in the Appendix.

We complete the analysis of the transformation $e^{\mathcal{K}}$ in Section 8, where we prove that $H'_{\rm kin}$ and $\mathcal{N}_E$ are stable under the transformation $e^{\mathcal{K}}$. In Section 9, we introduce the second unitary transformation $e^{\mathcal{J}}$. The analysis of this transformation is essentially similar to the first one, except that we require new one-body operator estimates which are somewhat more difficult. Finally, we conclude the proofs of the main theorems in Section 10.

2. Removal of the non-bosonizable terms

In this section, we collect several basic estimates concerning the operator H_N which can be obtained without using Bogolubov transformations. Recall the decomposition (1.22)

$$H_N = E_{\rm FS} + H'_{\rm kin} + k_F^{-1} H'_{\rm int}, \quad E_{\rm FS} = \langle \psi_{\rm FS}, H_N \psi_{\rm FS} \rangle . \tag{2.1}$$

We will bound the interaction operator $H'_{\rm int}$ in terms of the kinetic operator $H'_{\rm kin}$ and then prove a priori estimates for eigenstates of H_N which are parts of Theorem 1.2.

Recall the following result from [6, Lemma 2.4] concerning the kinetic operator $H'_{\rm kin}$ in (1.11).

Proposition 2.1. *We have $H'_{\rm kin} \geq \mathcal{N}_E$ with $\mathcal{N}_E$ given in (1.13).*

Proof. Since $|p|^2$ is an integer for $p \in \mathbb{Z}^3$, our assumption $|B_F| = N$ implies that

$$\inf_{p \in B_F^c} |p|^2 - \sup_{p \in B_F} |p|^2 \geq 1. \tag{2.2}$$

Therefore, in (1.14), we can choose ζ such that $||p|^2 - \zeta| \geq 1/2$ for all $p \in \mathbb{Z}^3$. □

Next, we consider the bosonizable terms in $H'_{\rm int}$. The following result is a minor extension of [24, Lemma 4.7] (see also [6, Appendix B] for a simplified proof).

Proposition 2.2. *For all $k \in \mathbb{Z}^3_*$, the operator $\tilde{B}_k$ in (1.24) satisfies that*

$$\tilde{B}_k^* \tilde{B}_k \leq C k_F H'_{\rm kin}, \quad \tilde{B}_k \tilde{B}_k^* \leq C k_F (H'_{\rm kin} + |k| k_F),$$

where the constant $C > 0$ is independent of k and k_F.

Proof. As argued in [24, 6], for any $\Psi \in \mathcal{H}_N$, it follows from the triangle and Cauchy-Schwarz inequalities that

$$\left\| \tilde{B}_k \Psi \right\| = \left\| \sum_{p \in L_k} c^*_{p-k} c_p \Psi \right\| \leq \sum_{p \in L_k} \left\| c^*_{p-k} c_p \Psi \right\| \leq \sqrt{\sum_{p \in L_k} \lambda_{k,p}^{-1}} \sqrt{\sum_{p \in L_k} \lambda_{k,p} \left\| c^*_{p-k} c_p \Psi \right\|^2}, \tag{2.3}$$

where $\lambda_{k,p} = \frac{1}{2}(|p|^2 - |p-k|^2)$. Using (1.14) and Pauli's exclusion principle $\|c_p\|_{\rm op} \leq 1$, $\|c_p^*\|_{\rm op} \leq 1$, we find that

$$\begin{aligned} \sum_{p \in L_k} \lambda_{k,p} \left\| c^*_{p-k} c_p \Psi \right\|^2 &= \frac{1}{2} \sum_{p \in L_k} \left(||p|^2 - \zeta| + ||p-k|^2 - \zeta| \right) \left\| c^*_{p-k} c_p \Psi \right\|^2 \\ &\leq \frac{1}{2} \sum_{p \in L_k} ||p|^2 - \zeta| \left\| c_p \Psi \right\|^2 + \frac{1}{2} \sum_{p \in L_k} ||p-k|^2 - \zeta| \left\| c^*_{p-k} \Psi \right\|^2 \\ &\leq \frac{1}{2} \sum_{p \in B_F^c} ||p|^2 - \zeta| \left\| c_p \Psi \right\|^2 + \frac{1}{2} \sum_{p \in B_F} ||p|^2 - \zeta| \left\| c^*_p \Psi \right\|^2 = \frac{1}{2} \left\langle \Psi, H'_{\rm kin} \Psi \right\rangle . \end{aligned} \tag{2.4}$$

Thus, it remains to show that $\sum_{p \in L_k} \lambda_{k,p}^{-1} \leq C k_F$. For $|k| \sim O(1)$, this bound was already proved in [24, 6]. For completeness, we will establish this bound for all $k \in \mathbb{Z}^3_*$ in the Appendix (Proposition A.2).

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Thus, in summary,

$$\tilde{B}_k^* \tilde{B}_k \le \frac{1}{2} \left(\sum_{p \in L_k} \lambda_{k,p}^{-1} \right) H'_{\rm kin} \le C k_F H'_{\rm kin}. \tag{2.5}$$

Then the bound for $\tilde{B}_k \tilde{B}_k^*$ follows from the fact that

$$\left[\tilde{B}_k, \tilde{B}_k^*\right] = |L_k| - \sum_{p \in L_k} c_p^* c_p - \sum_{p \in L_k} c_{p-k} c_{p-k}^* \le |L_k| \le C|k| k_F^2. \tag{2.6}$$

In the last estimate, we used $|L_k| \le C k_F^2 |k|$ for all $k \in \mathbb{Z}_*^3$ (see Proposition A.1 for details). □

For the non-bosonizable terms in $H'_{\rm int}$, it was proved in [5, Eq. (5.1)] that

$$D_k^* D_k \le 4 \mathcal{N}_E^2. \tag{2.7}$$

However, this bound is not optimal for low-lying eigenfunctions (for which $\mathcal{N}_E \sim k_F$). In order to remove the non-bosonizable terms completely, we need the following improvement.

Proposition 2.3. *For all $k \in \mathbb{Z}_*^3$ and any $0 < \lambda \le \frac{1}{6} k_F^2$, the operator D_k in (1.23) satisfies*

$$D_k^* D_k \le C \left(|k|^{-1} \lambda + |k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}} \right) (\lambda + |k|) \mathcal{N}_E + C \lambda^{-\frac{1}{2}} \mathcal{N}_E H'_{\rm kin}$$

for a constant $C > 0$ independent of k, k_F and λ.

In applications, we will eventually choose $\lambda = k_F^{2\gamma}/|k|^4$ for some constant $\gamma \in (0, 1/9)$.

Proof. For $k \in \mathbb{Z}_*^3$, we write $D_k = D_k^1 + D_k^2$ as in (1.24), namely,

$$D_k^1 = \sum_{q \in B_F \cap (B_F + k)} c_{q-k}^* c_q, \quad D_k^2 = \sum_{q \in B_F^c \cap \left(B_F^c + k\right)} c_{q-k}^* c_q. \tag{2.8}$$

By the Cauchy–Schwarz inequality,

$$D_k^* D_k \le 2 \left((D_k^1)^* D_k^1 + (D_k^2)^* D_k^2 \right). \tag{2.9}$$

We will estimate $(D_k^1)^* D_k^1$ in detail, with the estimate of $(D_k^2)^* D_k^2$ being similar. We have

$$\begin{aligned} (D_k^1)^* D_k^1 &= \sum_{p,q \in B_F \cap (B_F + k)} c_p^* c_{p-k} c_{q-k}^* c_q = \sum_{p,q \in B_F \cap (B_F + k)} (\delta_{p,q} c_{p-k} c_{p-k}^* - c_{p-k} c_q c_p^* c_{q-k}^*) \\ &= \sum_{p \in B_F \cap (B_F + k)} c_{p-k} c_{p-k}^* - \frac{1}{2} \sum_{p,q \in B_F \cap (B_F + k)} (c_{p-k} c_q c_p^* c_{q-k}^* + h.c.). \end{aligned} \tag{2.10}$$

Here, we used $k \neq 0$ so that c_{p-k} and c_p^* anti-commute. By the definition of $\mathcal{N}_E$ in (1.13),

$$\sum_{p \in B_F \cap (B_F + k)} c_{p-k} c_{p-k}^* \le \sum_{p \in B_F} c_p c_p^* = \mathcal{N}_E. \tag{2.11}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Moreover, by the Cauchy–Schwarz inequality, for all $\epsilon_p > 0$, we get

$$
\begin{aligned}
&\pm \frac{1}{2} \sum_{p,q \in B_F \cap (B_F+k)} (c_{p-k} c_q c_p^* c_{q-k}^* + h.c.) \\
&\leq \frac{1}{2} \sum_{p,q \in B_F \cap (B_F+k)} (\epsilon_p c_{p-k} c_q c_q^* c_{p-k}^* + \epsilon_p^{-1} c_{q-k} c_p c_p^* c_{q-k}^*) \\
&\leq \frac{1}{2} \sum_{p,q \in B_F \cap (B_F+k)} (\epsilon_p c_{p-k} c_{p-k}^* c_q c_q^* + \epsilon_p^{-1} c_{q-k} c_{q-k}^* c_p c_p^*) \\
&\leq \frac{1}{2} \sum_{p \in B_F \cap (B_F+k)} (\epsilon_p c_{p-k} c_{p-k}^* + \epsilon_p^{-1} c_p c_p^*) \mathcal{N}_E .
\end{aligned} \tag{2.12}
$$

By taking $\epsilon_p \equiv 1$, we obtain immediately $(D_k^1)^* D_k^1 \leq \mathcal{N}_E^2$, which, together with a similar bound for D_k^2, leads to (2.7). To improve on this, we have to choose ϵ_p differently.

Recall that in (1.14) we can choose $\zeta \in [\sup_{p \in B_F} |p|^2, \inf_{p \in B_F^c} |p|^2]$ such that $||p|^2 - \zeta| \geq 1/2$ for all $p \in \mathbb{Z}^3$. For any $\lambda > 0$, we can split

$$
B_F \cap (B_F + k) = S_{k,\lambda}^1 \cup S_{k,\geq\lambda}^1, \tag{2.13}
$$

where

$$
\begin{aligned}
S_{k,\lambda}^1 &= \left\{ p \in B_F \cap (B_F + k) \mid \max \left\{ ||p|^2 - \zeta|, ||p-k|^2 - \zeta| \right\} < \lambda \right\}, \\
S_{k,\geq\lambda}^1 &= \left\{ p \in B_F \cap (B_F + k) \mid \max \left\{ ||p|^2 - \zeta|, ||p-k|^2 - \zeta| \right\} \geq \lambda \right\}.
\end{aligned} \tag{2.14}
$$

Choosing $\epsilon_p = 1$ for $p \in S_{k,\lambda}^1$ and using $\|c_p^*\|_{\mathrm{op}} \leq 1$, we get

$$
\frac{1}{2} \sum_{p \in S_{k,\lambda}^1} (\epsilon_p c_{p-k} c_{p-k}^* + \epsilon_p^{-1} c_p c_p^*) \leq |S_{k,\lambda}^1|. \tag{2.15}
$$

Choosing $\epsilon_p = \sqrt{||p-k|^2 - \zeta|} / \sqrt{||p|^2 - \zeta|}$ for $p \in S_{k,\geq\lambda}^1$, we have

$$
\begin{aligned}
&\sum_{p \in S_{k,\geq\lambda}^1} (\epsilon_p c_{p-k} c_{p-k}^* + \epsilon_p^{-1} c_p c_p^*) \\
&= \sum_{p \in S_{k,\geq\lambda}^1} \frac{1}{\sqrt{||p|^2 - \zeta| \cdot ||p-k|^2 - \zeta|}} (||p-k|^2 - \zeta| c_{p-k} c_{p-k}^* + ||p|^2 - \zeta| c_p c_p^*) \\
&\leq \sum_{p \in S_{k,\geq\lambda}^1} \frac{1}{\sqrt{\lambda/2}} (||p-k|^2 - \zeta| c_{p-k} c_{p-k}^* + ||p|^2 - \zeta| c_p c_p^*) \leq \frac{2\sqrt{2}}{\sqrt{\lambda}} H'_{\mathrm{kin}} .
\end{aligned} \tag{2.16}
$$

Here, we used that among two factors $||p|^2 - \zeta|$ and $||p-k|^2 - \zeta|$, there is at least one $\geq \lambda$ due to the assumption $p \in S_{k,\geq\lambda}^1$, and the other one is trivially $\geq 1/2$. In summary,

$$
(D_k^1)^* D_k^1 \leq |S_{k,\lambda}^1| \mathcal{N}_E + C \lambda^{-\frac{1}{2}} H'_{\mathrm{kin}} \mathcal{N}_E . \tag{2.17}
$$

Similarly, we have

$$
(D_k^2)^* D_k^2 \leq |S_{k,\lambda}^2| \mathcal{N}_E + C \lambda^{-\frac{1}{2}} H'_{\mathrm{kin}} \mathcal{N}_E , \tag{2.18}
$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where

$$S^2_{k,\lambda} = \left\{p \in B_F^c \cap \left(B_F^c + k\right) \mid \max\left\{||p|^2 - \zeta|, ||p-k|^2 - \zeta|\right\} < \lambda\right\}. \tag{2.19}$$

The desired conclusion of $D_k^* D_k$ follows from the bound

$$|S^1_{k,\lambda}| + |S^2_{k,\lambda}| \le C\left(|k|^{-1}\lambda + |k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}}\right)(\lambda + |k|) \tag{2.20}$$

whose proof can be found in Proposition A.4 in the Appendix. □

2.1. Estimation of the non-bosonizable terms

Now we are ready to remove the non-bosonizable terms, namely, the terms involving operators D_k in the decomposition (1.28) of the interaction operator:

$$k_F^{-1} H'_{\text{int}} = \sum_{k\in\mathbb{Z}^3_+}\left(H^k_{\text{int}} - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|\right) + \frac{k_F^{-1}}{(2\pi)^3}\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\left(\tilde{B}_k^* D_k + D_k^*\tilde{B}_k + \frac{1}{2}D_k^* D_k\right), \tag{2.21}$$

where H^k_{int} is defined in (1.29). Moreover, for technical reasons, we will also impose a momentum cutoff in the bosonizable terms. Recall the set S_C in (1.76). Define

$$\mathcal{E}_{\text{NB}} = k_F^{-1} H'_{\text{int}} - \sum_{k\in S_C}\left(H^k_{\text{int}} - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|\right). \tag{2.22}$$

Proposition 2.4. *Let $\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k| < \infty$. Then for all $\gamma \in (0, 1/9)$ in S_C, we have*

$$\pm\mathcal{E}_{\text{NB}} \le C k_F^{-\gamma/2}\left(H'_{\text{kin}} + k_F^{-1}\mathcal{N}_E H'_{\text{kin}} + k_F\right).$$

Here, the constant $C > 0$ depends only on V (in particular, it is independent of k, k_F and λ).

We write $\pm X \le Y$ for two operator inequalities $X \le Y$ and $-X \le Y$.

Proof. For the bosonizable terms, by (2.6), Proposition 2.2 and Proposition 2.1, we can bound

$$\pm(\{\tilde{B}_k^*, \tilde{B}_k\} - |L_k|) = \pm\left(2\tilde{B}_k^*\tilde{B}_k - \sum_{p\in L_k} c_p^* c_p - \sum_{p\in L_k} c_{p-k} c_{p-k}^*\right) \le 2\tilde{B}_k^*\tilde{B}_k + \mathcal{N}_E \le C k_F H'_{\text{kin}} \tag{2.23}$$

for all $k \in \mathbb{Z}^3_*$. Moreover, by the Cauchy–Schwarz inequality,

$$\pm(\tilde{B}_k^*\tilde{B}_{-k}^* + \tilde{B}_{-k}\tilde{B}_k) \le |k|^{-1/2}\tilde{B}_{-k}\tilde{B}_{-k}^* + |k|^{1/2}\tilde{B}_k^*\tilde{B}_k \le C|k|^{1/2} k_F (H'_{\text{kin}} + k_F) \tag{2.24}$$

for all $k \in \mathbb{Z}^3_*$. Combining (2.23) and (2.24), we find that

$$\begin{aligned} \pm\sum_{k\in\mathbb{Z}^3_+\setminus S_C}\left(H^k_{\text{int}} - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|\right) &\le C(H'_{\text{kin}} + k_F)\sum_{k\in\mathbb{Z}^3_+\setminus S_C}\hat{V}_k|k|^{1/2} \\ &\le C(H'_{\text{kin}} + k_F) k_F^{-\gamma/2}\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k|. \end{aligned} \tag{2.25}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

For the non-bosonizable terms, by the Cauchy–Schwarz inequality and Proposition 2.2, we have

$$\pm\left(\tilde{B}_k^* D_k + D_k^* \tilde{B}_k\right) \le k_F^{-\gamma/2}|k|\tilde{B}_k^*\tilde{B}_k + k_F^{\gamma/2}|k|^{-1}D_k^*D_k \le Ck_F^{1-\gamma/2}|k|H'_{\mathrm{kin}} + k_F^{\gamma/2}|k|^{-1}D_k^*D_k, \quad (2.26)$$

and hence,

$$\pm\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\left(\tilde{B}_k^*D_k + D_k^*\tilde{B}_k + D_k^*D_k\right) \le Ck_F^{1-\gamma/2}H'_{\mathrm{kin}} + \sum_{k\in\mathbb{Z}^3_*}\hat{V}_k(k_F^{\gamma/2}|k|^{-1}+1)D_k^*D_k. \quad (2.27)$$

Let us decompose the sum on the right-hand side of (2.27) into the high-momenta $|k| > k_F^{\gamma/2}$ and the low-momenta $|k| \le k_F^{\gamma/2}$. For the high-momenta, from the simple bound (2.7), we get

$$\sum_{k\in\mathbb{Z}^3_*,|k|>k_F^{\gamma/2}}\hat{V}_k(k_F^{\gamma/2}|k|^{-1}+1)D_k^*D_k \le Ck_F^{-\gamma/2}\mathcal{N}_E^2\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k|. \quad (2.28)$$

For the low-momenta, using Proposition 2.3 with $\lambda = k_F^{\gamma}/|k|^2$, we have

$$D_k^*D_k \le C\left(k_F^{2\gamma} + k_F^{\gamma}|k|^{2/3}(\log k_F)^{\frac{2}{3}}k_F^{2/3} + |k|^{3+2/3}(\log k_F)^{\frac{2}{3}}k_F^{2/3}\right)|k|\mathcal{N}_E + Ck_F^{-\gamma/2}|k|\mathcal{N}_E H'_{\mathrm{kin}}, \quad (2.29)$$

and hence,

$$\begin{aligned}\sum_{k\in\mathbb{Z}^3_*,|k|\le k_F^{\gamma/2}}\hat{V}_kD_k^*D_k &\le C\left(\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k|\right)\left(k_F^{(3+\frac{2}{3})\frac{\gamma}{2}+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}\mathcal{N}_E + k_F^{-\frac{\gamma}{2}}\mathcal{N}_E H'_{\mathrm{kin}}\right)\\ &\le Ck_F^{-\gamma/2}\left(k_F\mathcal{N}_E + \mathcal{N}_E H'_{\mathrm{kin}}\right)\end{aligned} \quad (2.30)$$

for all $\gamma \in (0, 1/7)$. Moreover, using Proposition 2.3 with $\lambda = k_F^{2\gamma}/|k|^4$, we have

$$\begin{aligned}k_F^{\gamma/2}|k|^{-1}D_k^*D_k \le\ & C\left(k_F^{4\gamma} + k_F^{(2+\frac{1}{2})\gamma}(\log k_F)^{\frac{2}{3}}k_F^{2/3} + k_F^{\frac{\gamma}{2}}|k|^{2+2/3}(\log k_F)^{\frac{2}{3}}k_F^{2/3}\right)|k|\mathcal{N}_E\\ &+ Ck_F^{-\gamma/2}|k|\mathcal{N}_E H'_{\mathrm{kin}},\end{aligned} \quad (2.31)$$

and hence,

$$\begin{aligned}\sum_{k\in\mathbb{Z}^3_*,|k|\le k_F^{\gamma/2}}\hat{V}_k k_F^{\gamma/2}|k|^{-1}D_k^*D_k &\le C\left(\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k|\right)\left(k_F^{(2+\frac{1}{2})\gamma+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}\mathcal{N}_E + k_F^{-\frac{\gamma}{2}}\mathcal{N}_E H'_{\mathrm{kin}}\right)\\ &\le Ck_F^{-\gamma/2}\left(k_F\mathcal{N}_E + \mathcal{N}_E H'_{\mathrm{kin}}\right)\end{aligned} \quad (2.32)$$

for all $\gamma \in (0, 1/9)$. Inserting (2.28), (2.30) and (2.32) in (2.27) and using Proposition 2.1, we conclude that

$$\pm\frac{k_F^{-1}}{(2\pi)^3}\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\left(\tilde{B}_k^*D_k + D_k^*\tilde{B}_k + D_k^*D_k\right) \le Ck_F^{-\gamma/2}\left(H'_{\mathrm{kin}} + k_F^{-1}\mathcal{N}_E H'_{\mathrm{kin}}\right) \quad (2.33)$$

for all $\gamma \in (0, 1/9)$. The conclusion follows from (2.25) and (2.33). □

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

3. Overview of bosonic Bogolubov transformations

In this section, we review the general theory of quadratic Hamiltonians and Bogolubov transformations in the *exact* bosonic setting. Later, in the remainder of the paper, the analysis here will be adapted to handle the *quasi-bosonic* case where error terms have to be estimated carefully.

The study of bosonic quadratic Hamiltonians goes back to Bogolubov's 1947 paper [10] where he proposed an effective Hamiltonian to describe the excitation spectrum of weakly interacting Bose gases. An important property of quadratic Hamiltonians is that they can be diagonalized by suitable Bogolubov transformations; see, for example, [2, 31, 16] for recent results in the infinite dimensional cases. For our application, we will only focus on the situation where the one-body Hilbert space is *real* and *finite dimensional*. Historically, the diagonalization problem in finite dimensions can be solved abstractly by using Williamson's theorem [43]. We refer to [27] and [16, Section 2] for systematic discussions on the finite dimensional case.

In the present paper, we will need an explicit construction of the diagonalizing transformations so that we can adapt this to the quasi-bosonic operators. Such an explicit construction can be found in [23], which was also used in the fermionic context in [5, 6] and will be recalled below. Here, we will offer a slightly different treatment of Bogolubov transformations, in that we will view *quadratic operators* on Fock spaces as the fundamental object of study rather than the creation and annihilation operators.

Notation. We will denote by V a finite-dimensional real Hilbert space and let $n = \dim(V)$. The bosonic Fock space associated to V is

$$\mathcal{F}^+(V) = \bigoplus_{N=0}^{\infty} \bigotimes_{\mathrm{Sym}}^{N} V, \tag{3.1}$$

where $\bigotimes_{\mathrm{Sym}}^{N} V$ denotes the space of symmetric N-fold tensor products of V. To any element $\varphi \in V$, there are associated two operators on $\mathcal{F}^+(V)$: the annihilation operator $a(\varphi)$ and the creation operator $a^*(\varphi)$. These are (formal) adjoints of one another and obey the canonical commutation relations (CCR): for any $\varphi, \psi \in V$,

$$[a(\varphi), a(\psi)] = [a^*(\varphi), a^*(\psi)] = 0, \quad [a(\varphi), a^*(\psi)] = \langle \varphi, \psi \rangle . \tag{3.2}$$

Additionally, the mappings $\varphi \mapsto a(\varphi)$, $\varphi \mapsto a^*(\varphi)$ are linear.[6]

3.1. *Quadratic Hamiltonians*

Similarly to how we can to any $\varphi \in V$ associate the two operators $a(\varphi)$ and $a^*(\varphi)$, we may also associate two types of symmetric operators on $\mathcal{F}^+(V)$ to any symmetric operator on V. For the definition, we let $(e_i)_{i=1}^n$ denote an orthonormal basis of V. Given any symmetric operator $A : V \to V$, we then define the operator $Q_1(A)$ on $\mathcal{F}^+(V)$ by

$$Q_1(A) = \sum_{i,j=1}^{n} \langle e_i, A e_j \rangle \left(a^*(e_i) a(e_j) + a(e_j) a^*(e_i) \right), \tag{3.3}$$

and likewise, for any symmetric operator $B : V \to V$, we define the operator $Q_2(B)$ by

$$Q_2(B) = \sum_{i,j=1}^{n} \langle e_i, B e_j \rangle \left(a^*(e_i) a^*(e_j) + a(e_j) a(e_i) \right). \tag{3.4}$$

[6]If V is a complex Hilbert space space, the mapping $\varphi \mapsto a(\varphi)$ is anti-linear which complicates the exposition. In our quasi-bosonic application, although the relevant Hilbert spaces are complex, all relevant operators have real matrix elements, and hence, it suffices to restrict to the case of real spaces as in this section.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

These definitions are independent of the basis chosen, and we can write equivalently

$$Q_1(A) = \sum_{i=1}^{n} \left(a^*(Ae_i)a(e_i) + a(e_i)a^*(Ae_i)\right) \tag{3.5}$$
$$Q_2(B) = \sum_{i=1}^{n} \left(a^*\left(Be_i\right) a^*(e_i) + a(e_i)a\left(Be_i\right)\right).$$

Thus, for real, symmetric $A, B : V \to V$, we can define a quadratic Hamiltonian on $\mathcal{F}^+(V)$ by

$$H = Q_1(A) + Q_2(B). \tag{3.6}$$

Note that by the CCR, we may express $Q_1(A)$ as

$$Q_1(A) = 2\sum_{i,j=1}^{n} \left\langle e_i, Ae_j\right\rangle a^*(e_i)a(e_j) + \operatorname{tr}(A) = 2\,\mathrm{d}\Gamma(A) + \operatorname{tr}(A), \tag{3.7}$$

where $\mathrm{d}\Gamma(A)$ denotes the second quantization of $A : V \to V$. Sometimes in the literature, in particular in infinite dimensions, quadratic Hamiltonians are defined by $\mathrm{d}\Gamma(A) + Q_2(B)$, which is the same to our definition up to the constant $\operatorname{tr}(A)$. Here, we prefer to use $Q_1(A)$ instead of $\mathrm{d}\Gamma(\cdot)$; the reason for this is that the relations of Proposition 3.4 below are symmetric in the Q's.

Note that the basis-independence is a nice property of the real space setting. In general, if V is a complex Hilbert space and B is symmetric, then the definition of $Q_2(B)$ in (3.4) may depend on the basis. In fact, we can obtain a basis-independent formulation in the complex case, but the mapping $B \mapsto Q_2(B)$ is not to be defined for symmetric linear operators B, but rather symmetric *anti-linear* operators B to make up for the fact that in the complex case the assignment $\varphi \mapsto a(\varphi)$ is also anti-linear. This is unimportant for our application, which is why we only consider real Hilbert spaces in this section, for the sake of simplicity.

3.2. *Bogolubov transformations*

In this subsection, we review an explicit construction of a *Bogolubov transformation* $\mathcal{U} : \mathcal{F}^+(V) \to \mathcal{F}^+(V)$ that diagonalizes the quadratic Hamiltonian $H = Q_1(A) + Q_2(B)$, namely,

$$\mathcal{U}H\mathcal{U}^* = Q_1(E) \tag{3.8}$$

for a real, symmetric operator $E : V \to V$. Such a construction is well known; see, for example, [16] for a recent review. We consider a unitary transformation $\mathcal{U} = e^{\mathcal{K}}$ where $\mathcal{K}$ is an anti-symmetric operator on $\mathcal{F}^+(V)$ of the following form:

$$\mathcal{K} = \frac{1}{2}\sum_{i,j=1}^{n} \langle e_i, Ke_j\rangle \left(a(e_i)a(e_j) - a^*(e_j)a^*(e_i)\right) = \frac{1}{2}\sum_{i=1}^{n} \left(a(Ke_i)a(e_i) - a^*(e_i)a^*(Ke_i)\right). \tag{3.9}$$

Here, $K : V \to V$ is a symmetric operator (called the *transformation kernel*) and $(e_i)_{i=1}^n$ denotes any orthonormal basis of V (as with $Q_1(\cdot)$ and $Q_2(\cdot)$ this definition is independent of the basis).

In this subsection, we discuss the following:

Theorem 3.1. *Let $A, B : V \to V$ be real, symmetric operators such that $A \pm B > 0$ (namely, $A + B > 0$ and $A - B > 0$). Consider the Bogolubov transformation $e^{\mathcal{K}}$ where $\mathcal{K}$ is given in (3.9) with*

$$K = -\frac{1}{2}\log\left((A-B)^{-\frac{1}{2}}\left((A-B)^{\frac{1}{2}}(A+B)(A-B)^{\frac{1}{2}}\right)^{\frac{1}{2}}(A-B)^{-\frac{1}{2}}\right).$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Then

$$e^{\mathcal{K}}(Q_1(A)+Q_2(B))e^{-\mathcal{K}} = Q_1(E) = 2\,\mathrm{d}\Gamma(E) + \mathrm{tr}(E),$$

where

$$E = e^{K}(A+B)e^{K} = e^{-K}(A-B)e^{-K}.$$

Moreover, the diagonalizing K is uniquely determined by this.

In the following, we will prove Theorem 3.1 by using a generalization and simplification of the argument used in [23, 5]. We will first discuss the action of the Bogolubov transformation with a general kernel K and then explain where the diagonalization condition comes from.

Let us start with some basic properties of $\mathcal{K}$.

Proposition 3.2. *For any symmetric operator $K : V \to V$, the operator $\mathcal{K}$ defined by (3.9) is an anti-symmetric operator on $\mathcal{F}^+(V)$ and obeys the commutators*

$$[\mathcal{K}, a(\varphi)] = a^*(K\varphi), \quad [\mathcal{K}, a^*(\varphi)] = a(K\varphi), \quad \forall \varphi \in V.$$

Thus, $[\mathcal{K}, \cdot]$ acts on the creation and annihilation operators by ‘swapping’ each type into the other and applying the operator K to their arguments. From this, one can now deduce that the unitary transformation $e^{\mathcal{K}}$ acts on the creation and annihilation operators according to

$$\begin{aligned} e^{\mathcal{K}}a(\varphi)e^{-\mathcal{K}} &= a(\cosh(K)\varphi) + a^*(\sinh(K)\varphi) \\ e^{\mathcal{K}}a^*(\varphi)e^{-\mathcal{K}} &= a^*(\cosh(K)\varphi) + a(\sinh(K)\varphi), \end{aligned} \tag{3.10}$$

since by the Baker-Campbell-Hausdorff formula,

$$\begin{aligned} e^{\mathcal{K}}a(\varphi)e^{-\mathcal{K}} &= a(\varphi) + \frac{1}{1!}[\mathcal{K}, a(\varphi)] + \frac{1}{2!}[\mathcal{K},[\mathcal{K}, a(\varphi)]] + \frac{1}{3!}[\mathcal{K},[\mathcal{K},[\mathcal{K}, a(\varphi)]]] + \cdots \\ &= a(\varphi) + \frac{1}{1!}a^*(K\varphi) + \frac{1}{2!}a\left(K^2\varphi\right) + \frac{1}{3!}a^*\left(K^3\varphi\right) + \cdots \\ &= a\left(\varphi + \frac{1}{2!}K^2\varphi + \cdots\right) + a^*\left(\frac{1}{1!}K\varphi + \frac{1}{3!}K^3\varphi + \cdots\right) \\ &= a(\cosh(K)\varphi) + a^*(\sinh(K)\varphi), \end{aligned} \tag{3.11}$$

and the identity for $e^{\mathcal{K}}a^*(\varphi)e^{-\mathcal{K}}$ then follows immediately by taking the adjoint.

Now let us consider $e^{\mathcal{K}}Q_1(\cdot)e^{-\mathcal{K}}$ and $e^{\mathcal{K}}Q_2(\cdot)e^{-\mathcal{K}}$. For this, we will first make an observation on their structure which will greatly simplify computations: namely, we note that the operators $Q_1(A)$ and $Q_2(B)$ are both of a ‘trace-form’ in the sense that we can write, say, $Q_1(A) = \sum_{i=1}^n q(e_i, Ae_i)$, where

$$q(x,y) = a^*(y)\,a(x) + a(x)a^*(y) \tag{3.12}$$

defines a bilinear mapping from $V \times V$ into the space of operators on $\mathcal{F}^+(V)$, similar to how the trace of an operator T is $\mathrm{tr}(T) = \sum_{i=1}^n q(e_i, Te_i)$ for $q(x,y) = \langle x, y\rangle$. This abstract viewpoint is worth noting because all such expressions are both basis-independent and obey an additional property, which for the trace is just the familiar cyclicity property. Since we will encounter such ‘trace-form’ expressions repeatedly during computations throughout this paper, we state this property in full generality. In the following, we take sesquilinear to mean anti-linear in the first argument and linear in the second (we note that in the present real case a sesquilinear mapping is of course just a bilinear mapping, but stating it in this generality will prove useful later).

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Lemma 3.3. *Let $(V, \langle\cdot,\cdot\rangle)$ be an n-dimensional Hilbert space and let $q : V\times V \to W$ be a sesquilinear mapping into a vector space W. Let $(e_i)_{i=1}^n$ be an orthonormal basis for V. Then for any linear operators $S,T : V \to V$, it holds that*

$$\sum_{i=1}^n q\left(Se_i, Te_i\right) = \sum_{i=1}^n q\left(ST^*e_i, e_i\right).$$

As a consequence, the expression $\sum_{i=1}^n q\left(e_i, e_i\right)$ is independent of the chosen basis.

Proof. By orthonormal expansion, we find that

$$\begin{aligned}\sum_{i=1}^n q\left(Se_i, Te_i\right) &= \sum_{i=1}^n q\left(Se_i, \sum_{j=1}^n \left\langle e_j, Te_i\right\rangle e_j\right) = \sum_{j=1}^n q\left(\sum_{i=1}^n \left\langle Te_i, e_j\right\rangle Se_i, e_j\right)\\ &= \sum_{j=1}^n q\left(S\sum_{i=1}^n \left\langle e_i, T^*e_j\right\rangle e_i, e_j\right) = \sum_{i=1}^n q\left(ST^*e_i, e_i\right).\end{aligned} \tag{3.13}$$

The basis independence follows from the fact that for all unitary transformation $U : V \to V$,

$$\sum_{i=1}^n q\left(Ue_i, Ue_i\right) = \sum_{i=1}^n q\left(UU^*e_i, e_i\right) = \sum_{i=1}^n q\left(e_i, e_i\right). \tag{3.14}$$

□

The lemma thus allows us to move a mapping from one argument to the other when under a sum, which will be immensely useful when simplifying expressions. As mentioned, this can indeed be seen as a generalization of the cyclicity property of the trace, since the lemma implies

$$\operatorname{tr}(ST) = \sum_{i=1}^n \langle e_i, STe_i\rangle = \sum_{i=1}^n \langle S^*e_i, Te_i\rangle = \sum_{i=1}^n \langle S^*T^*e_i, e_i\rangle = \sum_{i=1}^n \langle e_i, TSe_i\rangle = \operatorname{tr}(TS), \tag{3.15}$$

but it is important to note that cyclicity is not a general property of trace-form sums; the assignments $A \mapsto Q_1(A)$ and $B \mapsto Q_2(B)$ do not obey such a property.

With the lemma, we can now easily derive the commutator of $\mathcal{K}$ with $Q_1(\cdot)$ and $Q_2(\cdot)$:

Proposition 3.4. *For any real, symmetric operators $A, B, K : V \to V$, the operator $\mathcal{K}$ defined by equation (3.9) obeys the following commutators on $\mathcal{F}^+(V)$:*

$$\begin{aligned}[\mathcal{K}, Q_1(A)] &= Q_2\left(\{K, A\}\right)\\ [\mathcal{K}, Q_2(B)] &= Q_1\left(\{K, B\}\right).\end{aligned}$$

Proof. We compute using the commutators of Proposition 3.2 that

$$\begin{aligned}[\mathcal{K}, Q_1(A)] &= \sum_{i=1}^n \left([\mathcal{K}, a^*(Ae_i)a(e_i)] + [\mathcal{K}, a(e_i)a^*(Ae_i)]\right)\\ &= \sum_{i=1}^n (a^*(Ae_i)\,[\mathcal{K}, a(e_i)]\\ &\quad + [\mathcal{K}, a^*(Ae_i)]\,a(e_i) + a(e_i)\,[\mathcal{K}, a^*(Ae_i)] + [\mathcal{K}, a(e_i)]\,a^*(Ae_i))\\ &= \sum_{i=1}^n \left(a^*(Ae_i)a^*(Ke_i) + a(KAe_i)a(e_i) + a(e_i)a(KAe_i) + a^*(Ke_i)a^*(Ae_i)\right).\end{aligned} \tag{3.16}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

As the assignments $\varphi, \psi \mapsto a(\varphi)a(\psi), a^*(\varphi)a^*(\psi)$ are bilinear, we can apply Lemma 3.3 to see that

$$\begin{aligned}
[\mathcal{K}, Q_1(A)] &= \sum_{i=1}^{n} \left(a^*\left(AK^*e_i\right)a^*(e_i) + a(e_i)a\left((KA)^*e_i\right) + a(e_i)a(KAe_i) + a^*\left(KA^*e_i\right)a^*(e_i)\right) \\
&= \sum_{i=1}^{n} \left(a^*\left(AKe_i\right)a^*(e_i) + a(e_i)a\left(AKe_i\right) + a(e_i)a(KAe_i) + a^*(KAe_i)a^*(e_i)\right) \\
&= \sum_{i=1}^{n} \left(a^*\left((AK + KA)e_i\right)a^*(e_i) + a(e_i)a\left((AK + KA)e_i\right)\right) = Q_2\left(\{K, A\}\right),
\end{aligned} \tag{3.17}$$

where we also used that A and K are symmetric. The computation of $[\mathcal{K}, Q_2(B)]$ is similar. □

Note the similarity between this result and that of Proposition 3.2. Again we see that that $[\mathcal{K}, \cdot]$ acts by 'swapping the types and applying K to the argument', although now the relevant types are $Q_1(\cdot)$, and $Q_2(\cdot)$ and the application of K is taking the anticommutator.

We can now appeal to the Baker-Campbell-Hausdorff formula again to conclude that

$$\begin{aligned}
&e^{\mathcal{K}}Q_1(A)e^{-\mathcal{K}} = Q_1(A) + \frac{1}{1!}[\mathcal{K}, Q_1(A)] + \frac{1}{2!}[\mathcal{K}, [\mathcal{K}, Q_1(A)]] + \frac{1}{3!}[\mathcal{K}, [\mathcal{K}, [\mathcal{K}, Q_1(A)]]] + \cdots \\
&= Q_1(A) + \frac{1}{1!}Q_2\left(\{K, A\}\right) + \frac{1}{2!}Q_1\left(\{K, \{K, A\}\}\right) + \frac{1}{3!}Q_2\left(\{K, \{K, \{K, A\}\}\}\right) + \cdots \\
&= Q_1\left(A + \frac{1}{2!}\{K, \{K, A\}\} + \cdots\right) + Q_2\left(\frac{1}{1!}\{K, A\} + \frac{1}{3!}\{K, \{K, \{K, A\}\}\} + \cdots\right),
\end{aligned} \tag{3.18}$$

but to succeed, we must identify the sums of these iterated anticommutators. First, we note that we can rephrase this in a manner closer to that of equation (3.10) for $e^{\mathcal{K}}a(\varphi)e^{-\mathcal{K}}$. One may view the anticommutator with K as a linear mapping $A \mapsto \{K, A\}$ on the space of operators on V, $\mathcal{B}(V)$ – denote this mapping by $\mathcal{A}_K : \mathcal{B}(V) \to \mathcal{B}(V)$ (i.e., $\mathcal{A}_K(\cdot) = \{K, \cdot\}$). Then we may phrase the above identity as

$$e^{\mathcal{K}}Q_1(A)e^{-\mathcal{K}} = Q_1\left(\cosh\left(\mathcal{A}_K\right)(A)\right) + Q_2\left(\sinh\left(\mathcal{A}_K\right)(A)\right) \tag{3.19}$$

and likewise

$$e^{\mathcal{K}}Q_2(B)e^{-\mathcal{K}} = Q_2\left(\cosh\left(\mathcal{A}_K\right)(B)\right) + Q_1\left(\sinh\left(\mathcal{A}_K\right)(B)\right) \tag{3.20}$$

so that the arguments again involve hyperbolic functions of linear operators, but now acting on $\mathcal{B}(V)$ rather than V itself. We then note the following 'anticommutator Baker-Campbell-Hausdorff formula':

Proposition 3.5. *Let $(V, \langle\cdot,\cdot\rangle)$ be an n-dimensional Hilbert space, let $K : V \to V$ be a self-adjoint operator and let $\mathcal{A}_K(\cdot) = \{K, \cdot\} : \mathcal{B}(V) \to \mathcal{B}(V)$ denote the anticommutator with K. Then for any linear operator $T : V \to V$,*

$$e^{\mathcal{A}_K}(T) = \sum_{m=0}^{\infty} \frac{1}{m!}\mathcal{A}_K^m(T) = e^K T e^K.$$

Consequently,

$$\begin{aligned}
\cosh\left(\mathcal{A}_K\right)(T) &= \frac{1}{2}(e^K T e^K + e^{-K} T e^{-K}), \\
\sinh(\mathcal{A}_K)(T) &= \frac{1}{2}(e^K T e^K - e^{-K} T e^{-K}).
\end{aligned}$$

Proof. Let $(x_i)_{i=1}^n$ be an eigenbasis for K with associated eigenvalues $(\lambda_i)_{i=1}^n$. Denote $P_{i,j} = |x_j\rangle\langle x_i|$, namely, $P_{i,j}x = \langle x_i, x\rangle x_j$ for all $x \in V$. It is well known that for any orthonormal basis $(x_i)_{i=1}^n$ of V, the collection $\left(P_{i,j}\right)_{i,j=1}^n$ form an orthonormal basis for $(\mathcal{B}(V), \langle\cdot,\cdot\rangle_{\mathrm{HS}})$. Moreover, for any $x \in V$ and $1 \le i, j \le n$, by self-adjointness of K,

$$\begin{aligned}\mathcal{A}_K\left(P_{i,j}\right)x &= \left\{K, P_{i,j}\right\}x = \langle x_i, x\rangle Kx_j + \langle x_i, Kx\rangle x_j = \langle x_i, x\rangle \lambda_j x_j + \langle \lambda_i x_i, x\rangle x_j \\ &= \left(\lambda_i + \lambda_j\right)\langle x_i, x\rangle x_j = \left(\lambda_i + \lambda_j\right)P_{i,j}x.\end{aligned} \tag{3.21}$$

Thus, $\{P_{i,j}\}_{i,j=1}^n$ an eigenbasis for $\mathcal{A}_K$ with associated eigenvalues $\left(\lambda_i + \lambda_j\right)_{i,j=1}^n$.

Hence, it suffices to verify the identity $e^{\mathcal{A}_K}(T) = e^K T e^K$ with the eigenbasis $\left(P_{i,j}\right)_{i,j=1}^n$:

$$\begin{aligned}e^{\mathcal{A}_K}\left(P_{i,j}\right)x &= e^{\lambda_i+\lambda_j}P_{i,j} = e^{\lambda_i+\lambda_j}\langle x_i, x\rangle x_j = \left\langle e^{\lambda_i}x_i, x\right\rangle e^{\lambda_j}x_j = \left\langle e^K x_i, x\right\rangle e^K x_j \\ &= \left\langle x_i, e^K x\right\rangle e^K x_j = e^K P_{i,j} e^K x.\end{aligned} \tag{3.22}$$

The statements regarding $\cosh(\mathcal{A}_K)$ and $\sinh(\mathcal{A}_K)$ follow from the identities

$$\cosh(x) = \frac{1}{2}\left(e^x + e^{-x}\right), \quad \sinh(x) = \frac{1}{2}\left(e^x - e^{-x}\right), \quad \text{and } (-\mathcal{A}_K) = \mathcal{A}_{-K}. \tag{3.23}$$

□

By these formulas, we thus deduce the quadratic operator analogue of equation (3.10):

$$\begin{aligned}e^{\mathcal{K}}Q_1(A)e^{-\mathcal{K}} &= \frac{1}{2}Q_1(e^K A e^K + e^{-K}Ae^{-K}) + \frac{1}{2}Q_2(e^K A e^K - e^{-K}Ae^{-K}) \\ e^{\mathcal{K}}Q_2(B)e^{-\mathcal{K}} &= \frac{1}{2}Q_1(e^K B e^K - e^{-K}Be^{-K}) + \frac{1}{2}Q_2(e^K B e^K + e^{-K}Be^{-K}).\end{aligned} \tag{3.24}$$

Diagonalization condition

We can now finally describe how to diagonalize a quadratic Hamiltonian using a Bogolubov transformation of the form $e^{\mathcal{K}}$. By the transformation identities above, we find that under $e^{\mathcal{K}}$, the quadratic Hamiltonian $H = Q_1(A) + Q_2(B)$ transforms as

$$\begin{aligned}e^{\mathcal{K}}He^{-\mathcal{K}} &= \frac{1}{2}Q_1(e^K A e^K + e^{-K}Ae^{-K}) + \frac{1}{2}Q_2(e^K A e^K - e^{-K}Ae^{-K}) \\ &\quad + \frac{1}{2}Q_1(e^K B e^K - e^{-K}Be^{-K}) + \frac{1}{2}Q_2(e^K B e^K + e^{-K}Be^{-K}) \\ &= \frac{1}{2}Q_1(e^K(A+B)e^K + e^{-K}(A-B)e^{-K}) + \frac{1}{2}Q_2(e^K(A+B)e^K - e^{-K}(A-B)e^{-K}).\end{aligned} \tag{3.25}$$

Therefore, the *diagonalization condition* on K is

$$e^K(A+B)e^K = e^{-K}(A-B)e^{-K}. \tag{3.26}$$

If we can find such a K, then

$$e^{\mathcal{K}}He^{-\mathcal{K}} = Q_1(E) = 2\,\mathrm{d}\Gamma(E) + \mathrm{tr}(E), \tag{3.27}$$

where

$$E = e^K(A+B)e^K = e^{-K}(A-B)e^{-K}. \tag{3.28}$$

There remains the question of existence and uniqueness of such a K:

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Conclusion of the proof of Theorem 3.1. Write $A_\pm = A \pm B > 0$ for brevity. Then we may write the diagonalization condition as

$$e^{-2K} A_- e^{-2K} = A_+. \tag{3.29}$$

Multiplying by $A_-^{\frac{1}{2}}$ on both sides yields

$$\left(A_-^{\frac{1}{2}} e^{-2K} A_-^{\frac{1}{2}}\right)^2 = A_-^{\frac{1}{2}} e^{-2K} A_- e^{-2K} A_-^{\frac{1}{2}} = A_-^{\frac{1}{2}} A_+ A_-^{\frac{1}{2}}, \tag{3.30}$$

which is equivalent to

$$A_-^{\frac{1}{2}} e^{-2K} A_-^{\frac{1}{2}} = \left(A_-^{\frac{1}{2}} A_+ A_-^{\frac{1}{2}}\right)^{\frac{1}{2}}, \text{ namely } e^{-2K} = A_-^{-\frac{1}{2}} \left(A_-^{\frac{1}{2}} A_+ A_-^{\frac{1}{2}}\right)^{\frac{1}{2}} A_-^{-\frac{1}{2}}. \tag{3.31}$$

This implies the existence and uniqueness of the diagonalizing K as the operator exponential is a bijection between the real, symmetric operators and the real, symmetric, positive-definite operators. □

4. The quasi-bosonic quadratic Hamiltonian

Now we turn to the quasi-bosonic setting. We start by casting the bosonizable terms $H'_{\text{kin}} + \sum_{k \in S_C} H^k_{\text{int}}$, which we encountered in Section 2.1, into a form which closely mirrors the form of the bosonic quadratic Hamiltonians that we considered in the preceding section.

4.1. Quadratic Hamiltonian

Let us define the pair excitation operators

$$b_{k,p} = c^*_{p-k} c_p, \quad b^*_{k,p} = c^*_p c_{p-k}, \quad k \in \mathbb{Z}^3_*, \quad p \in L_k. \tag{4.1}$$

We remark that in contrast to the bosonic case, the fermionic creation and annihilation operators are bounded (in fact, $\|c_{p,\sigma}\|_{\text{Op}} = \|c^*_{p,\sigma}\|_{\text{Op}} = 1$), and therefore so are the operators $b^*_{k,p}$, $b_{k,p}$.

Then H^k_{int} in (1.34) is exactly given by

$$\begin{aligned} H^k_{\text{int}} = &\sum_{p,q \in L_k} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{k,p} b_{k,q} + b_{k,q} b^*_{k,p}\right) + \sum_{p,q \in L_{-k}} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{-k,p} b_{-k,q} + b_{-k,q} b^*_{-k,p}\right) \\ &+ \sum_{p \in L_k} \sum_{q \in L_{-k}} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{k,p} b^*_{-k,q} + b_{-k,q} b_{k,p}\right) \\ &+ \sum_{p \in L_{-k}} \sum_{q \in L_k} \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \left(b^*_{-k,p} b^*_{k,q} + b_{k,q} b_{-k,p}\right). \end{aligned} \tag{4.2}$$

Thus, the natural one-body Hilbert space associated to H^k_{int} is $\ell^2(L_k \cup L_{-k})$. To free us from having to explicitly write sums over L_k and L_{-k} separately, we introduce some more notation. First, we will denote this union of lunes by

$$L^\pm_k = L_k \cup L_{-k}, \quad \ell^2\left(L^\pm_k\right) = \ell^2\left(L_k \cup L_{-k}\right) = \ell^2(L_k) \oplus \ell^2(L_{-k}), \quad k \in \mathbb{Z}^3_+. \tag{4.3}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Here, we used the fact that $L_k \cap L_{-k} = \emptyset$ for any $k \in \mathbb{Z}^3_+$, since if $p \in L_k \cap L_{-k}$, then

$$2|p|^2 \geq |p-k|^2 + |p+k|^2 = 2|p|^2 + 2|k|^2 > 2|p|^2,$$

which is a contradiction. It is also convenient to introduce the 'bar-notation'

$$\overline{k,p} = \begin{cases} k,p & p \in L_k \\ -k,p & p \in L_{-k} \end{cases}, \quad \overline{p-k} = \begin{cases} p-k & p \in L_k \\ p+k & p \in L_{-k} \end{cases} \tag{4.4}$$

to automatically encode the appropriate sign of k depending on $p \in L_k^{\pm} = L_k \cup L_{-k}$ (this will allow us to avoid expanding all our terms on a case-by-case basis when this is irrelevant).

In analogy with the definitions (3.3) and (3.4) we now define, for any $k \in \mathbb{Z}^3_+$ and symmetric operators $A, B : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$, the quadratic operators $Q_1^k(A), Q_2^k(B) : \mathcal{H}_N \to \mathcal{H}_N$ by

$$\begin{aligned} Q_1^k(A) &= \sum_{p,q \in L_k^{\pm}} \langle e_p, A e_q \rangle \left(b^*_{\overline{k,p}} b_{\overline{k,q}} + b_{\overline{k,q}} b^*_{\overline{k,p}} \right), \\ Q_2^k(B) &= \sum_{p,q \in L_k^{\pm}} \langle e_p, B e_q \rangle \left(b^*_{\overline{k,p}} b^*_{\overline{k,q}} + b_{\overline{k,q}} b_{\overline{k,p}} \right). \end{aligned} \tag{4.5}$$

In order to cast H^k_{int} as given by equation (4.2) into this form, we must identify the relevant operators A and B. Define the (un-normalized) rank-one projection $P_{v_k} : \ell^2(L_k) \to \ell^2(L_k)$ by

$$P_{v_k} = |v_k\rangle\langle v_k|, \quad v_k = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}} \sum_{p \in L_k} e_p \in \ell^2(L_k), \tag{4.6}$$

where $(e_p)_{p \in L_k}$ denotes the standard orthonormal basis of $\ell^2(L_k)$. Put differently, the matrix elements of P_{v_k} are $\langle e_p, P_{v_k} e_q \rangle = \frac{1}{2(2\pi)^3} \hat{V}_k k_F^{-1}$ for all $p, q \in L_k$. Next, we define the operators

$$A_k^{\oplus}, B_k^{\oplus} : \ell^2\left(L_k^{\pm}\right) \to \ell^2\left(L_k^{\pm}\right), \quad A_k^{\oplus} = \begin{pmatrix} P_{v_k} & 0 \\ 0 & P_{v_k} \end{pmatrix}, \quad B_k^{\oplus} = \begin{pmatrix} 0 & P_{v_k} \\ P_{v_k} & 0 \end{pmatrix} \tag{4.7}$$

with respect to the decomposition $\ell^2(L_k^{\pm}) = \ell^2(L_k) \oplus \ell^2(L_{-k})$ and the identification $\ell^2(L_k) \cong \ell^2(L_{-k})$ (under $e_p \mapsto e_{-p}$).

Thus, the operator H^k_{int} is concisely expressed as

$$H^k_{\text{int}} = Q_k^1\left(A_k^{\oplus}\right) + Q_k^2\left(B_k^{\oplus}\right). \tag{4.8}$$

It remains to consider the kinetic operator. The equality (1.74) bids us to think of H'_{kin} as it were

$$\begin{aligned} H'_{\text{kin}} &\sim \sum_{k \in \mathbb{Z}^3_*} \sum_{p \in L_k} \left(|p|^2 - |p-k|^2\right) b^*_{k,p} b_{k,p} \\ &= \sum_{k \in \mathbb{Z}^3_+} \left(\sum_{p \in L_k} \left(|p|^2 - |p-k|^2\right) b^*_{k,p} b_{k,p} + \sum_{p \in L_{-k}} \left(|p|^2 - |p+k|^2\right) b^*_{-k,p} b_{-k,p} \right) \end{aligned} \tag{4.9}$$

in an appropriate sense. To put this in the same framework as H^k_{int}, let us introduce (for every $k \in \mathbb{Z}^3_+$) the operator $h_k : \ell^2(L_k) \to \ell^2(L_k)$ by

$$h_k e_p = \lambda_{k,p} e_p, \quad \lambda_{k,p} = \frac{1}{2}(|p|^2 - |p-k|^2). \tag{4.10}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Using again the identification $\ell^2(L_k) \cong \ell^2(L_{-k})$ (under $e_p \mapsto e_{-p}$), we define the operators $h_k^\oplus : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$ by

$$h_k^\oplus = \begin{pmatrix} h_k & 0 \\ 0 & h_k \end{pmatrix}. \tag{4.11}$$

Then we can rewrite (4.9) as

$$H'_{\mathrm{kin}} \sim \sum_{k\in\mathbb{Z}^3_+} \Big(Q_1^k(h_k^\oplus) - 2\,\mathrm{tr}(h_k) \Big). \tag{4.12}$$

Recall that $S_C = \overline{B}\left(0, k_F^\gamma\right) \cap \mathbb{Z}^3_+$ for an exponent $1 \geq \gamma > 0$ which is to be optimized over at the end. As far as the lower bound is concerned, we may replace $\sum_{k\in\mathbb{Z}_+}$ by $\sum_{k\in S_C}$ (the upper bound is easier and will be explained separately). In summary, we arrive at the following quasi-bosonic expression for the bosonizable terms:

$$H'_{\mathrm{kin}} + \sum_{k\in S_C} H^k_{\mathrm{int}} \sim \sum_{k\in S_C} \Big(Q_1^k(h_k^\oplus + A_k^\oplus) + Q_2^k(B_k^\oplus) - 2\,\mathrm{tr}(h_k) \Big). \tag{4.13}$$

Note that unlike the bosonic case, the operators on the right side of (4.13) are bounded.

4.2. Generalized pair operators

For every $k \in \mathbb{Z}^3_+$ and $\varphi \in \ell^2(L_k^\pm)$, we define the operators

$$b_k(\varphi) = \sum_{p\in L_k^\pm} \langle \varphi, e_p \rangle\, b_{\overline{k,p}}, \quad b_k^*(\varphi) = \sum_{p\in L_k^\pm} \langle e_p, \varphi \rangle\, b^*_{\overline{k,p}}. \tag{4.14}$$

They obey the quasi-bosonic commutation relations (for $k, l \in \mathbb{Z}^3_+$ and $\varphi \in \ell^2(L_k^\pm)$, $\psi \in \ell^2\left(L_l^\pm\right)$)

$$\begin{aligned} [b_k(\varphi), b_l(\psi)] &= \left[b_k^*(\varphi), b_l^*(\psi)\right] = 0, \\ \left[b_k(\varphi), b_l^*(\psi)\right] &= \delta_{k,l}\,\langle \varphi, \psi\rangle + \varepsilon_{k,l}(\varphi;\psi), \end{aligned} \tag{4.15}$$

where the correction term is

$$\begin{aligned} \varepsilon_{k,l}(\varphi;\psi) &= \sum_{p\in L_k^\pm} \sum_{q\in L_l^\pm} \langle \varphi, e_p\rangle \langle e_q, \psi\rangle\, \varepsilon\left(\overline{k,p}; \overline{l,q}\right), \\ \varepsilon\left(\overline{k,p}; \overline{l,q}\right) &= -\left(\delta_{p,q}\, c_{\overline{q-l}} c^*_{\overline{p-k}} + \delta_{\overline{p-k},\overline{q-l}}\, c_q^* c_p \right). \end{aligned} \tag{4.16}$$

We simply have $b_k(e_p) = b_{\overline{k,p}}$ and the quadratic operators in (4.5) can be expressed as

$$\begin{aligned} Q_1^k(A) &= \sum_{p\in L_k^\pm} \left(b_k^*(Ae_p) b_k(e_p) + b_k(e_p) b_k^*(Ae_p) \right) \\ Q_2^k(B) &= \sum_{p\in L_k^\pm} \left(b_k^*\left(Be_p\right) b_k^*(e_p) + b_k(e_p) b_k\left(Be_p\right) \right) \end{aligned} \tag{4.17}$$

in analogy with equation (3.5). In order to justify the quasi-bosonic interpretation, we need rigorous estimates for the correction term in (4.16). Let us start with the following:

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proposition 4.1. *For all $k \in \mathbb{Z}^3_+$ and $\varphi \in \ell^2(L^\pm_k)$, it holds that $\varepsilon_{k,k}(\varphi,\varphi) \le 0$, namely,*

$$b_k(\varphi) b^*_k(\varphi) \le b^*_k(\varphi) b_k(\varphi) + \|\varphi\|^2 .$$

Note that the observation of the error term $\varepsilon_{k,k}$ being non-positive also appeared in [5, Proof of Lemma 4.2] in the context of different bosonic operators.

Proof. We expand the term

$$\begin{aligned}
\varepsilon_{k,k}(\varphi;\varphi) &= \sum_{p,q\in L^\pm_k} \langle \varphi, e_p\rangle \langle e_q,\varphi\rangle \varepsilon\left(\overline{k,p};\overline{k,q}\right) \\
&= -\sum_{p,q\in L^\pm_k} \langle \varphi, e_p\rangle \langle e_q,\varphi\rangle \left(\delta_{p,q} c_{\overline{q-k}} c^*_{\overline{p-k}} + \delta_{\overline{p-k},\overline{q-k}} c^*_q c_p\right) \\
&= -\sum_{p\in L^\pm_k} \left|\langle e_p,\varphi\rangle\right|^2 c_{\overline{p-k}} c^*_{\overline{p-k}} - \sum_{p,q\in L^\pm_k} \delta_{\overline{p-k},\overline{q-k}} \langle \varphi, e_p\rangle \langle e_q,\varphi\rangle c^*_q c_p \\
&\le -\sum_{p,q\in L^\pm_k} \delta_{\overline{p-k},\overline{q-k}} \langle \varphi, e_p\rangle \langle e_q,\varphi\rangle c^*_q c_p .
\end{aligned} \tag{4.18}$$

We treat the terms of the last sum on a case-by-case basis according to which of L_k and L_{-k}, p and q lie in: if p and q lie in the same lune, then $\delta_{\overline{p-k},\overline{q-k}} = \delta_{p\mp k,q\mp k} = \delta_{p,q}$ and so

$$A = \left(\sum_{p,q\in L_k} + \sum_{p,q\in L_{-k}}\right) \delta_{\overline{p-k},\overline{q-k}} \langle \varphi, e_p\rangle \langle e_q,\varphi\rangle c^*_q c_p = \sum_{p\in L^\pm_k} \left|\langle e_p,\varphi\rangle\right|^2 c^*_p c_p \ge 0. \tag{4.19}$$

However, by the Cauchy–Schwarz inequality,

$$\begin{aligned}
&\pm\left(\sum_{p\in L_k}\sum_{q\in L_{-k}} + \sum_{p\in L_{-k}}\sum_{q\in L_k}\right) \delta_{\overline{p-k},\overline{q-k}} \langle \varphi, e_p\rangle \langle e_q,\varphi\rangle c^*_q c_p \\
&\le \left(\sum_{p\in L_k}\sum_{q\in L_{-k}} + \sum_{p\in L_{-k}}\sum_{q\in L_k}\right) \frac{1}{2}\left(\delta_{\overline{p-k},\overline{q-k}} |\langle \varphi, e_p\rangle|^2 c^*_p c_p + |\langle e_q,\varphi\rangle|^2 c^*_q c_q\right) \\
&\le \sum_{p\in L^\pm_k} \left|\langle e_p,\varphi\rangle\right|^2 c^*_p c_p = A.
\end{aligned} \tag{4.20}$$

We thus conclude that $\varepsilon_{k,k}(\varphi,\varphi) \le 0$ as claimed. □

Next, we have the following:

Proposition 4.2. *For all $k \in \mathbb{Z}^3_+$, $\varphi \in \ell^2(L^\pm_k)$ and $\Psi \in \mathcal{H}_N$, it holds that*

$$\|b_k(\varphi)\Psi\| \le \|\varphi\| \sqrt{\langle \Psi, \mathcal{N}_E \Psi\rangle}, \quad \left\|b^*_k(\varphi)\Psi\right\| \le \|\varphi\| \sqrt{\langle \Psi, (1+\mathcal{N}_E)\Psi\rangle}.$$

The bounds here are similar to [5, Lemma 4.2]. Recall that in our quasi-bosonic setting the excitation number operator

$$\mathcal{N}_E = \sum_{p\in B^c_F} c^*_p c_p = \sum_{p\in B_F} c_p c^*_p \tag{4.21}$$

plays the role that the usual number operator $\mathcal{N}$ does in the exact bosonic case. Thus, Proposition 4.2 is the analogue of the well-known bosonic estimate

$$\|a(\varphi)\Psi\| \leq \|\varphi\| \sqrt{\langle \Psi, \mathcal{N}\Psi\rangle}, \quad \|a^*(\varphi)\Psi\| \leq \|\varphi\| \sqrt{\langle \Psi, (1+\mathcal{N})\,\Psi\rangle}. \tag{4.22}$$

Proof. By the Cauchy-Schwarz inequality,

$$\begin{aligned}\|b_k(\varphi)\Psi\| &= \left\| \sum_{p\in L_k^\pm} \langle \varphi, e_p\rangle\, b_{\overline{k,p}}\Psi \right\| \leq \sqrt{\sum_{p\in L_k^\pm} |\langle \varphi, e_p\rangle|^2} \sqrt{\sum_{p\in L_k^\pm} \left\| b_{\overline{k,p}}\Psi\right\|^2} \\ &\leq \|\varphi\| \sqrt{\sum_{p\in L_k^\pm} \left\| c_p\Psi\right\|^2} \leq \|\varphi\| \sqrt{\langle \Psi, \mathcal{N}_E\Psi\rangle}. \end{aligned} \tag{4.23}$$

The second bound follows from the first and Proposition 4.1. □

We remark that the above estimate is also valid for $\Psi \in \mathcal{H}_M$ when $M \neq N$, provided $\mathcal{N}_E$ is understood as $\sum_{p\in B_F^c} c_p^* c_p$ acting on $\mathcal{H}_M$ (in (4.23) we used $L_k^\pm \subset B_F^c$). One must be precise here as the identity $\mathcal{N}_E = \sum_{p\in B_F} c_p c_p^*$ does not hold on $\mathcal{H}_M$. In fact, the estimate also holds if $\mathcal{N}_E$ is understood as $\sum_{p\in B_F} c_p c_p^*$, up to an additional factor of $\sqrt{2}$ due to the necessary overcounting of the holes,[7] namely, from $\left\| b_{\overline{k,p}}\Psi\right\| = \left\| c^*_{\overline{p-k}} c_p \Psi\right\| \leq \left\| c^*_{\overline{p-k}}\Psi\right\|$ with $\overline{p-k} \in B_F$, we get

$$\|b_k(\varphi)\Psi\| \leq \|\varphi\| \sqrt{\sum_{p\in L_k^\pm} \left\| c^*_{\overline{p-k}}\Psi\right\|^2} \leq \sqrt{2}\, \|\varphi\| \sqrt{\left\langle \Psi, \left(\sum_{p\in B_F} c_p c_p^*\right) \Psi \right\rangle}. \tag{4.24}$$

This is a point that we must consider, since below we will also encounter expressions such as $\left\| b_k(\varphi) c_p \Psi\right\|$ for $\Psi \in \mathcal{H}_N$ (so that $c_p\Psi \in \mathcal{H}_{N-1}$). For this, we denote by $\mathcal{N}_E^{(-1)} : \mathcal{H}_{N-1} \to \mathcal{H}_{N-1}$ and $\mathcal{N}_E^{(+1)} : \mathcal{H}_{N+1} \to \mathcal{H}_{N+1}$ the operators

$$\mathcal{N}_E^{(-1)} = \sum_{p\in B_F^c} c_p^* c_p, \quad \mathcal{N}_E^{(+1)} = \sum_{p\in B_F} c_p c_p^*. \tag{4.25}$$

This choice is motivated by the following identities:

Lemma 4.3. *For all $p \in B_F^c$ and $q \in B_F$, it holds that*

$$\begin{aligned} \mathcal{N}_E c_p^* &= c_p^* \mathcal{N}_E^{(-1)} + c_p^*, \quad & c_p \mathcal{N}_E^{(-1)} c_p^* &\leq \mathcal{N}_E, \\ \mathcal{N}_E c_q &= c_q \mathcal{N}_E^{(+1)} + c_q, \quad & c_q^* \mathcal{N}_E^{(+1)} c_q &\leq \mathcal{N}_E. \end{aligned}$$

Consequently,

$$\sum_{p\in B_F^c} c_p^* \mathcal{N}_E^{(-1)} c_p = \mathcal{N}_E^2 - \mathcal{N}_E = \sum_{p\in B_F} c_p \mathcal{N}_E^{(+1)} c_p^*.$$

Proof. This follows directly by the CAR, as for all $p \in B_F^c$,

$$\mathcal{N}_E c_p^* = \sum_{q\in B_F^c} c_q^* c_q c_p^* = \sum_{q\in B_F^c} c_p^* c_q^* c_q + \sum_{q\in B_F^c} \left(c_q^* \left\{ c_q, c_p^* \right\} - \left\{ c_q^*, c_p^* \right\} c_q \right) = c_p^* \mathcal{N}_E^{(-1)} + c_p^*. \tag{4.26}$$

[7] While $L_k \cap L_{-k} = \emptyset$, it is generally the case that $(L_k - k) \cap (L_{-k} + k) \neq \emptyset$ (a single hole state may be ‘shared’ by both lunes), so when estimating in terms of a single sum over $p \in B_F$, a factor of 2 is often necessary.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Consequently, using $\|c_p\|_{\mathrm{Op}} = 1$ and $[\mathcal{N}_E, c_p^* c_p] = 0$, we have

$$\mathcal{N}_E \geq \mathcal{N}_E c_p^* c_p = c_p^* \mathcal{N}_E^{(-1)} c_p + c_p^* c_p \geq c_p^* \mathcal{N}_E^{(-1)} c_p. \tag{4.27}$$

Likewise, for all $q \in B_F$,

$$\mathcal{N}_E c_q = \sum_{p \in B_F} c_p c_p^* c_q = \sum_{p \in B_F} c_q c_p c_p^* + \sum_{p \in B_F} \left(c_p \left\{ c_p^*, c_q \right\} - \left\{ c_p, c_q \right\} c_p^* \right) = c_q \mathcal{N}_E^{(+1)} + c_q, \tag{4.28}$$

and hence, $\mathcal{N}_E \geq c_q^* \mathcal{N}_E^{(+1)} c_q$. Moreover,

$$\begin{aligned} \sum_{p \in B_F^c} c_p^* \mathcal{N}_E^{(-1)} c_p &= \sum_{p \in B_F^c} \left(\mathcal{N}_E c_p^* - c_p^* \right) c_p = \mathcal{N}_E^2 - \mathcal{N}_E = \sum_{p \in B_F} \left(\mathcal{N}_E c_p - c_p \right) c_p^* \\ &= \sum_{p \in B_F} c_p \mathcal{N}_E^{(+1)} c_p^*. \end{aligned} \tag{4.29}$$

□

In some cases, it is important to refine error estimates by using the kinetic operator H'_{kin} rather than $\mathcal{N}_E$. We can implement the kinetic estimate of Proposition 2.2 in the generalized setting:

Proposition 4.4. *For all $k \in \mathbb{Z}^3_+$, $\varphi \in \ell^2(L_k^\pm)$ and $\Psi \in D\left(H'_{\mathrm{kin}}\right)$, it holds that*

$$\|b_k(\varphi)\Psi\| \leq \left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} \varphi \right\| \sqrt{\left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle}, \quad \|b_k^*(\varphi)\Psi\| \leq \left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} \varphi \right\| \sqrt{\left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle} + \|\varphi\| \, \|\Psi\| \, .$$

Proof. We start by applying the Cauchy-Schwarz inequality

$$\|b_k(\varphi)\Psi\| = \left\| \sum_{p \in L_k^\pm} \left\langle \varphi, e_p \right\rangle b_{\overline{k,p}} \Psi \right\| \leq \sqrt{\sum_{p \in L_k^\pm} \lambda_{\overline{k,p}}^{-1} \left| \left\langle \varphi, e_p \right\rangle \right|^2} \sqrt{\sum_{p \in L_k^\pm} \lambda_{\overline{k,p}} \left\| b_{\overline{k,p}} \Psi \right\|^2}. \tag{4.30}$$

As the vectors $(e_p)_{p \in L_k^\pm}$ obey $h_k^\oplus e_p = \lambda_{\overline{k,p}} e_p$, we recognize the first sum on the right-hand side as

$$\sum_{p \in L_k^\pm} \lambda_{\overline{k,p}}^{-1} \left| \left\langle \varphi, e_p \right\rangle \right|^2 = \left\langle \varphi, \left(h_k^\oplus\right)^{-1} \varphi \right\rangle = \left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} \varphi \right\|^2. \tag{4.31}$$

For the second sum, we have by equation (2.4) that

$$\sum_{p \in L_k^\pm} \lambda_{\overline{k,p}} \left\| b_{\overline{k,p}} \Psi \right\|^2 = \sum_{p \in L_k} \lambda_{k,p} \left\| c_{p-k}^* c_p \Psi \right\|^2 + \sum_{p \in L_{-k}} \lambda_{-k,p} \left\| c_{p+k}^* c_p \Psi \right\|^2 \leq \left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle, \tag{4.32}$$

which implies the first claim. The second bound follows from the first and Proposition 4.1:

$$\|b_k^*(\varphi)\Psi\| \leq \sqrt{\left\langle \Psi, \left(\left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} \varphi \right\|^2 H'_{\mathrm{kin}} + \|\varphi\|^2 \right) \Psi \right\rangle} \leq \left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} \varphi \right\| \sqrt{\left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle} + \|\varphi\| \, \|\Psi\| \, . \tag{4.33}$$

□

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

4.3. Preliminary estimates for quadratic operators

In this subsection, we provide some basic bounds on the quadratic operators $Q_1^k(A)$ and $Q_1^k(B)$ defined in (4.5) for any $k \in \mathbb{Z}_+^3$. First, for $Q_1^k(A)$, we can normal order as follows:

$$\begin{aligned} Q_1^k(A) &= \sum_{p\in L_k^\pm} \left(2\, b_k^*(Ae_p)b_k(e_p) + \left[b_k(e_p), b_k^*(Ae_p)\right]\right) \\ &= 2\sum_{p\in L_k^\pm} b_k^*(Ae_p)b_k(e_p) + \sum_{p\in L_k^\pm}\left\langle e_p, Ae_p\right\rangle + \sum_{p\in L_k^\pm}\varepsilon_{k,k}\left(e_p; Ae_p\right) \\ &= 2\,\tilde{Q}_1^k(A) + \mathrm{tr}(A) + \varepsilon_k(A), \end{aligned} \tag{4.34}$$

where for brevity, we have defined the notation

$$\tilde{Q}_1^k(A) = \sum_{p\in L_k^\pm} b_k^*(Ae_p)b_k(e_p), \quad \varepsilon_k(A) = \sum_{p\in L_k^\pm}\varepsilon_{k,k}\left(e_p; Ae_p\right). \tag{4.35}$$

The term $\tilde{Q}_1^k(A)$ plays the same role of $\mathrm{d}\Gamma(A)$ in the exact bosonic case, whereas $\varepsilon_k(A)$ is a correction term in the quasi-bosonic case.

Proposition 4.5. *For all $k \in \mathbb{Z}_+^3$, symmetric $A : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$ and $\Psi \in \mathcal{H}_N$, it holds that*

$$\begin{aligned} \left|\left\langle \Psi, \tilde{Q}_1^k(A)\Psi\right\rangle\right| &\le \|A\|_{\mathrm{Op}}\left\langle \Psi, \mathcal{N}_E\Psi\right\rangle, \\ \left|\left\langle \Psi, \varepsilon_k(A)\Psi\right\rangle\right| &\le 3\,\|A\|_{\mathrm{Op}}\left\langle \Psi, \mathcal{N}_E\Psi\right\rangle. \end{aligned}$$

If furthermore, $A \ge 0$, then also $\tilde{Q}_1^k(A) \ge 0$.

Proof. Let $(x_i)_i$ be an eigenbasis for A with eigenvalues $(\lambda_i)_i$. Noting that the mapping $x, y \mapsto b_k^*(Ax)\, b_k(y)$ is bilinear, we may invoke Lemma 3.3 (the part of basis independence) to write

$$\tilde{Q}_1^k(A) = \sum_i b_k^*(Ax_i)\, b_k(x_i) = \sum_i \lambda_i b_k^*(x_i) b_k(x_i). \tag{4.36}$$

Clearly, if $A \ge 0$, then all $\lambda_i \ge 0$, and hence, $\tilde{Q}_1^k(A) \ge 0$. In general, we always have $|\lambda_i| \le \|A\|_{\mathrm{Op}}$ for all i. Hence, using Lemma 3.3 again and $b^*_{\overline{k,p}} b_{\overline{k,p}} \le c_p^* c_p$, we have

$$\pm\tilde{Q}_1^k(A) \le \|A\|_{\mathrm{Op}}\sum_i b_k^*(x_i)b_k(x_i) = \|A\|_{\mathrm{Op}}\sum_{p\in L_k^\pm} b^*_{\overline{k,p}} b_{\overline{k,p}} \le \|A\|_{\mathrm{Op}}\sum_{p\in L_k^\pm} c_p^* c_p \le \|A\|_{\mathrm{Op}}\mathcal{N}_E. \tag{4.37}$$

Similarly,

$$\begin{aligned} \pm\varepsilon_k(A) &= \pm\sum_i \varepsilon_{k,k}\left(x_i; Ax_i\right) = \pm\sum_i \lambda_i\varepsilon_{k,k}\left(x_i; x_i\right) \\ &\le -\|A\|_{\mathrm{Op}}\sum_i \varepsilon_{k,k}\left(x_i; x_i\right) = -\|A\|_{\mathrm{Op}}\sum_{p\in L_k^\pm}\varepsilon_{k,k}\left(e_p; e_p\right), \end{aligned} \tag{4.38}$$

where in the first inequality we used the fact that $\varepsilon_{k,k}(x_i; x_i) \leq 0$ as shown in the proof of Proposition 4.1. Using $\varepsilon_{k,k}(e_p; e_p) = \varepsilon(\overline{k,p}; \overline{l,q})$ and the definition (4.16), we get

$$-\sum_{p \in L_k^{\pm}} \varepsilon\left(\overline{k,p}; \overline{k,p}\right) = \sum_{p \in L_k^{\pm}} \left(c_{\overline{p-k}} c_{\overline{p-k}}^* + c_p^* c_p\right) \leq 2 \sum_{p \in B_F} c_p c_p^* + \sum_{p \in B_F^c} c_p^* c_p = 3\mathcal{N}_E, \tag{4.39}$$

which implies the desired claim. □

From these results and equation (4.34), we immediately obtain the following:

Proposition 4.6. *For all $k \in \mathbb{Z}_+^3$, symmetric $A : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$ and $\Psi \in \mathcal{H}_N$, it holds that*

$$\left|\left\langle \Psi, \left(Q_1^k(A) - \operatorname{tr}(A)\right) \Psi \right\rangle\right| \leq 5 \, \|A\|_{\mathrm{Op}} \langle \Psi, \mathcal{N}_E \Psi \rangle .$$

Next, we turn to $Q_2^k(B)$.

Proposition 4.7. *For all $k \in \mathbb{Z}_+^3$, symmetric $B : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$ and $\Psi \in \mathcal{H}_N$, it holds that*

$$\left|\left\langle \Psi, Q_2^k(B) \Psi \right\rangle\right| \leq 2 \, \|B\|_{\mathrm{HS}} \sqrt{\langle \Psi, (1 + \mathcal{N}_E) \Psi \rangle \langle \Psi, \mathcal{N}_E \Psi \rangle} \leq 2 \, \|B\|_{\mathrm{HS}} \langle \Psi, (1 + \mathcal{N}_E) \Psi \rangle .$$

Proof. We have (using that the b_k operators commute)

$$\begin{aligned}
\left\langle \Psi, Q_2^k(B) \Psi \right\rangle &= \sum_{p \in L_k^{\pm}} \left\langle \Psi, \left(b_k^*\left(B e_p\right) b_k^*(e_p) + b_k(e_p) b_k\left(B e_p\right)\right) \Psi \right\rangle \\
&= 2 \sum_{p \in L_k^{\pm}} \operatorname{Re} \left\langle b_k^*\left(B e_p\right) \Psi, b_k(e_p) \Psi \right\rangle ,
\end{aligned} \tag{4.40}$$

so using the estimates of Proposition 4.2 and the Cauchy-Schwarz inequality, we conclude that

$$\begin{aligned}
\left|\left\langle \Psi, Q_2^k(B) \Psi \right\rangle\right| &\leq 2 \sum_{p \in L_k^{\pm}} \left\|b_k^*\left(B e_p\right) \Psi\right\| \left\|b_k(e_p) \Psi\right\| \leq 2\sqrt{\langle \Psi, (1 + \mathcal{N}_E) \Psi \rangle} \sum_{p \in L_k^{\pm}} \left\|B e_p\right\| \left\|b_{\overline{k,p}} \Psi\right\| \\
&\leq 2\sqrt{\langle \Psi, (1 + \mathcal{N}_E) \Psi \rangle} \sqrt{\sum_{p \in L_k^{\pm}} \left\|B e_p\right\|^2} \sqrt{\sum_{p \in L_k^{\pm}} \left\|b_{\overline{k,p}} \Psi\right\|^2} \\
&\leq 2 \, \|B\|_{\mathrm{HS}} \langle \Psi, (1 + \mathcal{N}_E) \Psi \rangle ,
\end{aligned} \tag{4.41}$$

where we again used that $\left\|b_{\overline{k,p}} \Psi\right\| \leq \left\|c_p \Psi\right\|$. □

Kinetic estimates for quadratic operators

Finally, let us improve the estimates in this subsection by using the kinetic operator H_{kin}' instead of the number operator $\mathcal{N}_E$.

Proposition 4.8. *For all $k \in \mathbb{Z}_+^3$, symmetric $A : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$ and $\Psi \in D\left(H_{\mathrm{kin}}'\right)$, it holds that*

$$\left|\left\langle \Psi, \tilde{Q}_1^k(A) \Psi \right\rangle\right| \leq \left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} A \left(h_k^{\oplus}\right)^{-\frac{1}{2}}\right\|_{\mathrm{Op}} \left\langle \Psi, H_{\mathrm{kin}}' \Psi \right\rangle .$$

Proof. Let $(x_i)_i$ be an eigenbasis for $(h_k^{\oplus})^{-\frac{1}{2}} A (h_k^{\oplus})^{-\frac{1}{2}}$ with eigenvalues $(\mu_i)_i$. By Lemma 3.3, we then see that we may write $\tilde{Q}_1^k(A)$ as

$$\begin{aligned}\tilde{Q}_1^k(A) &= \sum_i b_k^*(Ax_i)\,b_k(x_i) = \sum_i b_k^*\left(\left(h_k^\oplus\right)^{\frac12}\left(h_k^\oplus\right)^{-\frac12} A\left(h_k^\oplus\right)^{-\frac12}\left(h_k^\oplus\right)^{\frac12} x_i\right) b_k(x_i)\\ &= \sum_i b_k^*\left(\left(h_k^\oplus\right)^{\frac12}\left(h_k^\oplus\right)^{-\frac12} A\left(h_k^\oplus\right)^{-\frac12} x_i\right) b_k\left(\left(h_k^\oplus\right)^{\frac12} x_i\right)\\ &= \sum_i \mu_i b_k^*\left(\left(h_k^\oplus\right)^{\frac12} x_i\right) b_k\left(\left(h_k^\oplus\right)^{\frac12} x_i\right),\end{aligned}\tag{4.42}$$

and so we can estimate

$$\begin{aligned}\left|\left\langle\Psi,\tilde{Q}_1^k(A)\Psi\right\rangle\right| &\le \left(\max_{1\le i\le\left|L_k^\pm\right|}|\mu_i|\right)\sum_i\left\langle\Psi, b_k^*\left(\left(h_k^\oplus\right)^{\frac12} x_i\right) b_k\left(\left(h_k^\oplus\right)^{\frac12} x_i\right)\Psi\right\rangle\\ &= \left\|\left(h_k^\oplus\right)^{-\frac12} A\left(h_k^\oplus\right)^{-\frac12}\right\|_{\mathrm{Op}}\left\langle\Psi,\sum_i b_k^*\left(\left(h_k^\oplus\right)^{\frac12} x_i\right) b_k\left(\left(h_k^\oplus\right)^{\frac12} x_i\right)\Psi\right\rangle.\end{aligned}\tag{4.43}$$

Applying Lemma 3.3 again, we also see that

$$\begin{aligned}\sum_i b_k^*\left(\left(h_k^\oplus\right)^{\frac12} x_i\right) b_k\left(\left(h_k^\oplus\right)^{\frac12} x_i\right) &= \sum_{p\in L_k^\pm} b_k^*\left(\left(h_k^\oplus\right)^{\frac12} e_p\right) b_k\left(\left(h_k^\oplus\right)^{\frac12} e_p\right)\\ &= \sum_{p\in L_k^\pm}\lambda_{\overline{k,p}}\, b_{\overline{k,p}}^*\, b_{\overline{k,p}},\end{aligned}\tag{4.44}$$

so by equation (4.32), we obtain the desired bound of

$$\left|\left\langle\Psi,\tilde{Q}_1^k(A)\Psi\right\rangle\right| \le \left\|\left(h_k^\oplus\right)^{-\frac12} A\left(h_k^\oplus\right)^{-\frac12}\right\|_{\mathrm{Op}}\left\langle\Psi, H_{\mathrm{kin}}'\Psi\right\rangle.\tag{4.45}$$

□

Next are the $\varepsilon_k(A)$ terms. These we cannot estimate in terms of H_{kin}', but for A of diagonal form, we can still control them strongly:

Proposition 4.9. *For all $k\in\mathbb{Z}_+^3$, symmetric $A^\oplus = \begin{pmatrix} A & 0\\ 0 & A\end{pmatrix} : \ell^2(L_k^\pm)\to\ell^2(L_k^\pm)$ and $\Psi\in\mathcal{H}_N$, it holds that*

$$\left|\left\langle\Psi,\varepsilon_k\left(A^\oplus\right)\Psi\right\rangle\right| \le 3\left(\max_{p\in L_k}\left|\left\langle e_p, Ae_p\right\rangle\right|\right)\left\langle\Psi,\mathcal{N}_E\Psi\right\rangle.$$

Proof. By the assumed form of $A^\oplus$, we may write $\varepsilon_k\left(A^\oplus\right)$ as

$$\begin{aligned}\varepsilon_k\left(A^\oplus\right) &= \sum_{p\in L_k^\pm}\varepsilon_{k,k}\left(e_p;A^\oplus e_p\right) = \sum_{p,q\in L_k^\pm}\left\langle e_q, A^\oplus e_p\right\rangle\varepsilon_{k,k}\left(\overline{k,p};\overline{k,q}\right)\\ &= -\sum_{p,q\in L_k^\pm}\left\langle e_q, A^\oplus e_p\right\rangle\left(\delta_{p,q}c_{\overline{q-k}}c_{\overline{p-k}}^* + \delta_{\overline{p-k},\overline{q-k}}c_q^*c_p\right)\\ &= -\sum_{p,q\in L_k}\left\langle e_q, Ae_p\right\rangle\left(\delta_{p,q}c_{q-k}c_{p-k}^* + \delta_{p-k,q-k}c_q^*c_p\right)\\ &\quad -\sum_{p,q\in L_{-k}}\left\langle e_{-q}, Ae_{-p}\right\rangle\left(\delta_{p,q}c_{q+k}c_{p+k}^* + \delta_{p+k,q+k}c_q^*c_p\right)\\ &= -\sum_{p\in L_k}\left\langle e_p, Ae_p\right\rangle\left(c_{p-k}c_{p-k}^* + c_p^*c_p\right) - \sum_{p\in L_{-k}}\left\langle e_{-p}, Ae_{-p}\right\rangle\left(c_{p+k}c_{p+k}^* + c_p^*c_p\right)\end{aligned}\tag{4.46}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

since the terms with $p \in L_k, q \in L_{-k}$ or $p \in L_{-k}, q \in L_k$ vanish (because $L_k \cap L_{-k} = \emptyset$ and there are $\delta_{p,q}, \delta_{p-k,q-k}$ in the summand). We can thus estimate

$$\begin{aligned}\left|\left\langle \Psi, \varepsilon_k\left(A^{\oplus}\right)\Psi\right\rangle\right| &\leq \sum_{p\in L_k}\left|\left\langle e_p, Ae_p\right\rangle\right|\left\langle \Psi, \left(c_{p-k}c_{p-k}^* + c_p^* c_p\right)\Psi\right\rangle \\ &\quad + \sum_{p\in L_{-k}}\left|\left\langle e_{-p}, Ae_{-p}\right\rangle\right|\left\langle \Psi, \left(c_{p+k}c_{p+k}^* + c_p^* c_p\right)\Psi\right\rangle \\ &\leq \left(\max_{p\in L_k}\left|\left\langle e_p, Ae_p\right\rangle\right|\right)\left\langle \Psi, \left(\sum_{p\in L_k^{\pm}} c_p^* c_p + \sum_{p\in L_k - k} c_p c_p^* + \sum_{p\in L_{-k}+k} c_p c_p^*\right)\Psi\right\rangle \\ &\leq 3\left(\max_{p\in L_k}\left|\left\langle e_p, Ae_p\right\rangle\right|\right)\left\langle \Psi, \mathcal{N}_E\Psi\right\rangle .\end{aligned} \tag{4.47}$$

□

Lastly, we consider the $Q_2^k(B)$ terms:

Proposition 4.10. *For all $k \in \mathbb{Z}_+^3$, symmetric $B : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$ and $\Psi \in D\left(H'_{\mathrm{kin}}\right)$, it holds that*

$$\left|\left\langle \Psi, Q_2^k(B)\Psi\right\rangle\right| \leq 2\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} B\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle + 2\left\|B\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\sqrt{\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle}\,\|\Psi\| .$$

Proof. By the Cauchy-Schwarz inequality and Proposition 4.4, we have

$$\begin{aligned}\left|\left\langle \Psi, Q_2^k(B)\Psi\right\rangle\right| &= \left|2\sum_{p\in L_k^{\pm}}\mathrm{Re}\left(\left\langle \Psi, b_k\left(Be_p\right)b_k(e_p)\Psi\right\rangle\right)\right| \leq 2\sum_{p\in L_k^{\pm}}\left\|b_k^*\left(Be_p\right)\Psi\right\|\left\|b_k(e_p)\Psi\right\| \\ &\leq 2\sum_{p\in L_k^{\pm}}\left(\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Be_p\right\|\sqrt{\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle} + \left\|Be_p\right\|\|\Psi\|\right)\left\|b_k(e_p)\Psi\right\| \\ &\leq 2\sqrt{\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle}\sum_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Be_p\right\|\left\|b_k(e_p)\Psi\right\| + 2\|\Psi\|\sum_{p\in L_k^{\pm}}\left\|Be_p\right\|\left\|b_k(e_p)\Psi\right\| .\end{aligned} \tag{4.48}$$

For the first sum, we can again apply the Cauchy-Schwarz inequality and (4.32):

$$\begin{aligned}\sum_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Be_p\right\|\left\|b_k(e_p)\Psi\right\| &\leq \sqrt{\sum_{p\in L_k^{\pm}}\lambda_{k,p}^{-1}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Be_p\right\|^2}\sqrt{\sum_{p\in L_k^{\pm}}\lambda_{\overline{k,p}}\left\|b_{\overline{k,p}}\Psi\right\|^2} \\ &\leq \sqrt{\sum_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}B\left(h_k^{\oplus}\right)^{-\frac{1}{2}}e_p\right\|^2}\sqrt{\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle},\end{aligned} \tag{4.49}$$

and we likewise estimate the second sum as

$$\begin{aligned}\sum_{p\in L_k^{\pm}}\left\|Be_p\right\|\left\|b_k(e_p)\Psi\right\| &\leq \sqrt{\sum_{p\in L_k^{\pm}}\lambda_{k,p}^{-1}\left\|Be_p\right\|^2}\sqrt{\sum_{p\in L_k^{\pm}}\lambda_{\overline{k,p}}\left\|b_{\overline{k,p}}\Psi\right\|^2} \\ &\leq \sqrt{\sum_{p\in L_k^{\pm}}\left\|B\left(h_k^{\oplus}\right)^{-\frac{1}{2}}e_p\right\|^2}\sqrt{\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle}.\end{aligned} \tag{4.50}$$

The claim now follows by recognizing the Hilbert-Schmidt norms. □

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

5. The quasi-bosonic Bogolubov transformation

Now we are prepared to define the quasi-bosonic Bogolubov transformation that will approximately diagonalize the Hamiltonian in (4.13),

$$\sum_{k\in S_C}\left(Q_1^k(h_k^\oplus + A_k^\oplus) + Q_2^k(B_k^\oplus) - 2\,\mathrm{tr}(h_k)\right),\tag{5.1}$$

where $h_k^\oplus, A_k^\oplus, B_k^\oplus$ are defined in (4.11) and (4.7).

We define the generator $\mathcal{K}:\mathcal{H}_N\to\mathcal{H}_N$ of the Bogolubov transformation as follows. Let $(K_k^\oplus)_{k\in S_C}$ be a collection of symmetric operators $K_k^\oplus:\ell^2(L_k^\pm)\to\ell^2(L_k^\pm)$. Then we define

$$\begin{aligned}\mathcal{K} &= \frac{1}{2}\sum_{k\in S_C}\sum_{p,q\in L_k^\pm}\left\langle e_p, K_k^\oplus e_q\right\rangle\left(b_{\overline{k,p}}b_{\overline{k,q}} - b^*_{\overline{k,q}}b^*_{\overline{k,p}}\right)\\ &= \frac{1}{2}\sum_{k\in S_C}\sum_{p\in L_k^\pm}\left(b_k\left(K_k^\oplus e_p\right)b_k(e_p) - b_k^*(e_p)b_k^*\left(K_k^\oplus e_p\right)\right)\end{aligned}\tag{5.2}$$

in analogy with equation (3.9). As in the bosonic case, $\mathcal{K}$ is seen to be a skew-symmetric operator.[8] Moreover, unlike the bosonic case, $\mathcal{K}$ is now a bounded operator by the same argument that $Q_1^k(\cdot)$ and $Q_2^k(\cdot)$ are. Therefore, $\mathcal{K}$ generates a unitary transformation $e^{\mathcal{K}}:\mathcal{H}_N\to\mathcal{H}_N$, which is the quasi-bosonic Bogolubov transformation.

The specific kernels $K_k^\oplus$ we will use are those which diagonalize the corresponding bosonic Hamiltonian exactly, but first we will consider the action of $e^{\mathcal{K}}$ on quadratic operators and the localized kinetic operator more generally.

5.1. *Transformation of quadratic operators*

By exploiting the similarity of our quasi-bosonic definitions with the exact bosonic case, we can now easily deduce the analogues of Propositions 3.2 and 3.4:

Proposition 5.1. *For all $k\in S_C$, $\varphi\in\ell^2(L_k^\pm)$ and symmetric operators $(K_l^\oplus)_{l\in S_C}$, it holds that*

$$\begin{aligned}[\mathcal{K}, b_k(\varphi)] &= b_k^*\left(K_k^\oplus\varphi\right) + \mathcal{E}_k(\varphi),\\ \left[\mathcal{K}, b_k^*(\varphi)\right] &= b_k\left(K_k^\oplus\varphi\right) + \mathcal{E}_k(\varphi)^*,\end{aligned}$$

where

$$\mathcal{E}_k(\varphi) = \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left\{b_l^*\left(K_l^\oplus e_q\right), \varepsilon_{k,l}\left(\varphi; e_q\right)\right\}.$$

Proof. We calculate using the commutation relations of (4.15) that

$$\begin{aligned}[\mathcal{K}, b_k(\varphi)] &= \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left(\left[b_l\left(K_l^\oplus e_q\right)b_l(e_p), b_k(\varphi)\right] - \left[b_l^*\left(e_q\right)b_l^*\left(K_l^\oplus e_q\right), b_k(\varphi)\right]\right)\\ &= \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left(b_l^*\left(e_q\right)\left[b_k(\varphi), b_l^*\left(K_l^\oplus e_q\right)\right] + \left[b_k(\varphi), b_l^*\left(e_q\right)\right]b_l^*\left(K_l^\oplus e_q\right)\right)\end{aligned}$$

[8] In the case of complex spaces, $\mathcal{K}$ is skew-symmetric if the $K_k^\oplus$'s are symmetric and $\left\langle e_p, K_k^\oplus e_q\right\rangle$ are real. In our application, all relevant operators have real matrix elements, and hence, we can think of the case of real spaces.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

$$
\begin{aligned}
&= \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm} b_l^*\left(e_q\right)\left(\delta_{k,l}\left\langle \varphi, K_l^\oplus e_q\right\rangle + \varepsilon_{k,l}\left(\varphi; K_l^\oplus e_q\right)\right) \\
&\quad + \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left(\delta_{k,l}\left\langle \varphi, e_q\right\rangle + \varepsilon_{k,l}\left(\varphi; e_q\right)\right) b_l^*\left(K_l^\oplus e_q\right) \\
&= \frac{1}{2} b_k^*\left(\sum_{q\in L_k^\pm}\left\langle \varphi, K_k^\oplus e_q\right\rangle e_q\right) + \frac{1}{2} b_k^*\left(K_k^\oplus \sum_{q\in L_k^\pm}\left\langle \varphi, e_q\right\rangle e_q\right) + \mathcal{E}_k(\varphi) \\
&= b_k^*\left(K_k^\oplus \varphi\right) + \mathcal{E}_k(\varphi)
\end{aligned}
\tag{5.3}
$$

for $\mathcal{E}_k(\varphi)$ given by

$$
\begin{aligned}
\mathcal{E}_k(\varphi) &= \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left(b_l^*\left(e_q\right)\varepsilon_{k,l}\left(\varphi; K_l^\oplus e_q\right) + \varepsilon_{k,l}\left(\varphi; e_q\right) b_l^*\left(K_l^\oplus e_q\right)\right) \\
&= \frac{1}{2}\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left\{ b_l^*\left(K_l^\oplus e_q\right), \varepsilon_{k,l}\left(\varphi; e_q\right)\right\},
\end{aligned}
\tag{5.4}
$$

where we used Lemma 3.3 to simplify the expression (as $x, y \mapsto b_l^*(x)\varepsilon_{k,l}(\varphi; y)$ is bilinear for fixed φ and $K_k^\oplus$ is symmetric). The commutator $\left[\mathcal{K}, b_k^*(\varphi)\right]$ follows by taking the adjoint. □

From this, we easily deduce the commutator of $\mathcal{K}$ with quadratic operators:

Proposition 5.2. *For all $k \in S_C$ and symmetric operators $A, B : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$, it holds that*

$$
\begin{aligned}
\left[\mathcal{K}, Q_1^k(A)\right] &= Q_2^k\left(\left\{K_k^\oplus, A\right\}\right) + \mathcal{E}_1^k(A) \\
\left[\mathcal{K}, Q_2^k(B)\right] &= Q_1^k\left(\left\{K_k^\oplus, B\right\}\right) + \mathcal{E}_2^k(B),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{E}_1^k(A) &= \frac{1}{2}\sum_{l\in S_C}\sum_{p\in L_k^\pm}\sum_{q\in L_l^\pm}\left(\left\{b_k^*(Ae_p), \left\{b_l^*\left(K_l^\oplus e_q\right), \varepsilon_{k,l}(e_p; e_q)\right\}\right\}\right. \\
&\quad \left. + \left\{\left\{\varepsilon_{l,k}\left(e_q; e_p\right), b_l\left(K_l^\oplus e_q\right)\right\}, b_k(Ae_p)\right\}\right) \\
\mathcal{E}_2^k(B) &= \frac{1}{2}\sum_{l\in S_C}\sum_{p\in L_k^\pm}\sum_{q\in L_l^\pm}\left(\left\{b_k^*\left(Be_p\right), \left\{b_l\left(K_l^\oplus e_q\right), \varepsilon_{l,k}\left(e_q; e_p\right)\right\}\right\}\right. \\
&\quad \left. + \left\{\left\{\varepsilon_{k,l}(e_p; e_q), b_l^*\left(K_l^\oplus e_q\right)\right\}, b_k\left(Be_p\right)\right\}\right).
\end{aligned}
$$

Proof. We compute using the commutators of the previous proposition (and Lemma 3.3, to simplify the resulting expressions) that

$$
\begin{aligned}
\left[\mathcal{K}, Q_1^k(A)\right] &= \sum_{p\in L_k^\pm}\left(\left[\mathcal{K}, b_k^*(Ae_p) b_k(e_p)\right] + \left[\mathcal{K}, b_k(e_p) b_k^*(Ae_p)\right]\right) \\
&= \sum_{p\in L_k^\pm}\left(b_k^*(Ae_p)\left[\mathcal{K}, b_k(e_p)\right] + \left[\mathcal{K}, b_k^*(Ae_p)\right] b_k(e_p)\right) \\
&\quad + \sum_{p\in L_k^\pm}\left(b_k(e_p)\left[\mathcal{K}, b_k^*(Ae_p)\right] + \left[\mathcal{K}, b_k(e_p)\right] b_k^*(Ae_p)\right)
\end{aligned}
$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

$$
\begin{aligned}
&= \sum_{p\in L_k^\pm} \left(b_k^*(Ae_p)\left(b_k^*\left(K_k^\oplus e_p\right)+\mathcal{E}_k(e_p)\right)+\left(b_k\left(K_k^\oplus Ae_p\right)+\mathcal{E}_k(Ae_p)^*\right)b_k(e_p)\right)\\
&\quad+\sum_{p\in L_k^\pm}\left(b_k(e_p)\left(b_k\left(K_k^\oplus Ae_p\right)+\mathcal{E}_k(Ae_p)^*\right)+\left(b_k^*\left(K_k^\oplus e_p\right)+\mathcal{E}_k(e_p)\right)b_k^*(Ae_p)\right)\\
&= \sum_{p\in L_k^\pm}\left(b_k^*\left(\left(AK_k^\oplus+K_k^\oplus A\right)e_p\right)b_k^*(e_p)+b_k(e_p)b_k\left(\left(K_k^\oplus A+K_k^\oplus A\right)e_p\right)\right)\\
&\quad+\sum_{p\in L_k^\pm}\left(b_k^*(Ae_p)\mathcal{E}_k(e_p)+\mathcal{E}_k(e_p)^*b_k(Ae_p)+b_k(Ae_p)\mathcal{E}_k(e_p)^*+\mathcal{E}_k(e_p)b_k^*(Ae_p)\right)\\
&= Q_2^k\left(\left\{K_k^\oplus,A\right\}\right)+\sum_{p\in L_k^\pm}\left(\left\{b_k^*(Ae_p),\mathcal{E}_k(e_p)\right\}+\left\{\mathcal{E}_k(e_p)^*,b_k(Ae_p)\right\}\right)
\end{aligned}
\tag{5.5}
$$

and

$$
\begin{aligned}
&\sum_{p\in L_k^\pm}\left(\left\{b_k^*(Ae_p),\mathcal{E}_k(e_p)\right\}+\left\{\mathcal{E}_k(e_p)^*,b_k(Ae_p)\right\}\right)\\
&\quad=\frac{1}{2}\sum_{l\in S_C}\sum_{p\in L_k^\pm}\sum_{q\in L_l^\pm}\left(\left\{b_k^*(Ae_p),\left\{b_l^*\left(K_l^\oplus e_q\right),\varepsilon_{k,l}(e_p;e_q)\right\}\right\}\right.\\
&\qquad\left.+\left\{\left\{\varepsilon_{l,k}\left(e_q;e_p\right),b_l\left(K_l^\oplus e_q\right)\right\},b_k(Ae_p)\right\}\right)=\mathcal{E}_1^k(A)
\end{aligned}
\tag{5.6}
$$

as $\varepsilon_{k,l}(e_p;e_q)^*=\varepsilon_{l,k}\left(e_q;e_p\right)$. The computation of $Q_2^k(B)$ is similar. □

Action of $e^{\mathcal{K}}$ on quadratic operators

With the commutators calculated, we are now ready to determine the full action of $e^{\mathcal{K}}$ on the quadratic operators $Q_1^k(\cdot)$ and $Q_2^k(\cdot)$. Rather than appeal to the Baker-Campbell-Hausdorff formula, which would also require describing the commutators $\left[\mathcal{K},\mathcal{E}_1^k(A)\right]$, etc., we will employ a 'Duhamel-type' argument which allows us to more selectively expand the operator $e^{\mathcal{K}}$.

As in Section 3, we use the notation $\mathcal{A}_{K_k^\oplus}=\left\{K_k^\oplus,\cdot\right\}$ for anticommutators with $K_k^\oplus$.

Before stating the proposition, we must make a remark. To use these identities, we will need to take limits, and to justify those limits, we need some general estimates on operators of the form $Q_1^k(\cdot),Q_2^k(\cdot),\mathcal{E}_1^k(\cdot),\mathcal{E}_2^k(\cdot)$. The Propositions 4.6, 4.7 establish these for $Q_1^k(\cdot)$ and $Q_2^k(\cdot)$, while Proposition 6.4 will establish these for $\mathcal{E}_1^k(\cdot)$ and $\mathcal{E}_2^k(\cdot)$.

The statement follows:

Proposition 5.3. *For all $k\in S_C$ and symmetric $A,B:\ell^2(L_k^\pm)\to\ell^2(L_k^\pm)$, it holds that*

$$
\begin{aligned}
e^{\mathcal{K}}Q_1^k(A)e^{-\mathcal{K}}&=\frac{1}{2}Q_1^k\left(e^{K_k^\oplus}Ae^{K_k^\oplus}+e^{-K_k^\oplus}Ae^{-K_k^\oplus}\right)+\frac{1}{2}Q_2^k\left(e^{K_k^\oplus}Ae^{K_k^\oplus}-e^{-K_k^\oplus}Ae^{-K_k^\oplus}\right)\\
&\quad+\int_0^1 e^{t\mathcal{K}}\left(\mathcal{E}_1^k\left(\cosh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)(A)\right)+\mathcal{E}_2^k\left(\sinh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)(A)\right)\right)e^{-t\mathcal{K}}\,dt\\
e^{\mathcal{K}}Q_2^k(B)e^{-\mathcal{K}}&=\frac{1}{2}Q_1^k\left(e^{K_k^\oplus}Be^{K_k^\oplus}-e^{-K_k^\oplus}Be^{-K_k^\oplus}\right)+\frac{1}{2}Q_2^k\left(e^{K_k^\oplus}Be^{K_k^\oplus}+e^{-K_k^\oplus}Be^{-K_k^\oplus}\right)\\
&\quad+\int_0^1 e^{t\mathcal{K}}\left(\mathcal{E}_1^k\left(\sinh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)(B)\right)+\mathcal{E}_2^k\left(\cosh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)(B)\right)\right)e^{-t\mathcal{K}}\,dt,
\end{aligned}
$$

the integrals being Riemann integrals of bounded operators.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proof. We consider $e^{\mathcal{K}}Q_1^k(A)e^{-\mathcal{K}}$, with the argument for $e^{\mathcal{K}}Q_2^k(B)e^{-\mathcal{K}}$ being similar. We first claim that for any $n \in \mathbb{N}$,

$$\begin{aligned} e^{\mathcal{K}}Q_1^k(A)e^{-\mathcal{K}} &= Q_1^k\left(\sum_{m=0}^{n_1}\frac{1}{(2m)!}\mathcal{A}_{K_k^\oplus}^{2m}(A)\right) + Q_2^k\left(\sum_{m=0}^{n_2}\frac{1}{(2m+1)!}\mathcal{A}_{K_k^\oplus}^{2m+1}(A)\right) \\ &+ \int_0^1 e^{t\mathcal{K}}\left(\mathcal{E}_1^k\left(\sum_{m=0}^{n_1}\frac{1}{(2m)!}\mathcal{A}^{2m}_{(1-t)K_k^\oplus}(A)\right) + \mathcal{E}_2^k\left(\sum_{m=0}^{n_2}\frac{1}{(2m+1)!}\mathcal{A}^{2m+1}_{(1-t)K_k^\oplus}(A)\right)\right)e^{-t\mathcal{K}}\,dt \\ &+ \frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{K}}Q^k_{\overline{n-1}}\left(\mathcal{A}^n_{K_k^\oplus}(A)\right)e^{-t\mathcal{K}}(1-t)^{n-1}\,dt, \end{aligned} \tag{5.7}$$

where, for brevity, $\overline{n-1} = n-1 \mod 2$ and n_1, n_2 are the largest integers such that $2n_1 < n$ and $2n_2+1 < n$, respectively.

We proceed by induction. For $n = 1$, we find by the fundamental theorem of calculus that

$$\begin{aligned} e^{\mathcal{K}}Q_1^k(A)e^{-\mathcal{K}} &= Q_1^k(A) + \int_0^1 \frac{d}{dt}\left(e^{t\mathcal{K}}Q_1^k(A)e^{-t\mathcal{K}}\right)dt = Q_1^k(A) + \int_0^1 e^{t\mathcal{K}}\left[\mathcal{K}, Q_1^k(A)\right]e^{-t\mathcal{K}}\,dt \\ &= Q_1^k(A) + \int_0^1 e^{t\mathcal{K}}\left(Q_2^k\left(\{K_k^\oplus, A\}\right) + \mathcal{E}_1^k(A)\right)e^{-t\mathcal{K}}\,dt \\ &= Q_1^k(A) + \int_0^1 e^{t\mathcal{K}}\mathcal{E}_1^k(A)e^{-t\mathcal{K}}\,dt + \int_0^1 e^{t\mathcal{K}}Q_2^k\left(\mathcal{A}_{K_k^\oplus}(A)\right)e^{-t\mathcal{K}}\,dt \end{aligned} \tag{5.8}$$

by the commutator of Proposition 5.2, which is the statement for $n = 1$ (in this case, $n_1 = 0$ and $n_2 = -1$, so $\sum_{m=0}^{n_1}$ contains one term and $\sum_{m=0}^{n_2}$ is empty).

For the inductive step, we now assume that case n holds. Integrating the last term of equation (5.7) by parts, we find that

$$\begin{aligned} &\frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{K}}Q^k_{\overline{n-1}}\left(\mathcal{A}^n_{K_k^\oplus}(A)\right)e^{-t\mathcal{K}}(1-t)^{n-1}dt \\ &= \frac{1}{(n-1)!}\left[e^{t\mathcal{K}}Q^k_{\overline{n-1}}\left(\mathcal{A}^n_{K_k^\oplus}(A)\right)e^{-t\mathcal{K}}\left(-\frac{(1-t)^n}{n}\right)\right]_0^1 \\ &\quad - \frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{K}}\left[\mathcal{K}, Q^k_{\overline{n-1}}\left(\mathcal{A}^n_{K_k^\oplus}(A)\right)\right]e^{-t\mathcal{K}}\left(-\frac{(1-t)^n}{n}\right)dt \\ &= \frac{1}{n!}Q^k_{\overline{n-1}}\left(\mathcal{A}^n_{K_k^\oplus}(A)\right) + \frac{1}{n!}\int_0^1 e^{t\mathcal{K}}\left(Q^k_{\overline{n}}\left(\left\{K_k^\oplus, \mathcal{A}^n_{K_k^\oplus}(A)\right\}\right) + \mathcal{E}^k_{\overline{n-1}}\left(\mathcal{A}^n_{K_k^\oplus}(A)\right)\right)e^{-t\mathcal{K}}(1-t)^n dt \\ &= Q^k_{\overline{n-1}}\left(\frac{1}{n!}\mathcal{A}^n_{K_k^\oplus}(A)\right) + \int_0^1 e^{t\mathcal{K}}\mathcal{E}^k_{\overline{n-1}}\left(\frac{1}{n!}\mathcal{A}^n_{(1-t)K_k^\oplus}(A)\right)e^{-t\mathcal{K}}\,dt \\ &\quad + \frac{1}{n!}\int_0^1 e^{t\mathcal{K}}Q^k_{\overline{n}}\left(\mathcal{A}^{n+1}_{K_k^\oplus}(A)\right)e^{-t\mathcal{K}}(1-t)^n dt, \end{aligned} \tag{5.9}$$

where we also used that

$$(1-t)^n\mathcal{A}^n_{K_k^\oplus}(A) = \left((1-t)\mathcal{A}_{K_k^\oplus}(A)\right)^n = \mathcal{A}^n_{(1-t)K_k^\oplus}(A). \tag{5.10}$$

Inserting this into (5.7) and collecting like terms yields the statement for case $n+1$.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

We now deduce the statement from (5.7) by taking $n \to \infty$. Recall the identities

$$\cosh\left(\mathcal{A}_{K_k^\oplus}\right)(T) = \frac{1}{2}\left(e^{K_k^\oplus} T e^{K_k^\oplus} + e^{-K_k^\oplus} T e^{-K_k^\oplus}\right) \tag{5.11}$$
$$\sinh\left(\mathcal{A}_{K_k^\oplus}\right)(T) = \frac{1}{2}\left(e^{K_k^\oplus} T e^{K_k^\oplus} - e^{-K_k^\oplus} T e^{-K_k^\oplus}\right)$$

from Proposition 3.5 and note that $((n-1)!)^{-1} \mathcal{A}_{K_k^\oplus}^n(A) \to 0$ as $n \to \infty$. By Proposition 4.6,

$$Q_1^k\left(\sum_{m=0}^{n_1} \frac{1}{(2m)!} \mathcal{A}_{K_k^\oplus}^{2m}(A)\right) \to \frac{1}{2} Q_1^k\left(e^{K_k^\oplus} A e^{K_k^\oplus} + e^{-K_k^\oplus} A e^{-K_k^\oplus}\right)$$

and

$$\frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{K}} Q_1^k\left(\mathcal{A}_{K_k^\oplus}^n(A)\right) e^{-t\mathcal{K}} (1-t)^{n-1}\, dt \to 0.$$

Similar convergence for Q_2 is justified by Proposition 4.7. The convergence for $\mathcal{E}_1^k$ and $\mathcal{E}_2^k$ follows from Proposition 6.4. □

Remark on the transformation of excitation operators

Let us make a quick remark on why we choose to approach the Bogolubov transformation from the point of view of quadratic operators rather than the usual creation and annihilation operator approach. Recall that in the exact bosonic case the creation and annihilation operators transformed under a Bogolubov transformation as

$$e^{\mathcal{K}} a(\varphi) e^{-\mathcal{K}} = a\left(\cosh(K)\,\varphi\right) + a^*\left(\sinh(K)\,\varphi\right) \tag{5.12}$$
$$e^{\mathcal{K}} a^*(\varphi) e^{-\mathcal{K}} = a^*\left(\cosh(K)\,\varphi\right) + a\left(\sinh(K)\,\varphi\right).$$

In the quasi-bosonic setting, we can use the commutators of Proposition 5.1 and a similar Duhamel-type argument to what we just applied to conclude that

$$\begin{aligned} e^{\mathcal{K}} b_k(\varphi) e^{-\mathcal{K}} &= b_k\left(\cosh\left(K_k^\oplus\right)\varphi\right) + b_k^*\left(\sinh\left(K_k^\oplus\right)\varphi\right) \\ &\quad + \int_0^1 e^{t\mathcal{K}}\left(\mathcal{E}_k\left(\cosh\left((1-t)K_k^\oplus\right)\varphi\right) + \mathcal{E}_k\left(\sinh\left((1-t)K_k^\oplus\right)\varphi\right)^*\right) e^{-t\mathcal{K}}\, dt \end{aligned} \tag{5.13}$$

with a similar expression for $e^{\mathcal{K}} b_k^*(\varphi) e^{-\mathcal{K}}$. This is a more cumbersome expression to work with, and if we were to describe $e^{\mathcal{K}} Q_1^k(A) e^{-\mathcal{K}}$ by transforming the individual terms of $Q_1^k(A)$ like this rather than transforming $Q_1^k(A)$ as a whole, the error terms would not only go from being under a single integral to involving the product of two integrals, it would also involve cross terms between the bosonic terms and the error terms of equation (5.13). These cross terms, in particular, would severely reduce the quality of the final error estimate. Hence, we prefer the quadratic operator approach in the quasi-bosonic setting.

5.2. *Transformation of the kinetic operator*

There remains the task of describing the action of $e^{\mathcal{K}}$ on the localized kinetic operator H'_{kin}. For this, we must first formulate H'_{kin} – or rather the commutator $[H'_{\text{kin}}, b^*_{k,p}]$ calculated in (1.74) – within the general framework that we have introduced in this section. Recalling the operators $h_k^\oplus : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$ in (4.11), then by (1.74) and linearity it follows that

$$\left[H'_{\text{kin}}, b_k(\varphi)\right] = -2\, b_k\left(h_k^\oplus \varphi\right), \quad \left[H'_{\text{kin}}, b_k^*(\varphi)\right] = 2\, b_k^*\left(h_k^\oplus \varphi\right) \tag{5.14}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

for all $\varphi \in \ell^2(L_k^\pm)$. (The factor of 2 is introduced here because in the analogy of equation (4.9), H'_{kin} appears like a $\mathrm{d}\Gamma(\cdot) = \frac{1}{2}Q_1(\cdot) - \frac{1}{2}\mathrm{tr}(\cdot)$ term rather than a pure $Q_1(\cdot)$ term.)

We now calculate $\left[\mathcal{K}, H'_{\text{kin}}\right]$ as follows:

Proposition 5.4. H'_{kin} *obeys*

$$\left[\mathcal{K}, H'_{\text{kin}}\right] = \sum_{k \in S_C} Q_2^k\left(\left\{K_k^\oplus, h_k^\oplus\right\}\right).$$

Proof. We compute, using the commutators of equation (5.14) and Lemma 3.3, that

$$\begin{aligned}\left[\mathcal{K}, H'_{\text{kin}}\right] &= \frac{1}{2}\sum_{k \in S_C}\sum_{p \in L_k^\pm}\left(\left[b_k\left(K_k^\oplus e_p\right)b_k(e_p), H'_{\text{kin}}\right] - \left[b_k^*(e_p)b_k^*\left(K_k^\oplus e_p\right), H'_{\text{kin}}\right]\right)\\ &= \sum_{k \in S_C}\sum_{p \in L_k^\pm}\left(b_k^*(e_p)b_k^*\left(h_k^\oplus K_k^\oplus e_p\right) + b_k^*\left(h_k^\oplus e_p\right)b_k^*\left(K_k^\oplus e_p\right)\right)\\ &\quad + \sum_{k \in S_C}\sum_{p \in L_k^\pm}\left(b_k\left(K_k^\oplus e_p\right)b_k\left(h_k^\oplus e_p\right) + b_k\left(h_k^\oplus K_k^\oplus e_p\right)b_k(e_p)\right)\\ &= \sum_{k \in S_C} Q_2^k\left(\left\{K_k^\oplus, h_k^\oplus\right\}\right).\end{aligned} \tag{5.15}$$

□

Note that because the commutator $\left[H'_{\text{kin}}, b_k^*(\varphi)\right] = 2\, b_k^*\left(h_k^\oplus\varphi\right)$ exactly mirrors the bosonic case (in that there is no additional error term), the commutator $\left[\mathcal{K}, H'_{\text{kin}}\right]$ is likewise 'purely bosonic', being simply a sum of $Q_2^k(\cdot)$ terms without error terms such as those appearing in the statement of Proposition 5.2. With the groundwork laid, we can now easily deduce the following:

Proposition 5.5. H'_{kin} *obeys*

$$\begin{aligned}& e^{\mathcal{K}}H'_{\text{kin}}e^{-\mathcal{K}} = H'_{\text{kin}}\\ &+ \sum_{k \in S_C}\left(\frac{1}{2}Q_1^k\left(e^{K_k^\oplus}h_k^\oplus e^{K_k^\oplus} + e^{-K_k^\oplus}h_k^\oplus e^{-K_k^\oplus} - 2h_k^\oplus\right) + \frac{1}{2}Q_2^k\left(e^{K_k^\oplus}h_k^\oplus e^{K_k^\oplus} - e^{-K_k^\oplus}h_k^\oplus e^{-K_k^\oplus}\right)\right)\\ &+ \sum_{k \in S_C}\int_0^1 e^{t\mathcal{K}}\left(\mathcal{E}_1^k\left(\cosh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)\left(h_k^\oplus\right) - h_k^\oplus\right) + \mathcal{E}_2^k\left(\sinh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)\left(h_k^\oplus\right)\right)\right)e^{-t\mathcal{K}}\,dt.\end{aligned}$$

Proof. By adding and subtracting, we have

$$e^{\mathcal{K}}H'_{\text{kin}}e^{-\mathcal{K}} = \sum_{k \in S_C} e^{\mathcal{K}}Q_1^k\left(h_k^\oplus\right)e^{-\mathcal{K}} + e^{\mathcal{K}}\left(H'_{\text{kin}} - \sum_{k \in S_C}Q_1^k\left(h_k^\oplus\right)\right)e^{-\mathcal{K}}, \tag{5.16}$$

and the first term on the right-hand side is by Proposition 5.3,

$$\begin{aligned}&\sum_{k \in S_C} e^{\mathcal{K}}Q_1^k\left(h_k^\oplus\right)e^{-\mathcal{K}}\\ &= \sum_{k \in S_C}\left(\frac{1}{2}Q_1^k\left(e^{K_k^\oplus}h_k^\oplus e^{K_k^\oplus} + e^{-K_k^\oplus}h_k^\oplus e^{-K_k^\oplus}\right) + \frac{1}{2}Q_2^k\left(e^{K_k^\oplus}h_k^\oplus e^{K_k^\oplus} - e^{-K_k^\oplus}h_k^\oplus e^{-K_k^\oplus}\right)\right)\\ &\quad + \sum_{k \in S_C}\int_0^1 e^{t\mathcal{K}}\left(\mathcal{E}_1^k\left(\cosh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)\left(h_k^\oplus\right)\right) + \mathcal{E}_2^k\left(\sinh\left(\mathcal{A}_{(1-t)K_k^\oplus}\right)\left(h_k^\oplus\right)\right)\right)e^{-t\mathcal{K}}\,dt,\end{aligned} \tag{5.17}$$

while the second is calculated using the commutators of the Propositions 5.2 and 5.4 to be

$$\begin{aligned} & e^{\mathcal{K}}\left(H'_{\text{kin}} - \sum_{k\in S_C} Q_1^k\left(h_k^{\oplus}\right)\right)e^{-\mathcal{K}} - \left(H'_{\text{kin}} - \sum_{k\in S_C} Q_1^k\left(h_k^{\oplus}\right)\right) \\ &= \int_0^1 e^{t\mathcal{K}}\left[\mathcal{K}, H'_{\text{kin}} - \sum_{k\in S_C} Q_1^k\left(h_k^{\oplus}\right)\right]e^{-t\mathcal{K}}\,dt = -\sum_{k\in S_C}\int_0^1 e^{t\mathcal{K}}\mathcal{E}_1^k\left(h_k^{\oplus}\right)e^{-t\mathcal{K}}\,dt, \end{aligned} \tag{5.18}$$

which yields the claim. □

5.3. *Fixing the transformation kernels*

With all the transformation identities determined, we now choose the transformation kernels $(K_k^{\oplus})_{k\in S_C}$ such that $H'_{\text{kin}} + \sum_{k\in S_C} H_{\text{int}}^k$ is diagonalized. For any choice of $(K_k^{\oplus})_{k\in S_C}$, the Propositions 5.3 and 5.5 imply that

$$\begin{aligned} & e^{\mathcal{K}}\left(H'_{\text{kin}} + \sum_{k\in S_C} H_{\text{int}}^k\right)e^{-\mathcal{K}} \\ &= \frac{1}{2}\sum_{k\in S_C} Q_1^k\left(e^{K_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}+B_k^{\oplus}\right)e^{K_k^{\oplus}} + e^{-K_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)e^{-K_k^{\oplus}} - 2h_k^{\oplus}\right) \\ &\quad + \frac{1}{2}\sum_{k\in S_C} Q_2^k\left(e^{K_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}+B_k^{\oplus}\right)e^{K_k^{\oplus}} - e^{-K_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)e^{-K_k^{\oplus}}\right) + H'_{\text{kin}} + \text{error terms}. \end{aligned} \tag{5.19}$$

In analogy with the bosonic case, we consider this expression to be diagonalized provided the $Q_2^k(\cdot)$ terms vanish, whence the diagonalization condition is that

$$e^{K_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}+B_k^{\oplus}\right)e^{K_k^{\oplus}} = e^{-K_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)e^{-K_k^{\oplus}}, \tag{5.20}$$

which we note is the same as the diagonalization condition (equation (3.26)) of the exact bosonic quadratic Hamiltonian

$$H = Q_1\left(h_k^{\oplus}+A_k^{\oplus}\right) + Q_2\left(B_k^{\oplus}\right) \quad \text{on } \mathcal{F}^+\left(\ell^2\left(L_k^{\pm}\right)\right). \tag{5.21}$$

Recalling the definitions of $h_k^{\oplus}$, $A_k^{\oplus}$ and $B_k^{\oplus}$ from (4.11) and (4.7), we have

$$h_k^{\oplus}+A_k^{\oplus}\pm B_k^{\oplus} = \begin{pmatrix} h_k + P_{v_k} & \pm P_{v_k} \\ \pm P_k & h_k + P_{v_k} \end{pmatrix} > 0.$$

So by Theorem 3.1, the choice

$$K_k^{\oplus} = -\frac{1}{2}\log\left(\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)^{-\frac{1}{2}}\left(\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)^{\frac{1}{2}}\left(h_k^{\oplus}+A_k^{\oplus}+B_k^{\oplus}\right)\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)^{\frac{1}{2}}\right)^{\frac{1}{2}}\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)^{-\frac{1}{2}}\right)$$

is the unique diagonalizing kernel for the Hamiltonian. In this form, it is, however, not easy to see how $K_k^\oplus$ acts, so we will proceed slightly differently: we define $K_k^\oplus : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$ by

$$K_k^\oplus = \begin{pmatrix} 0 & K_k \\ K_k & 0 \end{pmatrix}, \tag{5.22}$$

where the operator $K_k : \ell^2(L_k) \to \ell^2(L_k)$ is given by

$$K_k = -\frac{1}{2}\log\left(h_k^{-\frac{1}{2}}\left(h_k^{\frac{1}{2}}\left(h_k + 2P_{v_k}\right)h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} h_k^{-\frac{1}{2}}\right) = -\frac{1}{2}\log\left(h_k^{-\frac{1}{2}}\left(h_k^2 + 2P_{h_k^{\frac{1}{2}} v_k}\right)^{\frac{1}{2}} h_k^{-\frac{1}{2}}\right). \tag{5.23}$$

A kernel similar to K_k also appeared in [5, 6]. Note that K_k is precisely the diagonalizer of Theorem 3.1 for the exact bosonic quadratic Hamiltonian

$$H = Q_1\left(h_k^\oplus + P_{v_k}\right) + Q_2(P_{v_k}) \quad \text{on } \mathcal{F}^+(\ell^2(L_k)), \tag{5.24}$$

rather than that of equation (5.21). Now we can verify that this $K_k^\oplus$ is, in fact, equal to the diagonalizing kernel:

Proposition 5.6. *The operator $K_k^\oplus$ defined by the equations (5.22) and (5.23) satisfies*

$$e^{K_k^\oplus}\left(h_k^\oplus + A_k^\oplus + B_k^\oplus\right)e^{K_k^\oplus} = e^{-K_k^\oplus}\left(h_k^\oplus + A_k^\oplus - B_k^\oplus\right)e^{-K_k^\oplus} = \begin{pmatrix} E_k & 0 \\ 0 & E_k \end{pmatrix}$$

for $E_k = e^{-K_k} h_k e^{-K_k}$.

Proof. It is easily verified that $e^{\pm K_k^\oplus}$ is given by

$$e^{\pm K_k^\oplus} = \begin{pmatrix} \cosh(K_k) & \pm\sinh(K_k) \\ \pm\sinh(K_k) & \cosh(K_k) \end{pmatrix}, \tag{5.25}$$

and so

$$\begin{aligned} & e^{\pm K_k^\oplus}\left(h_k^\oplus + A_k^\oplus \pm B_k^\oplus\right)e^{\pm K_k^\oplus} \\ &= \begin{pmatrix} \cosh(K_k) & \pm\sinh(K_k) \\ \pm\sinh(K_k) & \cosh(K_k) \end{pmatrix}\begin{pmatrix} h_k + P_{v_k} & \pm P_{v_k} \\ \pm P_{v_k} & h_k + P_{v_k} \end{pmatrix}\begin{pmatrix} \cosh(K_k) & \pm\sinh(K_k) \\ \pm\sinh(K_k) & \cosh(K_k) \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} e^{K_k}\left(h_k + 2P_{v_k}\right)e^{K_k} + e^{-K_k}h_k e^{-K_k} & \pm\left(e^{K_k}\left(h_k + 2P_{v_k}\right)e^{K_k} - e^{-K_k}h_k e^{-K_k}\right) \\ \pm\left(e^{K_k}\left(h_k + 2P_{v_k}\right)e^{K_k} - e^{-K_k}h_k e^{-K_k}\right) & e^{K_k}\left(h_k + 2P_{v_k}\right)e^{K_k} + e^{-K_k}h_k e^{-K_k} \end{pmatrix}. \end{aligned} \tag{5.26}$$

The condition

$$e^{K_k^\oplus}\left(h_k^\oplus + A_k^\oplus + B_k^\oplus\right)e^{K_k^\oplus} = e^{-K_k^\oplus}\left(h_k^\oplus + A_k^\oplus - B_k^\oplus\right)e^{-K_k^\oplus} \tag{5.27}$$

thus holds if and only if

$$e^{K_k}\left(h_k + 2P_{v_k}\right)e^{K_k} = e^{-K_k}h_k e^{-K_k}, \tag{5.28}$$

which is the diagonalization condition for the bosonic Hamiltonian of equation (5.24). Theorem 3.1 asserts that this condition is satisfied for our choice of K_k, and the claim follows. □

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

5.4. Full transformation of the bosonizable terms

With the above choice of transformation kernels, we thus conclude that

$$\begin{aligned} e^{\mathcal{K}}\left(H'_{\text{kin}}+\sum_{k\in S_C}H^k_{\text{int}}\right)e^{-\mathcal{K}} &= H'_{\text{kin}}+\text{error terms}\\ &\quad+\frac{1}{2}\sum_{k\in S_C}Q^k_1\left(e^{K^\oplus_k}\left(h^\oplus_k+A^\oplus_k+B^\oplus_k\right)e^{K^\oplus_k}+e^{-K^\oplus_k}\left(h^\oplus_k+A^\oplus_k-B^\oplus_k\right)e^{-K^\oplus_k}-2h^\oplus_k\right)\\ &= H'_{\text{kin}}+\sum_{k\in S_C}Q^k_1\begin{pmatrix}E_k-h_k & 0\\ 0 & E_k-h_k\end{pmatrix}+\text{error terms}, \end{aligned} \tag{5.29}$$

and so we have succeeded in diagonalizing $H'_{\text{kin}}+\sum_{k\in S_C}H^k_{\text{int}}$ while simultanously decoupling the spaces $\ell^2(L_{\pm k})\subset\ell^2(L^\pm_k)$ in a symmetric fashion. We still need to determine the exact form of the error terms, which we record in the following proposition:

Proposition 5.7. *Let $S_C=\overline{B}\left(0,k_F^\gamma\right)\cap\mathbb{Z}^3_+$ with $\gamma\in(0,1]$. Then the unitary transformation $e^{\mathcal{K}}:\mathcal{H}_N\to\mathcal{H}_N$ with $\mathcal{K}$ defined by (5.2), (5.22), (5.23) satisfies*

$$\begin{aligned} e^{\mathcal{K}}\left(H'_{\text{kin}}+\sum_{k\in S_C}H^k_{\text{int}}\right)e^{-\mathcal{K}} &= H'_{\text{kin}}+\sum_{k\in S_C}Q^k_1(E^\oplus_k-h^\oplus_k)\\ &\quad+\sum_{k\in S_C}\int_0^1 e^{(1-t)\mathcal{K}}\left(\mathcal{E}^k_1(A^\oplus_k(t))+\mathcal{E}^k_2(B^\oplus_k(t))\right)e^{-(1-t)\mathcal{K}}\,dt, \end{aligned}$$

where $\mathcal{E}_1(\cdot)$, $\mathcal{E}_2(\cdot)$ are defined in Proposition 5.2 and

$$E^\oplus_k-h^\oplus_k=\begin{pmatrix}E_k-h_k & 0\\ 0 & E_k-h_k\end{pmatrix},\quad A^\oplus_k(t)=\begin{pmatrix}A_k(t) & 0\\ 0 & A_k(t)\end{pmatrix},\quad B^\oplus_k(t)=\begin{pmatrix}0 & B_k(t)\\ B_k(t) & 0\end{pmatrix}$$

with $E_k=e^{-K_k}h_ke^{-K_k}$ and the operators $A_k(t),B_k(t):\ell^2(L_k)\to\ell^2(L_k)$ defined by

$$\begin{aligned} A_k(t) &= \frac{1}{2}\left(e^{tK_k}\left(h_k+2P_{v_k}\right)e^{tK_k}+e^{-tK_k}h_ke^{-tK_k}\right)-h_k\\ B_k(t) &= \frac{1}{2}\left(e^{tK_k}\left(h_k+2P_{v_k}\right)e^{tK_k}-e^{-tK_k}h_ke^{-tK_k}\right). \end{aligned}$$

Proof. By the Propositions 5.3 and 5.5, the error terms are

$$\begin{aligned} &\sum_{k\in S_C}\int_0^1 e^{(1-t)\mathcal{K}}\left(\mathcal{E}^k_1\left(\cosh\left(\mathcal{A}_{tK^\oplus_k}\right)\left(h^\oplus_k+A^\oplus_k\right)+\sinh\left(\mathcal{A}_{tK^\oplus_k}\right)\left(B^\oplus_k\right)-h^\oplus_k\right)\right)e^{-(1-t)\mathcal{K}}\,dt\\ &+\sum_{k\in S_C}\int_0^1 e^{(1-t)\mathcal{K}}\left(\mathcal{E}^k_2\left(\sinh\left(\mathcal{A}_{tK^\oplus_k}\right)\left(h^\oplus_k+A^\oplus_k\right)+\cosh\left(\mathcal{A}_{tK^\oplus_k}\right)\left(B^\oplus_k\right)\right)\right)e^{-(1-t)\mathcal{K}}\,dt, \end{aligned}$$

where we have reparametrized the integral by $t\mapsto 1-t$ to simplify the arguments of the $\mathcal{E}^k_1(\cdot)$ and $\mathcal{E}^k_2(\cdot)$ operators. By (5.11), the arguments of $\mathcal{E}^k_1$ and $\mathcal{E}^k_2$ in each term above equal

$$\frac{1}{2}\left(e^{tK^\oplus_k}\left(h^\oplus_k+A^\oplus_k+B^\oplus_k\right)e^{tK^\oplus_k}+e^{-tK^\oplus_k}\left(h^\oplus_k+A^\oplus_k-B^\oplus_k\right)e^{-tK^\oplus_k}\right)-h^\oplus_k \tag{5.30}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and

$$\frac{1}{2}\left(e^{tK_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}+B_k^{\oplus}\right)e^{tK_k^{\oplus}}-e^{-tK_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}-B_k^{\oplus}\right)e^{-tK_k^{\oplus}}\right), \tag{5.31}$$

respectively. By the same identities that we used in the preceding proposition, it holds that

$$\begin{aligned} &e^{\pm tK_k^{\oplus}}\left(h_k^{\oplus}+A_k^{\oplus}\pm B_k^{\oplus}\right)e^{\pm tK_k^{\oplus}} \\ &=\frac{1}{2}\begin{pmatrix} e^{tK_k}\left(h_k+2P_{v_k}\right)e^{tK_k}+e^{-tK_k}h_ke^{-tK_k} & \pm\left(e^{tK_k}\left(h_k+2P_{v_k}\right)e^{tK_k}-e^{-tK_k}h_ke^{-tK_k}\right) \\ \pm\left(e^{tK_k}\left(h_k+2P_{v_k}\right)e^{tK_k}-e^{-tK_k}h_ke^{-tK_k}\right) & e^{tK_k}\left(h_k+2P_{v_k}\right)e^{tK_k}+e^{-tK_k}h_ke^{-tK_k}\end{pmatrix}, \end{aligned} \tag{5.32}$$

and the claim follows. □

6. Analysis of the exchange terms

In the preceding section, we accomplished a major *qualitative* goal of this paper, which was diagonalizing the bosonizable terms $H'_{\mathrm{kin}}+\sum_{k\in S_C}H^k_{\mathrm{int}}$ in an explicit, quasi-bosonic fashion. In this section, we begin the *quantitative* study of the quasi-bosonic expression in Proposition 5.7.

The aim of this section is to estimate the $\mathcal{E}_1^k(\cdot)$, $\mathcal{E}_2^k(\cdot)$ operators, which enter in the error terms due to the presence of the exchange correction $\varepsilon_{k,l}(\varphi;\psi)$ in the quasi-bosonic commutation relations. We will therefore refer to them as *exchange terms*. Since these expressions are complicated, we thus devote three subsections to the analysis of them. In the first, we carry out a reduction procedure, in which we systematically consider the type of terms that can appear in the sums defining $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$ for given A, B, and reduce these to simpler expressions, or *schematic forms*. In doing so, we will see that every term appearing in $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$ can for the purpose of estimation be sorted into one of four schematic forms. In the second subsection, we provide some basic commutator estimates associated with the four schematic forms, and in the final subsection we then carry out the quantitative analysis of these four forms to obtain the desired estimates of $\mathcal{E}_1^k(\cdot)$ and $\mathcal{E}_2^k(\cdot)$.

6.1. Reduction to simpler expressions

Recall that for $k\in S_C$ and symmetric operators $A, B:\ell^2(L_k^{\pm})\to\ell^2(L_k^{\pm})$, we already defined $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$ in Proposition 5.2. Since these expressions are complicated, it is helpful to discuss the general structure of $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$. Consider the first term of $\mathcal{E}_1^k(A)$, which upon expansion is

$$\begin{aligned} &\left\{b_k^*(Ae_p),\left\{b_l^*\left(K_l^{\oplus}e_q\right),\varepsilon_{k,l}\left(e_p,e_q\right)\right\}\right\} \\ &\quad=b_k^*(Ae_p)\left\{b_l^*\left(K_l^{\oplus}e_q\right),\varepsilon_{k,l}\left(e_p,e_q\right)\right\}+\left\{b_l^*\left(K_l^{\oplus}e_q\right),\varepsilon_{k,l}\left(e_p,e_q\right)\right\}b_k^*(Ae_p) \\ &\quad=b_k^*(Ae_p)b_l^*\left(K_l^{\oplus}e_q\right)\varepsilon_{k,l}\left(e_p,e_q\right)+b_k^*(Ae_p)\varepsilon_{k,l}\left(e_p,e_q\right)b_l^*\left(K_l^{\oplus}e_q\right) \\ &\qquad+b_l^*\left(K_l^{\oplus}e_q\right)\varepsilon_{k,l}\left(e_p,e_q\right)b_k^*(Ae_p)+\varepsilon_{k,l}\left(e_p,e_q\right)b_l^*\left(K_l^{\oplus}e_q\right)b_k^*(Ae_p), \end{aligned} \tag{6.1}$$

which we may expand further using

$$\varepsilon_{k,l}\left(e_p,e_q\right)=\varepsilon\left(\overline{k,p};\overline{l,q}\right)=-\left(\delta_{p,q}c_{\overline{q-l}}c^*_{\overline{p-k}}+\delta_{\overline{p-k},\overline{q-l}}c_q^*c_p\right) \tag{6.2}$$

and then removing the delta on a case-by-case basis. This causes the sums over $p\in L_k^{\pm}$ and $q\in L_l^{\pm}$ of any of these terms to reduce to one of the schematic forms

$$\begin{aligned} &\sum_{p\in S}b_k^{\natural}\left(Te_{p_1}\right)b_l^{\natural}\left(K_l^{\oplus}e_{p_2}\right)\tilde{c}^*_{p_3}\tilde{c}_{p_4}, \quad \sum_{p\in S}b_k^{\natural}\left(Te_{p_1}\right)\tilde{c}^*_{p_3}\tilde{c}_{p_4}b_l^{\natural}\left(K_l^{\oplus}e_{p_2}\right), \\ &\quad\sum_{p\in S}\tilde{c}^*_{p_3}\tilde{c}_{p_4}b_k^{\natural}\left(Te_{p_1}\right)b_l^{\natural}\left(K_l^{\oplus}e_{p_2}\right) \end{aligned} \tag{6.3}$$

subject to the following: S is a subset of $L_k^{\pm} \cap L_l^{\pm}$, $b_k^{\natural}$ can denote either b_k or b_k^*, $\varepsilon_{k,l}(e_p, e_q)$ may instead be $\varepsilon_{l,k}(e_q, e_p) = \varepsilon_{k,l}(e_p, e_q)^*$, T denotes either A or B, the terms $b_k^{\natural}(Te_p)$ and $b_l^{\natural}(K_l^{\oplus} e_q)$ may be interchanged, the notation

$$\tilde{c}_p = \begin{cases} c_p & p \in B_F^c \\ c_p^* & p \in B_F \end{cases} \tag{6.4}$$

encodes the correct type of creation/annihilation operator depending on whether p corresponds to a hole state or an excited state, and p_1, p_2, p_3, p_4 denote indices which depend on p.

The same decomposition holds for every term appearing in either $\mathcal{E}_1^k(A)$ or $\mathcal{E}_2^k(B)$, so we must consider the forms of (6.3).

The only important feature of the dependency that the p_i have with respect to p is that regardless of the term, when summing over $p \in S$, p_i ranges either exclusively over excited states (i.e., $p_i \in L_k^{\pm}$) or exclusively over hole states (i.e., $p_i \in (L_k - k) \cup (L_{-k} + k)$ or the analogous set for $L_l^{\pm}$), and that the assignments $p \mapsto p_i$ (for a given term) are injective. (Additionally, p_1 and p_2 will always be excited states.)

Therefore, when estimating, we can always expand the sum to either all of B_F or all of B_F^c, which is why the exact identities of S and the p_i are of no importance to the estimation. For example,

$$\left| \sum_{p \in S} \left\langle \Psi, \tilde{c}_{p_3}^* \tilde{c}_{p_4} \Psi \right\rangle \right| \le \sum_{p \in S} \left\| \tilde{c}_{p_3} \Psi \right\| \left\| \tilde{c}_{p_4} \Psi \right\| \le \sqrt{\sum_{p \in S} \left\| \tilde{c}_{p_3} \Psi \right\|^2} \sqrt{\sum_{p \in S} \left\| \tilde{c}_{p_4} \Psi \right\|^2} \le \langle \Psi, \mathcal{N}_E \Psi \rangle \tag{6.5}$$

independently of S, p_3 and p_4. Here, the two situations when both p_3 and p_4 range over excited states, and when both p_3 and p_4 range over hole states, can be treated similarly thanks to the particle-hole symmetry (1.13).

Discussion of estimation strategy

We conclude that both $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$ reduce to sums over $l \in S_C$ of finitely many terms of the schematic forms of equation (6.3), so it suffices to estimate these. To this end, we must first perform some additional algebraic manipulation.

To motivate our goal, let us first derive a simple but insufficent estimate for one of these terms:

$$\sum_{p \in S} b_k^* \left(T e_{p_1}\right) \tilde{c}_{p_3}^* \tilde{c}_{p_4} b_l \left(K_l^{\oplus} e_{p_2}\right). \tag{6.6}$$

Using $\left\| c_p \right\|_{\mathrm{Op}} = 1$, Proposition 4.4 and the Cauchy–Schwarz inequality, we find that

$$\begin{aligned} \sum_{p \in S} \left| \left\langle \Psi, b_k^* \left(T e_{p_1}\right) \tilde{c}_{p_3}^* \tilde{c}_{p_4} b_l \left(K_l^{\oplus} e_{p_2}\right) \Psi \right\rangle \right| &\le \sum_{p \in S} \left\| b_k \left(T e_{p_1}\right) \Psi \right\| \left\| b_l \left(K_l^{\oplus} e_{p_2}\right) \Psi \right\| \\ \le \sum_{p \in S} \left\| \left(h_k^{\oplus}\right)^{-\frac{1}{2}} T e_{p_1} \right\| \left\| \left(h_l^{\oplus}\right)^{-\frac{1}{2}} K_l^{\oplus} e_{p_2} \right\| \left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle &\le \left\| \left(h_k^{\oplus}\right)^{-\frac{1}{2}} T \right\|_{\mathrm{HS}} \left\| \left(h_l^{\oplus}\right)^{-\frac{1}{2}} K_l^{\oplus} \right\|_{\mathrm{HS}} \left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle \end{aligned} \tag{6.7}$$

for any $\Psi \in \mathcal{H}_N$. To get a feeling for the quality of this estimate, we must know what to expect of the quantities on the right-hand side. We will see in the next sections that $\|(h_l^{\oplus})^{-\frac{1}{2}} K_l^{\oplus}\|_{\mathrm{HS}} \le O(k_F^{-\frac{1}{3}+\epsilon})$. In general, what will take the place of T will be the $A_k(t)$ and $B_k(t)$ operators we defined in the last section, but as a simple example we consider

$$T = \begin{pmatrix} P_{v_k} & 0 \\ 0 & P_{v_k} \end{pmatrix}, \quad P_{v_k} = |v_k\rangle\langle v_k|, \quad v_k = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}} \sum_{p \in L_k} e_p \in \ell^2(L_k) \tag{6.8}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

for which

$$\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}T\right\|_{\mathrm{HS}} = 2\left\|h_k^{-\frac{1}{2}}P_{v_k}\right\|_{\mathrm{HS}} = 2\sqrt{\mathrm{tr}\left(P_{v_k}h_k^{-1}P_{v_k}\right)} \tag{6.9}$$
$$= 2\|v_k\|\sqrt{\left\langle v_k, h_k^{-1}v_k\right\rangle} = \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|^{\frac{1}{2}}\sqrt{\sum_{p\in L_k}\lambda_{k,p}^{-1}} \leq O\left(k_F^{\frac{1}{2}}\right)$$

when $|k| \sim 1$. Here, we used $|L_k| \leq C|k|k_F^2$ and the bound $\sum_{p\in L_k}\lambda_{k,p}^{-1} \leq Ck_F$ from Proposition A.2. Thus, for any state satisfying $\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle \leq O(k_F)$ (c.f. Theorem 1.2), the overall estimate for the right side of (6.7) is $O(k_F^{\frac{7}{6}+\epsilon})$ which is insufficient as the correlation energy is of order k_F.

The technical issue with the estimation in (6.7) lies in only using that $\|c_p\|_{\mathrm{Op}} = 1$, for we may get better bounds by using $\left\langle\Psi, \mathcal{N}_E H'_{\mathrm{kin}}\Psi\right\rangle$ instead of $\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle$. For example,

$$\sum_{p\in S}\left|\left\langle\Psi, b_k^*\left(Te_{p_1}\right)\tilde{c}_{p_3}^*\tilde{c}_{p_4}b_l\left(K_l^{\oplus}e_{p_2}\right)\Psi\right\rangle\right| \tag{6.10}$$

$$= \sum_{p\in S}\left|\left\langle\Psi, \tilde{c}_{p_3}^*b_k^*\left(Te_{p_1}\right)b_l\left(K_l^{\oplus}e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\rangle\right| \leq \sum_{p\in S}\left\|b_k\left(Te_{p_1}\right)\tilde{c}_{p_3}\Psi\right\|\left\|b_l\left(K_l^{\oplus}e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\|$$
$$\leq \sum_{p\in S}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Te_{p_1}\right\|\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}}K_l^{\oplus}e_{p_2}\right\|\sqrt{\left\langle\tilde{c}_{p_3}\Psi, H_{\mathrm{kin}}^{\prime(\pm1)}\tilde{c}_{p_3}\Psi\right\rangle\left\langle\tilde{c}_{p_4}\Psi, H_{\mathrm{kin}}^{\prime(\pm1)}\tilde{c}_{p_4}\Psi\right\rangle} \tag{6.11}$$
$$\leq \left(\max_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Te_p\right\|\right)\sqrt{\sum_{p\in S}\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}}K_l^{\oplus}e_q\right\|^2}\sqrt{\sum_{p\in S}\left\langle\Psi, \tilde{c}_{p_3}^*H_{\mathrm{kin}}^{\prime(\pm1)}\tilde{c}_{p_3}\Psi\right\rangle}\sqrt{\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle}$$
$$\leq \left(\max_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Te_p\right\|\right)\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}}K_l^{\oplus}\right\|_{\mathrm{HS}}\sqrt{\left\langle\Psi, \mathcal{N}_E H'_{\mathrm{kin}}\Psi\right\rangle\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle},$$

where we used that $\left[\tilde{c}_p, b_k(\cdot)\right] = 0$ (as we will see in Proposition 6.1 below) and momentarily looked ahead to the definition (6.30) for $H_{\mathrm{kin}}^{\prime(\pm1)}$ and Lemma 6.6 (we take supremum over p_4 and sum over p_3 to get the second inequality). Considering again the example in (6.8), we find

$$\max_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}Te_p\right\| = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}}\left\|h_k^{-\frac{1}{2}}v_k\right\| = \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}\sqrt{\sum_{k\in L_k}\lambda_{k,p}^{-1}} \leq O(k_F^{-\frac{1}{2}}). \tag{6.12}$$

Thus, for any state satisfying $\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle \leq O(k_F)$ and $\left\langle\Psi, \mathcal{N}_E H'_{\mathrm{kin}}\Psi\right\rangle \leq O(k_F^2)$ (c.f. Theorem 1.2), the right side of (6.10) is thus bounded by $O(k_F^{\frac{2}{3}+\epsilon})$, which is much smaller than the correlation energy.

Our goal is, therefore, to reduce the schematic forms of equation (6.3) to those of the form $\sum_{p\in S}\tilde{c}_{p_3}^*b_k^{\natural}(Te_{p_1})b_l^{\natural}(K_l^{\oplus}e_{p_2})\tilde{c}_{p_4}$, which we may then estimate as above. While $\left[\tilde{c}_p, b_k(\cdot)\right] = 0$, it is generally the case that $\left[\tilde{c}_p, b_k^*(\cdot)\right] \neq 0$, so this will also introduce additional commutator terms which we must then estimate separately.

Taking into account whether $b_k^{\natural} = b_k$ or $b_k^{\natural} = b_k^*$, the schematic forms of equation (6.3) are either of the form (supressing the summation, the arguments and the subscripts for brevity)

$$b^*b^*\tilde{c}^*\tilde{c}, \quad b^*b\tilde{c}^*\tilde{c}, \quad bb^*\tilde{c}^*\tilde{c}, \quad bb\tilde{c}^*\tilde{c}, \quad b^*\tilde{c}^*\tilde{c}b, \quad b\tilde{c}^*\tilde{c}b, \quad b\tilde{c}^*\tilde{c}b^*, \tag{6.13}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

or reduce to one of these by taking the adjoint, which as we will estimate, $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$ as bilinear forms does not matter. Using that commutators of the form $[b,\tilde{c}]$, $[b^*,\tilde{c}^*]$ and $[b,[b,\tilde{c}^*]]$ vanish (verified below), these schematic forms reduce to

$$\begin{aligned} b^*b^*\tilde{c}^*\tilde{c} &= \tilde{c}^*b^*b^*\tilde{c}, \\ b^*b\tilde{c}^*\tilde{c} &= \tilde{c}^*b^*b\tilde{c} + [b,\tilde{c}^*]\,b^*\tilde{c} + [b^*,[b,\tilde{c}^*]]\,\tilde{c}, \\ bb^*\tilde{c}^*\tilde{c} &= \tilde{c}^*bb^*\tilde{c} + [b,\tilde{c}^*]\,b^*\tilde{c}, \\ bb\tilde{c}^*\tilde{c} &= \tilde{c}^*bb\tilde{c} + [b,\tilde{c}^*]\,b\tilde{c} + [b,\tilde{c}^*]\,b\tilde{c}, \\ b^*\tilde{c}^*\tilde{c}b &= \tilde{c}^*b^*b\tilde{c}, \\ b\tilde{c}^*\tilde{c}b &= \tilde{c}^*bb\tilde{c} + [b,\tilde{c}^*]\,b\tilde{c}, \\ b\tilde{c}^*\tilde{c}b^* &= \tilde{c}^*bb^*\tilde{c} + [b,\tilde{c}^*]\,b^*\tilde{c} + \tilde{c}^*b\,[\tilde{c},b^*] + [b,\tilde{c}^*]\,[\tilde{c},b^*]\,. \end{aligned} \tag{6.14}$$

Reintroducing the $b^\natural$ notation outside the commutators and using once more our freedom to take adjoints, we find that every term on the right-hand sides of the two equations above takes one of the four schematic forms

$$\tilde{c}^*b^\natural b^\natural\tilde{c}, \quad [\tilde{c},b^*]^*\,b^\natural\tilde{c}, \quad [[\tilde{c},b^*]\,,b]^*\,\tilde{c}, \quad [\tilde{c},b^*]^*\,[\tilde{c},b^*]\,. \tag{6.15}$$

These are the final forms which we will explicitly estimate.

6.2. Preliminary commutator estimates

In addition to the general estimates which we derived at the start of this section, we will also need estimates on the commutator terms which appear in the schematic forms of equation (6.23), which we now derive. First, we must, however, verify that the commutators $[b,\tilde{c}]$, $[b^*,\tilde{c}^*]$ and $[b,[b,\tilde{c}^*]]$ vanish, which we relied upon in our reduction procedure:

Proposition 6.1. *For all $k,l\in\mathbb{Z}^3_+$, $\varphi\in\ell^2(L_k^\pm)$, $\psi\in\ell^2\left(L_l^\pm\right)$ and $p\in\mathbb{Z}^3$ it holds that*

$$\left[b_k(\varphi),\tilde{c}_p\right] = \left[b_k^*(\varphi),\tilde{c}_p^*\right] = 0, \quad \left[b_l\left(\psi\right),\left[b_k(\varphi),\tilde{c}_p^*\right]\right] = 0.$$

Proof. We compute from the definitions that for any $q\in L_k^\pm$,

$$\begin{aligned} \left[b_{\overline{k,q}},\tilde{c}_p\right] &= \left[c^*_{\overline{q-k}}c_q,\tilde{c}_p\right] = c^*_{\overline{q-k}}\left\{c_q,\tilde{c}_p\right\} - \left\{c^*_{\overline{q-k}},\tilde{c}_p\right\}c_q \\ &= \begin{cases} c^*_{\overline{q-k}}\left\{c_q,c_p\right\} - \left\{c^*_{\overline{q-k}},c_p\right\}c_q, & p\in B_F^c \\ c^*_{\overline{q-k}}\left\{c_q,c_p^*\right\} - \left\{c^*_{\overline{q-k}},c_p^*\right\}c_q, & p\in B_F \end{cases} \\ &= 0 \end{aligned} \tag{6.16}$$

as all anticommutators on the second line vanish either directly by the CAR or by disjointness of B_F and B_F^c. By linearity, $\left[b_k(\varphi),\tilde{c}_p\right]=0$, and $\left[b_k^*(\varphi),\tilde{c}_p^*\right] = -\left[b_k(\varphi),\tilde{c}_p\right]^* = 0$.

For the double commutator, we first compute $[b_{\overline{k,q}},\tilde{c}_p^*]$. As above, we find

$$\left[b_{\overline{k,q}},\tilde{c}_p^*\right] = \begin{cases} c^*_{\overline{q-k}}\left\{c_q,c_p^*\right\} - \left\{c^*_{\overline{q-k}},c_p^*\right\}c_q, & p\in B_F^c \\ c^*_{\overline{q-k}}\left\{c_q,c_p\right\} - \left\{c^*_{\overline{q-k}},c_p\right\}c_q, & p\in B_F \end{cases} = \begin{cases} \delta_{q,p}\tilde{c}_{\overline{q-k}}, & p\in B_F^c \\ -\delta_{\overline{q-k},p}\tilde{c}_q, & p\in B_F \end{cases}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

so

$$\left[b_k(\varphi), \tilde{c}_p^*\right] = \sum_{q\in L_k^\pm} \langle \varphi, e_q\rangle \left[b_{\overline{k,q}}, \tilde{c}_p^*\right] = \sum_{q\in L_k^\pm} \langle \varphi, e_q\rangle \begin{cases} \delta_{q,p}\tilde{c}_{\overline{q-k}} & p\in B_F^c \\ -\delta_{\overline{q-k},p}\tilde{c}_q & p\in B_F \end{cases}$$
$$= 1_{L_k^\pm}(p)\langle \varphi, e_p\rangle \tilde{c}_{\overline{p-k}} - 1_{L_k-k}(p)\langle \varphi, e_{p+k}\rangle \tilde{c}_{p+k} - 1_{L_k+k}(p)\langle \varphi, e_{p-k}\rangle \tilde{c}_{p-k}, \quad (6.17)$$

where $1_S(\cdot)$ denotes the indicator function of a set S. Observing that $\left[b_k(\varphi), \tilde{c}_p^*\right]$ is a linear combination of $\tilde{c}_p$ terms, we conclude that $\left[b_l(\psi), \left[b_k(\varphi), \tilde{c}_p^*\right]\right] = 0$ by the first part. □

We now move into the estimation of the nonvanishing commutators. We begin with the single commutator; we state the estimate and make a remark:

Proposition 6.2. *For all $k\in\mathbb{Z}_+^3$, sequences $(\varphi_p)_{p\in\mathbb{Z}^3}\in\ell^2(L_k^\pm)$ and $\Psi\in\mathcal{H}_N$, it holds that*

$$\sum_{p\in\mathbb{Z}^3} \left\|\left[\tilde{c}_p, b_k^*(\varphi_p)\right]\Psi\right\|^2 \le 3\left(\sum_{p\in L_k^\pm}\max_{q\in\mathbb{Z}^3}\left|\langle e_p,\varphi_q\rangle\right|^2\right)\|\Psi\|^2$$
$$\sum_{p\in\mathbb{Z}^3} \left\|\left[\tilde{c}_p^*, b_k(\varphi_p)\right]\Psi\right\|^2 \le 4\left(\max_{p\in L_k^\pm, q\in\mathbb{Z}^3}\left|\langle e_p,\varphi_q\rangle\right|^2\right)\langle\Psi,\mathcal{N}_E\Psi\rangle.$$

Remark 6.1. The statement may appear overly general in that it involves general sequences $(\varphi_p)_{p\in\mathbb{Z}^3}\subset\ell^2(L_k^\pm)$ rather than the explicit vectors $(Te_{p_1})_{p\in S}\subset\ell^2(L_k^\pm)$ that we must consider. The point of the generality is, however, only to avoid having to explicitly state the dependencies of the set S and the p_i's of each possible schematic form, as independently of these it is easy to see that a sum such as $\sum_{p\in S}\left\|[\tilde{c}_{p_3}, b_k^*(Te_{p_1})]\Psi\right\|^2$ can always be cast into the form in the statement.

Proof. Taking the adjoint of equation (6.17) yields

$$\left[\tilde{c}_p, b_k^*(\varphi)\right] = 1_{L_k^\pm}(p)\langle e_p,\varphi\rangle \tilde{c}^*_{\overline{p-k}} - 1_{L_k-k}(p)\langle e_{p+k},\varphi\rangle \tilde{c}^*_{p+k} - 1_{L_k+k}(p)\langle e_{p-k},\varphi\rangle \tilde{c}^*_{p-k}, \quad (6.18)$$

and so we can for any $\Psi\in\mathcal{H}_N$ estimate by the (squared) triangle inequality, using also that $L_k^\pm$ and $(L_k-k)\cap(L_{-k}+k)$ are disjoint and $\left\|\tilde{c}_p^*\right\|_{\mathrm{Op}} = 1$, that

$$\sum_{p\in\mathbb{Z}^3}\left\|\left[\tilde{c}_p, b_k^*(\varphi_p)\right]\Psi\right\|^2$$
$$\le \sum_{p\in L_k^\pm}\left|\langle e_p,\varphi_p\rangle\right|^2\left\|\tilde{c}^*_{\overline{p-k}}\Psi\right\|^2 + 2\sum_{p\in L_k-k}\left|\langle e_{p+k},\varphi_p\rangle\right|^2\left\|\tilde{c}^*_{p+k}\Psi\right\|^2 + 2\sum_{p\in L_{-k}+k}\left|\langle e_{p-k},\varphi_p\rangle\right|^2\left\|\tilde{c}^*_{p-k}\Psi\right\|^2$$
$$\le \left(\sum_{p\in L_k^\pm}\left|\langle e_p,\varphi_p\rangle\right|^2 + 2\sum_{p\in L_k-k}\left|\langle e_{p+k},\varphi_p\rangle\right|^2 + 2\sum_{p\in L_{-k}+k}\left|\langle e_{p-k},\varphi_p\rangle\right|^2\right)\|\Psi\|^2$$
$$\le \left(\sum_{p\in L_k^\pm}\max_{q\in\mathbb{Z}^3}\left|\langle e_p,\varphi_q\rangle\right|^2 + 2\sum_{p\in L_k}\max_{q\in\mathbb{Z}^3}\left|\langle e_p,\varphi_q\rangle\right|^2 + 2\sum_{p\in L_{-k}}\max_{q\in\mathbb{Z}^3}\left|\langle e_p,\varphi_q\rangle\right|^2\right)\|\Psi\|^2$$
$$= 3\left(\sum_{p\in L_k^\pm}\max_{q\in\mathbb{Z}^3}\left|\langle e_p,\varphi_q\rangle\right|^2\right)\|\Psi\|^2, \quad (6.19)$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

which implies the first estimate. For the second estimate, we find in a similar manner (now directly from equation (6.17)) that

$$\begin{aligned}\sum_{p\in\mathbb{Z}^3}\left\|\left[\tilde{c}_p^*,b_k\left(\varphi_p\right)\right]\Psi\right\|^2 &\leq \sum_{p\in L_k^\pm}\left|\left\langle e_p,\varphi_p\right\rangle\right|^2\left\|\tilde{c}_{\overline{p-k}}\Psi\right\|^2+2\sum_{p\in L_k-k}\left|\left\langle e_{p+k},\varphi_p\right\rangle\right|^2\left\|\tilde{c}_{p+k}\Psi\right\|^2\\ &\quad+2\sum_{p\in L_{-k}+k}\left|\left\langle e_{p-k},\varphi_p\right\rangle\right|^2\left\|\tilde{c}_{p-k}\Psi\right\|^2\\ &\leq\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\varphi_q\right\rangle\right|^2\right)\left(\sum_{p\in L_k^\pm}\left\|\tilde{c}_{\overline{p-k}}\Psi\right\|^2+2\sum_{p\in L_k-k}\left\|\tilde{c}_{p+k}\Psi\right\|^2+2\sum_{p\in L_{-k}+k}\left\|\tilde{c}_{p-k}\Psi\right\|^2\right)\\ &=\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\varphi_q\right\rangle\right|^2\right)\left(\sum_{p\in L_k^\pm}\left\|c^*_{\overline{p-k}}\Psi\right\|^2+2\sum_{p\in L_k}\left\|c_p\Psi\right\|^2+2\sum_{p\in L_{-k}}\left\|c_p\Psi\right\|^2\right)\\ &\leq 4\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\varphi_q\right\rangle\right|^2\right)\langle\Psi,\mathcal{N}_E\Psi\rangle\,. \end{aligned}\tag{6.20}$$

□

Lastly, we estimate the double commutator:

Proposition 6.3. *For all $k,l\in\mathbb{Z}^3_+$, sequences $\left(\varphi_p\right)_{p\in\mathbb{Z}^3}\subset\ell^2(L_k^\pm)$ and $\left(\psi_p\right)_{p\in\mathbb{Z}^3}\subset\ell^2\left(L_l^\pm\right)$, and $\Psi\in\mathcal{H}_N$, it holds that*

$$\sum_{p\in\mathbb{Z}^3}\left\|\left[\left[\tilde{c}_p,b_k^*\left(\varphi_p\right)\right],b_l\left(\psi_p\right)\right]\Psi\right\|^2\leq 12\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\varphi_q\right\rangle\right|^2\right)\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\psi_q\right\rangle\right|^2\right)\langle\Psi,\mathcal{N}_E\Psi\rangle\,.$$

Proof. From (6.18), we have that

$$\begin{aligned}\left[\left[\tilde{c}_p,b_k^*\left(\varphi_p\right)\right],b_l\left(\psi_p\right)\right]&=1_{L_k^\pm}(p)\left\langle e_p,\varphi\right\rangle\left[\tilde{c}^*_{\overline{p-k}},b_l\left(\psi_p\right)\right]-1_{L_k-k}(p)\left\langle e_{p+k},\varphi\right\rangle\left[\tilde{c}^*_{p+k},b_l\left(\psi_p\right)\right]\\ &\quad-1_{L_k+k}(p)\left\langle e_{p-k},\varphi\right\rangle\left[\tilde{c}^*_{p-k},b_l\left(\psi_p\right)\right],\end{aligned}\tag{6.21}$$

and so, by the triangle inequality and the second estimate of Proposition 6.2,

$$\begin{aligned}\sum_{p\in\mathbb{Z}^n}\left\|\left[\left[\tilde{c}_p,b_k^*\left(\varphi_p\right)\right],b_l\left(\psi_p\right)\right]\Psi\right\|^2&\leq\sum_{p\in L_k^\pm}\left|\left\langle e_p,\varphi_p\right\rangle\right|^2\left\|\left[\tilde{c}^*_{\overline{p-k}},b_l\left(\psi_p\right)\right]\Psi\right\|^2\\ &\quad+2\sum_{p\in L_k-k}\left|\left\langle e_{p+k},\varphi_p\right\rangle\right|^2\left\|\left[\tilde{c}^*_{p+k},b_l\left(\psi_p\right)\right]\Psi\right\|^2+2\sum_{p\in L_{-k}+k}\left|\left\langle e_{p-k},\varphi_p\right\rangle\right|^2\left\|\left[\tilde{c}^*_{p-k},b_l\left(\psi_p\right)\right]\Psi\right\|^2\\ &\leq\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\varphi_q\right\rangle\right|^2\right)\left(\sum_{p\in L_k^\pm}\left\|\left[\tilde{c}^*_{\overline{p-k}},b_l\left(\psi_p\right)\right]\Psi\right\|^2+2\sum_{p\in L_k}\left\|\left[\tilde{c}^*_p,b_l\left(\psi_{p-k}\right)\right]\Psi\right\|^2\right.\\ &\qquad\left.+2\sum_{p\in L_{-k}}\left\|\left[\tilde{c}^*_p,b_l\left(\psi_{p+k}\right)\right]\Psi\right\|^2\right)\\ &\leq 12\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\varphi_q\right\rangle\right|^2\right)\left(\max_{p\in L_k^\pm,q\in\mathbb{Z}^3}\left|\left\langle e_p,\psi_q\right\rangle\right|^2\right)\langle\Psi,\mathcal{N}_E\Psi\rangle\,.\end{aligned}\tag{6.22}$$

□

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

6.3. Final estimation of the exchange terms

Now we are ready to derive bounds for the exchange terms $\mathcal{E}_1^k(A)$ and $\mathcal{E}_2^k(B)$ defined in Proposition 5.2. Recall that we have reduced the estimation of these complicated operators to the task of obtaining a uniform estimate for the four explicit forms

$$\sum_{l\in S_C}\sum_{p\in S}\tilde{c}_{p_3}^*b_k^\natural\left(Te_{p_1}\right)b_l^\natural\left(K_l^\oplus e_{p_2}\right)\tilde{c}_{p_4},\qquad \sum_{l\in S_C}\sum_{p\in S}\left[\tilde{c}_{p_3},b_k^*\left(Te_{p_1}\right)\right]^*b_l^\natural\left(K_l^\oplus e_{p_2}\right)\tilde{c}_{p_4}$$
$$\sum_{l\in S_C}\sum_{p\in S}\left[\left[\tilde{c}_{p_3},b_k^*\left(Te_{p_1}\right)\right],b_l\left(K_l^\oplus e_{p_2}\right)\right]^*\tilde{c}_{p_4},\quad \sum_{l\in S_C}\sum_{p\in S}\left[\tilde{c}_{p_3},b_k^*\left(Te_{p_1}\right)\right]^*\left[\tilde{c}_{p_4},b_l^*\left(K_l^\oplus e_{p_2}\right)\right], \tag{6.23}$$

subject to the following rules: $b_k^\natural$ denotes either b_k or b_k^*, T denotes either A or B, and $b_k^\natural\left(Te_{p_1}\right)$ and $b_l^\natural(K_l^\oplus e_{p_2})$ may be interchanged. Furthermore, the notation $\tilde{c}_p$ denotes either c_p or c_p^* as appropriate for p, and the set S is such that the assignments $p\mapsto p_1,p_2,p_3,p_4$ are injective and map exclusively into B_F or B_F^c.

Let us start by giving estimates in terms of $\mathcal{N}_E^2$. For the statement, we define the $\|\cdot\|_{\infty,2}$-norm of an operator $T:\ell^2(L_k^\pm)\to\ell^2(L_k^\pm)$ by

$$\|T\|_{\infty,2}=\sqrt{\sum_{p\in L_k^\pm}\max_{q\in L_k^\pm}\left|\left\langle e_p,Te_q\right\rangle\right|^2}. \tag{6.24}$$

This is a minor but necessary detail, as unlike the simple estimate of equation (6.10), we cannot take the maximum outside the sum for all schematic terms, so we need this slightly stronger norm. Note that

$$\max_{p,q\in L_k^\pm}\left|\left\langle e_p,Te_q\right\rangle\right|\leq\max_{p\in L_k^\pm}\left\|Te_p\right\|\leq\|T\|_{\infty,2}\,. \tag{6.25}$$

Now the estimate the follows.

Proposition 6.4. *For all $k\in\mathbb{Z}_+^3$, symmetric $T:\ell^2(L_k^\pm)\to\ell^2(L_k^\pm)$ and $\Psi\in\mathcal{H}_N$, it holds that*

$$\left|\left\langle\Psi,\mathcal{E}_i^k(T)\Psi\right\rangle\right|\leq C\,\|T\|_{\infty,2}\left(\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\right)\left\langle\Psi,\left(1+\mathcal{N}_E^2\right)\Psi\right\rangle$$

with $i=1,2$, for a constant $C>0$ independent of all relevant quantities.

Proof. We estimate each schematic form of (6.23) using the estimates of the Propositions 4.2, 6.2, 6.3 and Lemma 4.3, as well as the Cauchy-Schwarz inequality. First is $\tilde{c}^*b^\natural b^\natural\tilde{c}$:

$$\begin{aligned}
&\sum_{l\in S_C}\sum_{p\in S}\left|\left\langle\Psi,\tilde{c}_{p_3}^*b_k^\natural\left(Te_{p_1}\right)b_l^\natural\left(K_l^\oplus e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\rangle\right|\leq\sum_{l\in S_C}\sum_{p\in S}\left\|b_k^\sharp\left(Te_{p_1}\right)^*\tilde{c}_{p_3}\Psi\right\|\left\|b_l^\natural\left(K_l^\oplus e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\|\\
&\leq C\sum_{l\in S_C}\sum_{p\in S}\left\|Te_{p_1}\right\|\left\|K_l^\oplus e_{p_2}\right\|\sqrt{\left\langle\tilde{c}_{p_3}\Psi,\left(1+\mathcal{N}_E^{(\pm1)}\right)\tilde{c}_{p_3}\Psi\right\rangle\left\langle\tilde{c}_{p_4}\Psi,\left(1+\mathcal{N}_E^{(\pm1)}\right)\tilde{c}_{p_4}\Psi\right\rangle}\\
&\leq C\max_{p\in L_k^\pm}\left\|Te_p\right\|\sum_{l\in S_C}\max_{q\in L_l^\pm}\left\|K_l^\oplus e_q\right\|\sqrt{\sum_{p\in S}\left\langle\Psi,\left(\tilde{c}_{p_3}^*\mathcal{N}_E^{(\pm1)}\tilde{c}_{p_3}+\tilde{c}_{p_3}^*\tilde{c}_{p_3}\right)\Psi\right\rangle}\\
&\quad\cdot\sqrt{\sum_{p\in S}\left\langle\Psi,\left(\tilde{c}_{p_4}^*\mathcal{N}_E^{(\pm1)}\tilde{c}_{p_4}+\tilde{c}_{p_4}^*\tilde{c}_{p_4}\right)\Psi\right\rangle}\leq C\,\|T\|_{\infty,2}\left(\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\right)\left\langle\Psi,\mathcal{N}_E^2\Psi\right\rangle.
\end{aligned} \tag{6.26}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Then, $[\tilde{c},b^*]^*\, b^\natural \tilde{c}$:

$$\begin{aligned}
&\sum_{l\in S_C}\sum_{p\in S}\left|\left\langle \Psi, \left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]^* b_l^\natural\left(K_l^\oplus e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\rangle\right| \\
&\le \sum_{l\in S_C}\sum_{p\in S}\left\|\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]\Psi\right\|\left\|b_l^\natural\left(K_l^\oplus e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\| \\
&\le C\sum_{l\in S_C}\sum_{p\in S}\left\|\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]\Psi\right\|\left\|K_l^\oplus e_{p_2}\right\|\sqrt{\left\langle \tilde{c}_{p_4}\Psi, \left(1+\mathcal{N}_E^{(\pm1)}\right)\tilde{c}_{p_4}\Psi\right\rangle} \\
&\le C\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\sqrt{\sum_{p\in S}\left\|\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]\Psi\right\|^2}\sqrt{\sum_{p\in S}\left\langle \tilde{c}_{p_4}\Psi, \left(1+\mathcal{N}_E^{(\pm1)}\right)\tilde{c}_{p_4}\Psi\right\rangle} \\
&\le C\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\sqrt{\sum_{p\in L_k^\pm}\max_{q\in L_k^\pm}\left|\left\langle e_p, Te_q\right\rangle\right|^2\|\Psi\|^2}\sqrt{\left\langle \Psi, \mathcal{N}_E^2\Psi\right\rangle} \\
&\le C\|T\|_{\infty,2}\left(\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\right)\|\Psi\|\sqrt{\left\langle \Psi, \mathcal{N}_E^2\Psi\right\rangle}.
\end{aligned}\tag{6.27}$$

Now, $[[\tilde{c},b^*],b]^*\,\tilde{c}$:

$$\begin{aligned}
&\sum_{l\in S_C}\sum_{p\in S}\left|\left\langle \Psi, \left[\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right], b_l\left(K_l^\oplus e_{p_2}\right)\right]^*\tilde{c}_{p_4}\Psi\right\rangle\right| \\
&\le \sum_{l\in S_C}\sum_{p\in S}\left\|\left[\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right], b_l\left(K_l^\oplus e_{p_2}\right)\right]\Psi\right\|\left\|\tilde{c}_{p_4}\Psi\right\| \\
&\le \sum_{l\in S_C}\sqrt{\sum_{p\in S}\left\|\left[\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right], b_l\left(K_l^\oplus e_{p_2}\right)\right]\Psi\right\|^2}\sqrt{\sum_{p\in S}\left\|\tilde{c}_{p_4}\Psi\right\|^2} \\
&\le C\sum_{l\in S_C}\sqrt{\left(\max_{p,q\in L_k^\pm}\left|\left\langle e_p, Te_q\right\rangle\right|^2\right)\left(\max_{p,q\in L_l^\pm}\left|\left\langle e_p, K_l^\oplus e_q\right\rangle\right|^2\right)\left\langle \Psi, \mathcal{N}_E\Psi\right\rangle}\sqrt{\left\langle \Psi, \mathcal{N}_E\Psi\right\rangle} \\
&\le C\|T\|_{\infty,2}\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\left\langle \Psi, \mathcal{N}_E\Psi\right\rangle .
\end{aligned}\tag{6.28}$$

And finally, $[\tilde{c},b^*]^*\,[\tilde{c},b^*]$:

$$\begin{aligned}
&\sum_{l\in S_C}\sum_{p\in S}\left|\left\langle \Psi, \left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]^*\left[\tilde{c}_{p_4}, b_l^*\left(K_l^\oplus e_{p_2}\right)\right]\Psi\right\rangle\right| \\
&\le \sum_{l\in S_C}\sum_{p\in S}\left\|\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]\Psi\right\|\left\|\left[\tilde{c}_{p_4}, b_l^*\left(K_l^\oplus e_{p_2}\right)\right]\Psi\right\| \\
&\le \sum_{l\in S_C}\sqrt{\sum_{p\in S}\left\|\left[\tilde{c}_{p_3}, b_k^*\left(Te_{p_1}\right)\right]\Psi\right\|^2}\sqrt{\sum_{p\in S}\left\|\left[\tilde{c}_{p_4}, b_l^*\left(K_l^\oplus e_{p_2}\right)\right]\Psi\right\|^2} \\
&\le C\sum_{l\in S_C}\sqrt{\sum_{p\in L_k^\pm}\max_{q\in L_k^\pm}\left|\left\langle e_p, Te_q\right\rangle\right|^2\|\Psi\|^2}\sqrt{\sum_{p\in L_l^\pm}\max_{q\in L_l^\pm}\left|\left\langle e_p, K_l^\oplus e_q\right\rangle\right|^2\|\Psi\|^2} \\
&\le C\|T\|_{\infty,2}\sum_{l\in S_C}\left\|K_l^\oplus\right\|_{\infty,2}\|\Psi\|^2 .
\end{aligned}\tag{6.29}$$

□

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Now we derive a kinetic bound.

Proposition 6.5. *For all $k \in \mathbb{Z}^3_+$, symmetric $T : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$ and $\Psi \in D\left(H'_{\mathrm{kin}}\right)$,*

$$\begin{aligned}\left|\left\langle \Psi, \mathcal{E}_i^k(T)\Psi\right\rangle\right| \leq C \sum_{l\in S_C}\left(\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}} K_l^{\oplus}\right\|_{\mathrm{HS}} + \left\|K_l^{\oplus}\right\|_{\infty,2}\right)\\ \times\left[\left(\max_{p\in L_k^\pm}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} Te_p\right\|\right)\sqrt{\left\langle \Psi, H'_{\mathrm{kin}}\Psi\right\rangle\left\langle \Psi, \mathcal{N}_E H'_{\mathrm{kin}}\Psi\right\rangle}\right.\\ \left. + \|T\|_{\infty,2}\left(\left\langle \Psi, \left(1+H'_{\mathrm{kin}}\right)\Psi\right\rangle + \|\Psi\|\sqrt{\left\langle \Psi, \mathcal{N}_E H'_{\mathrm{kin}}\Psi\right\rangle}\right)\right]\end{aligned}$$

for $i = 1, 2$, for a constant $C > 0$ independent of all relevant quantities.

As a technical preparation, let us observe that from (1.14) we may associate to H'_{kin} the operators

$$H'^{(\pm 1)}_{\mathrm{kin}} = \sum_{p\in B_F^c} \left||p|^2 - \zeta\right| c_p^* c_p + \sum_{p\in B_F} \left||p|^2 - \zeta\right| c_p c_p^* \tag{6.30}$$

acting on $\mathcal{H}_{N\pm 1}$ (the expressions of $H'^{(+1)}_{\mathrm{kin}}$ and $H'^{(-1)}_{\mathrm{kin}}$ are the same, but the domains are different). With this interpretation, we have the following lemma (c.f. Lemma 4.3):

Lemma 6.6. *It holds that*

$$\tilde{c}_p^* H'^{(\pm 1)}_{\mathrm{kin}} \tilde{c}_p \leq H'_{\mathrm{kin}}$$

for all $p \in \mathbb{Z}^3$ and

$$\sum_{p\in B_F^c} c_p^* H'^{(-1)}_{\mathrm{kin}} c_p \leq \mathcal{N}_E H'_{\mathrm{kin}}, \qquad \sum_{p\in B_F} c_p H'^{(+1)}_{\mathrm{kin}} c_p^* \leq \mathcal{N}_E H'_{\mathrm{kin}}.$$

Proof. By the CAR, we have that

$$\begin{aligned}&\sum_{p\in B_F^c} c_p^* H'^{(-1)}_{\mathrm{kin}} c_p = \sum_{p\in B_F^c} c_p^*\left(\sum_{q\in B_F^c}\left||q|^2-\zeta_0\right| c_q^* c_q + \sum_{q\in B_F}\left||q|^2-\zeta_0\right| c_q c_q^*\right) c_p\\ &= \left(\sum_{p\in B_F^c} c_p^* c_p\right)\left(\sum_{q\in B_F^c}\left||q|^2-\zeta_0\right| c_q^* c_q + \sum_{q\in B_F}\left||q|^2-\zeta_0\right| c_q c_q^*\right) + \sum_{p\in B_F^c} c_p^*\left[\sum_{q\in B_F^c}\left||q|^2-\zeta_0\right| c_q^* c_q, c_p\right]\\ &= \mathcal{N}_E H'_{\mathrm{kin}} - \sum_{p,q\in B_F^c}\left||q|^2-\zeta_0\right| \delta_{p,q} c_p^* c_q \leq \mathcal{N}_E H'_{\mathrm{kin}},\end{aligned} \tag{6.31}$$

and the inequality for $c_p H'^{(+1)}_{\mathrm{kin}} c_p^*$ can be derived similarly. That $\tilde{c}_p^* H'^{(\pm 1)}_{\mathrm{kin}} \tilde{c}_p \leq H'_{\mathrm{kin}}$ follows exactly as the inequality $\tilde{c}_p^* \mathcal{N}_E^{(\pm 1)} \tilde{c}_p \leq \mathcal{N}_E$ did in Lemma 4.3. □

Now we are ready to give the

Proof of Proposition 6.5. For all schematic forms except

$$\sum_{l\in S_C}\sum_{p\in S} \tilde{c}_{p_3}^* b_k^{\natural}\left(Te_{p_1}\right) b_l^{\natural}\left(K_l^{\oplus} e_{p_2}\right) \tilde{c}_{p_4}, \tag{6.32}$$

we can use the estimates derived in Proposition 6.4, specifically the equations (6.27) through (6.29), and the fact that $\mathcal{N}_E \leq H'_{\text{kin}}$. For the schematic form in (6.32), we can by Proposition 4.4 estimate that

$$\begin{aligned}
&\sum_{l \in S_C} \sum_{p \in S} \left| \left\langle \Psi, \tilde{c}^*_{p_3} b^\natural_k \left(T e_{p_1}\right) b^\natural_l \left(K^\oplus_l e_{p_2}\right) \tilde{c}_{p_4} \Psi \right\rangle \right| \leq \sum_{l \in S_C} \sum_{p \in S} \left\| b^\natural_k \left(T e_{p_1}\right) \tilde{c}_{p_3} \Psi \right\| \left\| b^\natural_l \left(K^\oplus_l e_{p_2}\right) \tilde{c}_{p_4} \Psi \right\| \\
&\leq \sum_{l \in S_C} \sum_{p \in S} \left\| \left(h^\oplus_k\right)^{-\frac{1}{2}} T e_{p_1} \right\| \left\| \left(h^\oplus_l\right)^{-\frac{1}{2}} K^\oplus_l e_{p_2} \right\| \sqrt{\left\langle \tilde{c}_{p_3} \Psi, H'^{(\pm 1)}_{\text{kin}} \tilde{c}_{p_3} \Psi \right\rangle \left\langle \tilde{c}_{p_4} \Psi, H'^{(\pm 1)}_{\text{kin}} \tilde{c}_{p_4} \Psi \right\rangle} \\
&+ \sum_{l \in S_C} \sum_{p \in S} \left\| \left(h^\oplus_k\right)^{-\frac{1}{2}} T e_{p_1} \right\| \left\| K^\oplus_l e_{p_2} \right\| \sqrt{\left\langle \tilde{c}_{p_3} \Psi, H'^{(\pm 1)}_{\text{kin}} \tilde{c}_{p_3} \Psi \right\rangle} \left\| \tilde{c}_{p_4} \Psi \right\| \\
&+ \sum_{l \in S_C} \sum_{p \in S} \left\| T e_{p_1} \right\| \left\| \left(h^\oplus_l\right)^{-\frac{1}{2}} K^\oplus_l e_{p_2} \right\| \left\| \tilde{c}_{p_3} \Psi \right\| \sqrt{\left\langle \tilde{c}_{p_4} \Psi, H'^{(\pm 1)}_{\text{kin}} \tilde{c}_{p_4} \Psi \right\rangle} \\
&+ \sum_{l \in S_C} \sum_{p \in S} \left\| T e_{p_1} \right\| \left\| K^\oplus_l e_{p_2} \right\| \left\| \tilde{c}_{p_3} \Psi \right\| \left\| \tilde{c}_{p_4} \Psi \right\| \\
&=: A_1 + A_2 + A_3 + A_4.
\end{aligned} \tag{6.33}$$

The terms A_1 through A_4 can be estimated by the Cauchy-Schwarz inequality, Lemma 6.6, the inequality $\mathcal{N}_E \leq H'_{\text{kin}}$ and the fact that $\max_{p \in L^\pm_k} \left\| T e_p \right\| \leq \|T\|_{\infty,2}$ as

$$\begin{aligned}
A_1 &\leq \left(\max_{p \in L^\pm_k} \left\| \left(h^\oplus_k\right)^{-\frac{1}{2}} T e_p \right\| \right) \left(\sum_{l \in S_C} \left\| \left(h^\oplus_l\right)^{-\frac{1}{2}} K^\oplus_l \right\|_{\text{HS}} \right) \sqrt{\left\langle \Psi, H'_{\text{kin}} \Psi \right\rangle \left\langle \Psi, \mathcal{N}_E H'_{\text{kin}} \Psi \right\rangle}, \\
A_2 &\leq \left(\max_{p \in L^\pm_k} \left\| \left(h^\oplus_k\right)^{-\frac{1}{2}} T e_p \right\| \right) \left(\sum_{l \in S_C} \left\| K^\oplus_l \right\|_{\infty,2} \right) \sqrt{\left\langle \Psi, H'_{\text{kin}} \Psi \right\rangle \left\langle \Psi, \mathcal{N}_E H'_{\text{kin}} \Psi \right\rangle}, \\
A_3 &\leq \|T\|_{\infty,2} \left(\sum_{l \in S_C} \left\| \left(h^\oplus_l\right)^{-\frac{1}{2}} K^\oplus_l \right\|_{\text{HS}} \right) \left\langle \Psi, H'_{\text{kin}} \Psi \right\rangle, \\
A_4 &\leq \|T\|_{\infty,2} \left(\sum_{l \in S_C} \left\| K^\oplus_l \right\|_{\infty,2} \right) \left\langle \Psi, H'_{\text{kin}} \Psi \right\rangle,
\end{aligned} \tag{6.34}$$

all of which are also accounted for by the statement. □

7. Analysis of the one-body operators K, $A(t)$ and $B(t)$

In this section, we study the one-body operators on $\ell^2(L_k)$ defined in Section 5, including K_k introduced in (5.23) and A_k, B_k defined in Proposition 5.7:

$$\begin{aligned}
K_k &= -\frac{1}{2} \log \left(h_k^{-\frac{1}{2}} \left(h_k^2 + 2P_{h_k^{\frac{1}{2}} v_k} \right)^{\frac{1}{2}} h_k^{-\frac{1}{2}} \right), \\
A_k(t) &= \frac{1}{2} \left(e^{tK_k} (h_k + 2P_{v_k}) e^{tK_k} + e^{-tK_k} h_k e^{-tK_k} \right) - h_k, \\
B_k(t) &= \frac{1}{2} \left(e^{tK_k} (h_k + 2P_{v_k}) e^{tK_k} - e^{-tK_k} h_k e^{-tK_k} \right),
\end{aligned} \tag{7.1}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where

$$h_k e_p = \lambda_{k,p} e_p, \qquad \lambda_{k,p} = \frac{1}{2}\left(|p|^2 - |p-k|^2\right), \qquad P_{v_k} = |v_k\rangle\langle v_k|, \qquad v_k = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}} \sum_{p \in L_k} e_p, \quad (7.2)$$

and $(e_p)_{p\in L_k}$ is the standard orthonormal basis of $\ell^2(L_k)$. We will need precise estimates on these operators to control the quasi-bosonic Bogolubov transformation $e^{\mathcal{K}}$ diagonalizing the bosonizable terms. In particular, we will prove the following bounds.

Proposition 7.1 (Trace formulas). *For all $k \in \mathbb{Z}^3_*$, it holds that $K_k \le 0$ and*

$$\operatorname{tr}(K_k) = -\frac{1}{4}\log\left(1 + 2\hat{V}_k\left(\frac{k_F^{-1}}{2(2\pi)^3}\sum_{p\in L_k}\lambda_{k,p}^{-1}\right)\right) \ge -C\hat{V}_k.$$

Moreover, with $E_k = e^{-K_k} h_k e^{-K_k}$, we have

$$\operatorname{tr}(E_k - h_k) - \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}|L_k| = \frac{1}{\pi}\int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\right)dt,$$

with $F(x) = \log(1+x) - x$, and

$$\left|\operatorname{tr}(E_k - h_k) - \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}|L_k|\right| \le C k_F \hat{V}_k^2 |k|, \quad k_F \to \infty.$$

Here, $C > 0$ is a constant independent of k and k_F.

Proposition 7.2 (Matrix element estimates). *For all $k \in \overline{B}(0, 2k_F) \cap \mathbb{Z}^3_*$, it holds that*

$$\|K_k\|_{\infty,2} \le C\hat{V}_k \log(k_F)^{\frac{1}{3}} k_F^{-\frac{2}{3}} |k|^{1+\frac{5}{6}},$$

and for all $t \in [0,1]$, that

$$\|A_k(t)\|_{\infty,2}, \|B_k(t)\|_{\infty,2} \le C\hat{V}_k |k|^{\frac{1}{2}}\left(1 + \hat{V}_k\right).$$

Moreover, with $E_k = e^{-K_k} h_k e^{-K_k}$, we have

$$\max_{p\in L_k}\left|\left\langle e_p, (E_k - h_k) e_p\right\rangle\right| \le C k_F^{-1}\hat{V}_k\left(1 + \hat{V}_k\right).$$

Here, $C > 0$ is a constant independent of k and k_F.

Proposition 7.3 (Kinetic estimates). *For all $k \in \overline{B}(0, 2k_F)$, it holds as $k_F \to \infty$ that*

$$\begin{aligned}
\left\|h_k^{-\frac{1}{2}} K_k\right\|_{\mathrm{HS}} &\le C(\log k_F)^{\frac{2}{3}} k_F^{-\frac{1}{3}} \hat{V}_k |k|^{3+\frac{2}{3}} \\
\left\|\{K_k, h_k\} h_k^{-\frac{1}{2}}\right\|_{\mathrm{HS}} &\le C k_F^{\frac{1}{2}} \hat{V}_k |k|^{\frac{1}{2}} \\
\left\|h_k^{-\frac{1}{2}}\{K_k, h_k\} h_k^{-\frac{1}{2}}\right\|_{\mathrm{HS}} &\le C\hat{V}_k,
\end{aligned}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and for all $t \in [0, 1]$,

$$\max_{p\in L_k}\left\|h_k^{-\frac{1}{2}}A_k(t)e_p\right\|,\ \max_{p\in L_k}\left\|h_k^{-\frac{1}{2}}B_k(t)e_p\right\| \leq Ck_F^{-\frac{1}{2}}\hat{V}_k\left(1+\hat{V}_k^2\right).$$

Here, $C > 0$ is a constant independent of k and k_F.

Notation. In order to simplify the notation, we will throughout this section let $h : V \to V$ denote any positive self-adjoint operator acting on an n-dimensional Hilbert space V, let $(x_i)_{i=1}^n$ be an eigenbasis for h with eigenvalues $(\lambda_i)_{i=1}^n$ and let $v \in V$ be any vector satisfying $\langle x_i, v\rangle \geq 0$ for all $1 \leq i \leq n$. We will establish general results for the operators (c.f. (7.1))

$$\begin{aligned} K &= -\frac{1}{2}\log\left(h^{-\frac{1}{2}}\left(h^2+2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}h^{-\frac{1}{2}}\right), \\ A(t) &= \frac{1}{2}\left(e^{tK}\left(h+2P_v\right)e^{tK}+e^{-tK}he^{-tK}\right)-h, \\ B(t) &= \frac{1}{2}\left(e^{tK}\left(h+2P_v\right)e^{tK}-e^{-tK}he^{-tK}\right) \end{aligned} \tag{7.3}$$

and then at the end insert the specific choice (7.2) to get explicit estimates.

We will prove the trace formulas first. Then we derive general estimates for the matrix elements of the operators e^{-2K} and e^{2K} in terms of a single, simpler operator T. This allows us to show that all matrix elements of K are non-negative, which in turn implies that all matrix elements of e^{-tK}, $\sinh(-tK)$ and $\cosh(-tK)$ are convex with respect to t. With these estimates, we can then obtain the desired estimates of K, $A(t)$ and $B(t)$.

7.1. Trace formulas

In this section, we prove Proposition 7.1. We will prove some general results using the notation in (7.3), and then we insert the special choice of h_k, v_k in (7.2) to conclude. Let us start with the following:

Proposition 7.4. *The operator K in (7.3) satisfies $K \leq 0$ and*

$$\mathrm{tr}(K) = -\frac{1}{4}\log(1+2\langle v, h^{-1}v\rangle).$$

Proof. Since $h^2 + 2P_{h^{\frac{1}{2}}v} \geq h^2 > 0$ and $A \mapsto A^{\frac{1}{2}}$ is operator monotone, we find that

$$h^{-\frac{1}{2}}\left(h^2+2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}h^{-\frac{1}{2}} \geq h^{-\frac{1}{2}}(h^2)^{\frac{1}{2}}h^{-\frac{1}{2}} = 1. \tag{7.4}$$

Hence, K is well defined and $K \leq 0$. By the identity $\mathrm{tr}(\log(A)) = \log(\det(A))$ and multiplicativity of the determinant, we find

$$\begin{aligned} \mathrm{tr}(K) &= -\frac{1}{2}\log\left(\det\left(h^{-\frac{1}{2}}\left(h^2+2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}h^{-\frac{1}{2}}\right)\right) \\ &= -\frac{1}{4}\log\left(\det(h)^{-1}\det\left(h^2+2P_{h^{\frac{1}{2}}v}\right)\det(h)^{-1}\right) \\ &= -\frac{1}{4}\log\left(\det\left(h^{-1}\left(h^2+2P_{h^{\frac{1}{2}}v}\right)h^{-1}\right)\right) = -\frac{1}{4}\log\left(\det\left(1+2P_{h^{-\frac{1}{2}}v}\right)\right), \end{aligned} \tag{7.5}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and by Sylvester's determinant theorem [40], $\det(1 + \alpha P_x) = 1 + \alpha\, \|x\|^2$ for any $\alpha \in \mathbb{C}$; hence,

$$\operatorname{tr}(K) = -\frac{1}{4}\log\left(1 + 2\|h^{-\frac{1}{2}}v\|^2\right) = -\frac{1}{4}\log\left(1 + 2\langle v, h^{-1}v\rangle\right). \tag{7.6}$$

□

Another exact trace formula which we will need is the following integral representation of the square root of a rank one perturbation, first presented in [5].

Proposition 7.5. *Let $(H, \langle\cdot,\cdot\rangle)$ be a Hilbert space and let $A : H \to H$ be a positive self-adjoint operator. Then for any $x \in H$ and $g \in \mathbb{R}$ such that $A + gP_x > 0$, it holds that*

$$(A + gP_x)^{\frac{1}{2}} = A^{\frac{1}{2}} + \frac{2g}{\pi}\int_0^\infty \frac{t^2}{1 + g\left\langle x, (A + t^2)^{-1}x\right\rangle} P_{(A+t^2)^{-1}x}\, dt$$

and

$$\operatorname{tr}\left((A + gP_x)^{\frac{1}{2}}\right) = \operatorname{tr}\left(A^{\frac{1}{2}}\right) + \frac{1}{\pi}\int_0^\infty \log\left(1 + g\left\langle x, \left(A + t^2\right)^{-1} x\right\rangle\right) dt.$$

Note that Proposition 7.5 follows from the Sherman–Morrison formula [39]

$$\left(A + gP_{x,y}\right)^{-1} = A^{-1} - \frac{g}{1 + g\left\langle x, A^{-1}y\right\rangle} P_{(A^*)^{-1}x, A^{-1}y}, \tag{7.7}$$

with $P_{x,y} = |y\rangle\langle x| = \langle x, \cdot\rangle\, y$, and the functional calculus

$$\sqrt{A} = \frac{2}{\pi}\int_0^\infty \frac{A}{A + t^2}\, dt = \frac{2}{\pi}\int_0^\infty \left(1 - \frac{t^2}{A + t^2}\right) dt \tag{7.8}$$

for every self-adjoint non-negative operator A. Using this, we conclude the following:

Proposition 7.6. *The trace of $E - h$ where $E = e^{-K}he^{-K}$ is given by*

$$\operatorname{tr}(E - h) = \frac{1}{\pi}\int_0^\infty \log\left(1 + 2\left\langle v, h\left(h^2 + t^2\right)^{-1} v\right\rangle\right) dt.$$

Proof. By cyclicity of the trace and the definition of K,

$$\operatorname{tr}\left(e^{-K}he^{-K}\right) = \operatorname{tr}\left(he^{-2K}\right) = \operatorname{tr}\left(h\left(h^{-\frac{1}{2}}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}} h^{-\frac{1}{2}}\right)\right) = \operatorname{tr}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}, \tag{7.9}$$

so applying Proposition 7.5 with $A = h^2$, $x = h^{\frac{1}{2}}v$ and $g = 2$, we get the claim. □

Proof of Proposition 7.1. By inserting h_k and v_k in Proposition 7.4, we get $K_k \le 0$ and

$$\operatorname{tr}(K_k) = -\frac{1}{4}\log\left(1 + 2\left\langle v_k, h_k^{-1}v_k\right\rangle\right). \tag{7.10}$$

With the choice of h_k and v_k in (7.2), we have

$$0 \le \left\langle v_k, h_k^{-1}v_k\right\rangle = \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}\sum_{p,q\in L_k}\left\langle e_p, h_k^{-1}e_q\right\rangle = \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}\sum_{p\in L_k}\lambda_{k,p}^{-1} \le C\hat{V}_k, \tag{7.11}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where the last inequality is taken from Proposition A.2 in the Appendix. Combining with the bound $\log(1+x) \le x$ with $x > 0$, we find that

$$\operatorname{tr}(K_k) = -\frac{1}{4}\log\left(1 + 2\hat{V}_k\left(\frac{k_F^{-1}}{2(2\pi)^3}\sum_{p\in L_k}\lambda_{k,p}^{-1}\right)\right) \ge -C\hat{V}_k. \tag{7.12}$$

Next, using Proposition 7.6 and the identity (c.f. (7.8))

$$|L_k| = \sum_{p\in L_k} 1 = \frac{2}{\pi}\int_0^\infty \sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\,dt, \tag{7.13}$$

we conclude that

$$\operatorname{tr}(E_k - h_k) - \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}|L_k| = \frac{1}{\pi}\int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\right)dt \tag{7.14}$$

with $F(x) = \log(1+x) - x$. Since $|F(x)| \le \frac{1}{2}x^2$, we have

$$\begin{aligned}\left|\operatorname{tr}(E_k - h_k) - \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}|L_k|\right| &\le \frac{1}{\pi}\int_0^\infty \frac{1}{2}\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\right)^2 dt \\ &= \frac{\hat{V}_k^2 k_F^{-2}}{(2\pi)^7}\sum_{p,q\in L_k}\int_0^\infty \frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\frac{\lambda_{k,q}}{\lambda_{k,q}^2 + t^2}\,dt,\end{aligned} \tag{7.15}$$

and by the integral identity

$$\int_0^\infty \frac{a}{a^2 + t^2}\frac{b}{b^2 + t^2}\,dt = \frac{\pi}{2}\frac{1}{a+b}, \quad a, b > 0, \tag{7.16}$$

it holds that

$$\begin{aligned}\sum_{p,q\in L_k}\int_0^\infty \frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\frac{\lambda_{k,q}}{\lambda_{k,q}^2 + t^2}\,dt &= \frac{\pi}{2}\sum_{p,q\in L_k}\frac{1}{\lambda_{k,p} + \lambda_{k,q}} \le \frac{\pi}{2}\sum_{p,q\in L_k}\frac{1}{\sqrt{\lambda_{k,p}}\sqrt{\lambda_{k,q}}} \\ &= \frac{\pi}{2}\left(\sum_{p\in L_k}\lambda_{k,p}^{-\frac{1}{2}}\right)^2.\end{aligned} \tag{7.17}$$

By Proposition A.1, we have for any $k \in \mathbb{Z}_*^3$ that

$$\sum_{p\in L_k}\lambda_{k,p}^{-\frac{1}{2}} \le C\begin{cases} k_F^{\frac{3}{2}}\sqrt{|k|} & |k| < 2k_F \\ k_F^3|k|^{-1} & |k| \ge 2k_F \end{cases} \le C k_F^{\frac{3}{2}}\sqrt{|k|}, \quad k_F \to \infty \tag{7.18}$$

for a constant $C > 0$ independent of k and k_F, so we get the desired bound

$$\left|\operatorname{tr}(E_k - h_k) - \frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}|L_k|\right| \le C\hat{V}_k^2 k_F^{-2}\left(k_F^{\frac{3}{2}}\sqrt{|k|}\right)^2 = C k_F \hat{V}_k^2|k|. \tag{7.19}$$

□

7.2. Preliminary estimates for e^{-2K} and e^{2K}

The square root formula also yields the following exact representations of e^{-2K} and e^{2K}:

Proposition 7.7. *The operator K in (7.3) satisfies*

$$e^{-2K} = 1 + \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h\left(h^2+t^2\right)^{-1} v\right\rangle} P_{\left(h^2+t^2\right)^{-1}v}\, dt$$

$$e^{2K} = 1 - \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h^{-1}\left(h^{-2}+t^2\right)^{-1} v\right\rangle t^2} P_{h^{-1}\left(h^{-2}+t^2\right)^{-1}v}\, dt.$$

Proof. Let us consider

$$e^{-2K} = h^{-\frac{1}{2}}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}} h^{-\frac{1}{2}}. \tag{7.20}$$

first. Applying Proposition 7.5 with $A = h^2$, $x = h^{\frac{1}{2}}v$ and $g = 2$, again we find

$$\begin{aligned}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}} &= \left(h^2\right)^{\frac{1}{2}} + \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\langle h^{\frac{1}{2}}v, (h^2+t^2)^{-1}h^{\frac{1}{2}}v\rangle} P_{\left(h^2+t^2\right)^{-1}h^{\frac{1}{2}}v}\, dt \\ &= h + \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\langle v, h(h^2+t^2)^{-1}v\rangle} P_{h^{\frac{1}{2}}\left(h^2+t^2\right)^{-1}v}\, dt\end{aligned} \tag{7.21}$$

whence

$$\begin{aligned}e^{-2K} &= h^{-\frac{1}{2}}\left(h + \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h\left(h^2+t^2\right)^{-1}v\right\rangle} P_{h^{\frac{1}{2}}\left(h^2+t^2\right)^{-1}v}\, dt\right) h^{-\frac{1}{2}} \\ &= 1 + \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h\left(h^2+t^2\right)^{-1}v\right\rangle} P_{\left(h^2+t^2\right)^{-1}v}\, dt.\end{aligned} \tag{7.22}$$

For $e^{2K} = h^{\frac{1}{2}}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{-\frac{1}{2}} h^{\frac{1}{2}}$, we first use (7.7) to write

$$\begin{aligned}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{-1} &= \left(h^2\right)^{-1} - \frac{2}{1+2\left\langle h^{\frac{1}{2}}v, \left(h^2\right)^{-1} h^{\frac{1}{2}}v\right\rangle} P_{\left(h^2\right)^{-1}h^{\frac{1}{2}}v} \\ &= h^{-2} - \frac{2}{1+2\left\langle v, h^{-1}v\right\rangle} P_{h^{-\frac{3}{2}}v}.\end{aligned} \tag{7.23}$$

As this is an equality, the right-hand side is, in fact, positive (as the left-hand side is), so we may apply Proposition 7.5 with $A = h^{-2}$, $x = h^{-\frac{3}{2}}v$ and $g = -2\left(1+2\left\langle v, h^{-1}v\right\rangle\right)^{-1}$ for

$$\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{-\frac{1}{2}} = \left(h^{-2} - \frac{2}{1+2\left\langle v, h^{-1}v\right\rangle} P_{h^{-\frac{3}{2}}v}\right)^{\frac{1}{2}}$$

$$= \left(h^{-2}\right)^{\frac{1}{2}} - \frac{2}{1+2\left\langle v, h^{-1}v\right\rangle}\frac{2}{\pi}\int_0^\infty \frac{t^2}{1-\frac{2}{1+2\left\langle v, h^{-1}v\right\rangle}\left\langle h^{-\frac{3}{2}}v, \left(h^{-2}+t^2\right)^{-1} h^{-\frac{3}{2}}v\right\rangle} P_{\left(h^{-2}+t^2\right)^{-1}h^{-\frac{3}{2}}v}\, dt$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

$$= h^{-1} - \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h^{-1}v\right\rangle - 2\left\langle v, h^{-3}\left(h^{-2}+t^2\right)^{-1} v\right\rangle} P_{h^{-\frac{3}{2}}\left(h^{-2}+t^2\right)^{-1} v}\, dt$$

$$= h^{-1} - \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h^{-1}\left(h^{-2}+t^2\right)^{-1} v\right\rangle t^2} P_{h^{-\frac{3}{2}}\left(h^{-2}+t^2\right)^{-1} v}\, dt. \tag{7.24}$$

Hence,

$$e^{2K} = h^{\frac{1}{2}}\left(h^{-1} - \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h^{-1}\left(h^{-2}+t^2\right)^{-1} v\right\rangle t^2} P_{h^{-\frac{3}{2}}\left(h^{-2}+t^2\right)^{-1} v}\, dt\right) h^{\frac{1}{2}} \tag{7.25}$$

$$= 1 - \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h^{-1}\left(h^{-2}+t^2\right)^{-1} v\right\rangle t^2} P_{h^{-1}\left(h^{-2}+t^2\right)^{-1} v}\, dt.$$

□

These exact formulas now allow us to derive some simple estimates for $e^{-2K}-1$ and $1-e^{2K}$. To state these estimates, we first define a new operator T on $\ell^2(L_k)$ with matrix elements

$$\left\langle x_i, T x_j\right\rangle = 2\frac{\langle x_i, v\rangle \left\langle v, x_j\right\rangle}{\lambda_i+\lambda_j}, \quad \forall 1 \le i, j \le n. \tag{7.26}$$

Recall that $(x_i)_{i=1}^n$ are an eigenbasis of h with eigenvalues λ_i's and $\langle x_i, v\rangle \ge 0$ for all $1 \le i \le n$.

Proposition 7.8. *For K in (7.3) and T in (7.26), we have both the operator estimates*

$$0 \le e^{-2K} - 1 \le T \le (1+2\langle v, h^{-1}v\rangle)(e^{-2K}-1),$$
$$0 \le 1 - e^{2K} \le T \le (1+2\langle v, h^{-1}v\rangle)(1-e^{2K}),$$

and for all $1 \le i, j \le n$, the elementwise estimates

$$0 \le \langle x_i, (e^{-2K}-1)x_j\rangle \le \langle x_i, T x_j\rangle \le (1+2\langle v, h^{-1}v\rangle)\langle x_i, (e^{-2K}-1)x_j\rangle,$$
$$0 \le \langle x_i, (1-e^{2K})x_j\rangle \le \langle x_i, T x_j\rangle \le (1+2\langle v, h^{-1}v\rangle)\langle x_i, (1-e^{2K})x_j\rangle.$$

Proof. We first prove the bound $0 \le e^{-2K}-1 \le T$. Obviously, $0 \le e^{-2K}-1$ since $K \le 0$. Noting that $\langle v, h(h^2+t^2)^{-1}v\rangle \ge 0$ and $P_{(h^2+t^2)^{-1}v} \ge 0$ for all $t \in [0,\infty)$, we have by the first identity of Proposition 7.7 that

$$e^{-2K} - 1 = \frac{4}{\pi}\int_0^\infty \frac{t^2}{1+2\left\langle v, h\left(h^2+t^2\right)^{-1} v\right\rangle} P_{\left(h^2+t^2\right)^{-1} v}\, dt \le \frac{4}{\pi}\int_0^\infty t^2 P_{\left(h^2+t^2\right)^{-1} v}\, dt. \tag{7.27}$$

We claim that the right-hand side is precisely T. To see this, we compute the matrix elements with respect to $(x_i)_{i=1}^n$: For any $1 \le i, j \le n$, we have

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

$$\left\langle x_i, \left(\frac{4}{\pi}\int_0^\infty t^2 P_{(h^2+t^2)^{-1}v}\,dt\right)x_j\right\rangle = \frac{4}{\pi}\int_0^\infty t^2\left\langle x_i,\left(h^2+t^2\right)^{-1}v\right\rangle\left\langle\left(h^2+t^2\right)^{-1}v, x_j\right\rangle dt$$
$$= \frac{4}{\pi}\int_0^\infty t^2\frac{\langle x_i,v\rangle}{\lambda_i^2+t^2}\frac{\langle v,x_j\rangle}{\lambda_j^2+t^2}\,dt = \langle x_i,v\rangle\langle v,x_j\rangle\left(\frac{4}{\pi}\int_0^\infty\frac{t^2}{(\lambda_i^2+t^2)\left(\lambda_j^2+t^2\right)}\,dt\right)$$
$$= \langle x_i,v\rangle\langle v,x_j\rangle\left(\frac{4}{\pi}\frac{\pi}{2}\frac{1}{\lambda_i+\lambda_j}\right) = 2\frac{\langle x_i,v\rangle\langle v,x_j\rangle}{\lambda_i+\lambda_j} = \langle x_i, Tx_j\rangle, \tag{7.28}$$

where we used that $(x_i)_{i=1}^n$ is an eigenbasis for h as well as the integral identity (7.16).

The lower bound $T \leq \left(1+2\left\langle v,h^{-1}v\right\rangle\right)(e^{-2K}-1)$ follows by the same argument as

$$\left\langle v, h(h^2+t^2)^{-1}v\right\rangle \leq \left\langle v, h(h^2)^{-1}v\right\rangle = \left\langle v, h^{-1}v\right\rangle \tag{7.29}$$

for all $t \in [0,\infty)$, so

$$e^{-2K}-1 \geq \frac{1}{1+2\left\langle v,h^{-1}v\right\rangle}\left(\frac{4}{\pi}\int_0^\infty t^2 P_{(h^2+t^2)^{-1}v}\,dt\right) = \frac{1}{1+2\left\langle v,h^{-1}v\right\rangle}T. \tag{7.30}$$

The bounds

$$0 \leq 1-e^{2K} \leq T \leq \left(1+2\left\langle v,h^{-1}v\right\rangle\right)(1-e^{2K}) \tag{7.31}$$

follow by exactly the same argument, starting from the second identity of Proposition 7.7, using that

$$0 \leq \left\langle v, h^{-1}(h^{-2}+t^2)^{-1}v\right\rangle t^2 \leq \left\langle v, h^{-1}v\right\rangle \tag{7.32}$$

for all $t \in [0,\infty)$ as well as the integral identity (7.16).

The matrix element estimates likewise follow by the same argument as, for example,

$$0 \leq \left\langle x_i,\left(e^{-2K}-1\right)x_j\right\rangle = \frac{4}{\pi}\int_0^\infty\frac{t^2}{1+2\left\langle v,h\left(h^2+t^2\right)^{-1}v\right\rangle}\frac{\langle x_i,v\rangle}{\lambda_i^2+t^2}\frac{\langle v,x_j\rangle}{\lambda_j^2+t^2}\,dt$$
$$= \langle x_i,v\rangle\langle v,x_j\rangle\left(\frac{4}{\pi}\int_0^\infty\frac{1}{1+2\left\langle v,h\left(h^2+t^2\right)^{-1}v\right\rangle}\frac{t^2}{(\lambda_i^2+t^2)\left(\lambda_j^2+t^2\right)}\,dt\right) \tag{7.33}$$
$$\leq \langle x_i,v\rangle\langle v,x_j\rangle\left(\frac{4}{\pi}\int_0^\infty\frac{t^2}{(\lambda_i^2+t^2)\left(\lambda_j^2+t^2\right)}\,dt\right) = 2\frac{\langle x_i,v\rangle\langle v,x_j\rangle}{\lambda_i+\lambda_j} = \langle x_i, Tx_j\rangle$$

by the assumption that the inner products $\langle x_i,v\rangle$ and $\langle v,x_j\rangle$ are non-negative. □

Remark 7.1 (Optimality of the estimates). We may observe that the estimates for e^{-2K}, e^{2K} are, in general, optimal. To see this, let us add a small parameter $g \geq 0$ to the problem by substituting $\sqrt{g}v$ for v in equation (7.3) – that is, defining

$$K_g = -\frac{1}{2}\log\left(h^{-\frac{1}{2}}\left(h^2+2gP_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}h^{-\frac{1}{2}}\right), \quad T_g = gT. \tag{7.34}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Then the general bounds of the corollary read for K_g that

$$\frac{1}{1+2g\left\langle v,h^{-1}v\right\rangle}T_g \leq -2K_g \leq T_g. \tag{7.35}$$

Hence,

$$0 \geq -2K_g - T_g \geq -\left(1-\frac{1}{1+2g\left\langle v,h^{-1}v\right\rangle}\right)T_g = -\left(\frac{2\left\langle v,h^{-1}v\right\rangle}{1+2g\left\langle v,h^{-1}v\right\rangle}T\right)g^2 \geq -Cg^2, \tag{7.36}$$

which by self-adjointness of the operators involved implies that

$$-2K_g = T_g + O\left(g^2\right) = gT + O\left(g^2\right) \tag{7.37}$$

with respect to, say, operator norm. This shows that the operator $T_g = gT$ is, in fact, the first-order expansion of K_g with respect to the parameter g, which is then also the case for $e^{-2K_g}-1$, $1-e^{2K_g}$ as, for example, $e^{-2K_g}-1 = -2K_g + O\left(g^2\right) = T_g + O\left(g^2\right)$. The estimate

$$\left\langle x_i, (e^{-2K_g}-1)x_j\right\rangle \leq \left\langle x_i, T_g x_j\right\rangle \tag{7.38}$$

is therefore (asymptotically) optimal since T_g is precisely the small g limit of $e^{-2K_g}-1$.

This is relevant for our application, for although we do not have an explicit parameter g to consider, we do have $\hat{V}_k$ as an effective one. More precisely, the summability condition of $\hat{V}_k$ ensures that essentially all but finitely many coefficients $\hat{V}_k$ are small, even when the coefficients $(\hat{V}_k)_{k\in\mathbb{Z}^3}$ are not finitely supported.

7.3. Matrix element estimates for K, $A(t)$, $B(t)$

In this section, we prove Proposition 7.1. As before, we will prove some general results using the notation from (7.3), and then we insert h_k, v_k from (7.2) at the end. Recall that $(x_i)_i$ is an eigenbasis of h. We start with the following:

Proposition 7.9. *For all $1 \leq i,j \leq n$, we have $\langle x_i, -Kx_j\rangle \geq 0$ and the functions*

$$t \mapsto \left\langle x_i, (e^{-tK}-1)x_j\right\rangle,\ \left\langle x_i, \sinh\left(-tK\right)x_j\right\rangle,\ \left\langle x_i, \left(\cosh\left(-tK\right)-1\right)x_j\right\rangle$$

are non-negative and convex for $t \in [0,\infty)$.

Proof. By Proposition 7.8, the operator $S = 1-e^{2K}$ satisfies that $0 \leq S < 1$ and $\langle x_i, Sx_j\rangle \geq 0$ for all $1 \leq i,j \leq n$. By writing

$$-2K = -\log(1-S) = S + \frac{S^2}{2} + \frac{S^3}{3} + \frac{S^4}{4} + \dots, \tag{7.39}$$

we find that $-2K$ also has non-negative matrix elements. By using the series expansion again, we see that for any $1 \leq i,j \leq n$ and $t \in [0,\infty)$,

$$\begin{aligned}\left\langle x_i, (e^{-tK}-1)x_j\right\rangle &= \sum_{m=1}^{\infty}\frac{\left\langle x_i, (-K)^m x_j\right\rangle}{m!}t^m \geq 0,\\ \frac{d^2}{dt^2}\left\langle x_i, (e^{-tK}-1)x_j\right\rangle &= \sum_{m=3}^{\infty}\frac{\left\langle x_i, (-K)^m x_j\right\rangle}{(m-2)!}t^{m-2} \geq 0,\end{aligned} \tag{7.40}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

yielding the claim for $t \mapsto \left\langle x_i, \left(e^{-tK} - 1\right) x_j\right\rangle$. The functions $t \mapsto \left\langle x_i, \sinh\left(-tK\right) x_j\right\rangle$ and $t \mapsto \left\langle x_i, \left(\cosh\left(-tK\right) - 1\right) x_j\right\rangle$ can be treated similarly. □

Next, we have the key matrix element bounds.

Proposition 7.10. *For all* $1 \leq i, j \leq n$ *and* $t \in [0, 1]$, *we have the elementwise estimates*

$$\left|\left\langle x_i, K x_j\right\rangle\right|, \left|\left\langle x_i, \left(e^{-tK} - 1\right) x_j\right\rangle\right|, \left|\left\langle x_i, \left(1 - e^{tK}\right) x_j\right\rangle\right|,$$
$$\left|\left\langle x_i, \sinh\left(-tK\right) x_j\right\rangle\right|, \left|\left\langle x_i, \left(\cosh\left(-tK\right) - 1\right) x_j\right\rangle\right| \leq \frac{\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j}.$$

Proof. The arguments for $e^{-tK} - 1$, $\sinh\left(-tK\right)$ and $\cosh\left(-tK\right) - 1$ are again the same, so we focus on $e^{-tK} - 1$. By the convexity of Proposition 7.9 and the elementwise estimate of Proposition 7.8, we find for all $t \in [0, 1]$ that

$$\begin{aligned} 0 &\leq \left\langle x_i, \left(e^{-tK} - 1\right) x_j\right\rangle \leq \left(1 - \frac{t}{2}\right)\left\langle x_i, \left(e^{-0\cdot K} - 1\right) x_j\right\rangle + \frac{t}{2}\left\langle x_i, \left(e^{-2K} - 1\right) x_j\right\rangle \\ &= \frac{t}{2}\left\langle x_i, \left(e^{-2K} - 1\right) x_j\right\rangle \leq \frac{1}{2}\left|\left\langle x_i, \left(e^{-2K} - 1\right) x_j\right\rangle\right| \leq \frac{1}{2}\left\langle x_i, T x_j\right\rangle = \frac{\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j}. \end{aligned} \tag{7.41}$$

This also gives us the estimate for K as

$$0 \leq \left\langle x_i, (-K) x_j\right\rangle \leq \sum_{m=1}^{\infty} \frac{1}{m!}\left\langle x_i, (-K)^m x_j\right\rangle = \left\langle x_i, \left(e^{-K} - 1\right) x_j\right\rangle \leq \frac{\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j}, \tag{7.42}$$

where we used again the positivity of $\langle x_i, -K x_j\rangle$ from Proposition 7.9. Finally, the estimate for $1 - e^{-tK}$ is deduced from that of $\sinh\left(-tK\right)$ and $\cosh\left(-tK\right) - 1$ as

$$\begin{aligned} |\langle x_i, (1 - e^{tK}) x_j\rangle| &= \left|\left\langle x_i, \sinh\left(-tK\right) x_j\right\rangle - \left\langle x_i, \left(\cosh\left(-tK\right) - 1\right) x_j\right\rangle\right| \\ &\leq \max\left\{\left|\left\langle x_i, \sinh\left(-tK\right) x_j\right\rangle\right|, \left|\left\langle x_i, \left(\cosh\left(-tK\right) - 1\right) x_j\right\rangle\right|\right\} \leq \frac{\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j}, \end{aligned} \tag{7.43}$$

where we also used the positivity of $\left\langle x_i, \left(\cosh\left(-tK\right) - 1\right) x_j\right\rangle$ and $\left\langle x_i, \sinh\left(-tK\right) x_j\right\rangle$ from Proposition 7.9 to justify the first inequality. □

As a simple application of these estimates, we can easily obtain the following:

Proposition 7.11. *It holds that*

$$\|K\|_{\infty,2} \leq \alpha\sqrt{\left\langle v, h^{-2} v\right\rangle},$$

where $\alpha = \max_{1 \leq j \leq n}\left\langle v, x_j\right\rangle$.

Proof. We estimate using Proposition 7.10 that

$$\begin{aligned} \|K\|_{\infty,2}^2 &= \sum_{i=1}^{n} \max_{1 \leq j \leq n}\left|\left\langle x_i, K x_j\right\rangle\right|^2 \leq \sum_{i=1}^{n} \max_{1 \leq j \leq n}\left(\frac{\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j}\right)^2 \\ &\leq \left(\max_{1 \leq j \leq n}\left|\left\langle v, x_j\right\rangle\right|^2\right)\sum_{i=1}^{n} \frac{\left|\left\langle x_i, v\right\rangle\right|^2}{\lambda_i^2} = \alpha^2\left\langle v, h^{-2} v\right\rangle. \end{aligned} \tag{7.44}$$

□

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Now we consider $A(t)$ and $B(t)$, which can be written as

$$A(t) = A_h(t) + e^{tK} P_v e^{tK}, \quad B(t) = B_h(t) + e^{tK} P_v e^{tK}, \tag{7.45}$$

for

$$\begin{aligned} A_h(t) &= \frac{1}{2}\left(e^{tK} h e^{tK} + e^{-tK} h e^{-tK}\right) - h \\ &= \cosh(-tK)\, h \cosh(-tK) + \sinh(-tK)\, h \sinh(-tK) - h \\ &= \sinh(-tK)\, h \sinh(-tK) + (\cosh(-tK) - 1)\, h\, (\cosh(-tK) - 1) + \{h, \cosh(-tK) - 1\} \end{aligned} \tag{7.46}$$

and

$$\begin{aligned} B_h(t) &= \frac{1}{2}\left(e^{tK} h e^{tK} - e^{-tK} h e^{-tK}\right) \\ &= -(\cosh(-tK)\, h \sinh(-tK) + \sinh(-tK)\, h \cosh(-tK)) \\ &= -((\cosh(-tK) - 1)\, h \sinh(-tK) + \sinh(-tK)\, h\, (\cosh(-tK) - 1) + \{h, \sinh(-tK)\}) . \end{aligned} \tag{7.47}$$

Specifically, we must estimate the $\|\cdot\|_{\infty,2}$ norms of $A(t)$ and $B(t)$ with respect to $(x_i)_{i=1}^n$. We begin with the $e^{tK} P_v e^{tK}$ term:

Proposition 7.12. *It holds for all $t \in [0,1]$ that*

$$\left\|e^{tK} P_v e^{tK}\right\|_{\infty,2} \le \alpha \left(1 + \left\langle v, h^{-1} v\right\rangle\right) \|v\| ,$$

where $\alpha = \max_{1\le j\le n} \left\langle v, x_j\right\rangle$.

Proof. We first observe that

$$\begin{aligned} \left\|e^{tK} P_v e^{tK}\right\|_{\infty,2}^2 &= \sum_{i=1}^n \max_{1\le j\le n} \left|\left\langle x_i, e^{tK} P_v e^{tK} x_j\right\rangle\right|^2 = \sum_{i=1}^n \max_{1\le j\le n} \left|\left\langle x_i, e^{tK} v\right\rangle\right|^2 \left|\left\langle v, e^{tK} x_j\right\rangle\right|^2 \\ &= \left(\max_{1\le j\le n} \left|\left\langle v, e^{tK} x_j\right\rangle\right|^2\right) \left\|e^{tK} v\right\|^2 \le \left(\max_{1\le j\le n} \left|\left\langle v, e^{tK} x_j\right\rangle\right|^2\right) \|v\|^2 , \end{aligned} \tag{7.48}$$

where we used that by monotonicity of e^x and the fact that $K \le 0$, $\left\|e^{tK} v\right\|^2 = \left\langle v, e^{2tK} v\right\rangle \le \|v\|^2$. For the remaining factor, we first write

$$\left\langle v, e^{tK} x_j\right\rangle = \left\langle v, x_j\right\rangle + \left\langle v, \left(e^{tK} - 1\right) x_j\right\rangle = \left\langle v, x_j\right\rangle + \sum_{i=1}^n \langle v, x_i\rangle \left\langle x_i, \left(e^{tK} - 1\right) x_j\right\rangle, \quad 1 \le j \le n \tag{7.49}$$

and estimate using Proposition 7.10 that

$$\begin{aligned} \left|\sum_{i=1}^n \langle v, x_i\rangle \left\langle x_i, \left(e^{tK} - 1\right) x_j\right\rangle\right| &\le \sum_{i=1}^n |\langle v, x_i\rangle| \left|\left\langle x_i, \left(e^{tK} - 1\right) x_j\right\rangle\right| \le \sum_{i=1}^n |\langle v, x_i\rangle| \frac{\langle x_i, v\rangle \left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j} \\ &\le \left\langle v, x_j\right\rangle \sum_{i=1}^n \frac{|\langle x_i, v\rangle|^2}{\lambda_i} = \left\langle v, x_j\right\rangle \left\langle v, h^{-1} v\right\rangle, \quad 1 \le j \le n. \end{aligned} \tag{7.50}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Hence,

$$\left|\left\langle v, e^{tK}x_j\right\rangle\right| \le \left\langle v, x_j\right\rangle + \left|\sum_{i=1}^{n}\left\langle v, x_i\right\rangle\left\langle x_i, \left(e^{tK}-1\right)x_j\right\rangle\right| \le \left\langle v, x_j\right\rangle\left(1+\left\langle v, h^{-1}v\right\rangle\right), \quad 1 \le j \le n, \tag{7.51}$$

so returning to equation (7.48), we conclude that

$$\left\|e^{tK}P_v e^{tK}\right\|_\infty^2 \le \left(\max_{1\le j\le n}\left|\left\langle v, x_j\right\rangle\left(1+\left\langle v, h^{-1}v\right\rangle\right)\right|^2\right)\|v\|^2 = \alpha^2\left(1+\left\langle v, h^{-1}v\right\rangle\right)^2\|v\|^2, \tag{7.52}$$

implying the claim. □

For $A_h(t)$ and $B_h(t)$, we estimate the matrix elements of the operators appearing in the equations (7.46) and (7.47):

Proposition 7.13. *It holds for all* $1 \le i, j \le n$ *and* $t \in [0,1]$ *that, for* $C_t = \cosh(-tK) - 1$ *and* $S_t = \sinh(-tK)$*,*

$$\left|\left\langle x_i, C_t h C_t x_j\right\rangle\right|, \left|\left\langle x_i, C_t h S_t x_j\right\rangle\right|, \left|\left\langle x_i, S_t h S_t x_j\right\rangle\right| \le \left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle\left\langle v, h^{-1}v\right\rangle \tag{7.53}$$

and

$$\left|\left\langle x_i, \{h, C_t\}x_j\right\rangle\right|, \left|\left\langle x_i, \{h, S_t\}x_j\right\rangle\right| \le \left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle. \tag{7.54}$$

Proof. The arguments for the elements of the two groups are the same, so we focus on particular representatives. For the first, we have by the estimates of Proposition 7.10 that

$$\begin{aligned}\left|\left\langle x_i, \sinh(-tK)\,h\sinh(-tK), x_j\right\rangle\right| &= \left|\sum_{k=1}^{n}\lambda_k\left\langle x_i, \sinh(-tK), x_k\right\rangle\left\langle x_k, \sinh(-tK), x_j\right\rangle\right| \\ &\le \sum_{k=1}^{n}\lambda_k\frac{\left\langle x_i, v\right\rangle\left\langle v, x_k\right\rangle}{\lambda_i+\lambda_k}\frac{\left\langle x_k, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_k+\lambda_j} \\ &\le \left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle\sum_{k=1}^{n}\frac{\left|\left\langle x_k, v\right\rangle\right|^2}{\lambda_k} = \left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle\left\langle v, h^{-1}v\right\rangle\end{aligned} \tag{7.55}$$

and for the second, that

$$\begin{aligned}\left|\left\langle x_i, \{h, \sinh(-tK)\}, x_j\right\rangle\right| &= \left(\lambda_i+\lambda_j\right)\left|\left\langle x_i, \sinh(-tK), x_j\right\rangle\right| \\ &\le \left(\lambda_i+\lambda_j\right)\frac{\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_i+\lambda_j} = \left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle.\end{aligned} \tag{7.56}$$

□

We can now obtain the desired estimate:

Proposition 7.14. *It holds for all* $t \in [0,1]$ *that*

$$\|A(t)\|_{\infty,2}, \|B(t)\|_{\infty,2} \le 3\alpha\left(1+\left\langle v, h^{-1}v\right\rangle\right)\|v\|,$$

where $\alpha = \max_{1\le j\le n}\left\langle v, x_j\right\rangle$*.*

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proof. Again, the arguments for $A(t)$ and $B(t)$ are the same, so we focus on $A(t)$. Using that $\|\cdot\|_{\infty,2}$ is indeed a norm, and hence obeys the triangle inequality, we have for any $t \in [0,1]$ that

$$\begin{aligned}\|A(t)\|_{\infty,2} &\le \left\|e^{tK} P_v e^{tK}\right\|_{\infty,2} + \|A_h(t)\|_{\infty,2} \\ &\le \left\|e^{tK} P_v e^{tK}\right\|_{\infty,2} + \|\sinh(-tK)\, h \sinh(-tK)\|_{\infty,2} \\ &\quad + \|(\cosh(-tK)-1)\, h\, (\cosh(-tK)-1)\|_{\infty,2} + \|\{h, \cosh(-tK)-1\}\|_{\infty,2}. \end{aligned} \tag{7.57}$$

We estimate $\|\sinh(-tK)\, h \sinh(-tK)\|_{\infty,2}$ using Proposition 7.13 as

$$\begin{aligned}\|\sinh(-tK)\, h \sinh(-tK)\|_{\infty,2}^2 &= \sum_{i=1}^n \max_{1\le j\le n} \left|\left\langle x_i, \sinh(-tK)\, h \sinh(-tK)\, x_j\right\rangle\right|^2 \\ &\le \sum_{i=1}^n \max_{1\le j\le n} \left|\langle x_i, v\rangle \left\langle v, x_j\right\rangle \left\langle v, h^{-1} v\right\rangle\right|^2 \\ &= \alpha^2 \left\langle v, h^{-1} v\right\rangle^2 \sum_{i=1}^n |\langle x_i, v\rangle|^2 = \alpha^2 \left\langle v, h^{-1} v\right\rangle^2 \|v\|^2, \end{aligned} \tag{7.58}$$

the same bound holding also for $\|(\cosh(-tK)-1)\, h\, (\cosh(-tK)-1)\|_{\infty,2}^2$. We likewise find

$$\begin{aligned}\|\{h, \cosh(-tK)-1\}\|_{\infty,2}^2 &= \sum_{i=1}^n \max_{1\le j\le n} \left|\left\langle x_i, \{h, \cosh(-tK)-1\}\, x_j\right\rangle\right|^2 \\ &\le \sum_{i=1}^n \max_{1\le j\le n} \left|\langle x_i, v\rangle \left\langle v, x_j\right\rangle\right|^2 = \alpha^2 \|v\|^2, \end{aligned} \tag{7.59}$$

so recalling the estimate of Proposition 7.12, we conclude that

$$\|A(t)\|_{\infty,2} \le \alpha \left(1 + \left\langle v, h^{-1} v\right\rangle\right) \|v\| + 2\alpha \left\langle v, h^{-1} v\right\rangle \|v\| + \alpha \|v\| \le 3\alpha \left(1 + \left\langle v, h^{-1} v\right\rangle\right) \|v\|. \tag{7.60}$$

□

Now we come to the last ingredient of Proposition 7.1.

Proposition 7.15. *Let $E = e^{-K} h e^{-K}$. For all $1 \le i, j \le n$, it holds that*

$$\left|\left\langle x_i, (E-h) x_j\right\rangle\right| \le \left(1 + \left\langle v, h^{-1} v\right\rangle\right) \left\langle x_i, v\right\rangle \left\langle v, x_j\right\rangle.$$

Proof. Using the identity

$$e^{-K} h e^{-K} - h = \left\{h, e^{-K} - 1\right\} + \left(e^{-K} - 1\right) h \left(e^{-K} - 1\right), \tag{7.61}$$

we can write

$$\left\langle x_i, \left(e^{-K} h e^{-K} - h\right) x_j\right\rangle = (\lambda_i + \lambda_j) \left\langle x_i, \left(e^{-K} - 1\right) x_j\right\rangle + \left\langle x_i, \left(e^{-K} - 1\right) h \left(e^{-K} - 1\right) x_j\right\rangle. \tag{7.62}$$

We can apply Proposition 7.10 to estimate the first term of this equation as

$$\left|(\lambda_i + \lambda_j) \left\langle x_i, \left(e^{-K} - 1\right) x_j\right\rangle\right| \le (\lambda_i + \lambda_j) \frac{\left\langle x_i, v\right\rangle \left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j} = \left\langle x_i, v\right\rangle \left\langle v, x_j\right\rangle \tag{7.63}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and the second term as

$$
\begin{aligned}
\left|\left\langle x_i, (e^{-K}-1)h(e^{-K}-1)x_j\right\rangle\right| &= \left|\sum_{k=1}^{n} \lambda_k \left\langle x_i, (e^{-K}-1)x_k\right\rangle\left\langle x_k, (e^{-K}-1)x_j\right\rangle\right| \\
&\leq \sum_{k=1}^{n} \lambda_k \frac{\left\langle x_i, v\right\rangle\left\langle v, x_k\right\rangle}{\lambda_i+\lambda_k} \frac{\left\langle x_k, v\right\rangle\left\langle v, x_j\right\rangle}{\lambda_k+\lambda_j} \qquad (7.64) \\
&\leq \left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle \sum_{k=1}^{n} \frac{\left|\left\langle x_k, v\right\rangle\right|^2}{\lambda_k} = \left\langle v, h^{-1}v\right\rangle\left\langle x_i, v\right\rangle\left\langle v, x_j\right\rangle,
\end{aligned}
$$

which implies the claim. □

Proof of Proposition 7.2. Now we insert h_k and v_k to conclude. Using Proposition 7.11, and noting that 'α' of our problem is simply the constant

$$
\max_{p\in L_k} \left\langle v_k, e_p\right\rangle = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}}, \qquad (7.65)
$$

we find that

$$
\begin{aligned}
\|K_k\|_{\infty,2} &\leq \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}} \sqrt{\left\langle v_k, h_k^{-2} v_k\right\rangle} = \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}} \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} \sum_{p\in L_k} \frac{1}{\lambda_{k,p}^2}} \\
&\leq C\hat{V}_k k_F^{-1} \sqrt{\sum_{p\in L_k} \frac{1}{\lambda_{k,p}^2}}. \qquad (7.66)
\end{aligned}
$$

The desired upper bound

$$
\|K_k\|_{\infty,2} \leq C\hat{V}_k \log\left(k_F\right)^{\frac{1}{3}} k_F^{-\frac{2}{3}} |k|^{1+\frac{5}{6}} \qquad (7.67)
$$

then follows from an estimate from Proposition A.3 in the Appendix:

$$
\sum_{p\in L_k} \frac{1}{\lambda_{k,p}^2} \leq C|k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}}, \quad k_F \to \infty.
$$

However, by Proposition 7.14 and (7.11), we conclude that

$$
\begin{aligned}
\|A_k(t)\|_{\infty,2}, \|B_k(t)\|_{\infty,2} &\leq 3\sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}} \left(1+\left\langle v_k, h_k^{-1} v_k\right\rangle\right) \sqrt{\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3} |L_k|} \\
&\leq C\hat{V}_k k_F^{-1} \sqrt{|L_k|} \left(1+\hat{V}_k k_F^{-1} \sum_{p\in L_k} \frac{1}{\lambda_{k,p}}\right) \leq C\hat{V}_k |k|^{\frac{1}{2}} \left(1+\hat{V}_k\right), \qquad (7.68)
\end{aligned}
$$

where we used

$$
|L_k| \leq C|k| k_F^2, \quad \sum_{p\in L_k} \frac{1}{\lambda_{k,p}} \leq C k_F \qquad (7.69)
$$

from Proposition A.1 and Proposition A.2. Finally, from Proposition 7.15, we have

$$\begin{aligned}\max_{p\in L_k}\left|\left\langle e_p,(E_k-h_k)\,e_p\right\rangle\right| &\le \left(1+\left\langle v_k,h_k^{-1}v_k\right\rangle\right)\sup_{p\in L_k}|\left\langle e_p,v\right\rangle|^2\\ &= \left(1+\hat{V}_k k_F^{-1}\sum_{p\in L_k}\frac{1}{\lambda_{k,p}}\right)\frac{\hat{V}_k k_F^{-1}}{2(2\pi)^3}\le Ck_F^{-1}\hat{V}_k\left(1+\hat{V}_k\right),\end{aligned}\tag{7.70}$$

where we used (7.65) and (7.69) again in the last estimate. □

7.4. Kinetic estimates

Now we prove Proposition 7.3. Again, let us start with the notation (7.3). We have the following:

Proposition 7.16. *Under the notation (7.3), it holds that*

$$\begin{aligned}\left\|h^{-\frac{1}{2}}K\right\|_{\mathrm{HS}} &\le \left\langle v,h^{-\frac{3}{2}}v\right\rangle,\\ \left\|\{K,h\}\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}} &\le 2\,\|v\|\sqrt{\left\langle v,h^{-1}v\right\rangle},\\ \left\|h^{-\frac{1}{2}}\{K,h\}\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}} &\le 2\left\langle v,h^{-1}v\right\rangle.\end{aligned}$$

Proof. Using Proposition 7.10, we estimate

$$\left\|h^{-\frac{1}{2}}K\right\|_{\mathrm{HS}}^2=\sum_{i,j=1}^n\frac{1}{\lambda_i}\left|\left\langle x_i,Kx_j\right\rangle\right|^2\le\sum_{i,j=1}^n\frac{1}{\lambda_i}\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2\le\left(\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i^{\frac{3}{2}}}\right)^2=\left\langle v,h^{-\frac{3}{2}}v\right\rangle^2,\tag{7.71}$$

and for $\left\|\{K,h\}\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}$ use that $\left\|\{K,h\}\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\le\left\|Kh^{\frac{1}{2}}\right\|_{\mathrm{HS}}+\left\|hKh^{-\frac{1}{2}}\right\|_{\mathrm{HS}}$ to estimate

$$\begin{aligned}\left\|Kh^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^n\lambda_j\left|\left\langle x_i,Kx_j\right\rangle\right|^2\le\sum_{i,j=1}^n|\langle x_i,v\rangle|^2\frac{\left|\left\langle x_j,v\right\rangle\right|^2}{\lambda_j}=\|v\|^2\left\langle v,h^{-1}v\right\rangle\\ \left\|hKh^{-\frac{1}{2}}\right\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^n\frac{\lambda_i^2}{\lambda_j}\left|\left\langle x_i,Kx_j\right\rangle\right|^2\le\sum_{i,j=1}^n|\langle x_i,v\rangle|^2\frac{\left|\left\langle x_j,v\right\rangle\right|^2}{\lambda_j}=\|v\|^2\left\langle v,h^{-1}v\right\rangle\end{aligned}\tag{7.72}$$

for the claimed $\left\|\{K,h\}\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\le 2\,\|v\|\sqrt{\left\langle v,h^{-1}v\right\rangle}$. We likewise have that

$$\left\|h^{-\frac{1}{2}}\{K,h\}\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\le\left\|h^{-\frac{1}{2}}Kh^{\frac{1}{2}}\right\|_{\mathrm{HS}}+\left\|h^{\frac{1}{2}}Kh^{-\frac{1}{2}}\right\|_{\mathrm{HS}}=2\left\|h^{-\frac{1}{2}}Kh^{\frac{1}{2}}\right\|_{\mathrm{HS}},\tag{7.73}$$

so the bound

$$\left\|h^{-\frac{1}{2}}Kh^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2=\sum_{i,j=1}^n\frac{\lambda_i}{\lambda_j}\left|\left\langle x_i,Kx_j\right\rangle\right|^2\le\left(\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\right)^2=\left\langle v,h^{-1}v\right\rangle^2\tag{7.74}$$

implies the final claim. □

For $A(t)$ and $B(t)$, we recall the decompositions (7.45)-(7.47). Recall also that $(x_i)_i$ is an eigenbasis of h and $\langle x_i,v\rangle\ge 0$ for all $1\le i\le n$. We first estimate the $e^{tK}P_ve^{tK}$ term:

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proposition 7.17. *For all $t \in [0,1]$, it holds that*

$$\max_{1\le j\le n}\left\|h^{-\frac{1}{2}}e^{tK}P_v e^{tK}x_j\right\| \le \alpha\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sqrt{\left\langle v,h^{-1}v\right\rangle},$$

where $\alpha = \max_{1\le j\le n}\left\langle v,x_j\right\rangle$.

Proof. We write $e^{tK}P_v e^{tK}$ as

$$e^{tK}P_v e^{tK} = P_v + (e^{tK}-1)P_v + P_v(e^{tK}-1) + (e^{tK}-1)P_v(e^{tK}-1) \tag{7.75}$$

and estimate each term separately. By the definition of P_v, the first term is simply

$$\left\|h^{-\frac{1}{2}}P_v x_j\right\| = \left|\left\langle v,x_j\right\rangle\right|\left\|h^{-\frac{1}{2}}v\right\| \le \alpha\sqrt{\left\langle v,h^{-1}v\right\rangle}. \tag{7.76}$$

For the remaining terms, we use Proposition 7.10 to estimate that

$$\begin{aligned}\left\|h^{-\frac{1}{2}}\left(e^{tK}-1\right)P_v x_j\right\|^2 &= \sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k=1}^n\left\langle x_i,\left(e^{tK}-1\right)x_k\right\rangle\left\langle x_k,P_v x_j\right\rangle\right|^2\\ &\le \sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\langle x_k,v\rangle\left\langle v,x_j\right\rangle\right|^2 \qquad (7.77)\\ &\le \left|\left\langle v,x_j\right\rangle\right|^2\left(\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\right)^3 \le \alpha^2\left\langle v,h^{-1}v\right\rangle^3, \qquad (7.78)\end{aligned}$$

and

$$\begin{aligned}\left\|h^{-\frac{1}{2}}P_v\left(e^{tK}-1\right)x_j\right\|^2 &= \sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k=1}^n\langle x_i,P_v x_k\rangle\left\langle x_k,\left(e^{tK}-1\right)x_j\right\rangle\right|^2\\ &\le \sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k=1}^n\langle x_i,v\rangle\langle v,x_k\rangle\frac{\langle x_k,v\rangle\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2 \qquad (7.79)\\ &\le \left|\left\langle v,x_j\right\rangle\right|^2\left(\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\right)^3 \le \alpha^2\left\langle v,h^{-1}v\right\rangle^3\end{aligned}$$

and

$$\begin{aligned}&\left\|h^{-\frac{1}{2}}\left(e^{tK}-1\right)P_v\left(e^{tK}-1\right)x_j\right\|^2\\ &= \sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k,l=1}^n\left\langle x_i,\left(e^{tK}-1\right)x_k\right\rangle\langle x_k,P_v x_l\rangle\left\langle x_l,\left(e^{tK}-1\right)x_j\right\rangle\right|^2\\ &\le \sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k,l=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\langle x_k,v\rangle\langle v,x_l\rangle\frac{\langle x_l,v\rangle\left\langle v,x_j\right\rangle}{\lambda_l+\lambda_j}\right|^2 \qquad (7.80)\\ &\le \left|\left\langle v,x_j\right\rangle\right|^2\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left(\sum_{k,l=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_k}\frac{|\langle x_l,v\rangle|^2}{\lambda_l}\right)^2 \le \alpha^2\left\langle v,h^{-1}v\right\rangle^5,\end{aligned}$$

which imply the claim. ◻

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Finally, the full estimates on $A(t)$ and $B(t)$ are now easily obtained:

Proposition 7.18. *It holds for all* $t \in [0,1]$ *that*

$$\max_{1\le j\le n}\left\|h^{-\frac{1}{2}}A(t)x_j\right\|,\ \max_{1\le j\le n}\left\|h^{-\frac{1}{2}}B(t)x_j\right\| \le 2\alpha\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sqrt{\left\langle v,h^{-1}v\right\rangle},$$

where $\alpha = \max_{1\le j\le n}\left\langle v,x_j\right\rangle$.

Proof. The estimates for $A(t)$ and $B(t)$ are similar, so we focus on $A(t)$. We have

$$\begin{aligned}&\left\|h^{-\frac{1}{2}}A(t)x_j\right\|\\ &\quad\le\left\|h^{-\frac{1}{2}}\sinh(-tK)\,h\sinh(-tK)\,x_j\right\|+\left\|h^{-\frac{1}{2}}(\cosh(-tK)-1)\,h\,(\cosh(-tK)-1)\,x_j\right\|\\ &\qquad+\left\|h^{-\frac{1}{2}}\{h,\cosh(-tK)-1\}\,x_j\right\|+\left\|h^{-\frac{1}{2}}e^{tK}P_ve^{tK}x_j\right\|,\end{aligned}\tag{7.81}$$

and by Proposition 7.13, we can estimate that

$$\begin{aligned}\left\|h^{-\frac{1}{2}}\sinh(-tK)\,h\sinh(-tK)\,x_j\right\|^2&=\sum_{i=1}^n\frac{1}{\lambda_i}\left|\left\langle x_i,\sinh(-tK)\,h\sinh(-tK)\,x_j\right\rangle\right|^2\\ &\le\left|\left\langle v,x_j\right\rangle\right|^2\left\langle v,h^{-1}v\right\rangle^2\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\le\alpha^2\left\langle v,h^{-1}v\right\rangle^3,\end{aligned}\tag{7.82}$$

the same estimate holding also for $\left\|h^{-\frac{1}{2}}(\cosh(-tK)-1)\,h\,(\cosh(-tK)-1)\,x_j\right\|$, and

$$\begin{aligned}\left\|h^{-\frac{1}{2}}\{h,\cosh(-tK)-1\}\,x_j\right\|^2&=\sum_{i=1}^n\frac{1}{\lambda_i}\left|\left\langle x_i,\{h,\cosh(-tK)-1\}\,x_j\right\rangle\right|^2\\ &\le\left|\left\langle v,x_j\right\rangle\right|^2\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\le\alpha^2\left\langle v,h^{-1}v\right\rangle.\end{aligned}\tag{7.83}$$

Inserting also the estimate of Proposition 7.17, we thus obtain

$$\begin{aligned}\max_{1\le j\le n}\left\|h^{-\frac{1}{2}}A(t)x_j\right\|&\le 2\alpha\left\langle v,h^{-1}v\right\rangle^{\frac{3}{2}}+\alpha\sqrt{\left\langle v,h^{-1}v\right\rangle}+\alpha(1+\left\langle v,h^{-1}v\right\rangle)^2\sqrt{\left\langle v,h^{-1}v\right\rangle}\\ &\le 2\alpha(1+\langle v,h^{-1}v\rangle)^2\sqrt{\langle v,h^{-1}v\rangle}.\end{aligned}\tag{7.84}$$

□

Proof of Proposition 7.3. The desired bounds follow from applying the general estimates of this section to h_k and v_k, plus using the uniform bound on α in (7.65) and the estimates

$$\|v\|^2\le C\hat{V}_k|k|k_F,\quad\left\langle v_k,h_k^{-1}v_k\right\rangle\le C\hat{V}_k,\quad\left\langle v_k,h_k^{-\frac{3}{2}}v_k\right\rangle\le C\hat{V}_k|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{-\frac{1}{3}},\tag{7.85}$$

which hold for all $k\in\overline{B}(0,2k_F)$ due to Propositions A.1, A.2, and A.3. □

8. Gronwall estimates for the Bogolubov transformation

In the previous sections, we have bounded several error terms using the operators H'_{kin} and $\mathcal{N}_E$. In this section, we control the propagation of these operators under the Bogolubov transformation $e^{-\mathcal{K}}$ defined in Section 5. We have the following Gronwall-type estimates.

Proposition 8.1. *Let $\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k|<\infty$. Then for all $\Psi\in D\left(H'_{\text{kin}}\right)$ and $|t|\le 1$, it holds that*

$$\left\langle e^{-t\mathcal{K}}\Psi,(H'_{\text{kin}}+k_F)e^{-t\mathcal{K}}\Psi\right\rangle\le C\left\langle\Psi,(H'_{\text{kin}}+k_F)\Psi\right\rangle,$$
$$\left\langle e^{-t\mathcal{K}}\Psi,(k_F^{-1}\mathcal{N}_E H'_{\text{kin}}+H'_{\text{kin}}+k_F)e^{-t\mathcal{K}}\Psi\right\rangle\le C\left\langle\Psi,(k_F^{-1}\mathcal{N}_E H'_{\text{kin}}+H'_{\text{kin}}+k_F)\Psi\right\rangle$$

for a constant $C>0$ independent of k_F.

As a preparation, let us first prove the following:

Lemma 8.2. *Let X,Y,Z be self-adjoint operators on a Hilbert space such that*

$$X,Z>0,\qquad [X,Z]=0,\qquad \pm\left[[Y,X],X\right]\le Z.$$

Then,

$$\pm\left[\left[Y,\sqrt{X}\right],\sqrt{X}\right]\le\frac{Z}{4X}.$$

Proof of Lemma 8.2. Using (7.8), we can write

$$\left[Y,\sqrt{X}\right]=\frac{2}{\pi}\int_0^\infty\left[Y,\frac{X}{X+t^2}\right]dt=\frac{2}{\pi}\int_0^\infty\left[Y,\frac{-t^2}{X+t^2}\right]dt=\frac{2}{\pi}\int_0^\infty\frac{1}{X+t^2}\left[Y,X\right]\frac{1}{X+t^2}t^2dt, \tag{8.1}$$

and applying this identity twice, we get

$$\left[\left[Y,\sqrt{X}\right],\sqrt{X}\right]=\left(\frac{2}{\pi}\right)^2\int_0^\infty\int_0^\infty\frac{1}{X+t^2}\frac{1}{X+s^2}\left[[Y,X],X\right]\frac{1}{X+s^2}\frac{1}{X+t^2}s^2t^2dsdt. \tag{8.2}$$

Therefore, the assumptions $\pm\left[[Y,X],X\right]\le Z$ and $[X,Z]=0$ imply that

$$\pm\left[\left[Y,\sqrt{X}\right],\sqrt{X}\right]\le\left(\frac{2}{\pi}\right)^2\int_0^\infty\int_0^\infty\frac{1}{X+t^2}\frac{1}{X+s^2}Z\frac{1}{X+s^2}\frac{1}{X+t^2}s^2t^2dsdt=\frac{Z}{4X}. \tag{8.3}$$

□

Now we give the following:

Proof of Proposition 8.1. Write $\Psi_t=e^{t\mathcal{K}}\Psi$ for brevity. Recalling Proposition 5.4, we see that

$$\frac{d}{dt}\left\langle\Psi_t,(H'_{\text{kin}}+k_F)\Psi_t\right\rangle=\left\langle\Psi_t,\left[\mathcal{K},H'_{\text{kin}}\right]\Psi_t\right\rangle=\sum_{k\in S_C}\left\langle\Psi_t,Q_2^k\left(\left\{K_k^\oplus,h_k^\oplus\right\}\right)\Psi_t\right\rangle. \tag{8.4}$$

The right-hand side can be bounded by using Propositions 4.10 and 7.3 as

$$\begin{aligned}\sum_{k\in S_C}\left|\left\langle\Psi_t,Q_2^k\left(\left\{K_k^\oplus,h_k^\oplus\right\}\right)\Psi_t\right\rangle\right|&\le 2\sum_{k\in S_C}\left\|\left(h_k^\oplus\right)^{-\frac12}\left\{K_k^\oplus,h_k^\oplus\right\}\left(h_k^\oplus\right)^{-\frac12}\right\|_{\text{HS}}\left\langle\Psi_t,H'_{\text{kin}}\Psi_t\right\rangle\\&\quad+2\sum_{k\in S_C}\left\|\left\{K_k^\oplus,h_k^\oplus\right\}\left(h_k^\oplus\right)^{-\frac12}\right\|_{\text{HS}}\sqrt{\left\langle\Psi_t,H'_{\text{kin}}\Psi_t\right\rangle}\,\|\Psi_t\|\\&\le C\left(\sum_{k\in S_C}\hat{V}_k\right)\left\langle\Psi_t,H'_{\text{kin}}\Psi_t\right\rangle+Ck_F^{\frac12}\left(\sum_{k\in S_C}\hat{V}_k|k|^{\frac12}\right)\sqrt{\left\langle\Psi_t,H'_{\text{kin}}\Psi_t\right\rangle}\,\|\Psi_t\|\\&\le C\left\langle\Psi_t,(H'_{\text{kin}}+k_F)\Psi_t\right\rangle,\end{aligned} \tag{8.5}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where we also used the Cauchy–Schwarz inequality in the last step. Thus, the first estimate of Proposition 8.1 follows by Gronwall's lemma. For the second bound of Proposition 8.1, let us denote

$$X_1 = \mathcal{N}_E + k_F \geq k_F, \quad X_2 = k_F + H'_{\mathrm{kin}} \geq k_F, \quad Y_1 = [\mathcal{K}, X_2] \quad \text{and} \quad Y_2 = [\mathcal{K}, X_1]. \tag{8.6}$$

Note that Y_1, Y_2 are symmetric since X_1, X_2 are symmetric and $\mathcal{K}$ is skew-symmetric. Moreover, since $[X_1, X_2] = 0$, $[\mathcal{K}, X_1X_2]$ is also symmetric, and we can write

$$\begin{aligned} 2[\mathcal{K}, X_1X_2] &= 2\Big(X_1[\mathcal{K}, X_2] + [\mathcal{K}, X_1]X_2\Big) = 2(X_1Y_1 + Y_2X_2) \\ &= \sum_{i=1}^{2}(X_iY_i + Y_iX_i) = \sum_{i=1}^{2}\Big(2\sqrt{X_i}Y_i\sqrt{X_i} + [[Y_i, \sqrt{X_i}], \sqrt{X_i}]\Big). \end{aligned} \tag{8.7}$$

For $i = 1$, arguing similarly to (8.4) and (8.5), we have

$$\pm Y_1 = \pm[\mathcal{K}, H'_{\mathrm{kin}}] = \pm \sum_{k \in S_C} Q_2^k\left(\{K_k^{\oplus}, h_k^{\oplus}\}\right) \leq CX_2, \quad \pm\sqrt{X_1}Y_1\sqrt{X_1} \leq CX_1X_2. \tag{8.8}$$

Here, we used $[X_1, X_2] = 0$ in the last estimate. To apply Lemma 8.2, let us compute $[[Y_1, X_1], X_1]$. Note that for every symmetric operator B on $\ell^2(L_k^{\pm})$, we deduce from (1.75) that

$$\begin{aligned} \left[Q_2^k(B), \mathcal{N}_E\right] &= 2\sum_{p,q \in L_k^{\pm}} \langle e_p, Be_q\rangle \left(-b^*_{\overline{k,p}}b^*_{\overline{k,q}} + b_{\overline{k,q}}b_{\overline{k,p}}\right), \\ \left[\left[Q_2^k(B), \mathcal{N}_E\right], \mathcal{N}_E\right] &= 4\sum_{p,q \in L_k^{\pm}} \langle e_p, Be_q\rangle \left(b^*_{\overline{k,p}}b^*_{\overline{k,q}} + b_{\overline{k,q}}b_{\overline{k,p}}\right) = 4Q_2^k(B). \end{aligned} \tag{8.9}$$

Using (8.9) and (8.8), we have

$$\pm[[Y_1, X_1], X_1] = \pm 4\sum_{k \in S_C} Q_2^k\left(\{K_k^{\oplus}, h_k^{\oplus}\}\right) \leq CX_2, \tag{8.10}$$

which implies by Lemma 8.2 that

$$\pm[[Y_1, \sqrt{X_1}], \sqrt{X_1}] \leq CX_2X_1^{-1}. \tag{8.11}$$

Next, we consider the terms of $i = 2$ in (8.7). Let us compute the commutator $Y_2 = [\mathcal{K}, \mathcal{N}_E]$. By linearity, we deduce from (1.75) that $[b_k(\varphi), \mathcal{N}_E] = b_k(\varphi)$ for any $\varphi \in \ell^2(L_k^{\pm})$, and hence from the definition of $\mathcal{K}$ in (5.2),

$$Y_2 = [\mathcal{K}, \mathcal{N}_E] = \sum_{k \in S_C}\sum_{p \in L_k^{\pm}} \left(b_k^*(e_p)b_k^*\left(K_k^{\oplus}e_p\right) + b_k\left(K_k^{\oplus}e_p\right)b_k(e_p)\right) = \sum_{k \in S_C} Q_2^k\left(K_k^{\oplus}\right). \tag{8.12}$$

Note that

$$K_k^{\oplus} = \begin{pmatrix} 0 & K_k \\ K_k & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} K_k & 0 \\ 0 & -K_k \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \tag{8.13}$$

and hence by Proposition 7.1, we obtain

$$\sum_{k \in S_C} \left\|K_k^{\oplus}\right\|_{\mathrm{HS}} \leq \sum_{k \in S_C} \mathrm{tr}(|K_k^{\oplus}|) = 2\sum_{k \in S_C} \mathrm{tr}\left(|K_k|\right) \leq C\sum_{k \in \mathbb{Z}^3} \hat{V}_k. \tag{8.14}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Therefore, by Proposition 4.7,

$$\pm Y_2 \le 2\left(\sum_{k\in S_C} \|K_k^{\oplus}\|_{\mathrm{HS}}\right)(1+\mathcal{N}_E) \le CX_1, \quad \pm\sqrt{X_2}Y_2\sqrt{X_2} \le CX_1X_2. \tag{8.15}$$

Finally, consider

$$[[Y_2, X_2], X_2] = \sum_{k\in S_C} \left[\left[Q_2^k\left(K_k^{\oplus}\right), H'_{\mathrm{kin}}\right], H'_{\mathrm{kin}}\right]. \tag{8.16}$$

For every symmetric operator B on $\ell^2(L_k^{\pm})$, by (1.74), we compute

$$\begin{aligned}
\left[Q_2^k(B), H'_{\mathrm{kin}}\right] &= \sum_{p,q\in L_k^{\pm}} \langle e_p, Be_q\rangle \left[\left(b^*_{\overline{k,p}} b^*_{\overline{k,q}} + b_{\overline{k,q}} b_{\overline{k,p}}\right), H'_{\mathrm{kin}}\right] \\
&= \sum_{p,q\in L_k^{\pm}} \langle e_p, Be_q\rangle (\lambda_{\overline{k,p}} + \lambda_{\overline{k,q}}) \left(-b^*_{\overline{k,p}} b^*_{\overline{k,q}} + b_{\overline{k,q}} b_{\overline{k,p}}\right), \\
\left[\left[Q_2^k(B), H'_{\mathrm{kin}}\right], H'_{\mathrm{kin}}\right] &= \sum_{p,q\in L_k^{\pm}} \langle e_p, Be_q\rangle (\lambda_{\overline{k,p}} + \lambda_{\overline{k,q}}) \left[\left(-b^*_{\overline{k,p}} b^*_{\overline{k,q}} + b_{\overline{k,q}} b_{\overline{k,p}}\right), H'_{\mathrm{kin}}\right] \\
&= \sum_{p,q\in L_k^{\pm}} \langle e_p, Be_q\rangle (\lambda_{\overline{k,p}} + \lambda_{\overline{k,q}})^2 \left(b^*_{\overline{k,p}} b^*_{\overline{k,q}} + b_{\overline{k,q}} b_{\overline{k,p}}\right).
\end{aligned} \tag{8.17}$$

By the Cauchy–Schwarz inequality, we can estimate

$$\pm\left[\left[Q_2^k(B), H'_{\mathrm{kin}}\right], H'_{\mathrm{kin}}\right] \le \sum_{p,q\in L_k^{\pm}} \left(\epsilon\left(\lambda_{\overline{k,p}} + \lambda_{\overline{k,q}}\right)^4 b^*_{\overline{k,p}} b^*_{\overline{k,q}} b_{\overline{k,q}} b_{\overline{k,p}} + \epsilon^{-1} |\langle e_p, Be_q\rangle|^2\right) \tag{8.18}$$

for all $\epsilon > 0$. From Propositions 4.5, 4.8 and the commutation relations (1.74), (1.75), we have

$$\begin{aligned}
0 &\le \sum_{p,q\in L_k^{\pm}} \lambda_{\overline{k,p}} b^*_{\overline{k,p}} b^*_{\overline{k,q}} b_{\overline{k,q}} b_{\overline{k,p}} \le \sum_{p\in L_k^{\pm}} \lambda_{\overline{k,p}} b^*_{\overline{k,p}} \mathcal{N}_E b_{\overline{k,p}} \le H'_{\mathrm{kin}} \mathcal{N}_E, \\
0 &\le \sum_{p,q\in L_k^{\pm}} \lambda_{\overline{k,q}} b^*_{\overline{k,p}} b^*_{\overline{k,q}} b_{\overline{k,q}} b_{\overline{k,p}} \le \sum_{p\in L_k^{\pm}} b^*_{\overline{k,p}} H'_{\mathrm{kin}} b_{\overline{k,p}} \le H'_{\mathrm{kin}} \mathcal{N}_E.
\end{aligned} \tag{8.19}$$

Moreover, when $k \in S_C = \overline{B}\left(0, k_F^{\gamma}\right) \cap \mathbb{Z}^3_+$ with $1 \ge \gamma > 0$, we have

$$|\lambda_{\overline{k,p}}| \le C|k|k_F, \quad |\lambda_{\overline{k,p}}|^3 \le C|k|^3 k_F^3 \le C|k|^2 k_F^4, \quad \forall p \in L_k^{\pm}. \tag{8.20}$$

Hence, we conclude from (8.18) that

$$\pm\left[\left[Q_2^k(B), H'_{\mathrm{kin}}\right], H'_{\mathrm{kin}}\right] \le C\left(\epsilon|k|^2 k_F^4 H'_{\mathrm{kin}} \mathcal{N}_E + \epsilon^{-1}\|B\|^2_{\mathrm{HS}}\right) \tag{8.21}$$

for all $\epsilon > 0$. Optimizing over ϵ gives

$$\pm\left[\left[Q_2^k(B), H'_{\mathrm{kin}}\right], H'_{\mathrm{kin}}\right] \le C\|B\|_{\mathrm{HS}} |k| k_F \left(H'_{\mathrm{kin}} \mathcal{N}_E + k_F^2\right) \tag{8.22}$$

for all symmetric operators B on $\ell^2(L_k^{\pm})$. Inserting this in (8.16) and using

$$\sum_{k\in S_C} |k|\, \|K_k^{\oplus}\|_{\mathrm{HS}} \le \sum_{k\in S_C} |k| \mathrm{tr}(|K_k^{\oplus}|) = 2\sum_{k\in S_C} |k| \mathrm{tr}\left(|K_k|\right) \le C\sum_{k\in\mathbb{Z}^3} |k|\hat{V}_k \tag{8.23}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

(which is similar to (8.14)), we find that

$$\pm[[Y_1, X_2], X_2] = \pm \sum_{k\in S_C} \left[\left[Q_2^k\left(K_k^{\oplus}\right), H'_{\text{kin}}\right], H'_{\text{kin}}\right] \le C k_F \left(H'_{\text{kin}}\mathcal{N}_E + k_F^2\right) \le C X_1 X_2^2. \tag{8.24}$$

Applying Lemma 8.2, we obtain

$$\pm[[Y_1, \sqrt{X_2}], \sqrt{X_2}] \le C X_1 X_2. \tag{8.25}$$

Putting together (8.8), (8.11), (8.15) and (8.25), we conclude from (8.7) that

$$\pm[\mathcal{K}, X_1X_2] \le C X_1 X_2. \tag{8.26}$$

Thus,

$$\left|\frac{d}{dt}\langle\Psi_t, X_1X_2\Psi_t\rangle\right| = |\langle\Psi_t, [\mathcal{K}, X_1X_2]\Psi_t\rangle| \le C\langle\Psi_t, X_1X_2\Psi_t\rangle. \tag{8.27}$$

By Gronwall's lemma, we have

$$\langle\Psi_t, X_1X_2\Psi_t\rangle \le C\langle\Psi, X_1X_2\Psi\rangle, \quad \forall|t| \le 1. \tag{8.28}$$

This implies the desired bound since $\frac{1}{2}X_1X_2 \le \mathcal{N}_E H'_{\text{kin}} + k_F H'_{\text{kin}} + k_F^2 \le X_1X_2$. Here, we used again Proposition 2.1. □

9. The second Bogolubov transformation

Recall that after the conjugation by $e^{\mathcal{K}}$, up to negligible error terms, we obtain the correlation energy and the operator

$$H'_{\text{kin}} + 2\sum_{k\in S_C} \tilde{Q}_1^k\left(E_k^{\oplus} - h_k^{\oplus}\right). \tag{9.1}$$

In the bosonic analogy, where we informally consider $H'_{\text{kin}} \sim 2\sum_{k\in\mathbb{Z}^3_+} \tilde{Q}_1^k(h_k^{\oplus})$, this expression would be manifestly non-negative as H'_{kin} cancels the negative terms $2\sum_{k\in S_C} \tilde{Q}_1^k(-h_k^{\oplus})$ (and $E_k^{\oplus} > 0$ as $E_k = e^{-K_k}h_k e^{-K_k} > 0$), so this term could be neglected for the lower bound. This analogy is only formal, however. One might still hope that $E_k^{\oplus} - h_k^{\oplus} \ge 0$ since E_k is isospectral to $\widetilde{E}_k$ and $\widetilde{E}_k \ge h_k$, but this fails too; it can be shown that $E_k - h_k$ is indefinite. While these two ideas – the bosonic analogy and the fact that $E_k - h_k \ge 0$ – fail on their own, we will overcome this issue by combining them. In this section, we will carry out another unitary transformation which effectively replaces E_k by $\widetilde{E}_k$ in (9.1).

Consider the unitary transformation $e^{\mathcal{J}} : \mathcal{H}_N \to \mathcal{H}_N$, where $\mathcal{J} : \mathcal{H}_N \to \mathcal{H}_N$ is now of the form

$$\mathcal{J} = \sum_{k\in S_C}\sum_{p,q\in L_k^{\pm}} \left\langle e_p, J_k^{\oplus} e_q\right\rangle b_k^*(e_p) b_k\left(e_q\right) = \sum_{k\in S_C}\sum_{p\in L_k^{\pm}} b_k^*\left(J_k^{\oplus} e_p\right) b_k(e_p), \tag{9.2}$$

where $S_C = \overline{B}\left(0, k_F^{\gamma}\right) \cap \mathbb{Z}^3_+$ with $1 \ge \gamma > 0$ and

$$J_k^{\oplus} = \begin{pmatrix} J_k & 0 \\ 0 & J_k \end{pmatrix}, \quad J_k = \log\left(U_k\right), \quad U_k = \left(h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} h_k^{-\frac{1}{2}} e^{K_k}. \tag{9.3}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Here, $U_k : \ell^2(L_k) \to \ell^2(L_k)$ is the unitary transformation which takes E_k to $\widetilde{E}_k$, namely,

$$U_k E_k U_k^* = \left(h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} \left(h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}}\right)^{\frac{1}{2}} = h_k^{\frac{1}{2}} e^{-2K_k} h_k^{\frac{1}{2}} = \widetilde{E}_k, \tag{9.4}$$

and J_k is the (principal) logarithm of U_k, so that $e^{J_k} = U_k$. Since J_k is skew-symmetric, so are $J_k^{\oplus}$ and $\mathcal{J}$, and hence, $e^{\mathcal{J}}$ is a unitary operator on $\mathcal{H}_N$.

In the exact bosonic case, it is not difficult to see that for every skew-symmetric operator $J : V \to V$, the unitary operator $e^{\mathcal{J}}$ with $\mathcal{J} = \mathrm{d}\Gamma(J) = \sum_i a^*(Je_i) a(e_i)$ is a Bogolubov transformation on $\mathcal{F}^+(V)$ which acts on a second-quantized operator as

$$e^{\mathcal{J}} \mathrm{d}\Gamma(A) e^{-\mathcal{J}} = \mathrm{d}\Gamma\left(e^J A e^{-J}\right). \tag{9.5}$$

Returning to the quasi-bosonic case, we will show that

$$e^{\mathcal{J}} \left(\sum_{k \in S_C} \tilde{Q}_1^k(E_k^{\oplus})\right) e^{-\mathcal{J}} \approx \sum_{k \in S_C} \tilde{Q}_1^k \left(e^{J_k^{\oplus}} E_k^{\oplus} e^{-J_k^{\oplus}}\right) = \sum_{k \in S_C} \tilde{Q}_1^k \left(\widetilde{E}_k^{\oplus}\right) \tag{9.6}$$

up to error terms which are similar to the exchange terms coming from the first transformation. Moreover, although $H'_{\mathrm{kin}} \sim 2\sum_{k \in \mathbb{Z}^3_+} \tilde{Q}_1^k(h_k^{\oplus})$ does not hold precisely, it is valid from the point of view of commutators as explained in (1.72), which results in $H'_{\mathrm{kin}} - 2\sum_{k \in \mathbb{Z}^3_+} \tilde{Q}_1^k(h_k^{\oplus})$ being essentially invariant under the Bogolubov transformation $e^{\mathcal{J}}$. The overall transformation then takes the form

$$e^{\mathcal{J}} \left(H'_{\mathrm{kin}} + 2\sum_{k \in S_C} \tilde{Q}_1^k \left(E_k^{\oplus} - h_k^{\oplus}\right)\right) e^{-\mathcal{J}} \approx H'_{\mathrm{kin}} + 2\sum_{k \in S_C} \tilde{Q}_1^k \left(\widetilde{E}_k^{\oplus} - h_k^{\oplus}\right), \tag{9.7}$$

and we now have the desired non-negative operator $\widetilde{E}_k^{\oplus} - h_k^{\oplus} \geq 0$ on the right-hand side.

While the error terms in (9.7) are similar to those coming from the first transformation, they are in practice more difficult to estimate, for although we derived simple, optimal estimates for the transformation kernels $(K_k)_{k \in S_C}$ in Section 7, we cannot obtain the same for the transformation kernels $(J_k)_{k \in S_C}$. The justification that the second transformation works as claimed will therefore take more effort than was needed for the first transformation.

9.1. Actions on the bosonizable terms

The first step of justifying (9.7) is to prove the following exact equality.

Proposition 9.1. *The unitary transformation $e^{\mathcal{J}} : \mathcal{H}_N \to \mathcal{H}_N$ given in (9.2)-(9.3) satisfies*

$$\begin{aligned} &e^{\mathcal{J}} \left(H'_{\mathrm{kin}} + 2\sum_{k \in S_C} \tilde{Q}_1^k \left(E_k^{\oplus} - h_k^{\oplus}\right)\right) e^{-\mathcal{J}} \\ &= H'_{\mathrm{kin}} + 2\sum_{k \in S_C} \tilde{Q}_1^k \left(\widetilde{E}_k^{\oplus} - h_k^{\oplus}\right) + 2\sum_{k \in S_C} \int_0^1 e^{(1-t)\mathcal{J}} \mathcal{E}_3^k(F_k^{\oplus}(t)) e^{-(1-t)\mathcal{J}} dt, \end{aligned}$$

where for all $k \in \mathbb{Z}^3_*$ and $t \in [0,1]$, we defined the operator $F_k^{\oplus}(t) : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$ by

$$F_k^{\oplus}(t) = \begin{pmatrix} e^{tJ_k} E_k e^{-tJ_k} - h_k & 0 \\ 0 & e^{tJ_k} E_k e^{-tJ_k} - h_k \end{pmatrix},$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and for symmetric $A : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$*, we defined the new exchange operator*

$$\begin{aligned}\mathcal{E}_3^k(A) &= 2\sum_{l\in S_C}\sum_{p\in L_k^\pm}\sum_{q\in L_l^\pm}\operatorname{Re}\left(b_k^*(Ae_p)\varepsilon_{k,l}(e_p;e_q)b_l(J_l^\oplus e_q)\right)\\ &= 2\sum_{l\in S_C}\sum_{p\in L_k^\pm}\sum_{q\in L_l^\pm}\operatorname{Re}\left(b_k^*(Ae_p)\left(\delta_{p,q}c_{\overline{q-k}}c^*_{\overline{p-k}}+\delta_{\overline{p-k},\overline{q-k}}c_q^*c_p\right)b_l(J_l^\oplus e_q)\right).\end{aligned}$$

We will follow the same strategy as we did when we considered the action of the quasi-bosonic Bogolubov transformation on the $Q_1^k(A)$ and $Q_2^k(B)$ terms. First, we calculate the commutator:

Proposition 9.2. *For all* $k \in S_C$ *and symmetric* $A : \ell^2(L_k^\pm) \to \ell^2(L_k^\pm)$*, it holds that*

$$\left[\mathcal{J},\tilde{Q}_1^k(A)\right] = \tilde{Q}_1^k([J_k^\oplus, A]) + \mathcal{E}_3^k(A).$$

Proof. We first calculate, using the commutation relations of the excitation operators $b_k(\varphi)$ and $b_k^*(\varphi)$, that for any $k \in S_C$ and $\varphi \in \ell^2(L_k^\pm)$,

$$\begin{aligned}[\mathcal{J}, b_k(\varphi)] &= -\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left(b_l^*(J_l^\oplus e_q)\left[b_k(\varphi), b_l(e_q)\right] + \left[b_k(\varphi), b_l^*(J_l^\oplus e_q)\right]b_l(e_q)\right)\\ &= -\sum_{l\in S_C}\sum_{q\in L_l^\pm}\left(\delta_{k,l}\left\langle\varphi, J_l^\oplus e_q\right\rangle + \varepsilon_{k,l}\left(\varphi; J_l^\oplus e_q\right)\right)b_l(e_q) \qquad (9.8)\\ &= -\sum_{q\in L_k^\pm}\left\langle\varphi, J_k^\oplus e_q\right\rangle b_k(e_q) - \sum_{l\in S_C}\sum_{q\in L_l^\pm}\varepsilon_{k,l}\left(\varphi; J_l^\oplus e_q\right)b_l(e_q)\\ &= \sum_{q\in L_k^\pm}\left\langle J_k^\oplus\varphi, e_q\right\rangle b_k(e_q) + \sum_{l\in S_C}\sum_{q\in L_l^\pm}\varepsilon_{k,l}\left(\varphi; e_q\right)b_l(J_l^\oplus e_q)\\ &= b_k\left(J_k^\oplus\varphi\right) + \mathcal{E}_k^{\mathcal{J}}(\varphi)\end{aligned}$$

for

$$\mathcal{E}_k^{\mathcal{J}}(\varphi) = \sum_{l\in S_C}\sum_{q\in L_l^\pm}\varepsilon_{k,l}\left(\varphi; e_q\right)b_l(J_l^\oplus e_q), \qquad (9.9)$$

where we used the skew-symmetry of $J_k^\oplus$, anti-linearity of $\varphi \mapsto b_k(\varphi)$, and Lemma 3.3. Consequently, we compute for $\tilde{Q}_1^k(A)$ that

$$\begin{aligned}\left[\mathcal{J},\tilde{Q}_1^k(A)\right] &= \sum_{p\in L_k^\pm}\left[\mathcal{J}, b_k^*(Ae_p)b_k(e_p)\right]\\ &= \sum_{p\in L_k^\pm}\left(b_k^*(Ae_p)\left[\mathcal{J}, b_k(e_p)\right] + \left[\mathcal{J}, b_k(Ae_p)\right]^* b_k(e_p)\right)\\ &= \sum_{p\in L_k^\pm}\left(b_k^*(Ae_p)\left(b_k\left(J_k^\oplus e_p\right) + \mathcal{E}_k^{\mathcal{J}}(e_p)\right) + \left(b_k\left(J_k^\oplus Ae_p\right) + \mathcal{E}_k^{\mathcal{J}}(Ae_p)\right)^* b_k(e_p)\right)\\ &= \sum_{p\in L_k^\pm}\left(b_k^*(Ae_p)b_k\left(J_k^\oplus e_p\right) + b_k^*\left(J_k^\oplus Ae_p\right)b_k(e_p)\right) \qquad (9.10)\\ &\quad + \sum_{p\in L_k^\pm}\left(b_k^*(Ae_p)\mathcal{E}_k^{\mathcal{J}}(e_p) + \left(b_k^*(e_p)\mathcal{E}_k^{\mathcal{J}}(Ae_p)\right)^*\right)\end{aligned}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

$$= \sum_{p\in L_k^\pm} b_k^*\left(\left(J_k^\oplus A - AJ_k^\oplus\right)e_p\right)b_k(e_p) + 2\sum_{l\in S_C}\sum_{p\in L_k^\pm}\sum_{q\in L_l^\pm}\mathrm{Re}\left(b_k^*(Ae_p)\varepsilon_{k,l}(e_p;e_q)b_l(J_l^\oplus e_q)\right)$$
$$= \tilde{Q}_1^k([J_k^\oplus, A]) + \mathcal{E}_3^k(A).$$

□

To derive an expression for $e^{\mathcal{J}}\tilde{Q}_1^k(A)e^{-\mathcal{J}}$, we will use the Baker-Campbell-Hausdorff formula

$$\exp\left(\mathcal{C}_{J_k^\oplus}\right)(A) = \sum_{m=0}^{\infty}\frac{1}{m!}\mathcal{C}_{J_k^\oplus}^m(A) = e^{J_k^\oplus}Ae^{-J_k^\oplus}, \quad \text{with } \mathcal{C}_{J_k^\oplus}(A) = \left[J_k^\oplus, A\right]. \tag{9.11}$$

Imitating the proof of Proposition 5.3, we deduce the following:

Proposition 9.3. *For all $k\in S_C$ and symmetric $A:\ell^2(L_k^\pm)\to\ell^2(L_k^\pm)$, it holds that*

$$e^{\mathcal{J}}\tilde{Q}_1^k(A)e^{-\mathcal{J}} = \tilde{Q}_1^k\left(e^{J_k^\oplus}Ae^{-J_k^\oplus}\right) + \int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k\left(e^{(1-t)J_k^\oplus}Ae^{-(1-t)J_k^\oplus}\right)e^{-t\mathcal{J}}dt,$$

the integrals being Riemann integrals of bounded operators.

Proof. We claim that for any $n\in\mathbb{N}$, it holds that

$$e^{\mathcal{J}}\tilde{Q}_1^k(A)e^{-\mathcal{J}} = \tilde{Q}_1^k\left(\sum_{m=0}^{n-1}\frac{1}{m!}\mathcal{C}_{J_k^\oplus}^m(A)\right) + \int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k\left(\sum_{m=0}^{n-1}\frac{1}{m!}\mathcal{C}_{(1-t)J_k^\oplus}^m(A)\right)e^{-t\mathcal{J}}dt$$
$$+\frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{J}}\tilde{Q}_1^k\left(\mathcal{C}_{J_k^\oplus}^n(A)\right)e^{-t\mathcal{J}}(1-t)^{n-1}dt. \tag{9.12}$$

We proceed by induction. For $n=1$, we have by the fundamental theorem of calculus and Proposition 9.2 that

$$e^{\mathcal{J}}\tilde{Q}_1^k(A)e^{-\mathcal{J}} = \tilde{Q}_1^k(A) + \int_0^1 e^{t\mathcal{J}}\left[\mathcal{J},\tilde{Q}_1^k(A)\right]e^{-t\mathcal{J}}\,dt$$
$$= \tilde{Q}_1^k(A) + \int_0^1 e^{t\mathcal{J}}\tilde{Q}_1^k([J_k^\oplus,A])e^{-t\mathcal{J}}\,dt + \int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k(A)e^{-t\mathcal{J}}\,dt, \tag{9.13}$$

which is the claim. For the inductive step, we assume that case n holds and integrate the last term of equation (9.12) by parts:

$$\frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{J}}\tilde{Q}_1^k\left(\mathcal{C}_{J_k^\oplus}^n(A)\right)e^{-t\mathcal{J}}(1-t)^{n-1}dt$$
$$= \frac{1}{(n-1)!}\left[e^{t\mathcal{J}}\tilde{Q}_1^k\left(\mathcal{C}_{J_k^\oplus}^n(A)\right)e^{-t\mathcal{J}}\left(-\frac{(1-t)^n}{n}\right)\right]_0^1$$
$$-\frac{1}{(n-1)!}\int_0^1 e^{t\mathcal{J}}\left[\mathcal{J},\tilde{Q}_1^k\left(\mathcal{C}_{J_k^\oplus}^n(A)\right)\right]e^{-t\mathcal{J}}\left(-\frac{(1-t)^n}{n}\right)dt \tag{9.14}$$
$$= \frac{1}{n!}\tilde{Q}_1^k\left(\mathcal{C}_{J_k^\oplus}^n(A)\right) + \frac{1}{n!}\int_0^1 e^{t\mathcal{J}}\left(\tilde{Q}_1^k\left(\left[J_k^\oplus,\mathcal{C}_{J_k^\oplus}^n(A)\right]\right) + \mathcal{E}_3^k\left(\mathcal{C}_{J_k^\oplus}^n(A)\right)\right)e^{-t\mathcal{J}}(1-t)^n dt$$
$$= \tilde{Q}_1^k\left(\frac{1}{n!}\mathcal{C}_{J_k^\oplus}^n(A)\right) + \int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k\left(\frac{1}{n!}\mathcal{C}_{(1-t)J_k^\oplus}^n(A)\right)e^{-t\mathcal{J}}\,dt$$
$$+\frac{1}{n!}\int_0^1 e^{t\mathcal{J}}\tilde{Q}_1^k\left(\mathcal{C}_{J_k^\oplus}^{n+1}(A)\right)e^{-t\mathcal{J}}(1-t)^n dt.$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Insertion of this identity into equation (9.12) yields the statement for case $n+1$, so our claim (9.12) holds. We can now take $n \to \infty$ and appeal to equation (9.11) to get the claim. □

Proposition 9.2 also allows us to describe the action of $e^{\mathcal{J}}$ on H'_{kin}:

Proposition 9.4. *It holds that*

$$e^{\mathcal{J}}\left(H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right)\right)e^{-\mathcal{J}} = H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right) - 2\sum_{k\in S_C}\int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k\left(h_k^{\oplus}\right)e^{-t\mathcal{J}}\,dt.$$

Proof. By the fundemental theorem of calculus and the fact that $\partial_t(e^{tA}Be^{-tA}) = e^{tA}[A,B]e^{-tA}$, the left side is equal to

$$H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right) + \int_0^1 e^{t\mathcal{J}}\left[\mathcal{J}, H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right)\right]e^{-t\mathcal{J}}\,dt.$$

Recalling (1.74), we may compute using Lemma 3.3 that

$$\begin{aligned}\left[\mathcal{J}, H'_{\text{kin}}\right] &= -\sum_{k\in S_C}\sum_{p\in L_k^{\pm}}\left[H'_{\text{kin}}, b_k^*\left(J_k^{\oplus}e_p\right)b_k(e_p)\right] \\ &\quad - 2\sum_{k\in S_C}\sum_{p\in L_k^{\pm}}\left(-b_k^*\left(J_k^{\oplus}e_p\right)b_k\left(h_k^{\oplus}e_p\right) + b_k^*\left(h_k^{\oplus}J_k^{\oplus}e_p\right)b_k(e_p)\right) \\ &= 2\sum_{k\in S_C}\sum_{p\in L_k^{\pm}}b_k^*\left(\left(J_k^{\oplus}h_k^{\oplus} - h_k^{\oplus}J_k^{\oplus}\right)e_p\right)b_k(e_p) = 2\sum_{k\in S_C}\tilde{Q}_1^k\left(\left[J_k^{\oplus}, h_k^{\oplus}\right]\right).\end{aligned} \tag{9.15}$$

Combining with Proposition 9.2, we have that

$$\left[\mathcal{J}, H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right)\right] = \left[\mathcal{J}, H'_{\text{kin}}\right] - 2\sum_{k\in S_C}\left[\mathcal{J}, \tilde{Q}_1^k\left(h_k^{\oplus}\right)\right] = -2\sum_{k\in S_C}\mathcal{E}_3^k\left(h_k^{\oplus}\right), \tag{9.16}$$

which implies the claim. □

We can now conclude:

Proof of Proposition 9.1. By the Propositions 9.3 and 9.4, we see that

$$\begin{aligned}&e^{\mathcal{J}}\left(H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right) + 2\sum_{k\in S_C}\tilde{Q}_1^k\left(E_k^{\oplus}\right)\right)e^{-\mathcal{J}} \\ &= H'_{\text{kin}} - 2\sum_{k\in S_C}\tilde{Q}_1^k\left(h_k^{\oplus}\right) - 2\sum_{k\in S_C}\int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k\left(h_k^{\oplus}\right)e^{-t\mathcal{J}}\,dt \\ &\quad + 2\sum_{k\in S_C}\tilde{Q}_1^k\left(e^{J_k^{\oplus}}E_k^{\oplus}e^{-J_k^{\oplus}}\right) + 2\int_0^1 e^{t\mathcal{J}}\mathcal{E}_3^k\left(e^{(1-t)J_k^{\oplus}}E_k^{\oplus}e^{-(1-t)J_k^{\oplus}}\right)e^{-t\mathcal{J}}\,dt \\ &= H'_{\text{kin}} + 2\sum_{k\in S_C}\tilde{Q}_1^k\left(e^{J_k^{\oplus}}E_k^{\oplus}e^{-J_k^{\oplus}} - h_k^{\oplus}\right) + 2\int_0^1 e^{(1-t)\mathcal{J}}\mathcal{E}_3^k\left(e^{tJ_k^{\oplus}}E_k^{\oplus}e^{-tJ_k^{\oplus}} - h_k^{\oplus}\right)e^{-(1-t)\mathcal{J}}\,dt,\end{aligned} \tag{9.17}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where we also reparametrized the integral. From the choice of $J_k^{\oplus}$ in (9.3), we have

$$e^{tJ_k^{\oplus}} E_k^{\oplus} e^{-tJ_k^{\oplus}} - h_k^{\oplus} = \begin{pmatrix} e^{tJ_k} E_k e^{-tJ_k} - h_k & 0 \\ 0 & e^{tJ_k} E_k e^{-tJ_k} - h_k \end{pmatrix} = E_k^{\oplus}(t) \tag{9.18}$$

for all $t \in [0,1]$. Moreover, using $e^{J_k} = U_k$ and (9.4), we get $e^{J_k^{\oplus}} E_k^{\oplus} e^{-tJ_k^{\oplus}} = \widetilde{E}_k^{\oplus}$. □

9.2. Estimates for the exchange terms

Now we estimate the new exchange term $\mathcal{E}_3$ in Proposition 9.1. We have the following:

Proposition 9.5. *For all $k \in \mathbb{Z}^3_+$, symmetric $E : \ell^2(L_k^{\pm}) \to \ell^2(L_k^{\pm})$ and $\Psi \in D\left(H'_{\text{kin}}\right)$, it holds that*

$$\left|\left\langle \Psi, \mathcal{E}_3^k(E)\Psi\right\rangle\right| \le C\left(\max_{p\in L_k^{\pm}} \left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} E e_p\right\|\right)\left(\sum_{l\in S_C}\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}} J_l^{\oplus}\right\|_{\text{HS}}\right)\sqrt{\left\langle \Psi, H'_{\text{kin}}\Psi\right\rangle\left\langle \Psi, \mathcal{N}_E H'_{\text{kin}}\Psi\right\rangle}$$

for a constant $C > 0$ independent of all quantities.

Proof of Proposition 9.5. We can follow the analysis in Section 6. In particular, the same reduction in Section 6.1 applies to $\mathcal{E}_3^k(E)$, but in this case, it is significantly simpler. By definition, up to taking adjoints, every term of $\mathcal{E}_3^k(E)$ immediately reduces to the schematic form

$$\sum_{l\in S_C}\sum_{p\in S} b_k^*\left(E e_{p_1}\right)\tilde{c}_{p_3}^*\tilde{c}_{p_4} b_l\left(J_l^{\oplus} e_{p_2}\right), \tag{9.19}$$

and recalling that commutators of the forms $\left[\tilde{c}_p, b_k(\varphi)\right]$ and $\left[\tilde{c}_p^*, b_k^*(\varphi)\right]$ also vanish, we may normal-order this schematic form without introducing additional terms. Controlling $\mathcal{E}_3^k(E)$ thus reduces entirely to the estimation of the single schematic form

$$\sum_{l\in S_C}\sum_{p\in S} \tilde{c}_{p_3}^* b_k^*\left(E e_{p_1}\right) b_l\left(J_l^{\oplus} e_{p_2}\right)\tilde{c}_{p_4}. \tag{9.20}$$

We estimate the schematic form of equation (9.20) using Proposition 4.4, Lemma 6.6 and the Cauchy-Schwarz inequality:

$$\begin{aligned}
\sum_{l\in S_C}\sum_{p\in S}\left|\left\langle \Psi, \tilde{c}_{p_3}^* b_k^*\left(E e_{p_1}\right) b_l\left(J_l^{\oplus} e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\rangle\right| &\le \sum_{l\in S_C}\sum_{p\in S}\left\|b_k\left(E e_{p_1}\right)\tilde{c}_{p_3}\Psi\right\|\left\|b_l\left(J_l^{\oplus} e_{p_2}\right)\tilde{c}_{p_4}\Psi\right\| \\
&\le \sum_{l\in S_C}\sum_{p\in S}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} E e_{p_1}\right\|\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}} J_l^{\oplus} e_{p_2}\right\|\sqrt{\left\langle \tilde{c}_{p_3}\Psi, H'^{(\pm 1)}_{\text{kin}}\tilde{c}_{p_3}\Psi\right\rangle\left\langle \tilde{c}_{p_4}\Psi, H'^{(\pm 1)}_{\text{kin}}\tilde{c}_{p_4}\Psi\right\rangle} \\
&\le \left(\max_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} E e_p\right\|\right)\sqrt{\left\langle \Psi, H'_{\text{kin}}\Psi\right\rangle}\sum_{l\in S_C}\sqrt{\sum_{p\in S}\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}} J_l^{\oplus} e_{p_2}\right\|^2}\sqrt{\sum_{p\in S}\left\langle \Psi, \tilde{c}_{p_4}^* H'^{(\pm 1)}_{\text{kin}}\tilde{c}_{p_4}\Psi\right\rangle} \\
&\le \left(\max_{p\in L_k^{\pm}}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}} E e_{p_1}\right\|\right)\left(\sum_{l\in S_C}\left\|\left(h_l^{\oplus}\right)^{-\frac{1}{2}} J_l^{\oplus}\right\|_{\text{HS}}\right)\sqrt{\left\langle \Psi, H'_{\text{kin}}\Psi\right\rangle\left\langle \Psi, \mathcal{N}_E H'_{\text{kin}}\Psi\right\rangle}.
\end{aligned} \tag{9.21}$$

□

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

9.3. One-body operator estimates

In this subsection, we derive estimates on the one-body quantities

$$\max_{p\in L_k}\left\|h_k^{-\frac{1}{2}}E_k(t)e_p\right\|,\quad \left\|h_k^{-\frac{1}{2}}J_k\right\|_{\mathrm{HS}},\qquad \left\|h_k^{-\frac{1}{2}}\left[J_k,h_k\right]h_k^{-\frac{1}{2}}\right\|_{\mathrm{HS}},\quad \mathrm{tr}(h_k^{-1/2}(\widetilde{E}_k-h_k)h_k^{-1/2}).\tag{9.22}$$

The first two quantities arise from the analysis of the exchange terms in the previous subsection, while the third quantity will be needed in order to derive Gronwall-type estimates for the kinetic operator. The last one is useful to remove the cutoff S_C on the right-hand side of (9.7) at the end. The estimates we will establish are the following:

Proposition 9.6. *Assume $\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k|k|<\infty$. Then for all $k\in\mathbb{Z}^3_*$, we have*

$$\mathrm{tr}\left(h_k^{-1/2}(\widetilde{E}_k-h_k)h_k^{-1/2}\right)\le C\hat{V}_k.$$

Moreover, if $k\in\overline{B}\left(0,k_F^{\gamma}\right)\cap\mathbb{Z}^3_*$, $0<\gamma<\frac{1}{47}$ and $t\in[0,1]$, it holds that

$$\max_{p\in L_k}\left\|h_k^{-\frac{1}{2}}E_k(t)e_p\right\|\le Ck_F^{-\frac{1}{2}}\left(\hat{V}_k+\hat{V}_k^3|k|^6\log\left(k_F\right)\right),$$
$$\left\|h_k^{-\frac{1}{2}}J_k\right\|_{\mathrm{HS}}\le C(\log k_F)^{\frac{2}{3}}k_F^{-\frac{1}{3}}\hat{V}_k,$$
$$\left\|h_k^{-\frac{1}{2}}\left[J_k,h_k\right]h_k^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\le C\hat{V}_k.$$

Here, the constant $C>0$ is independent of k and k_F.

Proposition 9.6 is the main source of the technical restriction $\gamma<\frac{1}{47}$ which comes from the use of the first bound in Proposition A.3 (we need $\gamma<\frac{4+3\beta}{8-3\beta}$ with $\beta=-\frac{5}{4}$).

As in Section 7, in order to simplify the notation, let $h:V\longrightarrow V$ denote a self-adjoint operator acting on an n-dimensional Hilbert space V, let $(x_i)_{i=1}^n$ denote an eigenbasis for h with eigenvalues $(\lambda_i)_{i=1}^n$, and let $v\in V$ be any vector such that $\langle v,x_i\rangle\ge 0$ for all $1\le i\le n$. As before, we take

$$K=-\frac{1}{2}\log\left(h^{-\frac{1}{2}}\left(h^2+2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}h^{-\frac{1}{2}}\right).\tag{9.23}$$

We will establish general estimates for the operators

$$U=\left(h^{\frac{1}{2}}e^{-2K}h^{\frac{1}{2}}\right)^{\frac{1}{2}}h^{-\frac{1}{2}}e^{K},\quad J=\log(U),\quad E(t)=e^{tJ}e^{-K}he^{-K}e^{-tJ}-h\tag{9.24}$$

and then at the end insert the explicit choice (7.2) to get the desired estimates.

Unlike the case in Section 7, we will now also take V to be a complex Hilbert space. This is not a strictly necessary assumption, but it allows us to streamline the presentation significantly, as it implies that the unitary operator U is diagonalizable and so lets us describe the operators $J=\log(U)$ and e^{tJ} solely in terms of eigenvectors of U.

The main difficulty of the proof of Proposition 9.6 is that we cannot extend the argument leading to matrix element estimates for $e^{-2K}-1$ and $1-e^{2K}$ in Section 7 to handle the operators J and e^{tJ}. Instead, we will utilize a technique which effectively lets us replace relevant quantities of J by these of $U-1$, by exploiting the diagonalizability of U.

We start with the easy part of Proposition 9.6.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proposition 9.7. *With* $\widetilde{E} = \left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}$*, we have*

$$\operatorname{tr}\left(h^{-1/2}(\widetilde{E} - h)h^{-1/2}\right) \le \langle v, h^{-1}v\rangle.$$

Proof. Using (7.21) for $\widetilde{E} = \left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{2}}$, we can write

$$h^{-1/2}(\widetilde{E} - h)h^{-1/2} = \frac{4}{\pi}\int_0^\infty \frac{t^2}{1 + 2\left\langle v, h\left(h^2 + t^2\right)^{-1} v\right\rangle} P_{(h^2+t^2)^{-1}v}\, dt \le \frac{4}{\pi}\int_0^\infty P_{(h^2+t^2)^{-1}v} t^2\, dt. \tag{9.25}$$

Taking the trace and using (7.8), we complete the proof. □

Estimates for U

Let us consider the unitary operator $U : V \longrightarrow V$ defined by

$$U = (h^{\frac{1}{2}} e^{-2K} h^{\frac{1}{2}})^{\frac{1}{2}} h^{-\frac{1}{2}} e^K = (h^2 + 2P_{h^{\frac{1}{2}}v})^{\frac{1}{4}} h^{-\frac{1}{2}} e^K. \tag{9.26}$$

First, the analysis of $(h^2 + 2P_{h^{\frac{1}{2}}v})^{\frac{1}{2}}$ in Section 7 can be extended to $(h^2 + 2P_{h^{\frac{1}{2}}v})^{\frac{1}{4}}$. We have the following:

Proposition 9.8. *For all* $1 \le i, j \le n$*, it holds that*

$$\left|\left\langle x_i, \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{4}} - h^{\frac{1}{2}}\right) x_j\right\rangle\right| \le \frac{2\sqrt{\lambda_i\lambda_j}}{\sqrt{\lambda_i} + \sqrt{\lambda_j}} \frac{\langle x_i, v\rangle \left\langle v, x_j\right\rangle}{\lambda_i + \lambda_j}.$$

Note that by using the integral identity

$$A^{\frac{1}{4}} = \frac{2\sqrt{2}}{\pi}\int_0^\infty \left(1 - \frac{t^4}{A + t^4}\right) dt \tag{9.27}$$

for every self-adjoint non-negative operator A instead of (7.8), we obtain the following analogue of Proposition 7.5:

Proposition 9.9. *Let* $(H, \langle\cdot,\cdot\rangle)$ *be a Hilbert space and let* $A : H \longrightarrow H$ *be a positive self-adjoint operator. Then for any* $x \in H$ *and* $g \in \mathbb{R}$ *such that* $A + gP_x > 0$*, it holds that*

$$(A + gP_x)^{\frac{1}{4}} = A^{\frac{1}{4}} + \frac{2\sqrt{2}g}{\pi}\int_0^\infty \frac{t^4}{1 + g\left\langle v, \left(A + t^4\right)^{-1} v\right\rangle} P_{(A+t^4)^{-1}v}\, dt.$$

Proof of Proposition 9.8. Applying Proposition 9.9 with $A = h^2$, $x = h^{\frac{1}{2}}v$ and $g = 2$, we find

$$\begin{aligned}\left(h^2 + 2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{4}} &= \left(h^2\right)^{\frac{1}{4}} + \frac{4\sqrt{2}}{\pi}\int_0^\infty \frac{t^4}{1 + 2\left\langle h^{\frac{1}{2}}v, \left(h^2 + t^4\right)^{-1} h^{\frac{1}{2}}v\right\rangle} P_{(h^2+t^4)^{-1}h^{\frac{1}{2}}v}\, dt \qquad (9.28)\\ &= h^{\frac{1}{2}} + \frac{4\sqrt{2}}{\pi}\int_0^\infty \frac{t^4}{1 + 2\left\langle v, h\left(h^2 + t^4\right)^{-1} v\right\rangle} P_{h^{\frac{1}{2}}(h^2+t^4)^{-1}v}\, dt,\end{aligned}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and so we can estimate that

$$
\begin{aligned}
0 &\leq \left\langle x_i, \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} - h^{\frac{1}{2}} \right) x_j \right\rangle \\
&= \frac{4\sqrt{2}}{\pi} \int_0^\infty \frac{t^4}{1 + 2\left\langle v, h\left(h^2 + t^4\right)^{-1} v \right\rangle} \left\langle x_i, P_{h^{\frac{1}{2}}(h^2+t^4)^{-1}v} x_j \right\rangle dt \\
&= \frac{4\sqrt{2}}{\pi} \langle x_i, v\rangle \left\langle v, x_j \right\rangle \int_0^\infty \frac{t^4}{1 + 2\left\langle v, h\left(h^2 + t^4\right)^{-1} v \right\rangle} \frac{\sqrt{\lambda_i}}{\lambda_i^2 + t^4} \frac{\sqrt{\lambda_j}}{\lambda_j^2 + t^4} \, dt \qquad (9.29) \\
&\leq \frac{4\sqrt{2}}{\pi} \langle x_i, v\rangle \left\langle v, x_j \right\rangle \int_0^\infty \frac{\sqrt{\lambda_i}}{\lambda_i^2 + t^4} \frac{\sqrt{\lambda_j}}{\lambda_j^2 + t^4} t^4 \, dt = \frac{2\sqrt{\lambda_i \lambda_j}}{\sqrt{\lambda_i} + \sqrt{\lambda_j}} \frac{\langle x_i, v\rangle \left\langle v, x_j \right\rangle}{\lambda_i + \lambda_j},
\end{aligned}
$$

where we also applied the integral identity

$$
\int_0^\infty \frac{\sqrt{a}}{a^2 + t^4} \frac{\sqrt{b}}{b^2 + t^4} t^4 \, dt = \frac{\pi}{2\sqrt{2}} \frac{\sqrt{ab}}{\sqrt{a} + \sqrt{b}} \frac{1}{a + b}, \quad a, b > 0. \tag{9.30}
$$

□

We may then conclude the following:

Proposition 9.10. *For all* $1 \leq i, j \leq n$*, it holds that*

$$
\left| \left\langle x_i, (U - 1) x_j \right\rangle \right|, \left| \left\langle x_i, (U^* - 1) x_j \right\rangle \right| \leq 3 \left(1 + \left\langle v, h^{-1} v \right\rangle \right) \frac{\langle x_i, v\rangle \left\langle v, x_j \right\rangle}{\lambda_i + \lambda_j}.
$$

Proof. As $\left| \left\langle x_i, (U - 1) x_j \right\rangle \right| = \left| \left\langle x_j, (U^* - 1) x_i \right\rangle \right|$ and the claimed estimate is symmetric with respect to i and j, it suffices to consider $U - 1$. We write

$$
\begin{aligned}
U - 1 &= \left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} h^{-\frac{1}{2}} e^K - 1 = \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} - h^{\frac{1}{2}} \right) h^{-\frac{1}{2}} e^K + e^K - 1 \qquad (9.31) \\
&= e^K - 1 + \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} - h^{\frac{1}{2}} \right) h^{-\frac{1}{2}} + \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} - h^{\frac{1}{2}} \right) h^{-\frac{1}{2}} \left(e^K - 1 \right)
\end{aligned}
$$

and estimate each term separately. The first is directly covered by Proposition 7.10, with

$$
\left| \left\langle x_i, \left(e^K - 1 \right) x_j \right\rangle \right| \leq \frac{\langle x_i, v\rangle \left\langle v, x_j \right\rangle}{\lambda_i + \lambda_j}. \tag{9.32}
$$

For the second term, we can by Proposition 9.8 estimate that

$$
\begin{aligned}
\left| \left\langle x_i, \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} - h^{\frac{1}{2}} \right) h^{-\frac{1}{2}} x_j \right\rangle \right| &= \frac{1}{\sqrt{\lambda_j}} \left| \left\langle x_i, \left(\left(h^2 + 2P_{h^{\frac{1}{2}}v} \right)^{\frac{1}{4}} - h^{\frac{1}{2}} \right) x_j \right\rangle \right| \qquad (9.33) \\
&\leq \frac{1}{\sqrt{\lambda_j}} \frac{2\sqrt{\lambda_i \lambda_j}}{\sqrt{\lambda_i} + \sqrt{\lambda_j}} \frac{\langle x_i, v\rangle \left\langle v, x_j \right\rangle}{\lambda_i + \lambda_j} \leq 2 \frac{\langle x_i, v\rangle \left\langle v, x_j \right\rangle}{\lambda_i + \lambda_j}.
\end{aligned}
$$

For the final term, we carry out an orthonormal expansion and apply the previous two estimates to see that

$$\begin{aligned}
&\left|\left\langle x_i, \left(\left(h^2+2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{4}}-h^{\frac{1}{2}}\right)h^{-\frac{1}{2}}\left(e^K-1\right)x_j\right\rangle\right| \\
&\quad \leq \sum_{k=1}^{n}\left|\left\langle x_i, \left(\left(h^2+2P_{h^{\frac{1}{2}}v}\right)^{\frac{1}{4}}-h^{\frac{1}{2}}\right)h^{-\frac{1}{2}}x_k\right\rangle\right|\left|\left\langle x_k, \left(e^K-1\right)x_j\right\rangle\right| \\
&\quad \leq 2\sum_{k=1}^{n}\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\langle v,x_j\rangle}{\lambda_k+\lambda_j} = 2\frac{\langle x_i,v\rangle\langle v,x_j\rangle}{\lambda_i+\lambda_j}\sum_{k=1}^{n}\frac{\lambda_i+\lambda_j}{(\lambda_i+\lambda_k)(\lambda_k+\lambda_j)}|\langle x_k,v\rangle|^2 \\
&\quad \leq 2\frac{\langle x_i,v\rangle\langle v,x_j\rangle}{\lambda_i+\lambda_j}\sum_{k=1}^{n}\frac{|\langle x_k,v\rangle|^2}{\lambda_k} = 2\left\langle v, h^{-1}v\right\rangle\frac{\langle x_i,v\rangle\langle v,x_j\rangle}{\lambda_i+\lambda_j},
\end{aligned} \tag{9.34}$$

where we also applied the elementary inequality

$$\frac{a+b}{(a+c)(c+b)} = \frac{a+b}{c(a+b)+ab+c^2} < \frac{1}{c}, \quad \forall a,b,c>0. \tag{9.35}$$

Combining the estimates now yields the claim. □

Estimates for J

Recall that we defined $J : V \to V$ to be the principal logarithm of U. Since U is a unitary operator on the finite-dimensional complex Hilbert space V, by the spectral theorem it is diagonalizable – that is, there exists an orthonormal basis $(w_j)_{j=1}^n$ for V of eigenstates of U with eigenvalues $\left(e^{i\theta_j}\right)_{j=1}^n$, $\left(\theta_j\right)_{j=1}^n \subset (-\pi,\pi]$, (i.e., $Uw_j = e^{i\theta_j}w_j$ for all $1 \leq j \leq n$). Thus, J can be explicitly written as

$$Jw_j = i\theta_j w_j, \quad 1 \leq j \leq n. \tag{9.36}$$

To estimate the quantity $\left\|h^{-\frac{1}{2}}J\right\|_{\mathrm{HS}}$, we will apply the following:

Proposition 9.11. *It holds that*

$$JJ^* \leq \frac{\pi^2}{4}(U-1)^*(U-1).$$

Proof. We note the elementary inequality

$$|x| \leq \frac{\pi}{2}\sqrt{2(1-\cos(x))} = \frac{\pi}{2}\left|e^{ix}-1\right|, \quad x \in [-\pi,\pi], \tag{9.37}$$

which can be deduced from the fact that $x \mapsto \left|e^{ix}-1\right|$ is an even function and concave on $x \in [0,\pi]$. As the eigenbasis $(w_j)_{j=1}^n$ obeys

$$Uw_j = e^{i\theta_j}w_j, \quad U^*w_j = e^{-i\theta_j}w_j, \quad Jw_j = i\theta_j w_j, \quad J^*w_j = -i\theta_j w_j, \tag{9.38}$$

we can for any $w \in V$ perform an orthonormal expansion in terms of $(w_j)_{j=1}^n$ to see that

$$\begin{aligned}
\langle w, JJ^*w\rangle = \|J^*w\|^2 &= \sum_{j=1}^{n}\left|\theta_j\right|^2\left|\left\langle w_j,w\right\rangle\right|^2 \leq \sum_{j=1}^{n}\left(\frac{\pi}{2}\left|e^{i\theta_j}-1\right|\right)^2\left|\left\langle w_j,w\right\rangle\right|^2 \\
&= \frac{\pi^2}{4}\sum_{j=1}^{n}\left|\left\langle (U^*-1)w_j,w\right\rangle\right|^2 = \frac{\pi^2}{4}\|(U-1)w\|^2 = \frac{\pi^2}{4}\left\langle w,(U-1)^*(U-1)w\right\rangle,
\end{aligned} \tag{9.39}$$

which is the claim. □

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Corollary 9.12. *There exists a universal constant $C > 0$ such that*

$$\left\|h^{-\frac{1}{2}}J\right\|_{\mathrm{HS}} \leq C\left(1+\left\langle v, h^{-1}v\right\rangle\right)\left\langle v, h^{-\frac{3}{2}}v\right\rangle.$$

Proof. By cyclicity of the trace and the estimate of the previous proposition, we have that

$$\begin{aligned}\left\|h^{-\frac{1}{2}}J\right\|_{\mathrm{HS}}^2 &= \operatorname{tr}\left(J^*h^{-1}J\right) = \operatorname{tr}\left(h^{-\frac{1}{2}}JJ^*h^{-\frac{1}{2}}\right)\\ &\leq \frac{\pi^2}{4}\operatorname{tr}\left(h^{-\frac{1}{2}}(U-1)^*(U-1)h^{-\frac{1}{2}}\right) = \frac{\pi^2}{4}\left\|(U-1)h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}^2,\end{aligned} \tag{9.40}$$

and by the matrix element estimate of Proposition 9.10,

$$\begin{aligned}\left\|(U-1)h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^n \left|\left\langle x_i, (U-1)h^{-\frac{1}{2}}x_j\right\rangle\right|^2 = \sum_{i,j=1}^n \frac{1}{\lambda_j}\left|\left\langle x_i,(U-1)x_j\right\rangle\right|^2\\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i,j=1}^n\frac{1}{\lambda_j}\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2 \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i,j=1}^n\frac{1}{\lambda_j}\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i^{\frac{3}{4}}\lambda_j^{\frac{1}{4}}}\right|^2\\ &= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left(\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_j^{\frac{3}{2}}}\right)^2 = C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-\frac{3}{2}}v\right\rangle^2,\end{aligned} \tag{9.41}$$

which gives the claim. □

Next, consider $\|h^{-\frac{1}{2}}[J,h]h^{-\frac{1}{2}}\|_{\mathrm{HS}}$. By the triangle inequality, it suffices to bound $\|h^{-\frac{1}{2}}Jh^{\frac{1}{2}}\|_{\mathrm{HS}}$. Unlike $\|h^{-\frac{1}{2}}J\|_{\mathrm{HS}}$, this is more involved as the presence of factors of h on both sides of J prevents us from combining J and J^* in $\|h^{-\frac{1}{2}}Jh^{\frac{1}{2}}\|_{\mathrm{HS}}^2 = \operatorname{tr}(J^*h^{-1}Jh)$, and so we need to proceed differently. First, we note the following elementary estimate:

Lemma 9.13. *There exists a constant $C > 0$ such that*

$$\left|i\theta - \frac{1}{2}\left(e^{i\theta}-e^{-i\theta}\right)\right| \leq C\left|e^{i\theta}-1\right|^3, \quad \theta \in [-\pi,\pi].$$

Proof. The left-hand side is $|\theta - \sin(\theta)| = O(|\theta|^3)$, while $|\theta| \geq \left|e^{i\theta}-1\right| \geq C^{-1}\theta$. □

Proposition 9.14. *There exists a universal constant $C > 0$ such that*

$$\left\|h^{-\frac{1}{2}}[J,h]h^{-\frac{1}{2}}\right\|_{\mathrm{HS}} \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^3\left(\left\langle v,h^{-1}v\right\rangle + \left\langle v,h^{-\frac{1}{2}}v\right\rangle\left\langle v,h^{-\frac{5}{4}}h\right\rangle^2\right).$$

Proof. It suffices to bound $\|h^{-\frac{1}{2}}Jh^{\frac{1}{2}}\|_{\mathrm{HS}}$. By writing

$$J = \frac{1}{2}(U-1) + \frac{1}{2}(1-U^*) + \widetilde{J}, \quad \widetilde{J} = J - \frac{1}{2}(U-U^*), \tag{9.42}$$

we see by the triangle inequality that

$$\left\|h^{-\frac{1}{2}}Jh^{\frac{1}{2}}\right\|_{\mathrm{HS}} \leq \frac{1}{2}\left\|h^{-\frac{1}{2}}(U-1)h^{\frac{1}{2}}\right\|_{\mathrm{HS}} + \frac{1}{2}\left\|h^{-\frac{1}{2}}(1-U^*)h^{\frac{1}{2}}\right\|_{\mathrm{HS}} + \left\|h^{-\frac{1}{2}}\widetilde{J}h^{\frac{1}{2}}\right\|_{\mathrm{HS}}. \tag{9.43}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

By Proposition 9.10, we have

$$\begin{aligned}\left\|h^{-\frac{1}{2}}(U-1)\,h^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^n \left|\left\langle x_i, h^{-\frac{1}{2}}(U-1)\,h^{\frac{1}{2}} x_j\right\rangle\right|^2 = \sum_{i,j=1}^n \frac{\lambda_j}{\lambda_i}\left|\left\langle x_i,(U-1)\,x_j\right\rangle\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2 \sum_{i,j=1}^n \frac{\lambda_j}{\lambda_i}\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left(\sum_{i=1}^n \frac{|\langle x_i,v\rangle|^2}{\lambda_i}\right)^2 \\ &= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-1}v\right\rangle^2 \end{aligned} \tag{9.44}$$

and likewise for $\|h^{-\frac{1}{2}}(1-U^*)h^{\frac{1}{2}}\|_{\mathrm{HS}}^2$. For $h^{-\frac{1}{2}}\widetilde{J}h^{\frac{1}{2}}$, we instead apply Lemma 9.13 and the Cauchy-Schwarz inequality to see that for any $1 \leq i, j \leq n$,

$$\begin{aligned}\left|\left\langle x_i, h^{-\frac{1}{2}}\widetilde{J}h^{\frac{1}{2}}x_j\right\rangle\right|^2 &= \left|\sum_{k=1}^n\left(i\theta_k-\frac{1}{2}\left(e^{i\theta_k}-e^{-i\theta_k}\right)\right)\left\langle h^{-\frac{1}{2}}x_i,w_k\right\rangle\left\langle w_k,h^{\frac{1}{2}}x_j\right\rangle\right|^2 \\ &\leq C\left(\sum_{k=1}^n\left|e^{i\theta_k}-1\right|^3\left|\left\langle h^{-\frac{1}{2}}x_i,w_k\right\rangle\right|\left|\left\langle w_k,h^{\frac{1}{2}}x_j\right\rangle\right|\right)^2 \\ &\leq C\left(\sum_{k=1}^n\left|e^{i\theta_k}-1\right|^4\left|\left\langle w_k,h^{-\frac{1}{2}}x_i\right\rangle\right|^2\right)\left(\sum_{k=1}^n\left|e^{i\theta_k}-1\right|^2\left|\left\langle w_k,h^{\frac{1}{2}}x_j\right\rangle\right|^2\right) \\ &= C\left(\sum_{k=1}^n\left|\left\langle (U^*-1)^2\,w_k,h^{-\frac{1}{2}}x_i\right\rangle\right|^2\right)\left(\sum_{k=1}^n\left|\left\langle (U^*-1)\,w_k,h^{\frac{1}{2}}x_j\right\rangle\right|^2\right) \\ &= C\left\|(U-1)^2\,h^{-\frac{1}{2}}x_i\right\|^2\left\|(U-1)\,h^{\frac{1}{2}}x_j\right\|^2. \end{aligned} \tag{9.45}$$

Summing over i, j, we obtain

$$\left\|h^{-\frac{1}{2}}\widetilde{J}h^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 \leq C\left\|(U-1)^2\,h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}^2\left\|(U-1)\,h^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2. \tag{9.46}$$

We can now again apply Proposition 9.10 to estimate that

$$\begin{aligned}\left\|(U-1)\,h^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^n\left|\left\langle x_i,(U-1)\,h^{\frac{1}{2}}x_j\right\rangle\right|^2 = \sum_{i,j=1}^n \lambda_j\left|\left\langle x_i,(U-1)\,x_j\right\rangle\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i,j=1}^n\lambda_j\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i,j=1}^n\lambda_j\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i^{\frac{1}{4}}\lambda_j^{\frac{3}{4}}}\right|^2 = C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-\frac{1}{2}}v\right\rangle^2 \end{aligned} \tag{9.47}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and

$$\begin{aligned}\left\|(U-1)^2 h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}^2 &= \sum_{i,j=1}^{n}\left|\left\langle x_i,(U-1)^2 h^{-\frac{1}{2}} x_j\right\rangle\right|^2 \\ &= \sum_{i,j=1}^{n}\frac{1}{\lambda_j}\left|\sum_{k=1}^{n}\langle x_i,(U-1)x_k\rangle\left\langle x_k,(U-1)x_j\right\rangle\right|^2 \\ &\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i,j=1}^{n}\frac{1}{\lambda_j}\left(\sum_{k=1}^{n}\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right)^2 \\ &\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i,j=1}^{n}|\langle x_i,v\rangle|^2\frac{\left|\left\langle x_j,v\right\rangle\right|^2}{\lambda_j}\left(\sum_{k=1}^{n}\frac{\langle v,x_k\rangle}{\lambda_i^{\frac{5}{8}}\lambda_k^{\frac{3}{8}}}\frac{\langle x_k,v\rangle}{\lambda_k^{\frac{7}{8}}\lambda_j^{\frac{1}{8}}}\right)^2 \\ &= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left(\sum_{i=1}^{n}\frac{|\langle x_i,v\rangle|^2}{\lambda_i^{\frac{5}{4}}}\right)^4 = C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-\frac{5}{4}}v\right\rangle^4 \end{aligned}\tag{9.48}$$

so

$$\|h^{-\frac{1}{2}}\widetilde{J}h^{\frac{1}{2}}\|_{\mathrm{HS}} \le \left(1+\left\langle v,h^{-1}v\right\rangle\right)^3\left\langle v,h^{-\frac{1}{2}}v\right\rangle\left\langle v,h^{-\frac{5}{4}}v\right\rangle^2. \tag{9.49}$$

Combining the estimates yields the claim. □

Remark 9.1 (Remarks on the estimation technique). As we will use the same approach to obtain estimates on $E(t)$, let us consider the technique of the proof in detail. The idea is that, as we have a good estimate for the matrix elements of $U-1$ and U^*-1, we should attempt to express our operator solely in terms of these. The first step is therefore to decompose J as in (9.42). The error term $\widetilde{J}=J-\frac{1}{2}(U-U^*)$ cannot be simplified further in terms of U but by orthonormal expansion and Lemma 9.13, we can nonetheless estimate it solely in terms of $U-1$, despite being unable to apply an operator inequality, as we did for $\|h^{-\frac{1}{2}}J\|_{\mathrm{HS}}$, to 'substitute' $U-1$ for J directly. The utility of the estimate (9.46) is thus that it allows us to replace the unknown error operator with factors of $U-1$, which we can estimate well. The downside to this is that it simultanously 'decouples' the $h^{-\frac{1}{2}}$ and $h^{\frac{1}{2}}$ factors, which prevents us from exploiting the cancellation between these.

This decoupling is also the reason why it is important that in (9.46) we distribute two factors of $U-1$ to $h^{-\frac{1}{2}}$ rather than only one. One can by the same argument estimate that

$$\begin{aligned}\|h^{-\frac{1}{2}}\widetilde{J}h^{\frac{1}{2}}\|_{\mathrm{HS}} &\le C\|(U-1)h^{-\frac{1}{2}}\|_{\mathrm{HS}}\|(U-1)^2h^{\frac{1}{2}}\|_{\mathrm{HS}} \\ &\le C\left(1+\left\langle v,h^{-3}v\right\rangle\right)^3\left\langle v,h^{-\frac{3}{4}}v\right\rangle^2\left\langle v,h^{-\frac{3}{2}}v\right\rangle,\end{aligned}\tag{9.50}$$

but in Proposition A.3, we only have the good estimates $\left\langle v_k,h_k^{\alpha}v_k\right\rangle\sim Ck_F^{1+\alpha}$ for $\alpha>-\frac{4}{3}$, which makes (9.50) a worse estimate due to the $\left\langle v,h^{-\frac{3}{2}}v\right\rangle$ factor. There is therefore a limit to how low the exponent α can be without affecting our estimates, and so it is advantageous to distribute the factors of $U-1$ such that the overall minimal exponent is not too small.

Estimation of $E(t)$

We now estimate $\max_j\|h^{-\frac{1}{2}}E(t)x_j\|$ using the technique outlined above. First, we decompose

$$E(t)=e^{tJ}e^{-K}he^{-K}e^{-tJ}-h=(e^{tJ}he^{-tJ}-h)+e^{tJ}(e^{-K}he^{-K}-h)e^{-tJ}=:E_1(t)+E_2(t) \tag{9.51}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and using the algebraic identity

$$ABC = B + (A-1)B + B(C-1) + (A-1)B(C-1) \tag{9.52}$$

with $A = e^{tJ}$, $B = h$ and $C = e^{-tJ}$ further decompose $E_1(t)$ as

$$\begin{aligned} E_1(t) = e^{tJ}he^{-tJ} - h &= \big((e^{tJ}-1)h + h(e^{-tJ}-1)\big) + (e^{tJ}-1)h(e^{-tJ}-1) \\ &=: E_{1,1}(t) + E_{1,2}(t). \end{aligned} \tag{9.53}$$

Defining $E_0 = E(0) = e^{-K}he^{-K} - h$, we likewise decompose $E_2(t)$ according to

$$\begin{aligned} E_2(t) = e^{tJ}E_0e^{-tJ} &= E_0 + \big((e^{tJ}-1)E_0 + E_0(e^{-tJ}-1)\big) + (e^{tJ}-1)E_0(e^{-tJ}-1) \\ &=: E_0 + E_{2,1}(t) + E_{2,2}(t). \end{aligned} \tag{9.54}$$

The $E_{1,1}(t)$, $E_{1,2}(t)$ and $E_{2,1}(t)$, $E_{2,2}(t)$ terms differ only in replacing the operator h by E_0. We can therefore estimate these terms similarly, provided we have an estimate on E_0. This is given by the following:

Proposition 9.15. *For all $1 \le i, j \le n$, it holds that*

$$\left|\langle x_i, E_0 x_j\rangle\right| = \left|\langle x_i, (e^{-K}he^{-K} - h)x_j\rangle\right| \le \left(1 + \langle v, h^{-1}v\rangle\right)\langle x_i, v\rangle\langle v, x_j\rangle.$$

Consequently,

$$\max_{1\le j\le n} \left\|h^{-\frac{1}{2}}E_0x_j\right\| \le \alpha\left(1 + \langle v, h^{-1}v\rangle\right)\sqrt{\langle v, h^{-1}v\rangle},$$

where $\alpha = \max_{1\le j\le n}\langle v, x_j\rangle$.

Proof. Using the identity of equation (9.52) with $A = e^{-K} = C$ and $B = h$, we have that

$$e^{-K}he^{-K} - h = \{h, e^{-K} - 1\} + (e^{-K}-1)h(e^{-K}-1). \tag{9.55}$$

Hence,

$$\langle x_i, (e^{-K}he^{-K} - h)x_j\rangle = (\lambda_i + \lambda_j)\langle x_i, (e^{-K}-1)x_j\rangle + \langle x_i, (e^{-K}-1)h(e^{-K}-1)x_j\rangle. \tag{9.56}$$

We can apply Proposition 7.10 to estimate the first term of this equation as

$$\left|(\lambda_i + \lambda_j)\langle x_i, (e^{-K}-1)x_j\rangle\right| \le (\lambda_i + \lambda_j)\frac{\langle x_i, v\rangle\langle v, x_j\rangle}{\lambda_i + \lambda_j} = \langle x_i, v\rangle\langle v, x_j\rangle \tag{9.57}$$

and the second term as

$$\begin{aligned} \left|\langle x_i, (e^{-K}-1)h(e^{-K}-1)x_j\rangle\right| &= \left|\sum_{k=1}^n \lambda_k\langle x_i, (e^{-K}-1)x_k\rangle\langle x_k, (e^{-K}-1)x_j\rangle\right| \\ &\le \sum_{k=1}^n \lambda_k \frac{\langle x_i, v\rangle\langle v, x_k\rangle}{\lambda_i + \lambda_k}\frac{\langle x_k, v\rangle\langle v, x_j\rangle}{\lambda_k + \lambda_j} \\ &\le \langle x_i, v\rangle\langle v, x_j\rangle \sum_{k=1}^n \frac{\left|\langle x_k, v\rangle\right|^2}{\lambda_k} = \langle v, h^{-1}v\rangle\langle x_i, v\rangle\langle v, x_j\rangle, \end{aligned} \tag{9.58}$$

which implies the first claim. Consequently,

$$\left\|h^{-\frac{1}{2}}E_0x_j\right\|^2 = \sum_{i=1}^n \left|\left\langle x_i, h^{-\frac{1}{2}}\left(e^{-K}he^{-K}-h\right)x_j\right\rangle\right|^2 = \sum_{i=1}^n \frac{1}{\lambda_i}\left|\left\langle x_i, \left(e^{-K}he^{-K}-h\right)x_j\right\rangle\right|^2$$
$$\leq \left(1+\left\langle v, h^{-1}v\right\rangle\right)^2 \sum_{i=1}^n \frac{1}{\lambda_i}\left|\langle x_i, v\rangle\left\langle v, x_j\right\rangle\right|^2 \leq \alpha^2\left(1+\left\langle v, h^{-1}v\right\rangle\right)^2\left\langle v, h^{-1}v\right\rangle. \tag{9.59}$$

□

Now it remains to consider the operators $e^{tJ}-1$ and $e^{-tJ}-1=\left(e^{tJ}-1\right)^*$. To implement the above estimation technique, from the following analogue of Lemma 9.13,

$$\left|(e^{it\theta}-1)-t(e^{i\theta}-1)+\frac{t(1-t)}{2}(e^{i\theta}+e^{-i\theta}-2)\right| \leq C\left|e^{i\theta}-1\right|^3, \quad t\in[0,1]\,,\ \theta\in[-\pi,\pi]\,, \tag{9.60}$$

we are motivated in approximating $e^{tJ}-1$ by

$$F_t = t\,(U-1) - \frac{t(1-t)}{2}\,(U+U^*-2)\,, \quad t\in[0,1]\,, \tag{9.61}$$

with the error term being cubic with respect to $U-1$. We then have the following bounds for F_t and the associated error terms:

Proposition 9.16. *For any $T : V \to V$, $x \in V$, $m \in \{1,2\}$ and $t \in [0,1]$, it holds that*

$$\left\|T(e^{tJ}-1-F_t)x\right\|, \left\|T(e^{-tJ}-1-F_t^*)x\right\| \leq C\left\|T(U-1)^m\right\|_{\text{HS}}\left\|(U-1)^{3-m}x\right\|$$

and for all $1 \leq i, j \leq n$, $t \in [0,1]$,

$$\left|\left\langle x_i, F_t x_j\right\rangle\right|, \left|\left\langle x_i, F_t^* x_j\right\rangle\right| \leq C\left(1+\left\langle v, h^{-1}v\right\rangle\right)\frac{\langle x_i, v\rangle\left\langle v, x_j\right\rangle}{\lambda_i+\lambda_j}$$

for a constant $C>0$ independent of all quantities.

Proof. Recall that $(w_j)_{j=1}^n$ is an orthonormal eigenbasis of J, namely, $e^{tJ}w_j = e^{it\theta_j}w_j$ for all $1 \leq j \leq n$. Using (9.60) and the Cauchy-Schwarz inequality, we have that

$$\left\|T\left(e^{tJ}-1-F_t\right)x\right\|^2 = \sum_{j=1}^n \left|\left\langle w_j, T\left(e^{tJ}-1-F_t\right)x\right\rangle\right|^2$$
$$= \sum_{j=1}^n \left|\sum_{k=1}^n \left(\left(e^{it\theta_k}-1\right)-t\left(e^{i\theta_k}-1\right)+\frac{t(1-t)}{2}\left(e^{i\theta_k}+e^{-i\theta_k}-2\right)\right)\left\langle w_j, Tw_k\right\rangle\langle w_k, x\rangle\right|^2$$
$$\leq C\sum_{j=1}^n\left(\sum_{k=1}^n\left|e^{i\theta_k}-1\right|^3\left|\left\langle w_j, Tw_k\right\rangle\right|\left|\langle w_k, x\rangle\right|\right)^2$$
$$\leq C\sum_{j=1}^n\left(\sum_{k=1}^n\left|e^{i\theta_k}-1\right|^{2m}\left|\left\langle w_j, Tw_k\right\rangle\right|^2\right)\left(\sum_{k=1}^n\left|e^{i\theta_k}-1\right|^{2(3-m)}\left|\langle w_k, x\rangle\right|^2\right)$$
$$= C\left(\sum_{j,k=1}^n\left|\left\langle w_j, T(U-1)^m w_k\right\rangle\right|^2\right)\left(\sum_{k=1}^n\left|\left\langle (U^*-1)^{3-m}w_k, x\right\rangle\right|^2\right)$$
$$= C\left\|T(U-1)^m\right\|_{\text{HS}}^2\left\|(U-1)^{3-m}x\right\|^2, \tag{9.62}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

the same estimate holding also for $\left\|T\left(e^{-tJ}-1-F_t^*\right)x\right\|$. For the matrix element estimate of F_t, we have by Proposition 9.10 that

$$\begin{aligned}\left|\left\langle x_i, F_t x_j\right\rangle\right| &= \left|\left\langle x_i, \left(\frac{t\,(1+t)}{2}\,(U-1) - \frac{t(1-t)}{2}\,(U^*-1)\right)x_j\right\rangle\right| \\ &\leq \frac{t\,(1+t)}{2}\left|\left\langle x_i,(U-1)\,x_j\right\rangle\right| + \frac{t(1-t)}{2}\left|\left\langle x_i,(U^*-1)\,x_j\right\rangle\right| \\ &\leq C\left(1+\left\langle v, h^{-1}v\right\rangle\right)\frac{\langle x_i, v\rangle\left\langle v, x_j\right\rangle}{\lambda_i+\lambda_j} \end{aligned} \tag{9.63}$$

as we only consider $t\in[0,1]$, and likewise for $\left|\left\langle x_i, F_t^* x_j\right\rangle\right|$. □

Estimation of $E_1(t)$

We are now ready to estimate $E_1(t) = E_{1,1}(t)+E_{1,2}(t)$, starting with $E_{1,1}(t) = \left(e^{tJ}-1\right)h + h\left(e^{-tJ}-1\right)$. Recall that $(x_i)_i$ are an eigenbasis of h with $\langle x_i, v\rangle \geq 0$ for all $1\leq i\leq n$.

Proposition 9.17. *For all $t\in[0,1]$, it holds that*

$$\begin{aligned}\max_{1\leq j\leq n}\left\|h^{-\frac12}E_{1,1}(t)x_j\right\| &\leq C\alpha\left(1+\left\langle v,h^{-1}v\right\rangle\right)\sqrt{\left\langle v,h^{-1}v\right\rangle} \\ &\quad + C\alpha\left(1+\left\langle v,h^{-1}v\right\rangle\right)^3\left(\|v\|\left\langle v,h^{-\frac54}h\right\rangle^2+\left\langle v,h^{-\frac12}v\right\rangle\left\langle v,h^{-\frac43}v\right\rangle^{\frac32}\right),\end{aligned}$$

where $\alpha = \max_{1\leq j\leq n}\left\langle v, x_j\right\rangle$ and $C>0$ is a constant independent of all quantities.

Proof. We write

$$E_{1,1}(t) = F_t h + hF_t^* + (e^{tJ}-1-F_t)h + h(e^{-tJ}-1-F_t^*) \tag{9.64}$$

so that for any $1\leq j\leq n$, we can estimate by Proposition 9.16

$$\begin{aligned}\left\|h^{-\frac12}E_{1,1}(t)x_j\right\| &\leq \left\|h^{-\frac12}F_t h x_j\right\| + \left\|h^{\frac12}F_t^* x_j\right\| + C\left\|h^{-\frac12}\,(U-1)^2\right\|_{\mathrm{HS}}\left\|(U-1)\,hx_j\right\| \\ &\quad + C\left\|h^{\frac12}\,(U-1)\right\|_{\mathrm{HS}}\left\|(U-1)^2x_j\right\|.\end{aligned} \tag{9.65}$$

We consider each term above for the following. By Proposition 9.16, we see that independently of $1\leq j\leq n$,

$$\begin{aligned}\left\|h^{-\frac12}F_t h x_j\right\|^2 &= \sum_{i=1}^n\frac{\lambda_j^2}{\lambda_i}\left|\left\langle x_i, F_t x_j\right\rangle\right|^2 \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i=1}^n\frac{\lambda_j^2}{\lambda_i}\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2 \\ &\leq C\left|\left\langle v,x_j\right\rangle\right|^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i} \leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-1}v\right\rangle, \\ \left\|h^{\frac12}F_t^* x_j\right\|^2 &= \sum_{i=1}^n\lambda_i\left|\left\langle x_i,F_t^*x_j\right\rangle\right|^2 \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i=1}^n\lambda_i\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2 \\ &\leq C\left|\left\langle v,x_j\right\rangle\right|^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i} \leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-1}v\right\rangle.\end{aligned} \tag{9.66}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

For the remaining terms of equation (9.65), we recall that we already estimated $\|h^{-\frac{1}{2}}(U-1)^2\|_{\mathrm{HS}}$ and $\|h^{\frac{1}{2}}(U-1)\|_{\mathrm{HS}}$ in the equations (9.47) and (9.48) to be

$$\left\|h^{\frac{1}{2}}(U-1)\right\|_{\mathrm{HS}}=\left\|(U-1)h^{\frac{1}{2}}\right\|_{\mathrm{HS}}\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)\left\langle v,h^{-\frac{1}{2}}v\right\rangle, \tag{9.67}$$
$$\left\|h^{-\frac{1}{2}}(U-1)^2\right\|_{\mathrm{HS}}=\left\|(U-1)^2h^{-\frac{1}{2}}\right\|_{\mathrm{HS}}\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-\frac{5}{4}}v\right\rangle^2,$$

the equalities holding by normality of U. The only unknown quantities are thus $\|(U-1)hx_j\|$ and $\|(U-1)^2x_j\|$, which we estimate using Proposition 9.10 as

$$\begin{aligned}\left\|(U-1)hx_j\right\|^2&=\sum_{i=1}^n\lambda_j^2\left|\left\langle x_i,(U-1)x_j\right\rangle\right|^2\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i=1}^n\lambda_j^2\left|\frac{\langle x_i,v\rangle\left\langle v,x_j\right\rangle}{\lambda_i+\lambda_j}\right|^2\\&\le C\left|\left\langle v,x_j\right\rangle\right|^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\sum_{i=1}^n|\langle x_i,v\rangle|^2\le C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\|v\|^2,\end{aligned} \tag{9.68}$$

$$\begin{aligned}\left\|(U-1)^2x_j\right\|^2&=\sum_{i=1}^n\left|\sum_{k=1}^n\langle x_i,(U-1)x_k\rangle\left\langle x_k,(U-1)x_j\right\rangle\right|^2\\&\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i=1}^n\left|\sum_{k=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2\\&\le C\left|\left\langle v,x_j\right\rangle\right|^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i=1}^n|\langle x_i,v\rangle|^2\left(\sum_{k=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_i^{\frac{2}{3}}\lambda_k^{\frac{4}{3}}}\right)^2\\&\le C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-\frac{4}{3}}v\right\rangle^3.\end{aligned} \tag{9.69}$$

Thus,

$$\|h^{-\frac{1}{2}}(U-1)^2\|_{\mathrm{HS}}\|(U-1)hx_j\|\le C\alpha(1+\langle v,h^{-1}v\rangle)^3\|v\|\langle v,h^{-\frac{5}{4}}h\rangle^2$$
$$\|h^{\frac{1}{2}}(U-1)\|_{\mathrm{HS}}\|(U-1)^2x_j\|\le C\alpha(1+\langle v,h^{-1}v\rangle)^3\langle v,h^{-\frac{1}{2}}v\rangle\langle v,h^{-\frac{4}{3}}v\rangle^{\frac{3}{2}}, \tag{9.70}$$

which, upon combination with the estimates of equation (9.66), imply the claim. □

Proposition 9.18. *For all $t\in[0,1]$, it holds that*

$$\begin{aligned}&(C\alpha)^{-1}\max_{1\le j\le n}\left\|h^{-\frac{1}{2}}E_{1,2}(t)x_j\right\|\\&\le\left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-1}v\right\rangle^{\frac{3}{2}}+\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\left\langle v,h^{-\frac{1}{2}}v\right\rangle^2\left\langle v,h^{-\frac{5}{4}}h\right\rangle^2\left\langle v,h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}}\\&+\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left(\sqrt{\left\langle v,h^{-1}v\right\rangle}\left\langle v,h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}}\left\langle v,h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}}+\left\langle v,h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}}\left\langle v,h^{-\frac{5}{4}}v\right\rangle^2\right),\end{aligned}$$

where $\alpha=\max_{1\le j\le n}\left\langle v,x_j\right\rangle$ and $C>0$ is a constant independent of all quantities.

Proof. We write $E_{1,2}(t)=\left(e^{tJ}-1\right)h\left(e^{-tJ}-1\right)$ as

$$E_{1,2}(t)=F_thF_t^*+F_th\left(e^{-tJ}-1-F_t^*\right)+\left(e^{tJ}-1-F_t\right)hF_t^*+\left(e^{tJ}-1-F_t\right)h\left(e^{-tJ}-1-F_t^*\right) \tag{9.71}$$

and see by Proposition 9.16 that

$$\begin{aligned}\left\|h^{-\frac{1}{2}}E_{1,2}(t)x_j\right\| \leq \left\|h^{-\frac{1}{2}}F_t h F_t^* x_j\right\| + C\left\|h^{-\frac{1}{2}}F_t h\,(U-1)\right\|_{\mathrm{HS}}\left\|(U-1)^2 x_j\right\| \\ + C\left\|h^{-\frac{1}{2}}\,(U-1)^2\right\|_{\mathrm{HS}}\left\|(U-1)\,hF_t^* x_j\right\| \\ + C\left\|h^{-\frac{1}{2}}\,(U-1)^2\right\|_{\mathrm{HS}}\left\|(U-1)\,h\,(U-1)\right\|_{\mathrm{HS}}\left\|(U-1)^2 x_j\right\|.\end{aligned} \tag{9.72}$$

We estimate by Propositions 9.10 and 9.16 that

$$\begin{aligned}\left\|h^{-\frac{1}{2}}F_t h F_t^* x_j\right\|^2 &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i=1}^n \frac{1}{\lambda_i}\left|\sum_{k=1}^n \lambda_k \frac{\langle x_i,v\rangle\,\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\,\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2 \\ &\leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i=1}^n \frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left(\sum_{k=1}^n \frac{|\langle x_k,v\rangle|^2}{\lambda_k}\right)^2 = C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-1}v\right\rangle^3,\end{aligned} \tag{9.73}$$

and

$$\begin{aligned}\left\|h^{-\frac{1}{2}}F_t h\,(U-1)\right\|_{\mathrm{HS}}^2 &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i,j=1}^n \frac{1}{\lambda_i}\left|\sum_{k=1}^n \lambda_k \frac{\langle x_i,v\rangle\,\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\,\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i,j=1}^n \frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left|\left\langle x_j,v\right\rangle\right|^2\left(\sum_{k=1}^n \frac{|\langle x_k,v\rangle|^2}{\lambda_k^{\frac{2}{3}}\lambda_j^{\frac{1}{3}}}\right)^2 \\ &= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-1}v\right\rangle\left\langle v,h^{-\frac{2}{3}}v\right\rangle^3,\end{aligned} \tag{9.74}$$

and

$$\begin{aligned}\left\|(U-1)\,hF_t^* x_j\right\|^2 &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i=1}^n \left|\sum_{k=1}^n \lambda_k \frac{\langle x_i,v\rangle\,\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\,\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2 \\ &\leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i=1}^n |\langle x_i,v\rangle|^2\left(\sum_{k=1}^n \frac{|\langle x_k,v\rangle|^2}{\lambda_i^{\frac{1}{3}}\lambda_k^{\frac{2}{3}}}\right)^2 = C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-\frac{2}{3}}v\right\rangle^3,\end{aligned} \tag{9.75}$$

and

$$\begin{aligned}\left\|(U-1)\,h\,(U-1)\right\|_{\mathrm{HS}}^2 &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i,j=1}^n \left|\sum_{k=1}^n \lambda_k \frac{\langle x_i,v\rangle\,\langle v,x_k\rangle}{\lambda_i+\lambda_k}\frac{\langle x_k,v\rangle\,\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4 \sum_{i,j=1}^n |\langle x_i,v\rangle|^2\left|\left\langle x_j,v\right\rangle\right|^2\left(\sum_{k=1}^n \frac{|\langle x_k,v\rangle|^2}{\lambda_i^{\frac{1}{4}}\lambda_j^{\frac{1}{4}}\lambda_k^{\frac{1}{2}}}\right)^2 \\ &= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-\frac{1}{2}}v\right\rangle^4.\end{aligned} \tag{9.76}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Combining these with the estimates of the equations (9.67) and (9.69) yields

$$\left\|h^{-\frac{1}{2}} F_t h\,(U-1)\right\|_{\mathrm{HS}} \left\|(U-1)^2 x_j\right\| \le C\alpha \left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \sqrt{\left\langle v, h^{-1}v\right\rangle} \left\langle v, h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}} \left\langle v, h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}},$$
$$\left\|h^{-\frac{1}{2}}\,(U-1)^2\right\|_{\mathrm{HS}} \left\|(U-1)\,hF_t^* x_j\right\| \le C\alpha \left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \left\langle v, h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}} \left\langle v, h^{-\frac{5}{4}}v\right\rangle^{2} \tag{9.77}$$

and

$$\left\|h^{-\frac{1}{2}}\,(U-1)^2\right\|_{\mathrm{HS}} \left\|(U-1)\,h\,(U-1)\right\|_{\mathrm{HS}} \left\|(U-1)^2 x_j\right\|$$
$$\le C\alpha \left(1+\left\langle v, h^{-1}v\right\rangle\right)^6 \left\langle v, h^{-\frac{1}{2}}v\right\rangle^{2} \left\langle v, h^{-\frac{5}{4}}v\right\rangle^{2} \left\langle v, h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}}, \tag{9.78}$$

which imply the claim. □

Estimation of $E_2(t)$

We now repeat the same steps for $E_2(t) = E_0 + E_{2,1}(t) + E_{2,2}(t)$ where

$$E_{2,1}(t) = F_t E_0 + E_0 F_t^* + \left(e^{tJ} - 1 - F_t\right) E_0 + E_0 \left(e^{-tJ} - 1 - F_t^*\right). \tag{9.79}$$

Proposition 9.19. *For all $t \in [0,1]$, it holds that*

$$\max_{1\le j\le n} \left\|h^{-\frac{1}{2}} E_{2,1}(t)x\right\| \le C\alpha \left(1+\left\langle v, h^{-1}v\right\rangle\right)^2 \left\langle v, h^{-1}v\right\rangle^{\frac{3}{2}}$$
$$+ C\alpha \left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \left\langle v, h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}} \left\langle v, h^{-\frac{5}{4}}v\right\rangle^{2}$$
$$+ C\alpha \left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \sqrt{\left\langle v, h^{-1}v\right\rangle} \left\langle v, h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}} \left\langle v, h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}},$$

where $\alpha = \max_{1\le j\le n} \left\langle v, x_j\right\rangle$ and $C > 0$ is a constant independent of all quantities.

Proof. By Proposition 9.16, we can estimate that

$$\left\|h^{-\frac{1}{2}} E_{2,1}(t)x_j\right\| \le \left\|h^{-\frac{1}{2}} F_t E_0 x_j\right\| + \left\|h^{-\frac{1}{2}} E_0 F_t^* x_j\right\| + C\left\|h^{-\frac{1}{2}}\,(U-1)^2\right\|_{\mathrm{HS}} \left\|(U-1)\,E_0 x_j\right\|$$
$$+ C\left\|h^{-\frac{1}{2}} E_0\,(U-1)\right\|_{\mathrm{HS}} \left\|(U-1)^2 x_j\right\|. \tag{9.80}$$

Then let us consider each term separately. By the Propositions 9.10, 9.15, 9.16, we have that

$$\left\|h^{-\frac{1}{2}} F_t E_0 x_j\right\|^2 \le C\left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \sum_{i=1}^n \frac{1}{\lambda_i} \left|\sum_{k=1}^n \frac{\langle x_i, v\rangle \langle v, x_k\rangle}{\lambda_i + \lambda_k} \langle x_k, v\rangle \left\langle v, x_j\right\rangle\right|^2$$
$$\le C\alpha^2 \left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \sum_{i=1}^n \frac{|\langle x_i, v\rangle|^2}{\lambda_i} \left(\sum_{k=1}^n \frac{|\langle x_k, v\rangle|^2}{\lambda_k}\right)^2$$
$$= C\alpha^2 \left(1+\left\langle v, h^{-1}v\right\rangle\right)^4 \left\langle v, h^{-1}v\right\rangle^3, \tag{9.81}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and

$$\left\|h^{-\frac{1}{2}}E_0F_tx_j\right\|^2 \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k=1}^n\langle x_i,v\rangle\langle v,x_k\rangle\frac{\langle x_k,v\rangle\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2$$
$$\leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left(\sum_{k=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_k}\right)^2 = C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-1}v\right\rangle^3, \tag{9.82}$$

and

$$\left\|(U-1)\,E_0x_j\right\|^2 \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i=1}^n\left|\sum_{k=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\langle x_k,v\rangle\left\langle v,x_j\right\rangle\right|^2$$
$$\leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i=1}^n|\langle x_i,v\rangle|^2\left(\sum_{k=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_i^{\frac{1}{3}}\lambda_k^{\frac{2}{3}}}\right)^2$$
$$= C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-\frac{2}{3}}v\right\rangle^3, \tag{9.83}$$

and

$$\left\|h^{-\frac{1}{2}}E_0\,(U-1)\right\|_{\mathrm{HS}}^2 \leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i,j=1}^n\frac{1}{\lambda_i}\left|\sum_{k=1}^n\langle x_i,v\rangle\langle v,x_k\rangle\frac{\langle x_k,v\rangle\left\langle v,x_j\right\rangle}{\lambda_k+\lambda_j}\right|^2$$
$$\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\sum_{i,j=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left|\left\langle x_j,v\right\rangle\right|^2\left(\sum_{k=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_k^{\frac{2}{3}}\lambda_j^{\frac{1}{3}}}\right)^2$$
$$= C(1+\langle v,h^{-1}v\rangle)^4\langle v,h^{-1}v\rangle\langle v,h^{-\frac{2}{3}}v\rangle^3. \tag{9.84}$$

Combining these with our prior estimates that

$$\left\|h^{-\frac{1}{2}}(U-1)^2\right\|_{\mathrm{HS}} \leq C\big(1+\left\langle v,h^{-1}v\right\rangle\big)^2\left\langle v,h^{-\frac{5}{4}}v\right\rangle^2,$$
$$\left\|(U-1)^2x_j\right\|_{\mathrm{HS}} \leq C\alpha\big(1+\left\langle v,h^{-1}v\right\rangle\big)^2\left\langle v,h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}}, \tag{9.85}$$

we obtain the claim. □

Proposition 9.20. *For all* $t\in[0,1]$, *it holds that*

$$(C\alpha)^{-1}\max_{1\leq j\leq n}\left\|h^{-\frac{1}{2}}E_{2,2}(t)x_j\right\|$$
$$\leq\big(1+\left\langle v,h^{-1}v\right\rangle\big)^3\left\langle v,h^{-1}v\right\rangle^{\frac{5}{2}}+\big(1+\left\langle v,h^{-1}v\right\rangle\big)^7\left\langle v,h^{-\frac{2}{3}}\right\rangle^3\left\langle v,h^{-\frac{5}{4}}v\right\rangle^2\left\langle v,h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}}$$
$$+\big(1+\left\langle v,h^{-1}v\right\rangle\big)^5\left\langle v,h^{-\frac{2}{3}}v\right\rangle^{\frac{3}{2}}\big(\left\langle v,h^{-1}v\right\rangle^{\frac{3}{2}}\left\langle v,h^{-\frac{4}{3}}v\right\rangle^{\frac{3}{2}}+\left\langle v,h^{-1}v\right\rangle\left\langle v,h^{-\frac{5}{4}}v\right\rangle^2\big),$$

where $\alpha=\max_{1\leq j\leq n}\left\langle v,x_j\right\rangle$ *and* $C>0$ *is a constant independent of all quantities.*

Proof. We decompose $E_{2,2}(t)=\big(e^{tJ}-1\big)E_0\big(e^{-tJ}-1\big)$ as

$$F_tE_0F_t^*+F_tE_0\big(e^{-tJ}-1-F_t^*\big)+\big(e^{tJ}-1-F_t\big)E_0F_t^*+\big(e^{tJ}-1-F_t\big)E_0\big(e^{-tJ}-1-F_t^*\big) \tag{9.86}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and estimate by Proposition 9.16 that

$$\begin{aligned}\left\|h^{-\frac{1}{2}}E_{2,2}(t)x_j\right\| \leq{}& \left\|h^{-\frac{1}{2}}F_tE_0F_t^*x_j\right\| + C\left\|h^{-\frac{1}{2}}F_tE_0\left(U-1\right)\right\|_{\mathrm{HS}}\left\|(U-1)^2x_j\right\| \\ &+ C\left\|h^{-\frac{1}{2}}\left(U-1\right)^2\right\|_{\mathrm{HS}}\left\|(U-1)\,E_0F_t^*x_j\right\| \\ &+ C\left\|h^{-\frac{1}{2}}\left(U-1\right)^2\right\|_{\mathrm{HS}}\left\|(U-1)\,E_0\left(U-1\right)\right\|_{\mathrm{HS}}\left\|(U-1)^2x_j\right\|. \end{aligned}\tag{9.87}$$

We estimate as in the previous proposition that

$$\begin{aligned}&\left\|h^{-\frac{1}{2}}F_tE_0F_t^*x_j\right\|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\sum_{i=1}^n\frac{1}{\lambda_i}\left|\sum_{k,l=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\langle x_k,v\rangle\langle v,x_l\rangle\frac{\langle x_l,v\rangle\left\langle v,x_j\right\rangle}{\lambda_l+\lambda_j}\right|^2 \\ &\leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\sum_{i=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left(\sum_{k=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_k}\frac{|\langle x_l,v\rangle|^2}{\lambda_l}\right)^2 \\ &= C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\left\langle v,h^{-1}v\right\rangle^5, \end{aligned}\tag{9.88}$$

and

$$\begin{aligned}&\left\|h^{-\frac{1}{2}}F_tE_0\left(U-1\right)\right\|_{\mathrm{HS}}^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\sum_{i,j=1}^n\frac{1}{\lambda_i}\left|\sum_{k,l=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\langle x_k,v\rangle\langle v,x_l\rangle\frac{\langle x_l,v\rangle\left\langle v,x_j\right\rangle}{\lambda_l+\lambda_j}\right|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\sum_{i,j=1}^n\frac{|\langle x_i,v\rangle|^2}{\lambda_i}\left|\left\langle x_j,v\right\rangle\right|^2\left(\sum_{k,l=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_k}\frac{|\langle x_l,v\rangle|^2}{\lambda_l^{\frac{2}{3}}\lambda_j^{\frac{1}{3}}}\right)^2 \\ &= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\left\langle v,h^{-1}v\right\rangle^3\left\langle v,h^{-\frac{2}{3}}v\right\rangle^3, \end{aligned}\tag{9.89}$$

and

$$\begin{aligned}&\left\|(U-1)\,E_0F_t^*x_j\right\|^2 \\ &\leq C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\sum_{i=1}^n\left|\sum_{k,l=1}^n\frac{\langle x_i,v\rangle\langle v,x_k\rangle}{\lambda_i+\lambda_k}\langle x_k,v\rangle\langle v,x_l\rangle\frac{\langle x_l,v\rangle\left\langle v,x_j\right\rangle}{\lambda_l+\lambda_j}\right|^2 \\ &\leq C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\sum_{i=1}^n|\langle x_i,v\rangle|^2\left(\sum_{k,l=1}^n\frac{|\langle x_k,v\rangle|^2}{\lambda_i^{\frac{1}{3}}\lambda_k^{\frac{2}{3}}}\frac{|\langle x_l,v\rangle|^2}{\lambda_l}\right)^2 \\ &= C\alpha^2\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\left\langle v,h^{-1}v\right\rangle^2\left\langle v,h^{-\frac{2}{3}}v\right\rangle^3, \end{aligned}\tag{9.90}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and finally that

$$
\begin{aligned}
&\|(U-1)\,E_0\,(U-1)\|_{\mathrm{HS}}^2 \\
&\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6 \sum_{i,j=1}^n \left|\sum_{k,l=1}^n \frac{\langle x_i,v\rangle\,\langle v,x_k\rangle}{\lambda_i+\lambda_k}\,\langle x_k,v\rangle\,\langle v,x_l\rangle\,\frac{\langle x_l,v\rangle\left\langle v,x_j\right\rangle}{\lambda_l+\lambda_j}\right|^2 \\
&\le C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6 \sum_{i,j=1}^n |\langle x_i,v\rangle|^2\left|\left\langle x_j,v\right\rangle\right|^2 \left|\sum_{k,l=1}^n \frac{|\langle x_k,v\rangle|^2}{\lambda_i^{\frac13}\lambda_k^{\frac23}}\frac{|\langle x_l,v\rangle|^2}{\lambda_l^{\frac23}\lambda_j^{\frac13}}\right|^2 \\
&= C\left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\left\langle v,h^{-\frac23}\right\rangle^6. \qquad (9.91)
\end{aligned}
$$

Combining these bounds with equation (9.85) yields the claim. □

Combining the estimates from Proposition 9.17 through 9.20 and the last bound of Proposition 9.15, we obtain

$$
\begin{aligned}
&(C\alpha)^{-1}\max_{1\le j\le n}\left\|h^{-\frac12}E(t)x_j\right\| \\
&\le \left(1+\left\langle v,h^{-1}v\right\rangle\right)\sqrt{\left\langle v,h^{-1}v\right\rangle} + \left(1+\left\langle v,h^{-1}v\right\rangle\right)^2\left\langle v,h^{-1}v\right\rangle^{\frac32} \\
&\quad + \left(1+\left\langle v,h^{-1}v\right\rangle\right)^3\left(\left\langle v,h^{-1}v\right\rangle^{\frac52} + \|v\|\left\langle v,h^{-\frac54}v\right\rangle^2 + \left\langle v,h^{-\frac12}v\right\rangle\left\langle v,h^{-\frac43}v\right\rangle^{\frac32}\right) \\
&\quad + \left(1+\left\langle v,h^{-1}v\right\rangle\right)^4\left\langle v,h^{-\frac23}v\right\rangle^{\frac32}\left(\sqrt{\left\langle v,h^{-1}v\right\rangle}\left\langle v,h^{-\frac43}v\right\rangle^{\frac32} + \left\langle v,h^{-\frac54}v\right\rangle^2\right) \\
&\quad + \left(1+\left\langle v,h^{-1}v\right\rangle\right)^5\left\langle v,h^{-\frac23}v\right\rangle^{\frac32}\left(\left\langle v,h^{-1}v\right\rangle^{\frac32}\left\langle v,h^{-\frac43}v\right\rangle^{\frac32} + \left\langle v,h^{-1}v\right\rangle\left\langle v,h^{-\frac54}v\right\rangle^2\right) \\
&\quad + \left(1+\left\langle v,h^{-1}v\right\rangle\right)^6\left\langle v,h^{-\frac12}v\right\rangle^2\left\langle v,h^{-\frac54}v\right\rangle^2\left\langle v,h^{-\frac43}v\right\rangle^{\frac32} \\
&\quad + \left(1+\left\langle v,h^{-1}v\right\rangle\right)^7\left\langle v,h^{-\frac23}v\right\rangle^3\left\langle v,h^{-\frac54}v\right\rangle^2\left\langle v,h^{-\frac43}v\right\rangle^{\frac32}. \qquad (9.92)
\end{aligned}
$$

The right hand can be simplified further using the Hölder estimates

$$
\left\langle v,h^{-\frac12}v\right\rangle \le \|v\|\left\langle v,h^{-1}v\right\rangle^{\frac12}, \quad \left\langle v,h^{-\frac23}v\right\rangle^{\frac32} \le \|v\|\left\langle v,h^{-1}v\right\rangle. \qquad (9.93)
$$

All this gives the following:

Proposition 9.21. *For all $t\in[0,1]$, it holds that*

$$
\max_{1\le j\le n}\left\|h^{-\frac12}E(t)x_j\right\| \le C\alpha\left(1+\left\langle v,h^{-1}v\right\rangle\right)^8\left(\langle v,h^{-1}v\rangle^{\frac12} + \|v\|\langle v,h^{-\frac54}v\rangle^2\right)\left(1+\|v\|\langle v,h^{-\frac43}v\rangle^{\frac32}\right),
$$

where $\alpha = \max_{1\le j\le n}\left\langle v,x_j\right\rangle$ and $C>0$ is a constant independent of all quantities.

Conclusion of Proposition 9.6: Inserting h_k and v_k in Proposition 9.7 and (7.11), we have immediately

$$
\operatorname{tr}\left|h_k^{-1/2}(\widetilde{E}_k - h_k)h_k^{-1/2}\right| \le \langle v_k, h_k^{-1}v_k\rangle \le C\hat{V}_k. \qquad (9.94)
$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Next, consider the corresponding expressions on the right-hand side of Proposition 9.21. Recall that $\alpha_k = \max_{p\in L_k}\left\langle v_k, e_p\right\rangle \le C(\hat{V}_k)^{\frac12}k_F^{-\frac12}$. Moreover, by Propositions A.1, A.2 and A.3, we get

$$\left\langle v_k, h_k^{\beta} v_k\right\rangle \le C\hat{V}_k k_F^{-1}\sum_{p\in L_k}\lambda_{k,p}^{\beta} \le C\hat{V}_k(|k|k_F)^{1+\beta}, \quad 0\ge\beta\ge-\frac54,$$
$$\left\langle v_k, h_k^{\beta} v_k\right\rangle^{\frac32} \le C(\hat{V}_k)^{\frac32}\left(|k|k_F\right)^{-\frac12}|k|^6\log\left(k_F\right), \quad \beta\le-\frac43. \tag{9.95}$$

Putting these bounds together, we deduce from Proposition 9.21 that

$$\max_{p\in L_k}\left\|h_k^{-\frac12}E_k(t)e_p\right\| \le C(\hat{V}_k)^{\frac12}k_F^{-\frac12}\left(1+\hat{V}_k\right)^8\left((\hat{V}_k)^{\frac12}+(\hat{V}_k)^{\frac52}\right)\left(1+\hat{V}_k^2|k|^6\log k_F\right)$$
$$\le Ck_F^{-\frac12}\left(\hat{V}_k+\hat{V}_k^3|k|^6\log\left(k_F\right)\right) \tag{9.96}$$

for $|k|\le k_F^{\gamma}$, as claimed. Similarly, inserting (9.95) in Corollary 9.12 and Proposition 9.14, we see that

$$\left\|h^{-\frac12}J\right\|_{\mathrm{HS}} \le C\left(1+\left\langle v_k, h_k^{-1}v_k\right\rangle\right)\left\langle v_k, h_k^{-\frac32}v_k\right\rangle \le C(\log k_F)^{\frac23}k_F^{-\frac13}\hat{V}_k\left(1+\hat{V}_k\right)|k|^{3+\frac23},$$
$$\left\|h^{-\frac12}\left[J,h\right]h^{-\frac12}\right\|_{\mathrm{HS}} \le C\left(1+\left\langle v_k, h_k^{-1}v_k\right\rangle\right)^3\left(\left\langle v_k, h_k^{-1}v_k\right\rangle+\left\langle v_k, h_k^{-\frac12}v_k\right\rangle\left\langle v_k, h_k^{-\frac54}v_k\right\rangle^2\right)$$
$$\le C\left(1+\hat{V}_k\right)^3\left(\hat{V}_k+\hat{V}_k\left(|k|k_F\right)^{\frac12}\left(\hat{V}_k\left(|k|k_F\right)^{-\frac14}\right)^2\right)\le C\hat{V}_k. \tag{9.97}$$

Here, we also note that $\hat{V}_k$ is uniformly bounded, and hence, the constant C may depend on V, but it is still independent of k and k_F.

9.4. Gronwall estimates for the kinetic operator

We now come to the kinetic Gronwall estimates for the transformation $e^{\mathcal{J}}$. We have the following:

Proposition 9.22. *Assume $\sum_{k\in\mathbb{Z}^3}\hat{V}_k|k|<\infty$ and $S_C=\mathbb{Z}_+^3\cap\overline{B}\left(0,k_F^{\gamma}\right)$ with $0<\gamma<\frac{1}{47}$. Then for all $\Psi\in D\left(H'_{\mathrm{kin}}\right)$ and $|t|\le1$, it holds that*

$$\left\langle e^{t\mathcal{J}}\Psi, H'_{\mathrm{kin}}e^{t\mathcal{J}}\Psi\right\rangle \le C\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle$$
$$\left\langle e^{t\mathcal{J}}\Psi, \mathcal{N}_E H'_{\mathrm{kin}}e^{t\mathcal{J}}\Psi\right\rangle \le C\left\langle\Psi, \mathcal{N}_E H'_{\mathrm{kin}}\Psi\right\rangle$$

for a constant $C>0$ independent of k_F.

Proof. Write $\Psi_t = e^{t\mathcal{J}}\Psi$ for brevity. By the commutator in (9.15), we have

$$-\frac{d}{dt}\left\langle\Psi_t, H'_{\mathrm{kin}}\Psi_t\right\rangle = \left\langle\Psi_t,\left[\mathcal{J},H'_{\mathrm{kin}}\right]\Psi_t\right\rangle = 2\sum_{k\in S_C}\left\langle\Psi_t,\tilde{Q}_1^k\left(\left[J_k^{\oplus},h_k^{\oplus}\right]\right)\Psi_t\right\rangle \tag{9.98}$$

with $\tilde{Q}_1^k$ defined in (4.35). Moreover, Proposition 4.8 allows us to estimate

$$\sum_{k\in S_C}\left|\left\langle\Psi_t,\tilde{Q}_1^k\left(\left[J_k^{\oplus},h_k^{\oplus}\right]\right)\Psi_t\right\rangle\right| \le \sum_{k\in S_C}\left\|\left(h_k^{\oplus}\right)^{-\frac12}\left[J_k^{\oplus},h_k^{\oplus}\right]\left(h_k^{\oplus}\right)^{-\frac12}\right\|_{\mathrm{Op}}\left\langle\Psi_t, H'_{\mathrm{kin}}\Psi_t\right\rangle. \tag{9.99}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Since

$$\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\left[J_k^{\oplus}, h_k^{\oplus}\right]\left(h_k^{\oplus}\right)^{-\frac{1}{2}} = \begin{pmatrix} h_k^{-\frac{1}{2}}\left[J_k, h_k\right] h_k^{-\frac{1}{2}} & 0 \\ 0 & h_k^{-\frac{1}{2}}\left[J_k, h_k\right] h_k^{-\frac{1}{2}} \end{pmatrix}, \tag{9.100}$$

by Proposition 9.6 we can estimate further that

$$\begin{aligned}\sum_{k\in S_C}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\left[J_k^{\oplus}, h_k^{\oplus}\right]\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\right\|_{\mathrm{Op}} &= \sum_{k\in S_C}\left\|h_k^{-\frac{1}{2}}\left[J_k, h_k\right] h_k^{-\frac{1}{2}}\right\|_{\mathrm{Op}} \leq \sum_{k\in S_C}\left\|h_k^{-\frac{1}{2}}\left[J_k, h_k\right] h_k^{-\frac{1}{2}}\right\|_{\mathrm{HS}} \\ &\leq C\sum_{k\in S_C}\hat{V}_k \leq C. \end{aligned}\tag{9.101}$$

Hence, $\left|\frac{d}{dt}\left\langle\Psi_t, H'_{\mathrm{kin}}\Psi_t\right\rangle\right| \leq C\left\langle\Psi_t, H'_{\mathrm{kin}}\Psi_t\right\rangle$, so by Gronwall's lemma

$$\left\langle\Psi_t, H'_{\mathrm{kin}}\Psi_t\right\rangle \leq \left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle e^{C|t|} \leq C\left\langle\Psi, H'_{\mathrm{kin}}\Psi\right\rangle, \quad |t| \leq 1. \tag{9.102}$$

For $\left\langle\Psi_t, \mathcal{N}_E H'_{\mathrm{kin}}\Psi_t\right\rangle$, besides the commutator in (9.15), we also note that

$$\begin{aligned}\left[\mathcal{J}, \mathcal{N}_E\right] &= \sum_{k\in S_C}\sum_{p\in L_k^{\pm}}\left[b_k^*\left(J_k^{\oplus}e_p\right)b_k(e_p), \mathcal{N}_E\right] \\ &= \sum_{k\in S_C}\sum_{p\in L_k^{\pm}}\left(b_k^*\left(J_k^{\oplus}e_p\right)\left[b_k(e_p), \mathcal{N}_E\right] + \left[b_k^*\left(J_k^{\oplus}e_p\right), \mathcal{N}_E\right]b_k(e_p)\right) \\ &= \sum_{k\in S_C}\sum_{p\in L_k^{\pm}}\left(b_k^*\left(J_k^{\oplus}e_p\right)b_k(e_p) - b_k^*\left(J_k^{\oplus}e_p\right)b_k(e_p)\right) = 0. \end{aligned}\tag{9.103}$$

Here again, we used $[\mathcal{N}_E, b_k(\varphi)] = -b_k$ for all $\varphi \in \ell^2(L_k^{\pm})$, which follows from (1.75) and linearity. Hence,

$$-\frac{d}{dt}\left\langle\Psi_t, \mathcal{N}_E H'_{\mathrm{kin}}\Psi_t\right\rangle = \left\langle\Psi_t, \mathcal{N}_E\left[\mathcal{J}, H'_{\mathrm{kin}}\right]\Psi_t\right\rangle = 2\sum_{k\in S_C}\left\langle\Psi_t, \mathcal{N}_E\tilde{Q}_1^k\left(\left[J_k^{\oplus}, h_k^{\oplus}\right]\right)\Psi_t\right\rangle.$$

Now, it holds that $[\mathcal{N}_E, \tilde{Q}_1^k([J_k^{\oplus}, h_k^{\oplus}])] = 0$ (as can be seen by a computation similar to that of equation (9.103)), so we may estimate as above for

$$\begin{aligned}\sum_{k\in S_C}\left|\left\langle\Psi_t, \mathcal{N}_E\tilde{Q}_1^k\left(\left[J_k^{\oplus}, h_k^{\oplus}\right]\right)\Psi_t\right\rangle\right| &= \sum_{k\in S_C}\left|\left\langle\mathcal{N}_E^{\frac{1}{2}}\Psi_t, \tilde{Q}_1^k\left(\left[J_k^{\oplus}, h_k^{\oplus}\right]\right)\mathcal{N}_E^{\frac{1}{2}}\Psi_t\right\rangle\right| \\ \leq \sum_{k\in S_C}\left\|\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\left[J_k^{\oplus}, h_k^{\oplus}\right]\left(h_k^{\oplus}\right)^{-\frac{1}{2}}\right\|_{\mathrm{Op}}&\left\langle\mathcal{N}_E^{\frac{1}{2}}\Psi_t, H'_{\mathrm{kin}}\mathcal{N}_E^{\frac{1}{2}}\Psi_t\right\rangle \leq C\left\langle\Psi_t, \mathcal{N}_E H'_{\mathrm{kin}}\Psi_t\right\rangle, \end{aligned}\tag{9.104}$$

where we also used that $[\mathcal{N}_E, H'_{\mathrm{kin}}] = 0$. The second claim now follows. □

10. Conclusion of the main results

Now we are ready to provide the proof of the main theorems stated in the introduction.

10.1. Proof of Theorem 1.1

The proof follows almost immediately by the analysis we have performed throughout the paper, for we will simply take $\mathcal{U} = e^{\mathcal{J}}e^{\mathcal{K}}$ where $e^{\mathcal{K}}$ is the quasi-bosonic Bogolubov transformation $e^{\mathcal{K}}$ of Section 4 and $e^{\mathcal{J}}$ is the second transformation of Section 9.

Step 1: Let us start from the decomposition (1.22):

$$H_N - E_{\mathrm{FS}} = H'_{\mathrm{kin}} + k_F^{-1} H'_{\mathrm{int}} = H'_{\mathrm{kin}} + \sum_{k\in S_C} \left(H^k_{\mathrm{int}} - \frac{\hat{V}_k k_F^{-1}}{(2\pi)^3} |L_k| \right) + \mathcal{E}_{\mathrm{NB}}, \tag{10.1}$$

where H^k_{int} is given in (1.29), $\mathcal{E}_{\mathrm{NB}}$ is given in (2.22), and $S_C = \overline{B}\left(0, k_F^\gamma\right) \cap \mathbb{Z}^3_+$ with $0 < \gamma < \frac{1}{47}$. From Proposition 2.4, the non-bosonizable term $\mathcal{E}_{\mathrm{NB}}$ is estimated as

$$\pm\mathcal{E}_{\mathrm{NB}} \le C k_F^{-\gamma/2} (H'_{\mathrm{kin}} + k_F^{-1} \mathcal{N}_E H'_{\mathrm{kin}} + k_F). \tag{10.2}$$

By the Gronwall estimates of Propositions 8.1, 9.22 and the choice $\mathcal{U} = e^{\mathcal{J}} e^{\mathcal{K}}$, we have

$$\pm\mathcal{U}\mathcal{E}_{\mathrm{NB}}\mathcal{U}^* \le C k_F^{-\gamma/2} (H'_{\mathrm{kin}} + k_F^{-1} \mathcal{N}_E H'_{\mathrm{kin}} + k_F). \tag{10.3}$$

Thus, it remains to apply the transformations $e^{\mathcal{K}}$ and $e^{\mathcal{J}}$ to the bosonizable terms.

Step 2: Now we apply the transformation $e^{\mathcal{K}}$. By Proposition 5.7, we have

$$\begin{aligned} & e^{\mathcal{K}} \left(H'_{\mathrm{kin}} + \sum_{k\in S_C} H^k_{\mathrm{int}} \right) e^{-\mathcal{K}} \\ & = H'_{\mathrm{kin}} + \sum_{k\in S_C} Q^k_1 (E^\oplus_k - h^\oplus_k) + \sum_{k\in S_C} \int_0^1 e^{(1-t)\mathcal{K}} \left(\mathcal{E}^k_1 (A^\oplus_k(t)) + \mathcal{E}^k_2 (B^\oplus_k(t)) \right) e^{-(1-t)\mathcal{K}}\, dt. \end{aligned} \tag{10.4}$$

We will use the kinetic estimate of Proposition 6.5 and the Gronwall estimates of Proposition 8.1 to bound the exchange terms in (10.4). Thanks to the one-body estimates in Propositions 7.3, 7.2 and our assumption $\sum_{k\in S_C} \hat{V}_k |k| < \infty$, we get

$$\begin{aligned} \sum_{k\in S_C} \max_{t\in[0,1]} \left\{ \max_{p\in L_k} \left\| h_k^{-\frac12} A^\oplus_k(t) e_p \right\|, \max_{p\in L_k} \left\| h_k^{-\frac12} B^\oplus_k(t) e_p \right\| \right\} &\le C \sum_{k\in S_C} k_F^{-\frac12} \hat{V}_k \left(1 + \hat{V}_k^2\right) \le C k_F^{-\frac12}, \\ \sum_{k\in S_C} \max_{t\in[0,1]} \left\{ \left\| A^\oplus_k(t) \right\|_{\infty,2}, \left\| B^\oplus_k(t) \right\|_{\infty,2} \right\} &\le C \sum_{k\in S_C} \hat{V}_k |k|^{\frac12} \left(1 + \hat{V}_k\right) \le C, \\ \sum_{k\in S_C} \left(\left\| (h^\oplus_k)^{-\frac12} K^\oplus_k \right\|_{\mathrm{HS}} + \left\| K^\oplus_k \right\|_{\infty,2} \right) &\le C (\log k_F)^{\frac23} k_F^{-\frac13} \sum_{k\in S_C} \hat{V}_k |k|^{3+\frac23}. \end{aligned} \tag{10.5}$$

All this gives that for every state $\Psi \in D\left(H'_{\mathrm{kin}}\right)$ and $\Psi_t = e^{-(1-t)\mathcal{K}}\Psi$,

$$\begin{aligned} & \sum_{k\in S_C} \int_0^1 \left| \left\langle \Psi_t, \left(\mathcal{E}^k_1 (A^\oplus_k(t)) + \mathcal{E}^k_2 (B^\oplus_k(t)) \right) \Psi_t \right\rangle \right| dt \le C (\log k_F)^{\frac23} \left(\sum_{k\in S_C} \hat{V}_k |k|^{3+\frac23} \right) \\ & \times \left(k_F^{-\frac56} \max_{t\in[0,1]} \sqrt{\left\langle \Psi_t, H'_{\mathrm{kin}} \Psi_t \right\rangle \left\langle \Psi_t, \mathcal{N}_E H'_{\mathrm{kin}} \Psi_t \right\rangle} \right. \\ & \left. + k_F^{-\frac13} \max_{t\in[0,1]} \left\langle \Psi_t, H'_{\mathrm{kin}} \Psi_t \right\rangle + k_F^{-\frac13} \max_{t\in[0,1]} \sqrt{\left\langle \Psi_t, \mathcal{N}_E H'_{\mathrm{kin}} \Psi_t \right\rangle} \right) \\ & \le C (\log k_F)^{\frac23} k_F^{-\frac13} \left(\sum_{k\in S_C} \hat{V}_k |k|^{3+\frac23} \right) \left(\left\langle \Psi, H'_{\mathrm{kin}} \Psi \right\rangle + k_F^{-1} \left\langle \Psi, \mathcal{N}_E H'_{\mathrm{kin}} \Psi \right\rangle + k_F \left\| \Psi \right\|^2 \right), \end{aligned} \tag{10.6}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

where we also used the Cauchy–Schwarz inequality to split the square roots at the end. Thus, the exchange terms in (10.4) can be estimated as

$$\begin{aligned}&\pm \sum_{k\in S_C}\int_0^1 e^{(1-t)\mathcal{K}}\left(\mathcal{E}_1^k(A_k^{\oplus}(t))+\mathcal{E}_2^k(B_k^{\oplus}(t))\right)e^{-(1-t)\mathcal{K}}dt\\ &\leq C(\log k_F)^{\frac{2}{3}}k_F^{-\frac{1}{3}}\left(\sum_{k\in S_C}\hat{V}_k|k|^{3+\frac{2}{3}}\right)\left(H'_{\text{kin}}+k_F^{-1}\mathcal{N}_E H'_{\text{kin}}+k_F\right).\end{aligned} \tag{10.7}$$

It remains to consider the main term $Q_1(E_k^{\oplus}-h_k^{\oplus})$ on the right side of (10.4). We use the normal order form in (4.34):

$$\sum_{k\in S_C}Q_1(E_k^{\oplus}-h_k^{\oplus})=\sum_{k\in S_C}2\widetilde{Q}_1(E_k^{\oplus}-h_k^{\oplus})+\sum_{k\in S_C}2\mathrm{tr}(E_k-h_k)+\sum_{k\in S_C}\varepsilon_k(E_k^{\oplus}-h_k^{\oplus}). \tag{10.8}$$

By Propositions 4.9, 7.2 and 2.1,

$$\pm\sum_{k\in S_C}\varepsilon_k\left(E_k^{\oplus}-h_k^{\oplus}\right)\leq C\sum_{k\in S_C}k_F^{-1}\hat{V}_k\left(1+\hat{V}_k\right)\mathcal{N}_E\leq Ck_F^{-1}H'_{\text{kin}}. \tag{10.9}$$

Moreover, by Proposition 7.1, we have

$$\sum_{k\in S_C}\left(2\,\mathrm{tr}\left(E_k-h_k\right)-\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|\right)=\sum_{k\in S_C}\frac{2}{\pi}\int_0^{\infty}F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2+t^2}\right)dt \tag{10.10}$$

with $F(x)=\log(1+x)-x$. Thus in summary, we conclude from (10.4) that

$$\begin{aligned}&e^{\mathcal{K}}\left(H'_{\text{kin}}+\sum_{k\in S_C}\left(H_{\text{int}}^k-\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}|L_k|\right)\right)e^{-\mathcal{K}}\\ &=H'_{\text{kin}}+2\sum_{k\in S_C}\widetilde{Q}_1^k(E_k^{\oplus}-h_k^{\oplus})+\sum_{k\in S_C}\frac{2}{\pi}\int_0^{\infty}F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2+t^2}\right)dt+\mathcal{E}_{\mathcal{K}},\end{aligned} \tag{10.11}$$

where

$$\pm\mathcal{E}_{\mathcal{K}}\leq C(\log k_F)^{\frac{2}{3}}k_F^{-\frac{1}{3}}\left(\sum_{k\in S_C}\hat{V}_k|k|^{3+\frac{2}{3}}\right)\left(H'_{\text{kin}}+k_F^{-1}\mathcal{N}_E H'_{\text{kin}}+k_F\right). \tag{10.12}$$

Step 3: Next, we apply the transformation $e^{\mathcal{J}}$ to the right-hand side of (10.11). From (10.12) and the Gronwall estimates of Proposition 9.22, we have

$$\pm e^{\mathcal{J}}\mathcal{E}_{\mathcal{K}}e^{-\mathcal{J}}\leq C(\log k_F)^{\frac{2}{3}}k_F^{-\frac{1}{3}}\left(\sum_{k\in S_C}\hat{V}_k|k|^{3+\frac{2}{3}}\right)\left(H'_{\text{kin}}+k_F^{-1}\mathcal{N}_E H'_{\text{kin}}+k_F\right). \tag{10.13}$$

For the main terms, by Proposition 9.1,

$$\begin{aligned}&e^{\mathcal{J}}\left(H'_{\text{kin}}+2\sum_{k\in S_C}\tilde{Q}_1^k\left(E_k^{\oplus}-h_k^{\oplus}\right)\right)e^{-\mathcal{J}}\\ &=H'_{\text{kin}}+2\sum_{k\in S_C}\tilde{Q}_1^k\left(\widetilde{E}_k^{\oplus}-h_k^{\oplus}\right)+2\sum_{k\in S_C}\int_0^1 e^{(1-t)\mathcal{J}}\mathcal{E}_3^k(F_k^{\oplus}(t))e^{-(1-t)\mathcal{J}}dt.\end{aligned} \tag{10.14}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Let us bound the exchange term $\mathcal{E}_3(\cdot)$. For all $k \in \overline{B}\left(0, k_F^\gamma\right) \cap \mathbb{Z}^3_*$ with $0 < \gamma < \frac{1}{47}$ and $t \in [0,1]$, by Proposition 9.6, we have

$$\max_{p \in L_k^\pm} \left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} E_k(t) e_p \right\| \le C k_F^{-\frac{1}{2}} \left(\hat{V}_k + \hat{V}_k^3 |k|^6 \log(k_F)\right) \le C k_F^{-\frac{1}{2}} \left(\hat{V}_k + \hat{V}_k^3 |k|^3 k_F^{\frac{3}{47}} \log(k_F)\right),$$

$$\sum_{l \in S_C} \left\| \left(h_l^\oplus\right)^{-\frac{1}{2}} J_l^\oplus \right\|_{\mathrm{HS}} \le C (\log k_F)^{\frac{2}{3}} k_F^{-\frac{1}{3}} \sum_{l \in S_C} \hat{V}_l \left(1 + \hat{V}_l\right) \le C (\log k_F)^{\frac{2}{3}} k_F^{-\frac{1}{3}}. \tag{10.15}$$

Hence, using the kinetic estimate of Proposition 9.5, Gronwall's bounds of Proposition 9.22 and the assumption $\sum_{k \in \mathbb{Z}^3_*} \hat{V}_k |k| < \infty$, we find that for every state $\Psi \in D\left(H'_{\mathrm{kin}}\right)$ and $\Psi_t = e^{-(1-t)\mathcal{J}} \Psi$,

$$\sum_{k \in S_C} \int_0^1 \left| \langle \Psi_t, \mathcal{E}_3^k (F_k^\oplus(t)) \Psi_t \rangle \right| dt \tag{10.16}$$
$$\le \sum_{k \in S_C} C (\log k_F)^{\frac{2}{3}} k_F^{-\frac{1}{3}} k_F^{-\frac{1}{2}} \left(\hat{V}_k + \hat{V}_k^3 |k|^3 k_F^{\frac{3}{47}} \log(k_F)\right) \max_{t \in [0,1]} \sqrt{\left\langle \Psi_t, H'_{\mathrm{kin}} \Psi_t \right\rangle \left\langle \Psi_t, \mathcal{N}_E H'_{\mathrm{kin}} \Psi_t \right\rangle}$$
$$\le C (\log k_F)^{\frac{5}{3}} k_F^{-\frac{1}{3}} \left\langle \Psi, (k_F^{-1} \mathcal{N}_E H'_{\mathrm{kin}} + H'_{\mathrm{kin}} + k_F) \Psi \right\rangle.$$

Here, we used $\sum_{k \in \mathbb{Z}^3_*} \hat{V}_k^3 |k|^3 \le \left(\sum_{k \in \mathbb{Z}^3_*} \hat{V}_k |k|\right)^3 < \infty$. Consequently,

$$\pm \sum_{k \in S_C} \int_0^1 e^{(1-t)\mathcal{J}} \mathcal{E}_3^k (F_k^\oplus(t)) e^{-(1-t)\mathcal{J}} dt \le C (\log k_F)^{\frac{5}{3}} k_F^{-\frac{1}{3}} (k_F^{-1} \mathcal{N}_E H'_{\mathrm{kin}} + H'_{\mathrm{kin}} + k_F). \tag{10.17}$$

In summary, we have for $\mathcal{U} = e^{\mathcal{J}} e^{\mathcal{K}}$ and $0 < \gamma < \frac{1}{47}$,

$$\mathcal{U} H_N \mathcal{U}^* = E_{\mathrm{FS}} + H'_{\mathrm{kin}} + 2 \sum_{k \in S_C} \tilde{Q}_1^k \left(\widetilde{E}_k^\oplus - h_k^\oplus\right) + \sum_{k \in S_C} \frac{2}{\pi} \int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3} \sum_{p \in L_k} \frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\right) dt + \mathcal{E}_{\mathcal{J}}, \tag{10.18}$$

where the error term is collected from (10.3), (10.13), (10.17) which satisfies

$$\pm \mathcal{E}_{\mathcal{J}} \le C k_F^{-\gamma/2} (k_F^{-1} \mathcal{N}_E H'_{\mathrm{kin}} + H'_{\mathrm{kin}} + k_F). \tag{10.19}$$

Step 4: Finally, let us remove the cutoff $S_C = \mathbb{Z}^3_+ \cap \overline{B}\left(0, k_F^\gamma\right)$ on the right-hand side of (10.18). By Proposition 7.1, we can bound

$$\left| \frac{1}{\pi} \sum_{k \in \mathbb{Z}^3_* \setminus S_C} \int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3} \sum_{p \in L_k} \frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\right) dt \right| \le C k_F \sum_{k \in \mathbb{Z}^3_* \setminus S_C} \hat{V}_k^2 |k| \le C k_F^{1-\gamma}. \tag{10.20}$$

Here, we used $\sum_{k \in \mathbb{Z}^3_*} \hat{V}_k^2 |k|^2 \le \left(\sum_{k \in \mathbb{Z}^3_*} \hat{V}_k |k|\right)^2 < \infty$. Moreover, by Propositions 4.8 and 9.6 (together with the fact that the trace norm dominates the operator norm), we can bound

$$\pm \tilde{Q}_1^k \left(\widetilde{E}_k^\oplus - h_k^\oplus\right) \le \left\| \left(h_k^\oplus\right)^{-\frac{1}{2}} \left(\widetilde{E}_k^\oplus - h_k^\oplus\right) \left(h_k^\oplus\right)^{-\frac{1}{2}} \right\|_{\mathrm{Op}} H'_{\mathrm{kin}}$$
$$= \left\| h_k^{-\frac{1}{2}} \left(\widetilde{E}_k - h_k\right) h_k^{-\frac{1}{2}} \right\|_{\mathrm{Op}} H'_{\mathrm{kin}} \le C \hat{V}_k H'_{\mathrm{kin}} \tag{10.21}$$

for all $k \in \mathbb{Z}^3_+$, and hence,

$$\pm \sum_{k\in\mathbb{Z}^3_+\backslash S_C} \tilde{Q}_1^k\left(\widetilde{E}_k^{\oplus} - h_k^{\oplus}\right) \le C\left(\sum_{k\in\mathbb{Z}^3_+\backslash S_C} \hat{V}_k\right) H'_{\mathrm{kin}} \le C k_F^{-\gamma} H'_{\mathrm{kin}}. \tag{10.22}$$

Therefore, we can deduce from (10.18) that for $\mathcal{U} = e^{\mathcal{J}} e^{\mathcal{K}}$ and $0 < \gamma < \frac{1}{47}$,

$$\mathcal{U} H_N \mathcal{U}^* = E_{\mathrm{FS}} + H'_{\mathrm{kin}} + 2\sum_{k\in\mathbb{Z}^3_+} \tilde{Q}_1^k\left(\widetilde{E}_k^{\oplus} - h_k^{\oplus}\right) + \sum_{k\in\mathbb{Z}^3_*} \frac{1}{\pi}\int_0^\infty F\left(\frac{\hat{V}_k k_F^{-1}}{(2\pi)^3}\sum_{p\in L_k}\frac{\lambda_{k,p}}{\lambda_{k,p}^2 + t^2}\right) dt + \mathcal{E}_{\mathcal{U}}, \tag{10.23}$$

where

$$\pm\mathcal{E}_{\mathcal{U}} \le C k_F^{-\gamma/2}(k_F^{-1}\mathcal{N}_E H'_{\mathrm{kin}} + H'_{\mathrm{kin}} + k_F). \tag{10.24}$$

The statement of Theorem 1.1 follows by recognizing the identity

$$2\sum_{k\in\mathbb{Z}^3_+} \tilde{Q}_1^k\left(\widetilde{E}_k^{\oplus} - h_k^{\oplus}\right) = 2\sum_{k\in\mathbb{Z}^3_*}\sum_{p,q\in L_k}\left\langle e_p, \left(\widetilde{E}_k - h_k\right) e_q\right\rangle b^*_{k,p} b_{k,q},$$

which follows from the definition of $\tilde{Q}_1^k$ in (4.35).

10.2. Proof of Theorem 1.2

Let $\Psi \in D\left(H'_{\mathrm{kin}}\right)$ be a normalized eigenstate of H_N with energy $\langle\Psi, H_N\Psi\rangle \le E_{\mathrm{FS}} + \kappa k_F$ for some $\kappa > 0$. Denoting $\tilde{H}_N = H_N - E_{\mathrm{FS}}$, we have $\tilde{H}_N\Psi = E'\Psi$ with $E' \le \kappa k_F$. Using (1.22) and the obvious inequality $A^*A \ge 0$, we obtain the Onsager-type estimate

$$\begin{aligned}\tilde{H}_N - H'_{\mathrm{kin}} &= \frac{k_F^{-1}}{2(2\pi)^3}\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\left(\mathrm{d}\Gamma(e^{-ik\cdot x})^*\mathrm{d}\Gamma(e^{-ik\cdot x}) - |L_k|\right) \\ &\ge -\frac{k_F^{-1}}{2(2\pi)^3}\sum_{k\in\mathbb{Z}^3_*}\hat{V}_k\,|L_k| \ge -Ck_F\sum_{k\in\mathbb{Z}^3}|k|\hat{V}_k.\end{aligned} \tag{10.25}$$

Here, we used $|L_k| \le Ck_F^2|k|$ for all $k \in \mathbb{Z}^3_*$ (see Proposition A.1). From (10.25) and the assumption $\tilde{H}_N\Psi = E'\Psi$ with $E' \le \kappa k_F$, we deduce immediately that

$$\langle\Psi, H'_{\mathrm{kin}}\Psi\rangle \le C(\kappa+1)k_F. \tag{10.26}$$

To prove the bound for $\mathcal{N}_E H'_{\mathrm{kin}}$, we use the operator inequality

$$\begin{aligned}\mathcal{N}_E^2 H'_{\mathrm{kin}} &= \mathcal{N}_E H'_{\mathrm{kin}}\mathcal{N}_E \le \mathcal{N}_E\tilde{H}_N\mathcal{N}_E + Ck_F\mathcal{N}_E^2 \\ &= \frac{1}{2}\left(\mathcal{N}_E^2\tilde{H}_N + \tilde{H}_N\mathcal{N}_E^2 - \left[\mathcal{N}_E, \left[\mathcal{N}_E, \tilde{H}_N\right]\right]\right) + Ck_F\mathcal{N}_E^2,\end{aligned} \tag{10.27}$$

which follows from (10.25) and the fact that $\left[\mathcal{N}_E, H'_{\mathrm{kin}}\right] = 0$. Thanks to the eigenvalue equation $\tilde{H}_N\Psi = E'\Psi$ with $E' \le \kappa k_F$, we deduce that

$$\left\langle\Psi, \mathcal{N}_E^2 H'_{\mathrm{kin}}\Psi\right\rangle \le C(\kappa+1)k_F\left\langle\Psi, \mathcal{N}_E^2\Psi\right\rangle - \frac{1}{2}\left\langle\Psi, \left[\mathcal{N}_E, \left[\mathcal{N}_E, \tilde{H}_N\right]\right]\Psi\right\rangle. \tag{10.28}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Using $\mathcal{N}_E = \sum_{s\in B_F^c} c_s^* c_s$ and

$$[c_s^* c_s, c_{p+k}^* c_{q-k}^* c_q c_p] = c_{p+k}^* c_{q-k}^* c_q c_p (\delta_{s,p+k} + \delta_{s,q-k} - \delta_{s,q} - \delta_{s,p}), \tag{10.29}$$

we deduce from (1.8) that

$$\big[\mathcal{N}_E, \big[\mathcal{N}_E, \tilde{H}_N\big]\big] = \frac{k_F^{-1}}{2(2\pi)^3} \sum_{k\in\mathbb{Z}^3} \sum_{p,q\in\mathbb{Z}^3} \hat{V}_k c_{p+k}^* c_{q-k}^* c_q c_p \Big(\sum_{s\in B_F^c} (\delta_{s,p+k} + \delta_{s,q-k} - \delta_{s,q} - \delta_{s,p}) \Big)^2. \tag{10.30}$$

Using the obvious bound

$$0 \le \Big(\sum_{s\in B_F^c} (\delta_{s,p+k} + \delta_{s,q-k} - \delta_{s,q} - \delta_{s,p}) \Big)^2 \le 4 \tag{10.31}$$

and the Cauchy–Schwarz inequality, we estimate

$$\begin{aligned} \big|\big\langle \Psi, \big[\mathcal{N}_E, \big[\mathcal{N}_E, \tilde{H}_N\big]\big]\, \Psi\big\rangle\big| &\le C k_F^{-1} \sum_{k\in\mathbb{Z}^3} \sum_{p,q\in\mathbb{Z}^3} \hat{V}_k \|c_{p+k} c_{q-k}\Psi\| \|c_q c_p \Psi\| \\ \le C k_F^{-1} \sum_{k\in\mathbb{Z}^3} \hat{V}_k \sum_{p,q\in\mathbb{Z}^3} (\|c_{p+k}c_{q-k}\Psi\|^2 + \|c_q c_p\Psi\|^2) &\le C k_F^{-1} \sum_{k\in\mathbb{Z}^3} \hat{V}_k \langle \Psi, \mathcal{N}_E^2 \Psi\rangle. \end{aligned} \tag{10.32}$$

Since $\hat{V}$ is summable, (10.28) and (10.32) imply that

$$\big\langle \Psi, \mathcal{N}_E^2 H'_{\text{kin}} \Psi \big\rangle \le C(\kappa+1) k_F \,\big\langle \Psi, \mathcal{N}_E^2 \Psi \big\rangle\,. \tag{10.33}$$

Combining with the inequality $H'_{\text{kin}} \ge \mathcal{N}_E$ from Proposition 2.1, we deduce by Hölder's inequality

$$\big\langle \Psi, \mathcal{N}_E^2 \Psi\big\rangle \le \big\langle \Psi, \mathcal{N}_E^3 \Psi\big\rangle^{2/3} \le \big\langle \Psi, \mathcal{N}_E^2 H'_{\text{kin}} \Psi\big\rangle^{2/3} \le \Big(C(\kappa+1) k_F \,\big\langle \Psi, \mathcal{N}_E^2 \Psi\big\rangle \Big)^{2/3}, \tag{10.34}$$

which implies that $\big\langle \Psi, \mathcal{N}_E^2 \Psi\big\rangle \le C(\kappa+1)^2 k_F^2$, and hence by (10.33) again,

$$\langle \Psi, \mathcal{N}_E^2 H'_{\text{kin}} \Psi\rangle \le C(\kappa+1)^3 k_F^3. \tag{10.35}$$

The bound $\big\langle \Psi, \mathcal{N}_E H'_{\text{kin}} \Psi\big\rangle \le C(\kappa+1)^2 k_F^2$ follows from (10.26) and (10.35). In summary, we have

$$\big\langle \Psi, (k_F^{-1} \mathcal{N}_E H'_{\text{kin}} + H'_{\text{kin}} + k_F)\Psi\big\rangle \le C(\kappa+1)^2 k_F. \tag{10.36}$$

By the Gronwall estimates of Propositions 8.1, 9.22 and the choice $\mathcal{U} = e^{\mathcal{J}} e^{\mathcal{K}}$, we also obtain

$$\big\langle \mathcal{U}\Psi, (k_F^{-1} \mathcal{N}_E H'_{\text{kin}} + H'_{\text{kin}} + k_F)\mathcal{U}\Psi\big\rangle \le C(\kappa+1)^2 k_F. \tag{10.37}$$

10.3. *Proof of Theorem 1.2*

Taking the expectation against Ψ_{FS} of the operator estimate in Theorem 1.1, we have

$$\inf \sigma(H_N) = \inf \sigma(\mathcal{U} H_N \mathcal{U}^*) \le \langle \Psi_{\text{FS}}, \mathcal{U} H_N \mathcal{U}^* \Psi_{\text{FS}}\rangle = E_{\text{FS}} + E_{\text{corr}} + O(k_F^{1-\frac{1}{94}+\epsilon}). \tag{10.38}$$

Here, we used the bound on $\mathcal{E}_U$ from Theorem 1.1 and the identities $H'_{\text{kin}}\Psi_{\text{FS}} = H_{\text{eff}}\Psi_{\text{FS}} = 0$.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

To see the lower bound, let $\Psi_{\mathrm{GS}} \in D\left(H'_{\mathrm{kin}}\right)$ be the normalized ground state of H_N. By the definition of Ψ_{GS} and the above upper bound, we have

$$\langle \Psi_{\mathrm{GS}}, H_N \Psi_{\mathrm{GS}} \rangle = \inf \sigma(H_N) \le E_{\mathrm{FS}} + C k_F, \tag{10.39}$$

and hence, Theorem 1.2 implies that the state $\Psi'_{\mathrm{GS}} = \mathcal{U}\Psi_{\mathrm{GS}}$ satisfies

$$\left\langle \Psi'_{\mathrm{GS}}, (k_F^{-1}\mathcal{N}_E H'_{\mathrm{kin}} + H'_{\mathrm{kin}} + k_F)\Psi'_{\mathrm{GS}} \right\rangle \le C k_F. \tag{10.40}$$

Taking the expectation against Ψ'_{GS} of the operator estimate in Theorem 1.1, we conclude that

$$\begin{aligned} \inf \sigma(H_N) &= \langle \Psi_{\mathrm{GS}}, H_N \Psi_{\mathrm{GS}} \rangle = \left\langle \Psi'_{\mathrm{GS}}, \mathcal{U} H_N \mathcal{U}^* \Psi'_{\mathrm{GS}} \right\rangle \\ &= E_{\mathrm{FS}} + E_{\mathrm{corr}} + \left\langle \Psi'_{\mathrm{GS}}, \left(H'_{\mathrm{kin}} + H_{\mathrm{eff}} + \mathcal{E}_{\mathcal{U}}\right) \Psi'_{\mathrm{GS}} \right\rangle \ge E_{\mathrm{FS}} + E_{\mathrm{corr}} + O(k_F^{1-\frac{1}{94}+\epsilon}). \end{aligned} \tag{10.41}$$

Here, we used the operator inequalities

$$H'_{\mathrm{kin}} \ge 0, \quad H_{\mathrm{eff}} \ge 0, \quad \mathcal{E}_{\mathcal{U}} \ge C k_F^{1-\frac{1}{94}+\epsilon} (k_F^{-1}\mathcal{N}_E H'_{\mathrm{kin}} + H'_{\mathrm{kin}} + k_F) \tag{10.42}$$

and the a priori estimate (10.40). This completes the proof of Theorem 1.3.

10.4. Proof of Theorems 1.4 and 1.5

In this subsection, we study the effective operator H_{eff} in Theorem 1.1 in more detail. First, we prove the following remarkable fact.

Proposition 10.1. *We have the operator identity on $D(H'_{\mathrm{kin}})$:*

$$2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p\in L_k} \lambda_{k,p} b^*_{k,p} b_{k,p} = \mathcal{N}_E H'_{\mathrm{kin}}.$$

Proof of Proposition 10.1. The idea is simply to interchange the summation on $k \in \mathbb{Z}^3_*$ and $p \in L_k$. By rephrasing the condition that $p \in L_k$, we have the equivalences

$$\begin{aligned} \left(k \in \mathbb{Z}^3_*\right) \wedge (p \in L_k) &\Leftrightarrow \left(k \in \mathbb{Z}^3_*\right) \wedge (|p-k| \le k_F < |p|) \\ &\Leftrightarrow \left(k \in \mathbb{Z}^3_*\right) \wedge \left(k \in \overline{B}(p, k_F)\right) \wedge \left(p \in B_F^c\right) \\ &\Leftrightarrow \left(p \in B_F^c\right) \wedge \left(k \in \overline{B}(p,k_F) \cap \mathbb{Z}^3\right), \end{aligned} \tag{10.43}$$

where we could replace $\mathbb{Z}^3_*$ by $\mathbb{Z}^3$ in the last line as the conditions $p \in B_F^c = \mathbb{Z}^3 \backslash \overline{B}(0,k_F)$ and $k \in \overline{B}(p,k_F)$ exclude $k = 0$ automatically. Recognizing that $\overline{B}(p,k_F) \cap \mathbb{Z}^3 = B_F + p$, we can now write

$$\begin{aligned} 2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p\in L_k} \lambda_{k,p} b^*_{k,p} b_{k,p} &= \sum_{k\in\mathbb{Z}^3_*} \sum_{p\in L_k} \left(|p|^2 - |p-k|^2\right) b^*_{k,p} b_{k,p} \\ &= \sum_{p\in B_F^c} \sum_{k\in(B_F+p)} |p|^2 b^*_{k,p} b_{k,p} - \sum_{p\in B_F^c} \sum_{k\in(B_F+p)} |p-k|^2 b^*_{k,p} b_{k,p}, \end{aligned} \tag{10.44}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and by expanding the excitation operators, we find for the first sum that

$$\begin{aligned}\sum_{p\in B_F^c}\sum_{k\in(B_F+p)}|p|^2 b_{k,p}^* b_{k,p} &= \sum_{p\in B_F^c}\sum_{k\in(B_F+p)}|p|^2 c_p^* c_{p-k} c_{p-k}^* c_p \\ &= \sum_{p\in B_F^c}\left(\sum_{k\in(B_F+p)} c_{p-k} c_{p-k}^*\right)|p|^2 c_p^* c_p \\ &= \sum_{p\in B_F^c}\left(\sum_{k\in B_F} c_{-k} c_{-k}^*\right)|p|^2 c_p^* c_p = \mathcal{N}_E \sum_{p\in B_F^c}|p|^2 c_p^* c_p \end{aligned} \tag{10.45}$$

as $\sum_{k\in B_F} c_{-k}c_{-k}^* = \sum_{k\in B_F} c_k c_k^* = \mathcal{N}_E$ by the particle-hole symmetry, and similarly,

$$\begin{aligned}\sum_{p\in B_F^c}\sum_{k\in(B_F+p)}|p-k|^2 b_{k,p}^* b_{k,p} &= \sum_{p\in B_F^c} c_p^* c_p \sum_{k\in(B_F+p)}|p-k|^2 c_{p-k} c_{p-k}^* \\ &= \sum_{p\in B_F^c} c_p^* c_p \sum_{k\in B_F}|k|^2 c_k c_k^* = \mathcal{N}_E \sum_{k\in B_F}|k|^2 c_k c_k^* \end{aligned} \tag{10.46}$$

for the claimed equality of

$$T = 2\sum_{k\in\mathbb{Z}_*^3}\sum_{p\in L_k}\lambda_{k,p} b_{k,p}^* b_{k,p} = \mathcal{N}_E\left(\sum_{p\in B_F^c}|p|^2 c_p^* c_p - \sum_{p\in B_F}|p|^2 c_p c_p^*\right) = \mathcal{N}_E H'_{\mathrm{kin}}. \tag{10.47}$$

To complete the proof, let us show that the relevant operators are well defined on the domain $D(H'_{\mathrm{kin}})$. This is clear for $\mathcal{N}_E H'_{\mathrm{kin}}$ since $\mathcal{N}_E$ is a bounded operator ($0 \le \mathcal{N}_E \le N$ on $\mathcal{H}_N$). For T, we can interchange the summations of k and p using the same observation in (10.43). This gives the quadratic form estimate

$$\begin{aligned} T = 2\sum_{k\in\mathbb{Z}_*^3}\sum_{p\in L_k}\lambda_{k,p} b_{k,p}^* b_{k,p} &= \sum_{k\in\mathbb{Z}_*^3}\sum_{p\in L_k}\left||p|^2-\zeta\right| b_{k,p}^* b_{k,p} + \sum_{k\in\mathbb{Z}_*^3}\sum_{p\in L_k}\left||p-k|^2-\zeta\right| b_{k,p}^* b_{k,p} \\ &\le \sum_{p\in B_F^c}\left(\sum_{k\in B_F} c_k c_k^*\right)\left||p|^2-\zeta\right| c_p^* c_p + \sum_{p\in B_F^c} c_p^* c_p \sum_{k\in B_F}\left||k|^2-\zeta\right| c_k c_k^* \le \mathcal{N}_E H'_{\mathrm{kin}}, \end{aligned} \tag{10.48}$$

where $\zeta > 0$ is the constant in (1.14). Moreover, it is easily seen that T commutes with both $\mathcal{N}_E$ and H'_{kin}. Therefore, the above quadratic form estimate also implies the stronger estimate

$$T^2 \le \left(\mathcal{N}_E H'_{\mathrm{kin}}\right)^2, \tag{10.49}$$

which justifies that $D(T) \subset D(\mathcal{N}_E H'_{\mathrm{kin}}) \subset D(H'_{\mathrm{kin}})$. □

Now we are ready to give the

Proof of Theorem 1.5. Thanks to Proposition 10.1 and the identity $\langle e_p, h_k e_q\rangle = \lambda_{k,p}\delta_{p,q}$, we have

$$\begin{aligned} H_{\mathrm{eff}} &= H'_{\mathrm{kin}} + 2\sum_{k\in\mathbb{Z}_*^3}\sum_{p,q\in L_k}\left\langle e_p, \left(\widetilde{E}_k - h_k\right) e_q\right\rangle b_{k,p}^* b_{k,q} \\ &= 2\sum_{k\in\mathbb{Z}_*^3}\sum_{p,q\in L_k}\langle e_p, \widetilde{E}_k e_q\rangle b_{k,p}^* b_{k,q} - (\mathcal{N}_E - 1)\, H'_{\mathrm{kin}}. \end{aligned} \tag{10.50}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Since $[H_{\text{eff}}, \mathcal{N}_E] = 0$, we can restrict H_{eff} to the eigenspaces of $\mathcal{N}_E$: for every $M = \{1, 2, \ldots\}$, we can write the restriction to $\{\mathcal{N}_E = M\}$ for $M \in \mathbb{N}$ in the quasi-bosonic form

$$H_{\text{eff}}|_{\mathcal{N}_E=M} = 2 \sum_{k\in\mathbb{Z}^3_*} \sum_{p,q\in L_k} \left\langle e_p, \left(\widetilde{E}_k - \left(1 - M^{-1}\right) h_k\right) e_q\right\rangle b^*_{k,p} b_{k,q}. \tag{10.51}$$

□

Proof of Theorem 1.4. We only need to verify the statement on the effective operator $H_{\text{eff}}|_{\mathcal{N}_E=M}$ with $M = 1$. In this case, it is convenient to introduce the total momentum $P = (P_1, P_2, P_3)$, where each P_j is given by $P_j = \sum_{p\in\mathbb{Z}^3} p_j c^*_p c_p$. It is easily checked that P_j obeys the commutators

$$\left[P_j, b_{k,p}\right] = -k_j b_{k,p}, \quad \left[P_j, b^*_{k,p}\right] = k_j b^*_{k,p}, \tag{10.52}$$

and additionally $\left[P_j, H'_{\text{kin}}\right] = 0$, whence the effective Hamiltonian H_{eff} also commutes with P_j, $j = 1, 2, 3$. It also holds that $\left[\mathcal{N}_E, P_j\right] = 0$, so we may restrict H_{eff} to the simultanous eigenspaces of $\mathcal{N}_E$ and P. It follows from $\left[P_j, b^*_{k,p}\right] = k_j b^*_{k,p}$ that this simultaneous eigenspace is precisely

$$\{\Psi \in \mathcal{H}_N \mid \mathcal{N}_E\Psi = \Psi,\ P\Psi = k\Psi\} = \operatorname{span}\left(b^*_{k,p}\psi_{\text{FS}}\right)_{p\in L_k} = \left\{b^*_k(\varphi)\psi_{\text{FS}} \mid \varphi \in L^2(L_k)\right\}. \tag{10.53}$$

In fact, the mapping $U : \varphi \mapsto b^*_k(\varphi)\psi_{\text{FS}}$ is an isomorphism. To see that, we compute, using the commutation relations of the excitation operators and the fact that $b_k(\phi)\psi_{\text{FS}} = 0 = \varepsilon_{k,k}(\phi; \varphi)\psi_{\text{FS}}$ for any $\phi, \varphi \in L^2(L_k)$, that

$$\begin{aligned}\langle U\phi, U\varphi\rangle &= \left\langle b^*_k(\phi)\psi_{\text{FS}}, b^*_k(\varphi)\psi_{\text{FS}}\right\rangle = \left\langle \psi_{\text{FS}}, \left(b^*_k(\varphi)b_k(\phi) + \langle\phi,\varphi\rangle + \varepsilon_{k,k}(\phi;\varphi)\right)\psi_{\text{FS}}\right\rangle \\ &= \langle\phi,\varphi\rangle \langle\psi_{\text{FS}}, \psi_{\text{FS}}\rangle = \langle\phi,\varphi\rangle,\end{aligned} \tag{10.54}$$

so U is a unitary embedding of $L^2(L_k)$ into $\{\Psi \in \mathcal{H}_N \mid \mathcal{N}_E\Psi = \Psi,\ P\Psi = k\Psi\}$ and hence an isomorphism for dimensional reasons.

Similarly, we find as $H_{\text{eff}}|_{\mathcal{N}_E=1} = 2\sum_{l\in\mathbb{Z}^3_*}\sum_{p,q\in L_l}\left\langle e_p, \widetilde{E}_l e_q\right\rangle b^*_{l,p} b_{l,q}$ that for any $\phi, \varphi \in L^2(L_k)$,

$$\begin{aligned}\langle U\phi, H_{\text{eff}}U\varphi\rangle &= 2\sum_{l\in\mathbb{Z}^3_*}\sum_{p,q\in L_l}\left\langle e_p, \widetilde{E}_l e_q\right\rangle \left\langle b_{l,p}b^*_k(\phi)\psi_{\text{FS}}, b_{l,q}b^*_k(\varphi)\psi_{\text{FS}}\right\rangle \\ &= 2\sum_{k\in\mathbb{Z}^3_*}\sum_{p,q\in L_k}\left\langle e_p, \widetilde{E}_l e_q\right\rangle \delta_{k,l}\left\langle\phi, e_p\right\rangle\left\langle e_q,\varphi\right\rangle = 2\left\langle\phi, \widetilde{E}_k\varphi\right\rangle,\end{aligned} \tag{10.55}$$

whence $U^* H_{\text{eff}} U = 2\widetilde{E}_k$. By elaborating the above argument slightly, one finds that the mapping

$$\tilde{U} : \bigoplus_{k\in\mathbb{Z}^3_*} L^2(L_k) \to \{\Psi \in \mathcal{H}_N \mid \mathcal{N}_E\Psi = \Psi\} \tag{10.56}$$

defined by

$$\tilde{U}\bigoplus_{k\in\mathbb{Z}^3_*}\varphi_k = \sum_{k\in\mathbb{Z}^3_*} b^*_k(\varphi_k)\psi_{\text{FS}} \tag{10.57}$$

is likewise a unitary isomorphism under which $\tilde{U}^* H_{\text{eff}}\tilde{U} = \bigoplus_{k\in\mathbb{Z}^3_*}\widetilde{E}_k$. □

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

A. Appendix: Lattice estimates and Riemann sums

In this appendix, we collect several useful estimates for the lattice points and Riemann sums. In particular, we want to obtain estimates on the sum $\sum_{p\in L_k}\lambda_{k,p}^{\beta}$, where $\beta \le 0$ and

$$L_k = (B_F + k)\backslash B_F = \left(\overline{B}\,(k,k_F)\backslash \overline{B}(0,k_F)\right)\cap \mathbb{Z}^3, \quad \lambda_{k,p} = \frac{1}{2}\left(|p|^2 - |p-k|^2\right) = k\cdot p - \frac{1}{2}|k|^2.$$

It is natural to expect the sum to be approximated by the corresponding integrals – that is,

$$\sum_{p\in L_k} f(\lambda_{k,p}) \sim \int_{\overline{B}(k,k_F)\backslash \overline{B}(0,k_F)} f\left(k\cdot p - \frac{1}{2}|k|^2\right) dp, \tag{A.1}$$

with $f(t) = t^{\beta}$. Indeed, when $-1 < \beta \le 0$, the Riemann sum is well behaved, and using general estimation methods based on (A.1), we have the following:

Proposition A.1. *For all $k \in \mathbb{Z}^3_*$ and $-1 < \beta \le 0$, it holds that*

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta} \le C\begin{cases} k_F^{2+\beta}|k|^{1+\beta} & |k| < 2k_F \\ k_F^3|k|^{2\beta} & |k| \ge 2k_F\end{cases}$$

for a constant $C > 0$ depending only on β.

For $\beta \le -1$, the summands are, however, too divergent to obtain good estimates using only general methods. For example, when $\beta = -1$, using standard estimates based on (A.1), we obtain

$$\sum_{p\in L_k}\lambda_{k,p}^{-1} \le C\begin{cases} \left(1 + |k|^{-1}\log\left(k_F\right)\right)k_F & |k| < 2k_F \\ k_F^3|k|^{-2} & |k| \ge 2k_F,\end{cases} \tag{A.2}$$

which is non-optimal when $|k| < 2k_F$. To obtain good estimates on the sums $\sum_{p\in L_k} f\left(\lambda_{k,p}\right)$ for more singular f, we will instead derive a summation formula which reduces the 3-dimensional Riemann sum to two 1-dimensional Riemann sums plus an error term. The utility of this summation formula, apart from reducing the dimensionality of the sums, is that the 1-dimensional Riemann sums contain weighting factors which explicitly cancel the divergent behaviour of the summands. To derive this summation formula, we need to carry out a detailed analysis of the structure of the lunes L_k, which is related to a lattice point counting problem in the plane and can be handled by classical results from analytic number theory.

With the summation formula at our disposal, we can improve (A.2) to the following:

Proposition A.2. *For all $k \in \mathbb{Z}^3_*$, it holds that*

$$\sum_{p\in L_k}\lambda_{k,p}^{-1} \le Ck_F, \quad k_F \to \infty,$$

for a constant $C > 0$ independent of k and k_F.

We refer to [24, Lemma 4.7] and [6, Eq. B.1] for results similar to Proposition A.2. However, the k-independence of the constant C was not completely clear in these previous results.

For more singular functions, we have the following:

Proposition A.3. *For $-\frac{4}{3} < \beta < -1$ and $k \in \overline{B}\left(0,k_F^{\gamma}\right)$ with $0 < \gamma < \frac{4+3\beta}{8-3\beta}$, we have*

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta} \le Ck_F^{2+\beta}|k|^{1+\beta}, \quad k_F \to \infty.$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Moreover, for $\beta \le -\frac{4}{3}$ and $k \in \overline{B}(0, 2k_F)$, we have

$$\sum_{p \in L_k} \lambda_{k,p}^{\beta} \le C|k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}}, \quad k_F \to \infty.$$

Here, the constant $C > 0$ is independent of k and k_F.

In Proposition A.3, the first bound is optimal in terms of both $k_F^{2+\beta}$ and $|k|^{\beta}$. The second bound is unlikely to be optimal but is sufficient in applications if $|k|$ is relatively small.

Finally, for the kinetic estimate in Proposition 2.3, we need the following proposition, which can be obtained by the same argument of the above results.

Proposition A.4. *Let $S_{k,\lambda}^1, S_{k,\lambda}^2$ as in (2.14), (2.19) with $k \in \overline{B}(0, k_F) \cap \mathbb{Z}_*^3$ and $0 < \lambda = \lambda(k_F, k) \le \frac{1}{6} k_F^2$. Then there exists a constant $C > 0$ independent of k, k_F, λ such that*

$$|S_{k,\lambda}^1| + |S_{k,\lambda}^2| \le C \left(|k|^{-1}\lambda + |k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}} \right) (\lambda + |k|), \quad k_F \to \infty.$$

In the rest of the appendix, we will discuss some preliminary results in Sections A.1 and A.2 and then turn to the proofs of Propositions A.1, A.2, A.3 and A.4.

A.1. Some lattice concepts

Let V be a real n-dimensional vector space. The lattice $\Lambda \subset V$ generated by $(v_i)_{i=1}^n$ is

$$\Lambda = \Lambda(v_1, \cdots, v_n) = \left\{ \sum_{i=1}^n m_i v_i \mid m_1, \ldots, m_n \in \mathbb{Z} \right\}. \tag{A.3}$$

Given two bases $(v_i)_{i=1}^n$ and $(w_i)_{i=1}^n$, it may happen that $\Lambda(v_1, \ldots, v_n) = \Lambda(w_1, \ldots, w_n)$ even if the bases are not equal. The following is well known (see, for example, [28, p. 4])

Proposition A.5. *Let $(v_i)_{i=1}^n$ and $(w_i)_{i=1}^n$ be bases of V. Then $\Lambda(v_1, \ldots, v_n) = \Lambda(w_1, \ldots, w_n)$ if and only if the transition matrix $T = (T_{i,j})_{i,j=1}^n$ defined by*

$$w_i = \sum_{j=1}^n T_{i,j} v_j, \quad 1 \le i \le n$$

has integer entries and determinant ± 1.

This result has an important consequence when V is endowed with an inner product.

Proposition A.6. *Let Λ be a lattice in $(V, \langle \cdot, \cdot \rangle)$ and let $(v_i)_{i=1}^n$ generate Λ. Then the quantity*

$$d(\Lambda) = \left| \det \begin{pmatrix} \langle e_1, v_1 \rangle & \cdots & \langle e_n, v_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle e_1, v_n \rangle & \cdots & \langle e_n, v_n \rangle \end{pmatrix} \right| = \sqrt{\det \begin{pmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_n, v_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle v_1, v_n \rangle & \cdots & \langle v_n, v_n \rangle \end{pmatrix}}$$

is independent of the choice of generators $(v_i)_{i=1}^n$. Here, $(e_i)_{i=1}^n$ is any orthonormal basis for V.

Here, $d(\Lambda)$ is referred to as the covolume (or simply determinant) of Λ. The fact that $d(\Lambda)$ is independent of $(e_i)_{i=1}^n$ follows by a standard orthonormal expansion, while the fact that $d(\Lambda)$ is independent

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

of $(v_i)_{i=1}^n$ follows from the previous proposition: if $(v_i)_{i=1}^n$ and $(w_i)_{i=1}^n$ are two bases with transition matrix T, then

$$\left|\det\begin{pmatrix} \langle e_1, w_1\rangle & \cdots & \langle e_n, w_1\rangle \\ \vdots & \ddots & \vdots \\ \langle e_1, w_n\rangle & \cdots & \langle e_n, w_n\rangle \end{pmatrix}\right| = |\det(T)| \left|\det\begin{pmatrix} \langle e_1, v_1\rangle & \cdots & \langle e_n, v_1\rangle \\ \vdots & \ddots & \vdots \\ \langle e_1, v_n\rangle & \cdots & \langle e_n, v_n\rangle \end{pmatrix}\right|. \tag{A.4}$$

Given a lattice Λ in an n-dimensional inner product space V, one defines the successive minima $(\lambda_i)_{i=1}^n$ (relative to the closed unit ball $\overline{B}\,(0,1)$) by

$$\lambda_i = \inf\left\{\lambda \mid \overline{B}\,(0,\lambda) \cap \Lambda \text{ contains } i \text{ linearly independent vectors}\right\}, \quad 1 \le i \le n. \tag{A.5}$$

A well-known theorem due to Minkowski provides an inequality relating the successive minima of a lattice Λ to its covolume:

Theorem A.7 (Minkowski's second theorem). *Let Λ be a lattice in an n-dimensional inner product space V. Then it holds that*

$$\frac{2^n d(\Lambda)}{n!\,\mathrm{Vol}(\overline{B}(0,1))} \le \lambda_1 \cdots \lambda_n \le \frac{2^n d(\Lambda)}{\mathrm{Vol}(\overline{B}(0,1))}.$$

Note that although $\overline{B}\,(0,\lambda_n) \cap \Lambda$ contains n linearly independent vectors, it is not ensured that these n vectors can be chosen to generate Λ. For $n = 2$, this is nonetheless the case:

Corollary A.8. *Let Λ be a lattice in a 2-dimensional inner product space V. Then there exist vectors $v_1, v_2 \in \Lambda$ which generate Λ such that*

$$|v_1||v_2| \le \frac{4}{\pi} d\,(\Lambda)\,.$$

Proof. By definition of λ_2, there exists linearly independent vectors $v_1, v_2 \in \Lambda$ such that $|v_1|, |v_2| \le \lambda_2$ and by Minkowski's second theorem $|v_1||v_2| \le \frac{4}{\pi} d(\Lambda)$. We argue that v_1 and v_2 must necessarily generate Λ. Suppose otherwise (i.e., that there exists a $v \in \Lambda$ such that $v \neq m_1 v_1 + m_2 v_2$ for $m_1, m_2 \in \mathbb{Z}$). As v_1 and v_2 are linearly independent and $\dim(V) = 2$, these do nonetheless span V (i.e., there must exist $c_1, c_2 \in \mathbb{R}$ such that $v = c_1 v_1 + c_2 v_2$).

Now we can assume that $|c_1|, |c_2| \le \frac{1}{2}$, since as Λ is a lattice and $v_1, v_2, v \in \Lambda$, we may subtract multiples of v_1 and v_2 from v until this is the case. Then, since $|\langle v_1, v_2\rangle| < |v_1||v_2|$ by the Cauchy-Schwarz inequality (strict inequality being a consequence of the linear independence of v_1 and v_2), we can estimate that

$$\begin{aligned} |v|^2 &= |v_1|^2 c_1^2 + |v_2|^2 c_2^2 + 2\,\langle v_1, v_2\rangle\, c_1 c_2 < |v_1|^2 c_1^2 + |v_2|^2 c_2^2 + 2|v_1||v_2|\,|c_1|\,|c_2| \\ &= (|c_1|\,|v_1| + |c_2|\,|v_2|)^2 \le \left(\frac{1}{2}\lambda_2 + \frac{1}{2}\lambda_2\right)^2 = \lambda_2^2, \end{aligned} \tag{A.6}$$

or $|v| < \lambda_2$. But this contradicts the minimality of λ_2 as $v \neq 0$, and at least one of $\{v_1, v\}$ and $\{v_2, v\}$ must be a linearly independent set, so such a v cannot exist. □

The sublattice orthogonal to a vector $k \in \mathbb{Z}^3$

Consider $\mathbb{Z}^3$ as a lattice in $\mathbb{R}^3$ endowed with the usual dot product. Let $k = (k_1, k_2, k_3) \in \mathbb{Z}^3 \setminus \{0\}$ be arbitrary and write $\hat{k} = |k|^{-1} k$. Now we consider the set $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\right\}$, namely, the sublattice orthogonal to k. Let us recall the following well-known result.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Theorem A.9. *For $(k_1, k_2, k_3) \in \mathbb{Z}^3 \setminus \{0\}$ and $c \in \mathbb{Z}$, the linear Diophantine equation*

$$k_1 m_1 + k_2 m_2 + k_3 m_3 = c$$

is solvable with $(m_1, m_2, m_3) \in \mathbb{Z}^3$ if and only if c is a multiple of $\gcd(k_1, k_2, k_3)$. Moreover, in this case, there exist linearly independent vectors $v_1, v_2 \in \mathbb{Z}^3$, which do not depend on c, such that if (m_1^*, m_2^*, m_3^*) is any particular solution of the equation, then all solutions are given by

$$\left\{(m_1, m_2, m_3) \in \mathbb{Z}^3 \mid k_1 m_1 + k_2 m_2 + k_3 m_3 = c\right\} = \left(m_1^*, m_2^*, m_3^*\right) + \{a_1 v_1 + a_2 v_2 \mid a_1, a_2 \in \mathbb{Z}\}.$$

Note that the second part of the proposition states that (up to translation by a particular solution) the solution set of a linear Diophantine equation forms a lattice, much as the solution set of a real-variable linear equation forms a linear subspace. This result implies the following:

Proposition A.10. *Let $k = (k_1, k_2, k_3) \in \mathbb{Z}^3 \setminus \{0\}$ be given. Then with $l = |k|^{-1} \gcd(k_1, k_2, k_3)$, the following disjoint union of nonempty sets holds:*

$$\mathbb{Z}^3 = \bigcup_{m \in \mathbb{Z}} \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}.$$

Additionally, there exist linearly independent vectors $v_1, v_2 \in \mathbb{Z}^3$, which span $\left\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\right\}$, such that for any $m \in \mathbb{Z}$, it holds for all $q \in \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}$ that

$$\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\} = q + \{a_1 v_1 + a_2 v_2 \mid a_1, a_2 \in \mathbb{Z}\}.$$

Proof. Clearly, $\mathbb{Z}^3 = \bigcup_{t \in \mathbb{R}} \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = t\right\}$, so we must determine for which values of t it holds that $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = t\right\} \neq \emptyset$. The equation $\hat{k} \cdot p = t$ is equivalent to

$$k_1 p_1 + k_2 p_2 + k_3 p_3 = |k| t, \tag{A.7}$$

where $p = (p_1, p_2, p_3) \in \mathbb{Z}^3$, and as the left-hand side is an integer, we must have $t = |k|^{-1} c$ for some $c \in \mathbb{Z}$. Theorem A.9 now furthermore implies that $c = \gcd(k_1, k_2, k_3) \cdot m$ for some $m \in \mathbb{Z}$, so that $t = |k|^{-1} \gcd(k_1, k_2, k_3) \cdot m = lm$, and as p was arbitrary, we see that $\mathbb{Z}^3 = \bigcup_{m \in \mathbb{Z}} \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}$ as claimed.

That all the sets $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}$, $m \in \mathbb{Z}$, are also nonempty similarly follows from the 'only if' part of Theorem A.9, and the representation

$$\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\} = q + \{a_1 v_1 + a_2 v_2 \mid a_1, a_2 \in \mathbb{Z}\} \tag{A.8}$$

for linearly independent $v_1, v_2 \in \mathbb{Z}^3$ is likewise a simple restatement of the second part of the theorem. Finally, that v_1 and v_2 span $\left\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\right\}$ follows by noting that $q = (0, 0, 0)$ is a particular solution of $\hat{k} \cdot p = 0$, whence by the previous part

$$\{v_1, v_2\} \subset q + \{a_1 v_1 + a_2 v_2 \mid a_1, a_2 \in \mathbb{Z}\} = \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\right\} \subset \left\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\right\}, \tag{A.9}$$

so we find that $\operatorname{span}(\{v_1, v_2\}) = \left\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\right\}$ by linear independence of $\{v_1, v_2\}$ and dimensionality consideration. □

Proposition A.10 implies that $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\right\}$ is a lattice in $\{k\}^\perp = \left\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\right\}$. Since $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\right\}$ is a lattice, it has a well-defined covolume

$$\sqrt{\det \begin{pmatrix} v_1 \cdot v_1 & v_2 \cdot v_1 \\ v_1 \cdot v_2 & v_2 \cdot v_2 \end{pmatrix}} = \sqrt{|v_1|^2 |v_2|^2 - (v_1 \cdot v_2)^2} \tag{A.10}$$

for any choice of generators v_1 and v_2. This covolume is explicitly given by the following:

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proposition A.11. *For any* $v_1, v_2 \in \mathbb{Z}^3$ *generating* $\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\}$*, it holds that*

$$|v_1|^2|v_2|^2 - (v_1 \cdot v_2)^2 = l^{-2}$$

with $l = |k|^{-1} \gcd(k_1, k_2, k_3)$*. Additionally,* v_1 *and* v_2 *can be chosen such that* $|v_1|^2 + |v_2|^2 \le \frac{8}{\pi^2 l^2}$*.*

Proof. Let v_1 and v_2 generate $\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\}$ and let $w \in \{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = l\}$ be arbitrary. By linearity, it holds that

$$\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\} = mw + \{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\}, \quad m \in \mathbb{Z}, \tag{A.11}$$

so by the Proposition A.10,

$$\mathbb{Z}^3 = \bigcup_{m \in \mathbb{Z}} \left(mw + \{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\}\right) = \{m_1 v_1 + m_2 v_2 + m_3 w \mid m_1, m_2, m_3 \in \mathbb{Z}\} \tag{A.12}$$

(i.e., (v_1, v_2, w) is a set of generators for $\mathbb{Z}^3$). Now, let $\{k\}^\perp = \{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\}$ be the orthogonal complement of $\{k\}$. Let (e_1, e_2) be an orthonormal basis for $\{k\}^\perp$ so that $\left(e_1, e_2, \hat{k}\right)$ forms an orthonormal basis for $\mathbb{R}^3$. Then $d\left(\mathbb{Z}^3\right)$ is equal to

$$\begin{aligned}
&\left|\det\begin{pmatrix} e_1 \cdot v_1 & e_2 \cdot v_1 & \hat{k} \cdot v_1 \\ e_1 \cdot v_2 & e_2 \cdot v_2 & \hat{k} \cdot v_2 \\ e_1 \cdot w & e_2 \cdot w & \hat{k} \cdot w \end{pmatrix}\right| = \left|\det\begin{pmatrix} e_1 \cdot v_1 & e_2 \cdot v_1 & 0 \\ e_1 \cdot v_2 & e_2 \cdot v_2 & 0 \\ e_1 \cdot w & e_2 \cdot w & l \end{pmatrix}\right| = l\left|\det\begin{pmatrix} e_1 \cdot v_1 & e_2 \cdot v_1 \\ e_1 \cdot v_2 & e_2 \cdot v_2 \end{pmatrix}\right| \\
&= l\sqrt{\det\begin{pmatrix} v_1 \cdot v_1 & v_2 \cdot v_1 \\ v_1 \cdot v_2 & v_2 \cdot v_2 \end{pmatrix}} = l\sqrt{|v_1|^2|v_2|^2 - (v_1 \cdot v_2)^2},
\end{aligned} \tag{A.13}$$

but it is also clear that $d\left(\mathbb{Z}^3\right) = 1$, so the first result follows. From this result, (A.10) and Corollary A.8, we deduce that there exist generators v_1 and v_2 such that

$$|v_1||v_2| \le \frac{4}{\pi} d\left(\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\}\right) = \frac{4}{\pi} l^{-1}. \tag{A.14}$$

Since $v_1, v_2 \in \mathbb{Z}^3 \setminus \{0\}$, we have $|v_1|, |v_2| \ge 1$, and hence,

$$|v_1|^2 + |v_2|^2 \le 2|v_1|^2|v_2|^2 \le \frac{8}{\pi^2} l^{-2}. \tag{A.15}$$

□

A.2. Plane decomposition of L_k and the summation formula

Now we turn to consider the lune $L_k = \{p \in \mathbb{Z}^3 \mid |p - k| \le k_F < |p|\}$. Throughout this subsection, we let $k = (k_1, k_2, k_3) \in \mathbb{Z}^3 \setminus \{0\}$ be fixed and write $\hat{k} = |k|^{-1} k$ and $l = |k|^{-1} \gcd(k_1, k_2, k_3)$ for the sake of brevity. The integrands of the Riemann sums we must consider only depend on the quantity $\lambda_{k,p} = k \cdot p - \frac{1}{2}|k|^2 = |k|\left(\hat{k} \cdot p - \frac{1}{2}|k|\right)$, so we begin by decomposing L_k along the $\hat{k} \cdot p =$ constant planes. By the definition of L_k, it easily follows that

$$L_k \subset \left\{p \in \mathbb{R}^3 \mid \frac{1}{2}|k| < \hat{k} \cdot p \le k_F + |k|\right\}. \tag{A.16}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Letting m^* be the least integer and M^* the greatest integer such that

$$\frac{1}{2}|k| < lm^*, \quad lM^* \leq k_F + |k|, \tag{A.17}$$

we see that the lune L_k can be expressed as the disjoint union

$$L_k = \bigcup_{m=m^*}^{M^*} L_k^m, \quad L_k^m = \left\{p \in L_k \mid \hat{k} \cdot p = lm\right\}. \tag{A.18}$$

So for any function $f : \mathbb{R} \to \mathbb{R}$, we may express a sum of the form $\sum_{p \in L_k} f\left(\lambda_{k,p}\right)$ as

$$\sum_{p \in L_k} f(\lambda_{k,p}) = \sum_{m=m^*}^{M^*} \sum_{p \in L_k^m} f\left(|k|\left(\hat{k} \cdot p - \frac{1}{2}|k|\right)\right) = \sum_{m=m^*}^{M^*} f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right) |L_k^m|. \tag{A.19}$$

Rewriting L_k^m

To proceed, we must analyze $|L_k^m|$, the number of points contained in L_k^m. For this we first rewrite

$$L_k = \left\{p \in \mathbb{Z}^3 \mid |p - k| \leq k_F < |p|\right\} = \left\{p \in \mathbb{Z}^3 \mid k_F^2 < |p|^2 \leq k_F^2 - |k|^2 + 2k \cdot p\right\}. \tag{A.20}$$

Now let $P_\perp : \mathbb{R}^3 \to \{k\}^\perp$ denote the orthogonal projection onto $\{k\}^\perp$. Then for any $p \in \mathbb{R}^3$, $|p|^2 = |P_\perp p|^2 + \left(\hat{k} \cdot p\right)^2$, whence

$$\begin{aligned} L_k &= \left\{p \in \mathbb{Z}^3 \mid k_F^2 - \left(\hat{k} \cdot p\right)^2 < |P_\perp p|^2 \leq k_F^2 - |k|^2 + 2k \cdot p - \left(\hat{k} \cdot p\right)^2\right\} \\ &= \left\{p \in \mathbb{Z}^3 \mid k_F^2 - \left(\hat{k} \cdot p\right)^2 < |P_\perp p|^2 \leq k_F^2 - \left(\hat{k} \cdot p - |k|\right)^2\right\}, \end{aligned} \tag{A.21}$$

and so the sets $L_k^m = L_k \cap \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}$ may be written as

$$\begin{aligned} L_k^m &= \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm,\ k_F^2 - (lm)^2 < |P_\perp p|^2 \leq k_F^2 - (lm - |k|)^2\right\} \\ &= \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm,\ \left(R_1^m\right)^2 < |P_\perp p|^2 \leq \left(R_2^m\right)^2\right\}, \end{aligned} \tag{A.22}$$

where the real numbers R_1^m and R_2^m are

$$R_1^m = \sqrt{k_F^2 - (lm)^2}, \quad R_2^m = \sqrt{k_F^2 - (lm - |k|)^2}, \quad m^* < m \leq M^*, \tag{A.23}$$

which are well defined by definition of m^* and M^*.

Now by Proposition A.10, we can find the generators $v_1, v_2 \in \mathbb{Z}^3$ of $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = 0\right\}$. Moreover, a fixed $m^* \leq m \leq M^*$, there exists $q \in \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}$, and any $p \in \mathbb{Z}^3$ is an element of $\left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\}$ if and only if it can be written as

$$p = a_1 v_1 + a_2 v_2 + q \tag{A.24}$$

for some $a_1, a_2 \in \mathbb{Z}$. Since $P_\perp q \in \{k\}^\perp$ by definition and the proposition likewise asserts that v_1 and v_2 span $\{k\}^\perp$, there must also exist $b_1, b_2 \in \mathbb{R}$ such that $P_\perp q = b_1 v_1 + b_2 v_2$. Consequently, $P_\perp p$ for our arbitrary element p takes the form

$$P_\perp p = a_1 P_\perp v_1 + a_2 P_\perp v_2 + P_\perp q = (a_1 + b_1)\, v_1 + (a_2 + b_2)\, v_2 \tag{A.25}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

whence

$$|P_\perp p|^2 = (a_1+b_1)^2 |v_1|^2 + (a_2+b_2)^2 |v_2|^2 + 2(a_1+b_1)(a_2+b_2)(v_1\cdot v_2), \tag{A.26}$$

so by equation (A.22) we conclude that

$$\begin{aligned}|L_k^m| &= \Big|\Big\{ (a_1,a_2)\in\mathbb{Z}^2 \mid \left(R_1^m\right)^2 < (a_1+b_1)^2|v_1|^2 + (a_2+b_2)^2|v_2|^2 \\ &\qquad + 2(a_1+b_1)(a_2+b_2)(v_1\cdot v_2) \le \left(R_2^m\right)^2 \Big\}\Big| \\ &= \left|\left(E_2^m\backslash E_1^m - (b_1,b_2)\right)\cap\mathbb{Z}^2\right|, \end{aligned} \tag{A.27}$$

where the sets E_1^m and E_2^m, defined by

$$E_i^m = \left\{(x,y)\in\mathbb{R}^2 \mid |v_1|^2x^2 + |v_2|^2y^2 + 2(v_1\cdot v_2)xy \le \left(R_i^m\right)^2\right\}, \quad i=1,2, \tag{A.28}$$

are seen to be (the closed interiors of) ellipses. The analysis of $|L_k^m|$ thus reduces to the estimation of the number of lattice points enclosed by these.

Lattice point estimation

To estimate $|L_k^m| = \left|\left(E_2^m\backslash E_1^m - (b_1,b_2)\right)\cap\mathbb{Z}^2\right|$, we will use the following result on the number of lattice points contained in compact, strictly convex regions in the plane:

Theorem A.12 [19]. *Let $K\subset\mathbb{R}^2$ be a compact, strictly convex set with C^2 boundary and let ∂K have minimal and maximal radii of curvature $0<r_1\le r_2$. If $r_2\ge 1$, then*

$$\left|\left|K\cap\mathbb{Z}^2\right| - \mathrm{Area}\,(K)\right| \le C\frac{r_2}{r_1} r_2^{\frac{2}{3}} \log\left(1+2\sqrt{2r_2}\right)^{\frac{2}{3}}$$

for a constant $C>0$ independent of K, r_1 and r_2.

This result follows from the techniques of Chapter 8 of [19].

From the theorem, we deduce the following practical corollary:

Corollary A.13. *Let $E\subset\mathbb{R}^2$ be an ellipse with radii of curvature $0<r_1\le r_2$. Then,*

$$\left|\left|E\cap\mathbb{Z}^2\right| - \mathrm{Area}(E)\right| \le C\left(1+\frac{r_2}{r_1} r_2^{\frac{2}{3}} \log\left(1+2\sqrt{2r_2}\right)^{\frac{2}{3}}\right)$$

for a constant $C>0$ independent of E, r_1 and r_2.

Proof. The theorem gives the case that $r_2\ge 1$. If $r_2<1$, then we can circumscribe some disk D of radius 1 around E, and trivially

$$\left|\left|E\cap\mathbb{Z}^2\right| - \mathrm{Area}(E)\right| \le \max\left(\left|E\cap\mathbb{Z}^2\right|, \mathrm{Area}(E)\right) \le \max\left(\left|D\cap\mathbb{Z}^2\right|, \mathrm{Area}\,(D)\right) \le C \tag{A.29}$$

as the right-hand side is seen to be bounded irrespective of the exact position of D. □

This corollary lets us estimate that

$$|L_k^m| = \mathrm{Area}\left(E_2^m\backslash E_1^m\right) + O\left(1+\frac{r_2}{r_1} r_2^{\frac{2}{3}} \log\left(1+2\sqrt{2r_2}\right)^{\frac{2}{3}} + \frac{r_2'}{r_1'}\left(r_2'\right)^{\frac{2}{3}} \log\left(1+2\sqrt{2r_2'}\right)^{\frac{2}{3}}\right), \tag{A.30}$$

where r_i and r_i', $i=1,2$ denote the radii of curvature of E_1^m and E_2^m, as the translation by (b_1,b_2) affects neither the areas nor the radii of curvature of the ellipses.

To proceed, we must obtain some information on the geometry of the ellipses E_i^m. By the definition (A.28), the semi-axes $a_i \geq b_i > 0$ of E_i^m are given by

$$a_i = \sqrt{2} R_i^m \left(|v_1|^2 + |v_2|^2 - \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2} \right)^{-\frac{1}{2}},$$
$$b_i = \sqrt{2} R_i^m \left(|v_1|^2 + |v_2|^2 + \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2} \right)^{-\frac{1}{2}}. \tag{A.31}$$

We can now describe the geometry of the ellipses E_i^m in terms of k and m:

Proposition A.14. *If $|k| \leq 2k_F$, then*

$$\text{Area}\left(E_2^m \backslash E_1^m\right) = \begin{cases} 2\pi|k|\left(lm - \frac{1}{2}|k|\right) l & \text{if } lm^* \leq lm \leq k_F, \\ \pi\left(k_F^2 - (lm - |k|)^2\right) l & \text{if } k_F < lm \leq lM^*, \end{cases}$$

and the radii of curvature $0 < r_1 \leq r_2$ of both E_1^m, E_2^m obey

$$\frac{r_2}{r_1} \leq C l^{-3}, \quad r_2 \leq C l^{-1} k_F,$$

for a constant $C > 0$ independent of k and m.

(The condition $|k| \leq 2k_F$ ensures that the lune does not degenerate into a ball, in which case the area formula must be modified.)

Proof. Let v_1 and v_2 be the generators given by Proposition A.11. The area enclosed by an ellipse with semi-axes a and b is πab, so as $E_1^m \subset E_2^m$ for any $m^* \leq m \leq M^*$ and $E_1^m \neq \emptyset$ when $lm \leq k_F$, we find in this case that

$$\text{Area}\left(E_2^m \backslash E_1^m\right) = \pi\left(a_2 b_2 - a_1 b_1\right) = \frac{2\pi\left(\left(R_2^m\right)^2 - \left(R_1^m\right)^2\right)}{\sqrt{\left(|v_1|^2 + |v_2|^2\right)^2 - \left(\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2\right)}}$$
$$= \frac{2\pi\left(k_F^2 - (lm - |k|)^2 - \left(k_F^2 - (lm)^2\right)\right)}{\sqrt{4|v_1|^2|v_2|^2 + 4\left(v_1 \cdot v_2\right)^2}} = 2\pi|k|\left(lm - \frac{1}{2}|k|\right) l \tag{A.32}$$

and similarly in the case $k_F < lm$ that

$$\text{Area}\left(E_2^m \backslash E_1^m\right) = \text{Area}\left(E_2^m\right) = \pi a_2 b_2 = \frac{2\pi\left(R_2^m\right)^2}{2l^{-1}} = \pi\left(k_F^2 - (lm - |k|)^2\right) l. \tag{A.33}$$

For the radii of curvature, we note that for an ellipse with semi-axes $a \geq b > 0$, these are given by $r_1 = a^{-1} b^2$ and $r_2 = b^{-1} a^2$, respectively, so for the ratio $r_1^{-1} r_2$, we can for either of E_1^m and E_2^m estimate using equation (A.31) that

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

$$\frac{r_2}{r_1} = \left(\frac{a_i}{b_i}\right)^3 = \left(\frac{|v_1|^2 + |v_2|^2 + \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2}}{|v_1|^2 + |v_2|^2 - \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2}}\right)^{\frac{3}{2}}$$

$$= \left(\frac{\left(|v_1|^2 + |v_2|^2 + \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + \left(v_1 \cdot v_2\right)^2}\right)^2}{\left(|v_1|^2 + |v_2|^2\right)^2 - \left(\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2\right)}\right)^{\frac{3}{2}}$$

$$\leq \left(\frac{\left(2(|v_1|^2 + |v_2|^2)\right)^2}{4(|v_1|^2|v_2|^2 - (v_1 \cdot v_2)^2)}\right)^{\frac{3}{2}} \leq \left(\frac{(Cl^{-2})^2}{l^2}\right)^{3/2} \leq Cl^{-3} \tag{A.34}$$

and likewise estimate for r_2 that

$$r_2 = \frac{a_i^2}{b_i} = \sqrt{2}R_i^m \frac{\sqrt{|v_1|^2 + |v_2|^2 + \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2}}^{\,2}}{|v_1|^2 + |v_2|^2 - \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2}}$$

$$= \sqrt{2}R_i^m \frac{\left(|v_1|^2 + |v_2|^2 + \sqrt{\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2}\right)^{\frac{3}{2}}}{\left(|v_1|^2 + |v_2|^2\right)^2 - \left(\left(|v_1|^2 - |v_2|^2\right)^2 + 4\left(v_1 \cdot v_2\right)^2\right)}$$

$$\leq \sqrt{2}R_i^m \frac{\left(2\left(|v_1|^2 + |v_2|^2\right)\right)^{\frac{3}{2}}}{4(|v_1|^2|v_2|^2 - (v_1 \cdot v_2)^2)} \leq \left(Cl^{-2}\right)^{\frac{3}{2}} l^2 R_i^m \leq Cl^{-1}k_F. \tag{A.35}$$

Here, we also used that $R_1^m, R_2^m \leq k_F$ for all $m^* \leq m \leq M^*$. □

The summation formula

We can now present the summation formula that we will use to estimate the sums $\sum_{p \in L_k} f\left(\lambda_{k,p}\right)$. Noting that the quantity $l = |k|^{-1} \gcd\left(k_1, k_2, k_3\right)$ obeys the lower bound $l \geq |k|^{-1}$ independently of k, we can by equation (A.30) and Proposition A.14 estimate (provided $|k| \leq 2k_F$) that

$$\left||L_k^m| - \text{Area}\left(E_2^m \backslash E_1^m\right)\right| \leq C\left(1 + l^{-3}\left(l^{-1}k_F\right)^{\frac{2}{3}} \log\left(1 + 2\sqrt{2}\left(l^{-1}k_F\right)^{\frac{1}{2}}\right)^{\frac{2}{3}}\right)$$

$$\leq C\left(1 + |k|^{3+\frac{2}{3}} k_F^{\frac{2}{3}} \log\left(1 + \sqrt{|k|k_F}\right)^{\frac{2}{3}}\right) \leq C|k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}} \tag{A.36}$$

as $k_F \to \infty$, for a constant $C > 0$ independent of k and m. Inserting the expression for Area $\left(E_2^m \backslash E_1^m\right)$ that we determined in Proposition A.14, we then have

$$|L_k^m| = \begin{cases} 2\pi|k|\left(lm - \frac{1}{2}|k|\right)l & lm^* \leq lm \leq k_F \\ \pi\left(k_F^2 - (lm - |k|)^2\right)l & k_F < lm \leq lM^* \end{cases} + O\left(|k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}}\right). \tag{A.37}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Letting M denote the greatest integer such that $lM \leq k_F$, it now follows from equation (A.19) that for any $f : (0, \infty) \to \mathbb{R}$, it holds that

$$\sum_{p \in L_k} f(\lambda_{k,p}) = 2\pi|k| \sum_{m=m^*}^{M} f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right)\left(lm - \frac{1}{2}|k|\right) l + \pi \sum_{m=M+1}^{M^*} f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right)\left(k_F^2 - (lm - |k|)^2\right) l + O\left(|k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}} \sum_{m=m^*}^{M^*} \left| f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right)\right|\right), \tag{A.38}$$

so the 3-dimensional Riemann sum $\sum_{p \in L_k} f\left(\lambda_{k,p}\right)$ has been reduced to two 1-dimensional Riemann sums plus an error term. In fact, these two 1-dimensional Riemann sums are just what one would expect, since by 3D integrating along the $\hat{k}$ axis it is not difficult to show that, in general,

$$\int_{\overline{B}(k,k_F)\setminus\overline{B}(0,k_F)} f\left(k \cdot p - \frac{1}{2}|k|^2\right) dp = 2\pi|k| \int_{\frac{1}{2}|k|}^{k_F} f\left(|k|\left(t - \frac{1}{2}|k|\right)\right)\left(t - \frac{1}{2}|k|\right) dt + \pi \int_{k_F}^{k_F+|k|} f\left(|k|\left(t - \frac{1}{2}|k|\right)\right)\left(k_F^2 - (t - |k|)^2\right) dt, \tag{A.39}$$

and the two Riemann sums of equation (A.38) are seen to be Riemann sums for the two 1-dimensional integrals above.

In the statement in the following proposition, we make a minor adjustment: We expand the factor $k_F^2 - (lm - |k|)^2$ as

$$k_F^2 - (lm - |k|)^2 = k_F^2 - (lm)^2 - |k|^2 + 2|k|lm = \left(k_F^2 - (lm)^2\right) + 2|k|\left(lm - \frac{1}{2}|k|\right) \tag{A.40}$$

and collect the $2|k|(lm - \frac{1}{2}|k|)$ terms in the first sum. We have the summation formula:

Proposition A.15. *Let* $k = (k_1, k_2, k_3) \in \mathbb{Z}^3 \setminus \{0\}$ *with* $|k| \leq 2k_F$, $f : (0, \infty) \to \mathbb{R}$. *Let* $l = |k|^{-1} \gcd(k_1, k_2, k_3)$ *and* m^* *is the least integer and* M, M^* *the greatest integers for which*

$$\frac{1}{2}|k| < lm^*, \quad lM \leq k_F, \quad lM^* \leq k_F + |k|.$$

Then for all functions $f : (0, \infty) \to \mathbb{R}$, *it holds that*

$$\sum_{p \in L_k} f(\lambda_{k,p}) = 2\pi|k| \sum_{m=m^*}^{M^*} f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right)\left(lm - \frac{1}{2}|k|\right) l + \pi \sum_{m=M+1}^{M^*} f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right)\left(k_F^2 - (lm)^2\right) l + O\left(|k|^{3+\frac{2}{3}} (\log k_F)^{\frac{2}{3}} k_F^{\frac{2}{3}} \sum_{m=m^*}^{M^*} \left| f\left(|k|\left(lm - \frac{1}{2}|k|\right)\right)\right|\right), \quad k_F \to \infty.$$

A.3. *Proof of Proposition A.1*

Now we prove Proposition A.1 and (A.2). In this part, we do not use Proposition A.15.

Some Riemann sum estimation techniques

We must first establish some preliminary Riemann sum estimation results. Let $S \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be given, define for $k \in \mathbb{Z}^n$ the translated unit cube $\mathcal{C}_k$ by

$$\mathcal{C}_k = \left[-2^{-1}, 2^{-1}\right]^n + k \tag{A.41}$$

and let $\mathcal{C}_S = \bigcup_{k \in S \cap \mathbb{Z}^n} \mathcal{C}_k$ denote the union of the cubes centered at the lattice points contained in S. The first result we will establish is that for a convex function f, the integral $\int_{\mathcal{C}_S} f(p)dp$ always yields an upper bound to the Riemann sum $\sum_{k \in S \cap \mathbb{Z}^n} f(k)$:

Proposition A.16. *Let $f \in C(\mathcal{C}_S)$ be a function which is convex on $\mathcal{C}_k$ for all $k \in S \cap \mathbb{Z}^n$. Then,*

$$\sum_{k \in S \cap \mathbb{Z}^n} f(k) \leq \int_{\mathcal{C}_S} f(p)dp.$$

Proof. As a convex function admits a supporting hyperplane at every interior point of its domain, we see that for every $k \in S \cap \mathbb{Z}^n$, there exists a $c \in \mathbb{R}^n$ such that

$$f(p) \geq f(k) + c \cdot (p - k), \quad p \in \mathcal{C}_k, \tag{A.42}$$

which upon integration over $\mathcal{C}_k$ yields

$$\int_{\mathcal{C}_k} f(p)dp \geq \int_{\mathcal{C}_k} f(k)dp + \int_{\mathcal{C}_k} c \cdot (p - k)\, dp = f(k) \tag{A.43}$$

as $\int_{\mathcal{C}_S} f(k)dp = f(k)$ since $\mathrm{Vol}(\mathcal{C}_k) = 1$ and $\int_{\mathcal{C}_S} c \cdot (p - k)\, dp = 0$, as $\mathcal{C}_k$ is symmetric with respect to k but the integrand $p \mapsto c \cdot (p - k)$ is antisymmetric. Consequently,

$$\sum_{k \in S \cap \mathbb{Z}^n} f(k) \leq \sum_{k \in S \cap \mathbb{Z}^n} \int_{\mathcal{C}_k} f(p)dp = \int_{\mathcal{C}_S} f(p)dp. \tag{A.44}$$

□

This proposition lets us replace the sum by an integral but over an integration domain $\mathcal{C}_S$ which will generally be complicated. An exception is the $n = 1$ case which we record in the following (generalizing also the statement to any lattice spacing l):

Proposition A.17. *Let $a, b \in \mathbb{Z}$, $l > 0$, and $f \in C\left(\left[la - \frac{1}{2}l, lb + \frac{1}{2}l\right]\right)$ be a convex function. Then,*

$$\sum_{m=a}^{b} f(lm)l \leq \int_{la-\frac{1}{2}l}^{lb+\frac{1}{2}l} f(x)dx.$$

For $n \neq 1$, we instead require an additional result that lets us replace $\mathcal{C}_S$ by a simpler integration domain. We define a subset $S_+ \subset \mathbb{R}^n$ by

$$S_+ = \left\{p \in \mathbb{R}^n \mid \inf_{q \in S} |p - q| \leq \frac{\sqrt{n}}{2}\right\} \tag{A.45}$$

and observe the following:

Proposition A.18. *It holds that $\mathcal{C}_S \subset S_+$. Consequently,*

$$|S \cap \mathbb{Z}^n| \leq \mathrm{Vol}(S_+).$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proof. We first note that for any $p \in \mathbb{R}^n$, every point of the translated cube $\left(\left[-2^{-1}, 2^{-1}\right] + p\right)^n$ is a distance of at most $\frac{\sqrt{n}}{2}$ separated from p itself. Now, let $p \in \mathcal{C}_S$. Then by definition of $\mathcal{C}_S$ and the previous observation, there exists some $k \in S \cap \mathbb{Z}^n$ such that $|p - k| \le \frac{\sqrt{n}}{2}$, and hence, $p \in S_+$ since

$$\inf_{q \in S} |p - q| \le |p - k| \le \frac{\sqrt{n}}{2}. \tag{A.46}$$

Clearly, $|S \cap \mathbb{Z}^n| = \sum_{k \in S \cap \mathbb{Z}^n} 1 = \sum_{k \in S \cap \mathbb{Z}^n} \mathrm{Vol}\left(\mathcal{C}_k\right) = \mathrm{Vol}\left(\mathcal{C}_S\right)$, so the inclusion $\mathcal{C}_S \subset S_+$ immediately implies that $|S \cap \mathbb{Z}^n| \le \mathrm{Vol}\left(S_+\right)$. □

Lune geometry

Returning to Proposition A.1 and (A.2), we now let $k \in \mathbb{Z}^3_*$ and $-1 \le \beta \le 0$ be fixed. The Riemann sum ranges over $p \in L_k = (\overline{B}(k, k_F) \backslash \overline{B}(0, k_F)) \cap \mathbb{Z}^3$, so in the notation of the above discussion we must consider $S = \overline{B}(k, k_F) \backslash \overline{B}(0, k_F)$. The relevant integrand,

$$p \mapsto \lambda^{\beta}_{k,p} = \left(\frac{1}{2}\left(|p|^2 - |p-k|^2\right)\right)^{\beta} = |k|^{\beta}\left(\hat{k} \cdot p - \frac{1}{2}|k|\right)^{\beta}, \tag{A.47}$$

is convex on $\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p > \frac{1}{2}|k|\}$ but singular at $\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = \frac{1}{2}|k|\}$. For this reason, we must introduce a cutoff to the Riemann sum $\sum_{p \in L_k} \lambda^{\beta}_{k,p}$. We write $S = S^1 \cup S^2$

$$S^1 = \left\{p \in S \mid \hat{k} \cdot p \le \frac{1}{2}|k| + \frac{2 + \sqrt{3}}{2}\right\}, \quad S^2 = \left\{p \in S \mid \hat{k} \cdot p > \frac{1}{2}|k| + \frac{2 + \sqrt{3}}{2}\right\}, \tag{A.48}$$

so that likewise, $L_k = L^1_k \cup L^2_k$ where $L^1_k = L_k \cap S^1$, $L^2_k = L_k \cap S^2$. Hence, by Proposition A.18,

$$\begin{aligned}\sum_{p \in L_k} \lambda^{\beta}_{k,p} = \sum_{p \in L^1_k} \lambda^{\beta}_{k,p} + \sum_{p \in L^2_k} \lambda^{\beta}_{k,p} &\le \left(\inf_{p \in L_k} \lambda_{k,p}\right)^{\beta} \left|L^1_k\right| + \int_{\mathcal{C}_{S^2}} |k|^{\beta}\left(\hat{k} \cdot p - \frac{1}{2}|k|\right)^{\beta} dp \\ &\le \left(\inf_{p \in L_k} \lambda_{k,p}\right)^{\beta} \mathrm{Vol}\left(S^1_+\right) + |k|^{\beta} \int_{S^2_+} \left(\hat{k} \cdot p - \frac{1}{2}|k|\right)^{\beta} dp, \end{aligned} \tag{A.49}$$

where we also used that $p \mapsto (\hat{k} \cdot p - \frac{1}{2}|k|)^{\beta}$ is non-negative to expand the integration range of the integral. In order to apply this inequality, we will again replace the sets S^1_+, S^2_+ by ones which are easier to work with. We have the following:

Proposition A.19. *For all $k \in \mathbb{Z}^3$, it holds that*

$$\begin{aligned} S_+ &= \left\{p \in \mathbb{R}^3 \mid \inf_{q \in S} |p - q| \le \frac{\sqrt{3}}{2}\right\} \subset \widetilde{S} = \overline{B}\left(k, k_F + \frac{\sqrt{3}}{2}\right) \backslash B\left(0, k_F - \frac{\sqrt{3}}{2}\right), \\ S^1_+ &= \left\{p \in \mathbb{R}^3 \mid \inf_{q \in S^1} |p - q| \le \frac{\sqrt{3}}{2}\right\} \subset \widetilde{S}^1 = \left\{p \in \widetilde{S} \mid -\frac{\sqrt{3}}{2} \le \hat{k} \cdot p - \frac{1}{2}|k| \le 1 + \sqrt{3}\right\}, \\ S^2_+ &= \left\{p \in \mathbb{R}^3 \mid \inf_{q \in S^2} |p - q| \le \frac{\sqrt{3}}{2}\right\} \subset \widetilde{S}^2 = \left\{p \in \widetilde{S} \mid \hat{k} \cdot p - \frac{1}{2}|k| \ge 1\right\}. \end{aligned}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proof. We first show that $S_+ \subset \widetilde{S}$. For every $p \in S_+$ by the triangle inequality, we can estimate

$$|p| \geq \sup_{q\in S}\Big(|q| - |p-q|\Big) > k_F - \inf_{q\in S}|p-q| \geq k_F - \frac{\sqrt{3}}{2}, \tag{A.50}$$

$$|p-k| \leq \inf_{q\in S}\Big(|q-k| + |p-q|\Big) \leq k_F + \inf_{q\in S}|p-q| \leq k_F + \frac{\sqrt{3}}{2},$$

and hence, $p \in \widetilde{S}$. Next, we prove $S^1_+ \subset \widetilde{S}^1$: for every $p \in S^1_+$, we have

$$\hat{k}\cdot p - \frac{1}{2}|k| = \inf_{q\in S^1}\Big(\hat{k}\cdot q - \frac{1}{2}|k| + \hat{k}\cdot(p-q)\Big) \leq \frac{2+\sqrt{3}}{2} + \inf_{q\in S^1}|p-q| \leq 1+\sqrt{3}, \tag{A.51}$$

$$\hat{k}\cdot p - \frac{1}{2}|k| = \sup_{q\in S^1}\Big(\hat{k}\cdot q - \frac{1}{2}|k| + \hat{k}\cdot(p-q)\Big) \geq -\inf_{q\in S^1}|p-q| \geq -\frac{\sqrt{3}}{2} \tag{A.52}$$

and hence, $p \in \widetilde{S}^1$. Here, we used the definition of S^1 and $S^1 \subset S \subset \left\{q \in \mathbb{R}^3 \mid \hat{k}\cdot q > \frac{1}{2}|k|\right\}$. That $p \in S^2_+$ implies $\hat{k}\cdot p - \frac{1}{2}|k| \geq 1$ follows by the same argument. □

Thanks to the simple bound $\lambda_{k,p} \geq \frac{1}{2}$ for all $p \in L_k$, we can now conclude the inequality

$$\sum_{p\in L_k} \lambda^\beta_{k,p} \leq 2^{-\beta}\,\mathrm{Vol}(\widetilde{S}^1) + |k|^\beta \int_{\widetilde{S}^2}\left(\hat{k}\cdot p - \frac{1}{2}|k|\right)^\beta dp. \tag{A.53}$$

Hence, we need only consider the sets $\widetilde{S}^1$ and $\widetilde{S}^2$, which consist of 'slices' of $\widetilde{S}$:

$$\widetilde{S} = \bigcup_t \widetilde{S}_t, \quad \widetilde{S}_t = \{p \in \widetilde{S} \mid \hat{k}\cdot p = t\}. \tag{A.54}$$

Recalling the definition of $\widetilde{S}$ from Proposition A.19 and using elementary trigonometry, we can show that

$$\begin{aligned}\mathrm{Area}\left(\widetilde{S}_t\right) &= \pi\left(\left(k_F + \frac{\sqrt{3}}{2}\right)^2 - (t-|k|)^2\right) - \pi\left(\left(k_F - \frac{\sqrt{3}}{2}\right)^2 - |t|^2\right)\\ &= \pi\left(2\sqrt{3}k_F - \left(|k|^2 - 2|k|t\right)\right) = 2\pi\left(|k|\left(t - \frac{1}{2}|k|\right) + \sqrt{3}k_F\right)\end{aligned} \tag{A.55}$$

for $|k|/2 - \sqrt{3}/2 \leq t \leq k_F - \sqrt{3}/2$, and that

$$\begin{aligned}\mathrm{Area}\left(\widetilde{S}_t\right) &= \pi\left(\left(k_F + \frac{\sqrt{3}}{2}\right)^2 - (t-|k|)^2\right) = \pi\left(\left(k_F + \frac{\sqrt{3}}{2}\right)^2 - \left(t^2 - 2|k|\left(t - \frac{1}{2}|k|\right)\right)\right)\\ &= 2\pi\left(|k|\left(t - \frac{1}{2}|k|\right) + \sqrt{3}k_F\right) + \pi\left(\left(k_F - \frac{\sqrt{3}}{2}\right)^2 - t^2\right)\\ &\leq 2\pi\left(|k|\left(t - \frac{1}{2}|k|\right) + \sqrt{3}k_F\right)\end{aligned} \tag{A.56}$$

for $k_F - \sqrt{3}/2 \leq t \leq k_F + \sqrt{3}/2 + |k|$.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

With these formulas, we can now give the following:

Proof of the* $|k| < 2k_F$ *case of Proposition A.1 and (A.2). By equation (A.53), we have

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta} \leq 2^{-\beta}\,\mathrm{Vol}(\widetilde{S}^1) + |k|^{\beta}\int_{\widetilde{S}^2}\left(\hat{k}\cdot p - \frac{1}{2}|k|\right)^{\beta} dp, \tag{A.57}$$

and we can estimate

$$\begin{aligned}\mathrm{Vol}(\widetilde{S}^1) &= \int_{\frac{1}{2}|k|-\frac{\sqrt{3}}{2}}^{\frac{1}{2}|k|+1+\sqrt{3}} \mathrm{Area}(\widetilde{S}_t)dt = 2\pi\int_{\frac{1}{2}|k|-\frac{\sqrt{3}}{2}}^{\frac{1}{2}|k|+1+\sqrt{3}}\left(|k|\left(t-\frac{1}{2}|k|\right)+\sqrt{3}k_F\right)dt \\ &= 2\pi\int_{-\frac{\sqrt{3}}{2}}^{1+\sqrt{3}}\left(|k|t+\sqrt{3}k_F\right)dt \leq C\left(|k|+k_F\right) \leq Ck_F = O\left(k_F^{2+\beta}|k|^{1+\beta}\right),\end{aligned} \tag{A.58}$$

for all $-1 \leq \beta \leq 0$, and

$$\begin{aligned}\int_{\widetilde{S}^2}\left(\hat{k}\cdot p-\frac{1}{2}|k|\right)^{\beta}dp &= \int_{\frac{1}{2}|k|+1}^{k_F+\frac{\sqrt{3}}{2}+|k|}\left(t-\frac{1}{2}|k|\right)^{\beta}\mathrm{Area}\left(\widetilde{S}_t\right)dt \\ &\leq 2\pi\int_{\frac{1}{2}|k|+1}^{k_F+\frac{\sqrt{3}}{2}+|k|}\left(t-\frac{1}{2}|k|\right)^{\beta}\left(|k|\left(t-\frac{1}{2}|k|\right)+\sqrt{3}k_F\right)dt \\ &= 2\pi\left(|k|\int_1^{k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|}t^{1+\beta}\,dt+\sqrt{3}k_F\int_1^{k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|}t^{\beta}\,dt\right) \\ &\leq 2\pi\left(\frac{|k|}{2+\beta}\left(k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|\right)^{2+\beta}+\frac{\sqrt{3}}{1+\beta}k_F\left(k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|\right)^{1+\beta}\right) \leq Ck_F^{2+\beta}|k|\end{aligned} \tag{A.59}$$

for $-1 < \beta \leq 0$, and

$$\begin{aligned}\int_{\widetilde{S}^2}\left(\hat{k}\cdot p-\frac{1}{2}|k|\right)^{-1}dp &\leq 2\pi\left(|k|\int_1^{k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|}1\,dt+\sqrt{3}k_F\int_1^{k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|}t^{-1}\,dt\right) \\ &\leq C\left(|k|k_F+k_F\log\left(k_F+\frac{\sqrt{3}}{2}+\frac{1}{2}|k|\right)\right) \leq C|k|\left(1+|k|^{-1}\log\left(k_F\right)\right)k_F\end{aligned} \tag{A.60}$$

for $\beta = -1$. Combining the estimates yields the claim. □

Proof of the $|k| \geq 2k_F$ case of Proposition A.1. For $|k| \geq 2k_F$, the lune $S = \overline{B}(k, k_F)\setminus\overline{B}(0, k_F)$ degenerates into a ball, and so we must adapt our argument. Now it is simply the case that

$$S_+ = \widetilde{S} = \overline{B}\left(k, k_F + \frac{\sqrt{3}}{2}\right). \tag{A.61}$$

If $\frac{1}{2}|k| \geq k_F + \frac{2+\sqrt{3}}{2}$, then every $p \in \widetilde{S}$ satisfies $\hat{k}\cdot p - \frac{1}{2}|k| \geq 1$ and the cutoff set $\widetilde{S}^1$ is unnecessary. Otherwise, the equation (A.53),

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta} \leq 2^{-\beta}\,\mathrm{Vol}\left(\widetilde{S}^1\right) + |k|^{\beta}\int_{\widetilde{S}^2}\left(\hat{k}\cdot p - \frac{1}{2}|k|\right)^{\beta} dp, \tag{A.62}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

still holds for

$$\widetilde{S}^1 = \left\{p \in \widetilde{S} \mid \hat{k} \cdot p - \frac{1}{2}|k| \leq 1 + \sqrt{3}\right\}, \quad \widetilde{S}^2 = \left\{p \in \widetilde{S} \mid \hat{k} \cdot p - \frac{1}{2}|k| \geq +1\right\}, \tag{A.63}$$

where we simplified the description for $\widetilde{S}^1$ using that $\hat{k} \cdot p - \frac{1}{2}|k| \geq -\frac{\sqrt{3}}{2}$ holds for all $p \in \widetilde{S}$ when $|k| \geq 2k_F$. We can then easily estimate $\mathrm{Vol}\left(\widetilde{S}_1\right)$, as it is now seen to be a spherical cap of radius $k_F + \frac{\sqrt{3}}{2}$ and height

$$\left(\frac{1}{2}|k| + 1 + \sqrt{3}\right) - \left(|k| - k_F - \frac{\sqrt{3}}{2}\right) \leq k_F - \frac{1}{2}|k| + \frac{2 + 3\sqrt{3}}{2} \leq \frac{2 + 3\sqrt{3}}{2}, \tag{A.64}$$

whence

$$\mathrm{Vol}\left(\widetilde{S}^1\right) \leq \frac{\pi}{3}\left(\frac{2 + 3\sqrt{3}}{2}\right)^2 \left(3\left(k_F + \frac{\sqrt{3}}{2}\right) - \frac{2 + 3\sqrt{3}}{2}\right) \leq C k_F, \tag{A.65}$$

so as $k_F = O\left(k_F^3 |k|^{2\beta}\right)$ for all $-1 \leq \beta \leq 0$ when $2k_F \leq |k| \leq 2k_F + \frac{\sqrt{3}}{2}$, this is again negligible. □

We estimate the integrals to conclude the following:

Proof of the second part of Proposition A.1. We again note that the area of the slice $\widetilde{S}_t$ is given by

$$\mathrm{Area}\left(\widetilde{S}_t\right) = \pi\left(\left(k_F + \frac{\sqrt{3}}{2}\right)^2 - (t - |k|)^2\right), \tag{A.66}$$

now for $|k| - k_F - \frac{\sqrt{3}}{2} \leq t \leq |k| + k_F + \frac{\sqrt{3}}{2}$. If $|k| \leq 2k_F + 1 + \sqrt{3}$, we just saw that the contribution coming from the cutoff set $\widetilde{S}^1$ is negligible, while the integral term is

$$|k|^\beta \int_{\widetilde{S}^2} \left(\hat{k} \cdot p - \frac{1}{2}|k|\right)^\beta dp = |k|^\beta \int_{\frac{1}{2}|k|+1}^{k_F + \frac{\sqrt{3}}{2} + |k|} \left(t - \frac{1}{2}|k|\right)^\beta \mathrm{Area}\left(\widetilde{S}_t\right) dt \leq C k_F^{2+\beta} |k|^{1+\beta} \tag{A.67}$$

as calculated in equation (A.59), which is $O\left(k_F^3 |k|^{2\beta}\right)$ for $2k_F \leq |k| \leq 2k_F + 1 + \sqrt{3}$ (here we also use that for $\beta = -1$, the logarithmic term in the estimate of equation (A.60) is negligible when $|k| \geq 2k_F$ due to the additional factor of $|k|^{-1}$).

If $|k| > 2k_F + \frac{2+\sqrt{3}}{2}$, we simply have

$$\sum_{p \in L_k} \lambda_{k,p}^\beta \leq |k|^\beta \int_{\widetilde{S}} \left(\hat{k} \cdot p - \frac{1}{2}|k|\right)^\beta dp = |k|^\beta \int_{|k| - k_F - \frac{\sqrt{3}}{2}}^{|k| + k_F + \frac{\sqrt{3}}{2}} \left(t - \frac{1}{2}|k|\right)^\beta \mathrm{Area}\left(\widetilde{S}_t\right) dt, \tag{A.68}$$

and by writing $(t - |k|)^2 = \left(t - \frac{1}{2}|k|\right)^2 - |k|\left(t - \frac{1}{2}|k|\right) + \frac{1}{4}|k|^2$, we can furthermore estimate that

$$\begin{aligned} \mathrm{Area}\left(\widetilde{S}_t\right) &= \pi\left(\left(k_F + \frac{\sqrt{3}}{2}\right)^2 - \left(\left(t - \frac{1}{2}|k|\right)^2 - |k|\left(t - \frac{1}{2}|k|\right) + \frac{1}{4}|k|^2\right)\right) \\ &= \pi\left(|k|\left(t - \frac{1}{2}|k|\right) - \left(\frac{1}{4}|k|^2 - \left(k_F + \frac{\sqrt{3}}{2}\right)^2\right) - \left(t - \frac{1}{2}|k|\right)^2\right) \leq \pi|k|\left(t - \frac{1}{2}|k|\right), \end{aligned} \tag{A.69}$$

so

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta}\leq\pi|k|^{1+\beta}\int_{|k|-k_F-\frac{\sqrt{3}}{2}}^{|k|+k_F+\frac{\sqrt{3}}{2}}\left(t-\frac{1}{2}|k|\right)^{1+\beta}dt$$
$$=\frac{\pi|k|^{1+\beta}}{2+\beta}\left(\left(\frac{1}{2}|k|+k_F+\frac{\sqrt{3}}{2}\right)^{2+\beta}-\left(\frac{1}{2}|k|-k_F-\frac{\sqrt{3}}{2}\right)^{2+\beta}\right)$$
$$\leq C|k|^{1+\beta}\left(\frac{1}{2}|k|+k_F+\frac{\sqrt{3}}{2}\right)^{2+\beta}\leq C|k|^{3+2\beta}. \tag{A.70}$$

If additionally $|k|\leq 3k_F$ (say), then this is again $O\left(k_F^3|k|^{2\beta}\right)$. If this is not the case, however, then we can instead trivially estimate that

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta}\leq|k|^{\beta}\int_{\widetilde{S}}\left(\hat{k}\cdot p-\frac{1}{2}|k|\right)^{\beta}dp\leq|k|^{\beta}\left(\inf_{p\in\widetilde{S}}\left(\hat{k}\cdot p-\frac{1}{2}|k|\right)\right)^{\beta}\int_{\widetilde{S}}1\,dp$$
$$\leq|k|^{\beta}\left(\frac{1}{2}|k|-k_F-\frac{\sqrt{3}}{2}\right)^{\beta}\mathrm{Vol}\left(\overline{B}\left(0,k_F+\frac{\sqrt{3}}{2}\right)\right)$$
$$\leq Ck_F^3|k|^{\beta}\left(\frac{1}{2}|k|-\frac{1}{3}|k|-\frac{\sqrt{3}}{2}\right)^{\beta}\leq Ck_F^3|k|^{2\beta}. \tag{A.71}$$

□

A.4. Proof of Proposition A.2

In the cases $|k|\geq 2k_F$ and $2k_F\geq|k|\geq\log(k_F)$, the claim has been proved. Thus, it remains to consider the case $|k|\leq\log(k_F)$, for which we will apply the summation formula in Proposition A.15 to improve (A.2). By Proposition A.15, we have

$$\sum_{p\in L_k}\frac{1}{\lambda_{k,p}}=2\pi|k|\sum_{m=m^*}^{M^*}\frac{lm-\frac{1}{2}|k|}{|k|\left(lm-\frac{1}{2}|k|\right)}l+\pi\sum_{m=M+1}^{M^*}\frac{k_F^2-(lm)^2}{|k|\left(lm-\frac{1}{2}|k|\right)}l$$
$$+O\left(|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\sum_{m=m^*}^{M^*}\frac{1}{|k|\left(lm-\frac{1}{2}|k|\right)}\right)$$
$$\leq 2\pi\sum_{m=m^*}^{M^*}l+O\left(|k|^{2+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\sum_{m=m^*}^{M^*}\frac{1}{lm-\frac{1}{2}|k|}\right),\quad k_F\to\infty, \tag{A.72}$$

where we used that by definition of M, $(k_F^2-(lm)^2)<0$ for all $m\geq M+1$. As $|k|\leq 2k_F$,

$$\sum_{m=m^*}^{M^*}l=l\left(M^*-m^*+1\right)\leq k_F+|k|+l\leq Ck_F,\quad k_F\to\infty, \tag{A.73}$$

where we also used that

$$l=|k|^{-1}\gcd\left(k_1,k_2,k_3\right)\leq\frac{\max\left(|k_1|,|k_2|,|k_3|\right)}{\sqrt{k_1^2+k_2^2+k_3^2}}\leq 1. \tag{A.74}$$

We now consider the sum $\sum_{m=m^*}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{-1}$. To apply Proposition A.17, we must estimate the $m=m^*$ term separately, so that the integration range does not cross the point $x=\frac{1}{2}|k|$, where the integrand diverges. Note that using $\lambda_{k,p}\geq\frac{1}{2}$ for all $p\in L_k$, we have

$$lm^*-\frac{1}{2}|k|=\min_{p\in L_k}\left(\hat{k}\cdot p-\frac{1}{2}|k|\right)=|k|^{-1}\left(\min_{p\in L_k}\left(k\cdot p-\frac{1}{2}|k|^2\right)\right)\geq\frac{1}{2}|k|^{-1}. \tag{A.75}$$

Therefore,

$$\begin{aligned}\sum_{m=m^*}^{M^*}\frac{1}{lm-\frac{1}{2}|k|}&\leq 2|k|+\sum_{m=m^*+1}^{M^*}\frac{1}{lm-\frac{1}{2}|k|}\leq 2|k|+|k|\int_{lm^*+\frac{1}{2}l}^{lM^*+\frac{1}{2}l}\frac{1}{x-\frac{1}{2}|k|}\,dx\\&\leq C|k|\left(1+\log\left(\frac{lM^*+\frac{1}{2}l-\frac{1}{2}|k|}{lm^*+\frac{1}{2}l-\frac{1}{2}|k|}\right)\right)\leq C|k|\left(1+\log\left(\frac{k_F+|k|+\frac{1}{2}l-\frac{1}{2}|k|}{\frac{1}{2}l}\right)\right)\\&\leq C|k|\left(1+\log\left(|k|k_F\right)\right)\leq C|k|\log\left(k_F\right),\quad k_F\to\infty,\end{aligned} \tag{A.76}$$

yielding the total bound when $|k|\leq\log(k_F)$

$$\sum_{p\in L_k}\frac{1}{\lambda_{k,p}}\leq C\left(k_F+|k|^{3+\frac{2}{3}}\log\left(k_F\right)^{\frac{5}{3}}k_F^{\frac{2}{3}}\right)\leq Ck_F. \tag{A.77}$$

A.5. *Proof of Proposition A.3*

First, consider the case $-\frac{4}{3}\leq\beta<-1$ and $k\in\overline{B}(0,2k_F)$. By Proposition A.15, we can estimate using the argument leading to (A.77) that

$$\begin{aligned}\sum_{p\in L_k}\lambda_{k,p}^{\beta}&=2\pi|k|\sum_{m=m^*}^{M^*}\left(|k|\left(lm-\frac{1}{2}|k|\right)\right)^{\beta}\left(lm-\frac{1}{2}|k|\right)l\\&\quad+\pi\sum_{m=M+1}^{M^*}\left(|k|\left(lm-\frac{1}{2}|k|\right)\right)^{\beta}\left(k_F^2-(lm)^2\right)l\\&\quad+O\left(|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\sum_{m=m^*}^{M^*}\left(|k|\left(lm-\frac{1}{2}|k|\right)\right)^{\beta}\right)\\&\leq 2\pi|k|^{1+\beta}\sum_{m=m^*}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{1+\beta}l+O\left(|k|^{3+\frac{2}{3}+\beta}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\sum_{m=m^*}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{\beta}\right).\end{aligned} \tag{A.78}$$

Applying Proposition A.17 and A.75, again we have

$$\begin{aligned}\sum_{m=m^*}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{1+\beta}l&=\left(lm^*-\frac{1}{2}|k|\right)^{1+\beta}l+\sum_{m=m^*+1}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{1+\beta}l\\&\leq\left(\frac{1}{2}|k|^{-1}\right)^{1+\beta}+\int_{lm^*+\frac{1}{2}l}^{lM^*+\frac{1}{2}l}\left(lm-\frac{1}{2}|k|\right)^{1+\beta}dx\\&=\left(\frac{1}{2}|k|^{-1}\right)^{1+\beta}+\frac{1}{2+\beta}\left(\left(lM^*+\frac{1}{2}l-\frac{1}{2}|k|\right)^{2+\beta}-\left(lm^*+\frac{1}{2}l-\frac{1}{2}|k|\right)^{2+\beta}\right)\\&\leq C\left(|k|^{-(1+\beta)}+\left(k_F+\frac{1}{2}l+\frac{1}{2}|k|\right)^{2+\beta}\right)\leq Ck_F^{2+\beta},\quad k_F\to\infty,\end{aligned} \tag{A.79}$$

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

and likewise,

$$\begin{aligned}\sum_{m=m^*}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{\beta} &= \left(lm^*-\frac{1}{2}|k|\right)^{\beta}+l^{-1}\sum_{m=m^*+1}^{M^*}\left(lm-\frac{1}{2}|k|\right)^{\beta} l\\ &\leq \left(\frac{1}{2}|k|^{-1}\right)^{\beta}+|k|\int_{lm^*+\frac{1}{2}l}^{lM^*+\frac{1}{2}l}\left(lm-\frac{1}{2}|k|\right)^{\beta}dx\\ &= 2^{-\beta}|k|^{-\beta}+\frac{|k|}{1+\beta}\left(\left(lm^*+\frac{1}{2}l-\frac{1}{2}|k|\right)^{1+\beta}-\left(lM^*+\frac{1}{2}l-\frac{1}{2}|k|\right)^{1+\beta}\right)\\ &\leq C\left(|k|^{-\beta}+|k|\left(lm^*+\frac{1}{2}l-\frac{1}{2}|k|\right)^{1+\beta}\right)\geq C\left(|k|^{-\beta}+|k|\left(\frac{1}{2}|k|^{-1}\right)^{1+\beta}\right)\leq C|k|^{-\beta}. \end{aligned} \tag{A.80}$$

Combining these, we find that for all $-\frac{4}{3}\leq\beta<-1$ and $k\in\overline{B}(0,2k_F)$,

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta}\leq C\left(k_F^{2+\beta}|k|^{1+\beta}+|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\right),\quad k_F\to\infty. \tag{A.81}$$

Consequently, if $\beta\leq-\frac{4}{3}$ and $k\in\overline{B}(0,2k_F)$, then using $\lambda_{k,p}\geq\frac{1}{2}$, we have

$$\sum_{p\in L_k}\lambda_{k,p}^{\beta}\leq C\sum_{p\in L_k}\lambda_{k,p}^{-\frac{4}{3}}\leq C|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}. \tag{A.82}$$

Moreover, if $-\frac{4}{3}<\beta<-1$ and $|k|\leq k_F^{\gamma}$ with $\gamma<\frac{4+3\beta}{8-3\beta}$, then the right-hand side of (A.81) can be simplified to $Ck_F^{2+\beta}|k|^{1+\beta}$.

A.6. Proof of Proposition A.4

In this subsection, we prove Proposition A.4. We first establish a simple upper bound:

Proposition A.20. *For all $k\in\mathbb{Z}^3_*$ and any $\lambda>0$, it holds that $\left|S^1_{k,\lambda}\right|+\left|S^2_{k,\lambda}\right|\leq\left|S_{k,\lambda}\right|$, where*

$$S_{k,\lambda}=\left\{p\in\mathbb{Z}^3\mid ||p|^2-\zeta|<\lambda\ \ and\ \ \left|\hat{k}\cdot p-\frac{1}{2}|k|\right|<\frac{1}{2}|k|^{-1}\lambda\right\}.$$

Proof. As $S^1_{k,\lambda}\cap S^2_{k,\lambda}=\emptyset$, the claim will follow if we can show that $S^1_{k,\lambda},S^2_{k,\lambda}\subset S_{k,\lambda}$. Consider an arbitrary $p\in S^1_{k,\lambda}$. By definition of $S^1_{k,\lambda}$,

$$||p|^2-\zeta|\leq\max\left\{||p|^2-\zeta|,||p-k|^2-\zeta|\right\}<\lambda, \tag{A.83}$$

so the first condition for $S_{k,\lambda}$ is satisfied. For the other, we note that

$$\begin{aligned}\left|2k\cdot p-|k|^2\right| &= \left||p|^2-|p-k|^2\right|=\left|||p|^2-\zeta|-||p-k|^2-\zeta|\right|\\ &\leq\max\left\{||p|^2-\zeta|,||p-k|^2-\zeta|\right\}<\lambda,\end{aligned} \tag{A.84}$$

where in the second equality we used that both $p,(p-k)\in B_F$ if $p\in S_{k,\lambda}$. This now implies that $p\in S_{k,\lambda}$, so indeed, $S^1_{k,\lambda}\subset S_{k,\lambda}$. The inclusion $S^2_{k,\lambda}\subset S_{k,\lambda}$ follows similarly. □

The quantity $\left|S_{k,\lambda}\right|$ can in turn be estimated with exactly the same techniques which we used for the estimation of Riemann sums in the previous subsections. Let us start by using the arguments from

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

Proposition A.10. Now, the condition that $||p|^2 - \zeta| < \lambda$ is equivalent with $\zeta - \lambda < |p|^2 < \zeta + \lambda$, and writing $|p|^2 = (\hat{k} \cdot p)^2 + |P_\perp p|^2$ (where $P_\perp : \mathbb{R}^3 \to \mathbb{R}^3$ denotes the orthogonal projection onto $\{k\}^\perp = \left\{p \in \mathbb{R}^3 \mid \hat{k} \cdot p = 0\right\}$), this is equivalent with

$$\zeta - \left(\hat{k} \cdot p\right)^2 - \lambda < \left|P_\perp p\right|^2 < \zeta - \left(\hat{k} \cdot p\right)^2 + \lambda. \tag{A.85}$$

Consequently, if we let m_- and m_+ be the least and greatest integers, respectively, such that

$$\frac{1}{2}\left(|k| - |k|^{-1}\lambda\right) < lm_- \quad \text{and} \quad lm_+ < \frac{1}{2}\left(|k| + |k|^{-1}\lambda\right), \tag{A.86}$$

it follows that we can decompose $S_{k,\lambda} = \bigcup_{m=m_-}^{m_+} S^m_{k,\lambda}$, where

$$\begin{aligned} S^m_{k,\lambda} &= S_{k,\lambda} \cap \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm\right\} = \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm, ||p|^2 - \zeta| < \lambda\right\} \\ &= \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm,\ \zeta - (lm)^2 - \lambda < |P_\perp p|^2 < \zeta - (lm)^2 + \lambda\right\} \\ &= \left\{p \in \mathbb{Z}^3 \mid \hat{k} \cdot p = lm,\ (R^m_-)^2 < |P_\perp p|^2 < (R^m_-)^2\right\} \end{aligned} \tag{A.87}$$

for

$$\left(R^m_\pm\right)^2 = \zeta - (lm)^2 \pm \lambda, \quad m_- \le m \le m_+. \tag{A.88}$$

We see that the sets $S^m_{k,\lambda}$ are of the same form as the sets L^m_k which we considered in Section A.2. The arguments which we used to estimate $|L^m_k|$ thus immediately carry over, provided we can establish some basic estimates on R^m_- and R^m_+. We have the following:

Proposition A.21. *For all $k \in \overline{B}(0, k_F) \cap \mathbb{Z}^3_*$ and $0 < \lambda = \lambda\left(k_F, k\right) \le \frac{1}{6}k_F^2$, it holds that*

$$C^{-1}k_F < R^m_- < R^m_+ \le Ck_F, \quad \forall m_- \le m \le m_+, \tag{A.89}$$

as $k_F \to \infty$ for a constant $C > 0$ independent of k, k_F and λ.

Proof. First, recall that ζ is the midpoint of the interval $I = \left[\sup_{p\in B_F} |p|^2, \inf_{p\in B_F^c} |p|^2\right]$. Since $k_F^2 \in I$ by definition of the Fermi ball, we can bound

$$|\zeta - k_F^2| \le \frac{|I|}{2} = \frac{1}{2}\left(\inf_{q\in B_F^c} |p|^2 - \sup_{q\in B_F} |p|^2\right) \le k_F + 1. \tag{A.90}$$

Here, the last inequality can be seen by taking the trial points $p_- = (\lfloor k_F \rfloor, 0, 0) \in B_F$ and $p_+ = (\lfloor k_F \rfloor + 1, 0, 0) \in B_F^c$. Combining (A.90), the definitions of m_-, m_+ and the assumptions of the statement, we may estimate independently of m that

$$\begin{aligned} (R^m_-)^2 &\ge \zeta - \max\left\{(lm_-)^2, (lm_+)^2\right\} - \lambda \\ &\ge \zeta - \frac{1}{4}\left(\left(|k| - |k|^{-1}\lambda\right)^2 + \left(|k| + |k|^{-1}\lambda\right)^2\right) - \lambda \\ &\ge \zeta - \frac{1}{2}\left(|k|^2 + |k|^{-2}\lambda\right) - \lambda \ge \zeta - \frac{1}{2}k_F^2 - \frac{3}{2}\lambda \ge \frac{1}{4}k_F^2 - k_F - 1, \end{aligned} \tag{A.91}$$

and

$$(R^m_+)^2 = \zeta - (lm)^2 + \lambda \le \zeta + \frac{1}{6}k_F^2 \le \frac{7}{6}k_F^2 + k_F + 1. \tag{A.92}$$

□

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

This allows us to estimate $|S^m_{k,\lambda}|$ with the same error term as that of $|L^m_k|$, which is to say $C|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}$. We can now give the following:

Proof of Proposition A.4. By Proposition A.21 and the above arguments, we can estimate

$$\begin{aligned}\left|S^m_{k,\lambda}\right| &\leq \frac{2\pi\left((R^m_+)^2-(R^m_-)^2\right)}{2l^{-1}}+C|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\\ &= 2\pi\lambda\,(k_F,k)\,l+C|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\end{aligned} \tag{A.93}$$

for $m_-\leq m\leq m_+$. By the decomposition $S_{k,\lambda}=\bigcup_{m=m_-}^{m_+}S^m_{k,\lambda}$, we can then estimate further

$$\begin{aligned}\left|S_{k,\lambda}\right| &= \sum_{m=m_-}^{m_+}\left|S^m_{k,\lambda}\right|\leq 2\pi\lambda\sum_{m=m_-}^{m_+}l+C|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\sum_{m=m_-}^{m_+}1\\ &\leq C\left(\lambda+|k|^{4+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\right)(lm_+-lm_-+l)\\ &\leq C\left(\lambda+|k|^{4+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\right)\left(\frac{1}{2}\left(|k|+|k|^{-1}\lambda\right)-\frac{1}{2}\left(|k|-|k|^{-1}\lambda\right)+l\right)\\ &\leq C\left(|k|^{-1}\lambda+|k|^{3+\frac{2}{3}}(\log k_F)^{\frac{2}{3}}k_F^{\frac{2}{3}}\right)(\lambda+|k|)\,,\end{aligned} \tag{A.94}$$

where we also applied the estimate $|k|^{-1}\leq l\leq 1$. □

Acknowledgements. PTN thanks Niels Benedikter, Marcello Porta, Benjamin Schlein and Robert Seiringer for helpful discussions. We thank the referees for constructive remarks and suggestions. MRC and PTN acknowledge the support from the Deutsche Forschungsgemeinschaft (DFG project Nr. 426365943).

Competing interest. The authors have no competing interest to declare.

References

[1] V. Bach, 'Error bound for the Hartree-Fock energy of atoms and molecules', *Commun. Math. Phys.* **147** (1992), 527–548.
[2] V. Bach and J.-B. Bru, 'Diagonalizing quadratic bosonic operators by nonautonomous flowequation', *Mem. Amer. Math. Soc.* **240**(1138) (2016).
[3] J. Bardeen, 'Electron exchange in the theory of metals', *Phys. Rev.* **50** (1936), 1098.
[4] N. Benedikter, 'Bosonic collective excitations in fermi gases', *Rev. Math. Phys.* **32** (2020), 2060009.
[5] N. Benedikter, P. T. Nam, M. Porta, B. Schlein and R. Seiringer, 'Optimal upper bound for the correlation energy of a fermi gas in the mean-field regime', *Commun. Math. Phys.* **374** (2020), 2097–2150.
[6] N. N. Benedikter, P. T. Nam, M. Porta, B. Schlein and R. Seiringer, 'Correlation energy of a weakly interacting Fermi gas', *Invent. Math.* **225** (2021), 885–979.
[7] N. Benedikter, P. T. Nam, M. Porta, B. Schlein and R. Seiringer, 'Bosonization of fermionic many-body dynamics', *Ann. Henri Poincaré* **23** (2022), 1725–1764.
[8] N. Benedikter, M. Porta, B. Schlein and R. Seiringer, 'Correlation energy of a weakly interacting fermi gas with large interaction potential', *Arch. Rational Mech. Anal.* (to appear). arXiv:2106.13185.
[9] C. Boccato, C. Brennecke, S. Cenatiempo and B. Schlein. 'Bogoliubov theory in the Gross– Pitaevskii limit', *Acta Math.* **222** (2019), 219–335.
[10] N. N. Bogolubov, 'On the theory of superfluidity', *J. Phys. (USSR)* **11** (1947), 23.
[11] D. Bohm and D. Pines, 'A collective description of electron interactions. I. Magnetic interactions', *Phys. Rev.* **82** (1951), 625–634.
[12] D. Bohm and D. Pines, 'A collective description of electron interactions: II. Collective vs. Individual particle aspects of the interactions', *Phys. Rev.* **85** (1952), 338–353.
[13] D. Bohm and D. Pines, 'A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas', *Phys. Rev.* **92** (1953), 609–625.
[14] M. R. Christiansen, C. Hainzl and P. T. Nam, 'On the effective quasi-bosonic Hamiltonian of the electron gas: Collective excitations and plasmon modes', *Lett. Math. Phys.* **112**(Art. 114), 2022.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press

[15] M. R. Christiansen, C. Hainzl and P. T. Nam, 'The Gell-Mann–Brueckner formula for the correlation energy of the electron gas: A rigorous upper bound in the mean-field regime', *Commun. Math. Phys.* (to appear). arXiv:2208.01581

[16] J. Dereziński, 'Bosonic quadratic Hamiltonians', *J. Math. Phys.* **58** (2107), 121101.

[17] R. A. Ferrell, 'Characteristic energy loss of electrons passing through metal foils. II. Dispersion relation and short wavelength cutoff for plasma oscillations', *Phys. Rev.* **107** (1957), 450–460.

[18] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, 1971).

[19] A. O. Gel'fond and Y. V. Linnik, *Elementary Methods in the Analytic Theory of Numbers* (Pergamon Press, 1966). *Translation by I. N. Sneddon*.

[20] M. Gell-Mann and K. A. Brueckner, 'Correlation energy of an electron gas at high density', *Phys. Rev.* **106** (1957), 364.

[21] D. Gontier, C. Hainzl and M. Lewin, 'Lower bound on the Hartree-Fock energy of the electron gas', *Phys. Rev. A* **99** (2019), 052501.

[22] G. M. Graf and J. P. Solovej, 'A correlation estimate with applications to quantum systems with Coulomb interactions', *Rev. Math. Phys.* **6** (1994), 977–997.

[23] P. Grech and R. Seiringer, 'The excitation spectrum for weakly interacting bosons in a trap', *Commun. Math. Phys.* **322** (2013), 559–591.

[24] C. Hainzl, M. Porta and F. Rexze, 'On the correlation energy of interacting fermionic systems in the mean-field regime', *Commun. Math. Phys.* **524** (2020), 374–485.

[25] C. Hainzl, B. Schlein and A. Triay, 'Bogoliubov theory in the Gross–Pitaevskii limit: A simplified approach', *Forum Math. Sigma* **10**(e90) (2022), 1–39.

[26] H. Hansen, G. Chanfray, D. Davesne and P. Schuck, 'Random phase approximation and extensions applied to a bosonic field theory', *Eur. Phys. J. A* **14** (2002), 397–411.

[27] L. Hörmander, 'Sympletic classification of quadratic forms, and general Mehler formulas', *Math. Z.* **219** (1995), 413–449.

[28] D. Micciancio and S. Goldwasser, *Complexity of Lattice Problems – a Cryptographic Perspective* (Springer, 2002).

[29] P. T. Nam, 'Binding energy of homogeneous Bose gas', *Lett. Math. Phys.* **108** (2018), 141–159.

[30] P. T. Nam and M. Napiórkowski, 'Two-term expansion of the ground state one-body density matrix of a mean-field Bose gas', *Calc. Var. PDE* **60**(Art. 99), 2021.

[31] P. T. Nam, M. Napiórkowski and J. P. Solovej, 'Diagonalization of bosonic quadratic Hamiltonians by Bogolubov transformations', *J. Funct. Anal.* **270** (2016), 4340–4368.

[32] D. Pines, 'Electron interaction in metals,' *Phys. Rev.* **92** (1953), 626.

[33] D. Pines, *Elementary Excitations in Solids, Lectures on Protons, Electrons, and Plasmons* (Avalon Publishing, 1999).

[34] A. Repko, J. Kvasil, V. O. Nesterenko and P. G. Reinhard, 'Skyrme RPA for spherical and axially symmetric nuclei', *Proceedings of the ISTROS 2015 International Conference*, Casta-Papiernicka, Slovakia, May 1–6, 2015.

[35] S. Raimes, *Many-Electron Theory* (North-Holland Publishing Company, 1972).

[36] K. Sawada, 'Correlation energy of an electron gas at high density', *Phys. Rev.* **106** (1957), 372.

[37] K. Sawada, K. A. Brueckner, N. Fukuda and R. Brout, 'Correlation energy of an electron gas at high density: Plasma oscillations', *Phys. Rev.* **108** (1957), 507.

[38] R. Seiringer, 'The excitation spectrum for weakly interacting bosons', *Commun. Math. Phys.* **306** (2011), 565–578.

[39] J. Sherman and W. J. Morrison, 'Adjustment of an inverse matrix corresponding to a change in one element of a given matrix', *Ann. Math. Statist*. **21**(1) (195), 124–127.

[40] J. J. Sylvester, 'On the equation to the secular inequalities in the planetary theory', *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **16**(100) (1883), 267–269.

[41] J. Walecka, *Theoretical Nuclear and Sub-nuclear Physics* (World Scientific, 2004).

[42] H. Watanabe, 'Experimental evidence for the collective nature of the characteristic energy loss of electrons in solids. Studies on the dispersion relation of plasma frequencies', *J. Phys. Soc. Japan* **11** (1956), 112–119.

[43] J. Williamson, 'On the algebraic problem concerning the normal forms of linear dynamical systems', *Amer. J. Math.* **58** (1936), 141–163.

https://doi.org/10.1017/fmp.2023.31 Published online by Cambridge University Press