



Maciazek, T. (2021). Repulsively diverging gradient of the density functional in the Reduced Density Matrix Functional Theory. *New Journal of Physics*, 23(11), [113006]. <https://doi.org/10.1088/1367-2630/ac309c>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1088/1367-2630/ac309c](https://doi.org/10.1088/1367-2630/ac309c)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via IoP at <https://doi.org/10.1088/1367-2630/ac309c>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

PAPER • OPEN ACCESS

## Repulsively diverging gradient of the density functional in the reduced density matrix functional theory

To cite this article: Tomasz Maciek 2021 *New J. Phys.* **23** 113006

View the [article online](#) for updates and enhancements.

You may also like

- [Conceptual design of high resolution and reliable density measurement system on helical reactor FFHR-d1 and demonstration on LHD](#)  
T. Akiyama, R. Yasuhara, T. Tanaka et al.
- [STEM-EELS analysis reveals stable high-density He in nanopores of amorphous silicon coatings deposited by magnetron sputtering](#)  
Roland Schierholz, Bertrand Lacroix, Vanda Godinho et al.
- [Pair production in the collapse of a Hopf texture](#)  
Eric W Hirschmann and Steven L Liebling



## PAPER

# Repulsively diverging gradient of the density functional in the reduced density matrix functional theory

## OPEN ACCESS

RECEIVED  
9 July 2021REVISED  
24 September 2021ACCEPTED FOR PUBLICATION  
18 October 2021PUBLISHED  
1 November 2021

Original content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the  
title of the work, journal  
citation and DOI.



Tomasz Maciążek\*

School of Mathematics, University of Bristol, Fry Building, Woodland Road, Bristol BS8 1UG, United Kingdom

\* Author to whom any correspondence should be addressed.

E-mail: [tomasz.maciazek@bristol.ac.uk](mailto:tomasz.maciazek@bristol.ac.uk)**Keywords:** density matrix methods, *ab initio* calculations, density functional theory development, Bose–Einstein condensates, electronic structure

## Abstract

The reduced density matrix functional theory (RDMFT) is a remarkable tool for studying properties of ground states of strongly interacting quantum many body systems. As it gives access to the one-particle reduced density matrix of the ground state, it provides a perfectly tailored approach to studying the Bose–Einstein condensation (BEC) or systems of strongly correlated electrons. In particular, for homogeneous BECs as well as for the Bose–Hubbard dimer it has been recently shown that the relevant density functional exhibits a repulsive gradient (called the Bose–Einstein condensation force) which diverges when the fraction of non-condensed bosons tends to zero. In this paper, we show that the existence of the Bose–Einstein condensation force is completely universal for any type of pair-interaction and also in the non-homogeneous gases. To this end, we construct a universal family of variational trial states which allows us to suitably approximate the relevant density functional in a finite region around the set of the completely condensed states. We also show the existence of an analogous repulsive gradient in the fermionic RDMFT for the  $N$ -fermion singlet sector in the vicinity of the set of the Hartree–Fock states. Finally, we show that our approximate functional may perform well in electron transfer calculations involving low numbers of electrons. This is demonstrated numerically in the Fermi–Hubbard model in the strongly correlated limit where some other approximate functionals are known to fail.

## 1. Introduction

The Bose–Einstein condensate (BEC) has been theoretically predicted by Bose and Einstein [1, 2] in 1924. Seventy years passed before it was finally realised experimentally by Wieman and Cornell and independently by Ketterle in gases of ultracold atoms [3, 4]. Nowadays, the BEC provides a widely-used experimental testing ground for quantum many-body theories, especially superfluidity, superconductivity [7–9] and the laser phenomenon [5, 6]. In this work, we focus on the description of the BEC via the density functional theories. Our developed methods for the BEC are subsequently extended to systems of strongly interacting electrons in the singlet sector.

The most successful and renowned functional theory is perhaps the one which uses the Hohenberg–Kohn density functional [13] (see also [14]). The natural variable of the Hohenberg–Kohn density functional is the single-particle density which unfortunately is not suitable for the purpose of this paper [15]. One of the reasons is that the single-particle densities of the fully condensed quantum states can range through all possible densities. In other words, it is not possible to decide whether a given single-particle density describes a BEC condensate or not. Thus, one needs to use a more sophisticated generalization of the Hohenberg–Kohn density functional, namely the reduced density matrix functional theory (RDMFT) due to Gilbert [16] and Levy [17]. RDMFT uses the single-particle reduced density matrix as its natural variable and hence is able to recover quantum correlations exactly. All density functional theories suffer the issue that their respective functionals are extremely challenging to find and hence they

are usually not known. Remarkable exceptional cases in RDMFT where the functional is actually known analytically include the Bose–Hubbard dimer occupied by two bosons [18] and the singlet sector of the Fermi–Hubbard dimer for two spinful electrons [19, 20]. The common practical strategy is thus to find suitable approximations to the relevant density functional. In RDMFT it is often the case that approximate functionals for some specific systems often lead to progress in finding more complicated approximate functionals for more general systems. For instance, in the fermionic RDMFT this has been the case with the well-known Hartree–Fock functional [28] which in turn led to the discovery of the Müller and Power functionals [29, 30] which in turn inspired numerous more complicated hierarchies of functionals [31, 32]. It is also a crucial task to construct approximate functionals that are simple enough to make explicit calculations tractable, but accurate enough to capture some desired information about the exact functional. Our presented work makes effort in this direction. Our results build on the recent tremendous progress in the foundations of both bosonic and fermionic RDMFT [18, 21–23], especially in RDMFT for bosons and homogeneous BEC, where the remarkable concept of the repulsive *BEC force* has been introduced. In particular, it has been shown that the gradient of the relevant density functional in the space of the reduced density matrices diverges repulsively in the regime of the BEC as the inverse of the square root of the fraction of the non-condensed bosons. This provided an alternative and more fundamental explanation for the existence of the quantum depletion in the homogeneous condensate [23] (see also [25–27] for the underlying theory) and in the Bose–Hubbard dimer [18]. This is compared with the celebrated theories by Bogoliubov [10] and Gross–Pitaevskii [11] which work in the dilute or weakly interacting regime as well as in the high density regime for charged bosons and that have been recently verified experimentally [12].

Our work shows that the repulsive BEC force is present also in non-homogeneous BEC, showing that the presence of the BEC force is an universal phenomenon making the above alternative explanation of the quantum depletion complete for any system of pairwise interacting bosons. Our arguments are remarkably versatile and they can be extended to other physical systems. In particular, similar methods are employed to prove the existence of an analogous phenomenon for the singlet fermionic systems. Our construction coincides with the exact functional in the dimer case (two fermions occupying two sites/orbitals) known from previous works [19]. We numerically test the accuracy of our approximate functional in the Fermi–Hubbard model. For two fermions ( $N = 2$ ) we show that the approximate functional gives very accurate values of the ground state energy. For the Fermi–Hubbard chains, we employ the approximate functional for calculating the electron transfer. In particular, for the half-filled chain of length four, our functional recovers the exact electron transfer in the strongly correlated limit with a good accuracy. Remarkably, this is in contrast with the Müller and Power functionals [29, 30] which are known to fail even in the dimer case [19]. This shows that our proposed functional or its generalizations (see section 8) may be useful in electron transfer calculations, a task which has been pointed out as one of the current challenges of the RDMFT [32, 33].

For the sake of clarity, we would like to make precise what is the form of the most general bosonic Hamiltonian with pairwise interactions that we consider. We work in the discrete setting and consider quantum systems of  $N$  bosons occupying  $d$  sites interacting according to Hamiltonians of the form

$$\hat{H} = \hat{h} + \hat{W}, \quad (1)$$

where  $\hat{h}$  is the single-particle term which comprises of hopping and local potentials while  $\hat{W}$  is an arbitrary pair interaction. Thus, the most general form of  $\hat{h}$  is

$$\hat{h} = \sum_{i < j}^d \left( t_{ij} b_i b_j^\dagger + \bar{t}_{ij} b_j b_i^\dagger \right) + \sum_{i=1}^d v_i \hat{n}_i,$$

where  $t_{i,j}$  are (possibly complex) hopping amplitudes and  $v_i$  are on-site potentials. Recall that the creation and annihilation operators  $b_i, b_j^\dagger$  satisfy the bosonic commutation rules  $[b_i, b_j^\dagger] = \delta_{i,j}$ ,  $[b_i^\dagger, b_j^\dagger] = [b_i, b_j] = 0$  and  $\hat{n}_i := b_i^\dagger b_i$  is the occupation number operator associated with site  $i$ . Similarly, the most general form of the pair interaction reads

$$\hat{W} = \sum_{i,j,k,l=1}^d \Omega_{ijkl} b_i b_j b_k^\dagger b_l^\dagger + \text{h.c.}, \quad (2)$$

where  $\Omega_{ijkl}$  are complex parameters.

## 2. A brief recap of the RDMFT

The setting of the RDMFT is to fix the interaction term  $\hat{W}$  in equation (1) and consider a family of Hamiltonians  $\hat{H}(h) = \hat{h} + \hat{W}$ . One of the key objects in RDMFT is the one-particle reduced density matrix

(denoted here shortly by 1RDM) which in the case of  $N$  bosons occupying  $d$  sites is a  $d \times d$  Hermitian matrix assigned to any state  $|\Psi\rangle$  whose entries read

$$\gamma(|\Psi\rangle)_{ij} = \langle \Psi | b_j b_i^\dagger | \Psi \rangle. \quad (3)$$

For any one particle Hamiltonian  $\hat{h} = \sum_{ij} h_{ij} b_i b_j^\dagger$  we have

$$\langle \Psi | \hat{h} | \Psi \rangle = \sum_{ij} h_{ij} \gamma(|\Psi\rangle)_{ji} = \text{Tr}(h\gamma(|\Psi\rangle)),$$

where  $h$  is simply the matrix of coefficients  $[h_{ij}]$ . The set of all 1RDMs coming from reductions of the pure states (also called the  $N$ -representable 1RDMs) will be denoted by  $\Gamma$ , i.e.

$$\Gamma := \{\gamma \mid \gamma = \gamma(|\Psi\rangle) \text{ for some } |\Psi\rangle\}. \quad (4)$$

Note that not every  $\gamma \in \Gamma$  is the 1RDM of the ground state of  $\hat{H}(h)$  for some  $h$ . Thus, it is also necessary to consider the set of so-called  $v$ -representable 1RDMs associated with  $\hat{W}$  which is the set  $\Gamma_v \subset \Gamma$  of 1RDMs stemming from ground states of Hamiltonians  $\hat{H}(h)$ .

Following Levy [17], Gilbert [16] and Lieb [14] one defines the RDMFT functional  $\mathcal{F}(\gamma)$  defined on  $\Gamma_v$  as follows. Consider the ground state energy of  $\hat{H}(h)$  as a function of  $h$ ,  $E_{\text{gs}}(h)$ . For any  $\gamma \in \Gamma_v$  we define

$$\mathcal{F}(\gamma) := E_{\text{gs}}(h) - \text{Tr}(h\gamma).$$

Then, by the variational principle we have

$$E_{\text{gs}}(h) = \min_{\gamma \in \Gamma_v} (\text{Tr}(h\gamma) + \mathcal{F}(\gamma)). \quad (5)$$

The following two notable problems arise: (i) we do not know what the set  $\Gamma_v$  is and (ii) we do not know what  $\mathcal{F}(\gamma)$  is Levy has proposed to circumvent this problem by extending the domain of the RDMFT functional and the domain of the search (5) to all (possibly non-physical) 1RDMs from  $\Gamma$  [17] and define the extended functional  $\tilde{\mathcal{F}}$  as

$$\tilde{\mathcal{F}}(\gamma) := \min_{|\Psi\rangle \rightarrow \gamma} \langle \Psi | \hat{W} | \Psi \rangle, \quad (6)$$

where the minimization is done over all pure states whose 1RDM is  $\gamma$ . Functional  $\tilde{\mathcal{F}}$  has the following two properties: (P1)  $\tilde{\mathcal{F}}(\gamma) + \text{Tr}(h\gamma) \geq E_{\text{gs}}(h)$  for every  $h$  and  $\gamma \in \Gamma$ , (P2)  $\tilde{\mathcal{F}}(\gamma_{\text{gs}}) + \text{Tr}(h\gamma_{\text{gs}}) = E_{\text{gs}}(h)$ , if  $\gamma_{\text{gs}}$  is the 1RDM of the ground state of  $\hat{H}(h)$ . Properties P1 and P2 imply that for  $\gamma \in \Gamma_v$  we have  $\tilde{\mathcal{F}}(\gamma) = \mathcal{F}(\gamma)$ . Thus, functionals  $\mathcal{F}$  as well as  $\tilde{\mathcal{F}}$  are universal, i.e. they are independent of  $h$ . They do however depend on the interaction  $\hat{W}$ .

Finally, let us point out that by equation (6) the functional  $\tilde{\mathcal{F}}(\gamma)$  is upper bounded by  $\langle \Psi | \hat{W} | \Psi \rangle$  for any  $|\Psi\rangle$  whose 1RDM is  $\gamma$ . We will strongly rely on this elementary fact in the following parts of the paper.

### 3. The repulsive gradient in the Bose–Hubbard dimer

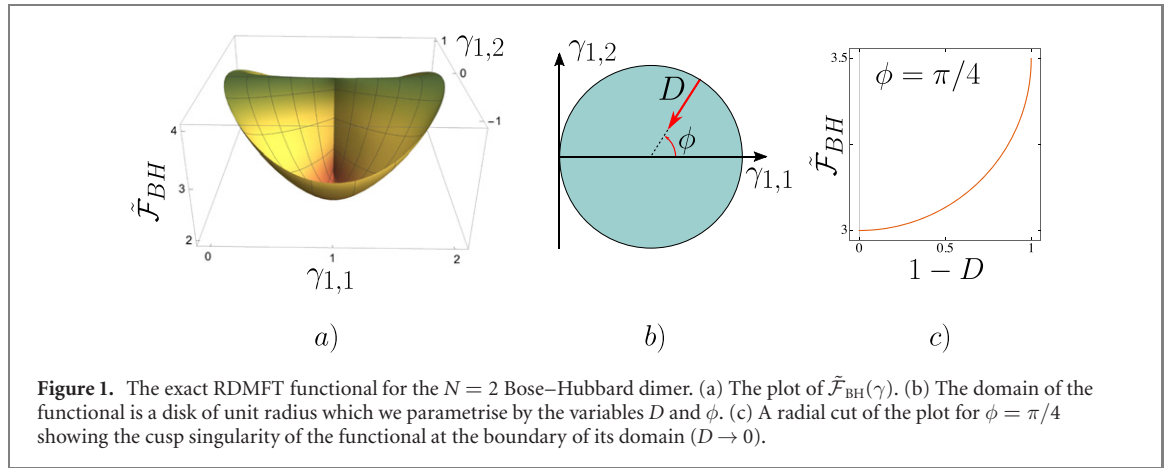
In this section we introduce the main concepts and build some key intuitions that we will develop further in the remaining parts of the paper. A simple and intuitive explanation of the relevant repulsive gradient in the RDMFT functional that we are going to focus on in this section comes from the Bose–Hubbard dimer. This is a system of two sites labelled by the numbers  $i = 1, 2$  with the corresponding creation operators  $b_1^\dagger, b_2^\dagger$  and the interaction given by  $\hat{W}_{\text{BH}} = \hat{n}_1^2 + \hat{n}_2^2$ , where  $\hat{n}_i$  is the particle number operator at the site  $i$ . To simplify things further, we will assume that the system at hand involves only  $N = 2$  bosons and consider only real wave functions. Then, the corresponding two-boson states can be written in the position basis as the superposition

$$|\Psi\rangle = a_1 |2_{(1)}\rangle + a_2 |1_{(1)}, 1_{(2)}\rangle + a_3 |2_{(2)}\rangle, \quad a_1^2 + a_2^2 + a_3^2 = 1,$$

where the basis states are  $|2_{(i)}\rangle = (b_i^\dagger)^2 |\text{vac}\rangle$ ,  $|1_{(1)}, 1_{(2)}\rangle = b_1^\dagger b_2^\dagger |\text{vac}\rangle$ . According to equation (3) the 1RDM of  $|\Psi\rangle$  is a  $2 \times 2$  real symmetric matrix whose coefficients are given by

$$\gamma(|\Psi\rangle) = \begin{pmatrix} \gamma_{1,1} & \gamma_{1,2} \\ \gamma_{1,2} & \gamma_{2,2} \end{pmatrix} = \begin{pmatrix} 2a_1^2 + a_2^2 & \sqrt{2}a_2(a_1 + a_3) \\ \sqrt{2}a_2(a_1 + a_3) & a_2^2 + 2a_3^2 \end{pmatrix}. \quad (7)$$

Remarkably, the above relations can be inverted in order to express the wave function's coefficients in terms of the variables  $\gamma_{ij}$  leading to an exact expression for the Levy's RDMFT functional according to



equation (6). This fact has been studied in the literature—see [18] for the complete derivation and [19] for a similar treatment of the two-electron Fermi–Hubbard dimer. The resulting expression for the exact RDMFT functional reads

$$\tilde{\mathcal{F}}_{\text{BH}}(\gamma) = 4 - \frac{1 + \sqrt{1 - \gamma_{1,2}^2 - (\gamma_{1,1} - 1)^2}}{\gamma_{1,2}^2 + (\gamma_{1,1} - 1)^2}, \quad \gamma_{1,2}^2 + (\gamma_{1,1} - 1)^2 \leq 1. \quad (8)$$

Obtaining exact expressions for the RDMFT functionals is generally not possible for larger systems as the multi-particle wave functions are parametrised by many more variables than the respective 1RDMs which makes realising the minimization in equation (6) in an exact way very difficult.

According to equation (8) the domain of the RDMFT functional for the  $N = 2$  dimer is the disk of unit radius  $\gamma_{1,2}^2 + (\gamma_{1,1} - 1)^2 \leq 1$ . This fact comes solely from the condition that the 1RDM  $\gamma$  has to be positive-semidefinite. Importantly, the 1RDMs which lie on the boundary of the disk have the property that  $\gamma^2 = 2\gamma$ , i.e. they have the eigenvalues  $\nu_1 = 2, \nu_2 = 0$ . In other words, such 1RDMs come from the states which are completely condensed in a single mode  $|\psi\rangle$  being some linear combination of the spatial modes  $|1\rangle$  and  $|2\rangle$ . In this paper, we will be mainly interested in the behaviour of the functional  $\tilde{\mathcal{F}}_{\text{BH}}(\gamma)$  when  $\gamma$  is close to the set of 1RDMs which come from such completely condensed states. In contrast to the dimer, in higher dimensions the 1RDMs of the completely condensed states form only a subset of the boundary of the domain of the RDMFT functional. However, in the simple example at hand we can visualise the exact functional (figure 1(a)) and see what exactly is happening in the vicinity of its boundary (figure 1(c)). Namely, we see that the gradient of  $\tilde{\mathcal{F}}_{\text{BH}}(\gamma)$  diverges at the boundary of its domain (this is often called a cusp singularity).

More precisely, we introduce the distance and angle variables,  $D \in [0, 1]$  and  $\phi \in [0, 2\pi]$  (figure 1(b)), which parametrise the disk such that

$$\gamma_{1,1} = 1 + (1 - D) \cos \phi, \quad \gamma_{1,2} = (1 - D) \sin \phi,$$

see figure 1(a). Using the above parametrisation in equation (8) we get that

$$\frac{\partial \tilde{\mathcal{F}}_{\text{BH}}(\gamma(D, \phi))}{\partial D} = -\frac{1 - D}{\sqrt{D(2 - D)}} \sin^2 \phi,$$

which diverges as  $-\frac{\sin^2 \phi}{\sqrt{2}} \frac{1}{\sqrt{D}}$  in the leading order when  $D \rightarrow 0$ . This repulsive divergence of the RDMFT functional at the boundary of its domain has been confirmed to exist in many more physical systems including the  $N$ -boson Bose–Hubbard dimer, homogeneous BECs (where it is called the *BEC force* [18, 23, 24]) and translationally invariant one-dimensional fermionic systems [21] (where it is called the *exchange force*). It has been conjectured that it is a property of generic interacting many-body systems. In this paper, we show that this is indeed the case. However, as we are not able to write down the exact functional for a general system and simply compute its derivative, we need to find an appropriate bound for the RDMFT functional whose diverging gradient will imply that the gradient of the exact RDMFT functional diverges as well when  $D \rightarrow 0$ . To this end, for any 1RDM  $\gamma$  we construct an ansatz state  $|\Psi(\gamma)\rangle$  such that  $\gamma$  is the 1RDM of the state  $|\Psi(\gamma)\rangle$ . With such an ansatz at hand, we simply bound the exact functional by

$$\tilde{\mathcal{F}}(\gamma) \leq \langle \Psi(\gamma) | \hat{W} | \Psi(\gamma) \rangle$$

and show that the bound diverges repulsively when the (appropriately defined) distance variable  $D$  tends to 0. Let us also remark that the distance variable  $D$  has the interpretation as the depletion parameter. To see this, we find the eigenvalues of  $\gamma$  to be  $\nu_1 = 2 - D$  and  $\nu_2 = D$ . Thus,  $D$  is the number of bosons that are outside of the highest occupied mode.

Let us next show how this strategy works out for the  $N$ -boson Bose–Hubbard dimer. We assume that the 1RDMs are normalised to  $\text{tr } \gamma = N$ . This implies that the set of 1RDMs is a disk of the radius  $N/2$  given by the inequality  $(\gamma_{1,1} - N/2)^2 + \gamma_{1,2}^2 \leq N^2/4$ . Thus, we use the parametrisation of  $\gamma(D, \phi)$  which reads

$$\gamma_{1,1} = \frac{N}{2} + \left(\frac{N}{2} - D\right) \cos \phi, \quad \gamma_{1,2} = \left(\frac{N}{2} - D\right) \sin \phi. \quad (9)$$

We will construct the ansatz state  $|\Psi(\gamma(D, \phi))\rangle \equiv |\Psi(D, \phi)\rangle$  in the basis of eigenmodes of  $\gamma(D, \phi) = (N - D) |\phi\rangle\langle\phi| + D|\phi^\perp\rangle\langle\phi^\perp|$ , where

$$|\phi\rangle = \cos\left(\frac{\phi}{2}\right) |1\rangle + \sin\left(\frac{\phi}{2}\right) |2\rangle, \quad |\phi^\perp\rangle = -\sin\left(\frac{\phi}{2}\right) |1\rangle + \cos\left(\frac{\phi}{2}\right) |2\rangle.$$

The ansatz is the following superposition of the completely condensed state in the mode  $|\phi\rangle$  with the doubly-excited state

$$|\Psi(D, \phi)\rangle = \sqrt{\frac{2-D}{2}} |N_{(\phi)}\rangle - \sqrt{\frac{D}{2}} |(N-2)_{(\phi), 2_{(\phi^\perp)}}\rangle.$$

It is straightforward to verify that the state  $|\Psi(D, \phi)\rangle$  gives the correct 1RDM (9) which is diagonal in the  $|\phi\rangle, |\phi^\perp\rangle$  basis and whose diagonal entries in this basis are  $N - D$  and  $D$ . By evaluating the expression  $\langle\Psi(D, \phi)|\hat{W}|\Psi(D, \phi)\rangle$  we get the following upper bound for the RDMFT functional for any  $D \leq 2$

$$\begin{aligned} \tilde{\mathcal{F}}_{\text{BH}}(D, \phi) &\leq \frac{2-D}{2} \langle N_{(\phi)} | \hat{W}_{\text{BH}} | N_{(\phi)} \rangle - \sqrt{D(2-D)} \langle (N-2)_{(\phi), 2_{(\phi^\perp)}} | \hat{W}_{\text{BH}} | N_{(\phi)} \rangle \\ &\quad + \frac{D}{2} \langle (N-2)_{(\phi), 2_{(\phi^\perp)}} | \hat{W}_{\text{BH}} | (N-2)_{(\phi), 2_{(\phi^\perp)}} \rangle. \end{aligned}$$

The upper bound is equal to the exact functional at  $D = 0$  and its gradient diverges repulsively when  $D \rightarrow 0$  proportionally to  $-1/\sqrt{D}$  in the leading order which implies the desired result. It is however crucial to be assured that  $\langle (N-2)_{(\phi), 2_{(\phi^\perp)}} | \hat{W} | N_{(\phi)} \rangle$  is generically strictly positive. Otherwise, the entire argument breaks down as there are no other terms in the bound that give the repulsive gradient. This is done by a straightforward but rather tedious calculation which relies on expressing the interaction  $\hat{W}_{\text{BH}} = \hat{n}_1^2 + \hat{n}_2^2$  in terms of the creation and annihilation operators  $b_\phi = \cos(\phi/2)b_1 + \sin(\phi/2)b_2$  and  $b_{\phi^\perp} = -\sin(\phi/2)b_1 + \cos(\phi/2)b_2$ . This way, we find that

$$\langle (N-2)_{(\phi), 2_{(\phi^\perp)}} | \hat{W} | N_{(\phi)} \rangle = \sqrt{\frac{N(N-1)}{2}} \sin^2 \phi \geq 0,$$

thus the above overlap vanishes only when  $\phi \in \{0, \pi\}$ . A similar calculation allows us to obtain expressions for the remaining expectation values of  $\hat{W}_{\text{BH}}$ . The resulting upper bound for the RDMFT functional is plotted for  $N = 4$  and  $N = 10$  in figure 2. It is given by the formula

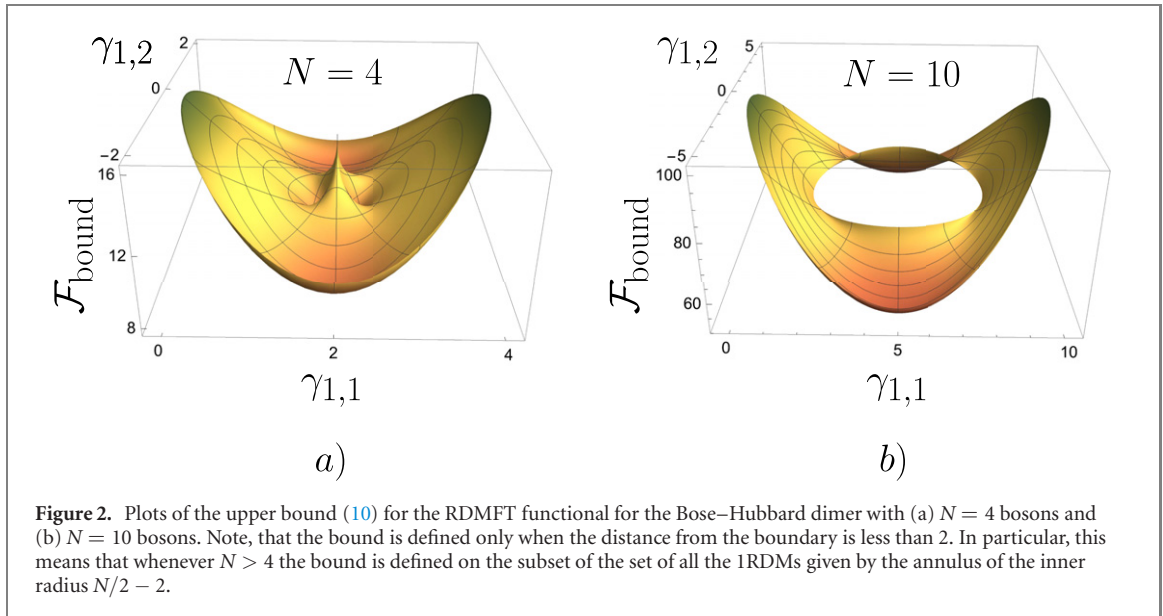
$$\begin{aligned} \tilde{\mathcal{F}}_{\text{BH}}(D, \phi) &\leq \frac{1}{4}N(1 + 3N + (N-1)\cos(2\phi)) - \sin^2 \phi \sqrt{\frac{N(N-1)}{2}} \sqrt{D(2-D)} + \\ &\quad - \frac{D}{2}(N-2)(1 + 3\cos(2\phi)). \end{aligned} \quad (10)$$

Finally, the resulting bound for the gradient of the functional reads

$$\frac{\partial \tilde{\mathcal{F}}_{\text{BH}}(\gamma(D, \phi))}{\partial D} \leq -\sin^2 \phi \sqrt{\frac{N(N-1)}{2}} \frac{1-D}{\sqrt{D(2-D)}} - \frac{1}{2}(N-2)(1 + 3\cos(2\phi)). \quad (11)$$

The above bound for the gradient can be compared with the results of [18] where the first term of the expansion of the gradient in the powers of  $D$  has been obtained. The results agree with equation (11) up to some factors of  $N$  which come from the different choices of the normalisation of 1RDMs.

To end this section, let us remark that analytic bounds for the exact RDMFT functional are potentially very useful tools in approximately solving complex physical systems. This is because the functional is universal for all the single-particle terms  $\hat{h}$  in equation (1). For the Hubbard model the single-particle terms include the external on-site potentials and all the possible hopping terms. Thus, quite remarkably, an approximate functional can be applied to the Hubbard model of any geometry. Moreover, finding the



ground state energy boils down to the easy task of finding the minimum of a surface (just like the surfaces on figures 2 and 1(a), but generally of a higher dimension) which is tilted by the single-particle terms of the Hamiltonian. In particular, for the dimer this boils down to the almost trivial task of finding the minimum of a two-dimensional surface—independently of the number of bosons  $N$ .

#### 4. The spectral simplex $\Delta_{N,d}$

This section is devoted to explaining some key geometric properties of the set  $\Gamma$  defined in equation (4). One of the main difficulties in RDMFT is the need of optimization over set  $\Gamma$  which is in general a complicated non-convex set. However, in contrast to its fermionic counterpart [37, 40–42], set  $\Gamma$  for bosons is convex and has a tractable description in terms of a geometric object which we call the spectral simplex and denote by  $\Delta_{N,d}$ . Before we define  $\Delta_{N,d}$ , we need to take a closer look at a particular set of operations called the single-particle unitaries. The single-particle unitaries are unitary operators generated by single-particle Hermitian operators, i.e. they are operations of the form  $\hat{U} = e^{i\hat{A}}$ , where  $\hat{A} = \sum_{i,j=1}^d A_{i,j} b_i b_j^\dagger$  with  $A_{i,j} = A_{j,i}^*$  (the star denotes the complex conjugate). The matrix of coefficients  $[A_{i,j}]$  will be denoted by  $A$ . It is a Hermitian  $d \times d$  matrix, thus  $U = e^{iA}$  is a unitary  $d \times d$  matrix. Hence, with a given matrix  $A$  we have associated a single-particle unitary  $\hat{U}$  and a  $d \times d$  unitary  $U$ . While  $\hat{U}$  acts on  $N$ -boson wave functions, matrix  $U$  acts on 1RDMs via conjugation. Importantly, these two actions are equivalent, namely

$$\gamma(\hat{U}|\Psi\rangle) = U\gamma(|\Psi\rangle)U^\dagger. \quad (12)$$

In particular, any 1RDM  $\gamma$  can be diagonalised by the above operations. This means that set  $\Gamma$  is parametrised by the spectra of 1RDMs coming from pure states of  $N$  bosons. Note that the spectrum of any 1RDM  $(n_1, n_2, \dots, n_d)$  has to satisfy

$$0 \leq n_i \leq N \quad \text{and} \quad n_1 + \dots + n_d = N. \quad (13)$$

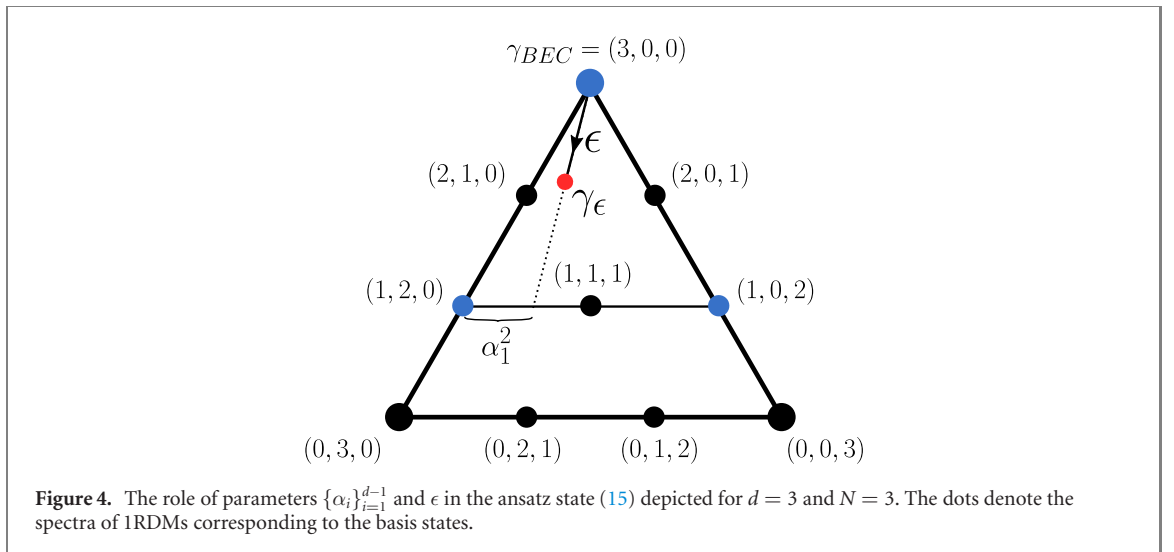
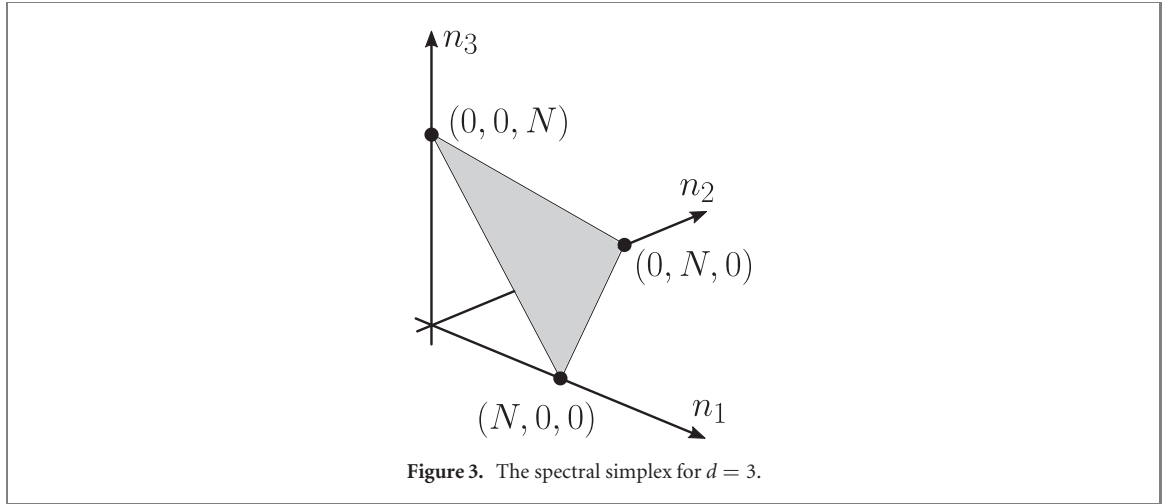
These constraints define the spectral simplex  $\Delta_{N,d}$  (see figure 3). Simplex  $\Delta_{N,d}$  is the convex hull of  $d$  vertices  $(N, 0, \dots, 0)$ ,  $(0, N, 0, \dots, 0)$ ,  $\dots$ ,  $(0, \dots, 0, N)$ . It turns out that any point of the simplex corresponds to the spectrum of a 1RDM. To see this, consider state

$$|\Psi_{\bar{\alpha}}\rangle := \sum_{i=1}^d \alpha_i |N_{(i)}\rangle,$$

where  $|N_{(i)}\rangle := \frac{1}{\sqrt{N!}} \left(b_i^\dagger\right)^N |\text{vac}\rangle$  and  $\bar{\alpha} = (\alpha_1, \dots, \alpha_d)$  is a vector of real numbers whose squares sum up to one. The 1RDM of  $|\Psi_{\bar{\alpha}}\rangle$  is diagonal and it reads

$$\gamma(|\Psi_{\bar{\alpha}}\rangle) = \text{diag}(N\alpha_1^2, N\alpha_2^2, \dots, N\alpha_d^2).$$





Thus, by changing the parameters  $\bar{\alpha}$  we can reach any convex combination of simplex's vertices. Stated another way, the above fact means that any  $d \times d$  density matrix  $\gamma$  (i.e. a positive-definite Hermitian matrix of trace  $N$ ) is the 1RDM of a pure bosonic state of the form  $\hat{U} |\Psi_{\bar{\alpha}}\rangle$  for some single-particle  $\hat{U}$  and some  $\bar{\alpha}$ .

For a given typical  $d \times d$  density matrix  $\gamma$  there exist many  $N$ -boson pure quantum states that have such a  $\gamma$  as their 1RDM. However, this changes diametrically when we consider a special subset of  $\Gamma$  which is defined as the set of 1RDMs of completely condensed states of  $N$  bosons. All such completely condensed states are of the forms  $\hat{U} |N_{(1)}\rangle$  where  $\hat{U}$  is a single-particle unitary. By the property (12) the 1RDM of a completely condensed state is of the form  $U\gamma_{\text{BEC}}U^\dagger$ , where

$$\gamma_{\text{BEC}} := \gamma(|N_{(1)}\rangle) = \text{diag}(N, 0, \dots, 0).$$

Thus, we define

$$\Gamma_{\text{BEC}} := \{U\gamma_{\text{BEC}}U^\dagger \mid U - d \times d \text{ unitary}\}.$$

Importantly, set  $\Gamma_{\text{BEC}}$  is in a one-to-one correspondence with the set of completely condensed states and the correspondence is  $U\gamma_{\text{BEC}}U^\dagger \leftrightarrow \hat{U} |N_{(1)}\rangle$ . Hence, when computing Levy's functional (6) no minimization is required and we have

$$\tilde{\mathcal{F}}(U\gamma_{\text{BEC}}U^\dagger) = \langle N_{(1)} | \hat{U}^\dagger \hat{W} \hat{U} | N_{(1)} \rangle. \quad (14)$$

## 5. The exposition of the main result

The proposed variational ansatz state is a combination of the completely condensed state  $|N_{(1)}\rangle$  and a weighted superposition of states with two bosons outside the condensate. Its precise normalised form reads

$$|\Psi_{\epsilon,\sigma}\rangle = \frac{1}{\sqrt{1+\epsilon^2}} \left( 1 + \frac{\sigma\epsilon}{\sqrt{2N(N-1)}} b_1^2 \sum_{i=1}^{d-1} \alpha_i (b_{i+1}^\dagger)^2 \right) |N_{(1)}\rangle, \quad (15)$$

where the variational parameters are  $\epsilon \geq 0$  and  $\sigma = \pm 1$ . Numbers  $\{\alpha_i\}_{i=1}^{d-1}$  are arbitrary fixed numbers whose role is to probe all directions inside the simplex. They are all nonnegative and satisfy  $\alpha_1^2 + \dots + \alpha_{d-1}^2 = 1$ . Parameter  $\epsilon$  determines the Euclidean distance of  $\gamma_\epsilon := \gamma(|\Psi_\epsilon\rangle)$  from the 1RDM of the completely condensed state,  $\gamma_{\text{BEC}} = \text{diag}(N, 0, \dots, 0)$ . To see this, recall that the distance between two matrices is given by the Hilbert–Schmidt measure

$$D(\gamma, \gamma') = \sqrt{\text{Tr}((\gamma - \gamma')^2)}. \quad (16)$$

It is straightforward to see that  $\gamma_\epsilon$  is diagonal and its precise form reads

$$\gamma_\epsilon = \frac{1}{1+\epsilon^2} \text{diag}(N(1+\epsilon^2) - 2\epsilon^2, 2\epsilon^2\alpha_1^2, \dots, 2\epsilon^2\alpha_{d-1}^2).$$

Thus, numbers  $\{\alpha_i\}_{i=1}^{d-1}$  are in a one-to-one correspondence with the possible directions from vertex  $\gamma_{\text{BEC}}$  to the interior of the simplex  $\Delta_{N,d}$  (see figure 4). A straightforward calculation shows that

$$D_\epsilon := D(\gamma_{\text{BEC}}, \gamma_\epsilon) = \beta_d \frac{\epsilon^2}{1+\epsilon^2}, \quad (17)$$

where  $\beta_d := 2\sqrt{1+\alpha_1^4+\dots+\alpha_{d-1}^4}$ . Note that the number of non-condensed bosons in  $\gamma_\epsilon$  is exactly  $N - N_{\text{BEC}} = 2\epsilon^2/(1+\epsilon^2)$ , thus  $D_\epsilon$  has the physical interpretation as the depletion parameter

$$D_\epsilon = \frac{1}{2}\beta_d(N - N_{\text{BEC}}).$$

Note that distance  $D$  is invariant with respect to conjugating its arguments by a  $d \times d$  unitary  $U$ . Hence,  $D_\epsilon$  is also equal to the distance from  $\gamma(\hat{U}|\Psi_\epsilon\rangle) = U\gamma_\epsilon U^\dagger$  to the 1RDM of the corresponding completely condensed state  $\gamma(\hat{U}|N_{(1)}\rangle) = U\gamma_{\text{BEC}}U^\dagger$  (see figure 5). The ansatz state (15) provides an upper bound for the RDMFT functional of the form

$$\tilde{\mathcal{F}}(U\gamma_\epsilon U^\dagger) \leq \langle \Psi_{\epsilon,\sigma} | \hat{U}^\dagger \hat{W} \hat{U} | \Psi_{\epsilon,\sigma} \rangle.$$

The above bound is defined for any 1RDM  $\gamma$  in the finite region around the set  $\Gamma_{\text{BEC}}$  consisting of the 1RDMs that satisfy the condition  $N - N_{\text{BEC}} \leq 2$ . This is because any such  $\gamma$  can be written as  $U\gamma_\epsilon U^\dagger$  for some  $U$  and  $\gamma_\epsilon$  that are determined via the diagonalisation of  $\gamma$ . What is more, the matrix  $U\gamma_{\text{BEC}}U^\dagger$  minimises the distance from  $\gamma$  to the set  $\Gamma_{\text{BEC}}$ . The above upper bound has a particularly simple form, namely

$$\langle \Psi_{\epsilon,\sigma} | \hat{U}^\dagger \hat{W} \hat{U} | \Psi_{\epsilon,\sigma} \rangle = \frac{\tilde{\mathcal{F}}(U\gamma_{\text{BEC}}U^\dagger) + c_1\sigma\epsilon + c_2\epsilon^2}{1+\epsilon^2}. \quad (18)$$

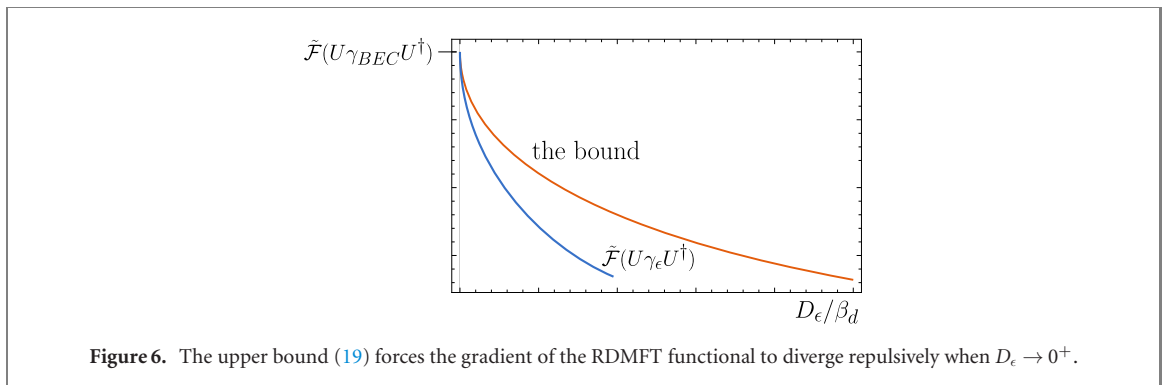
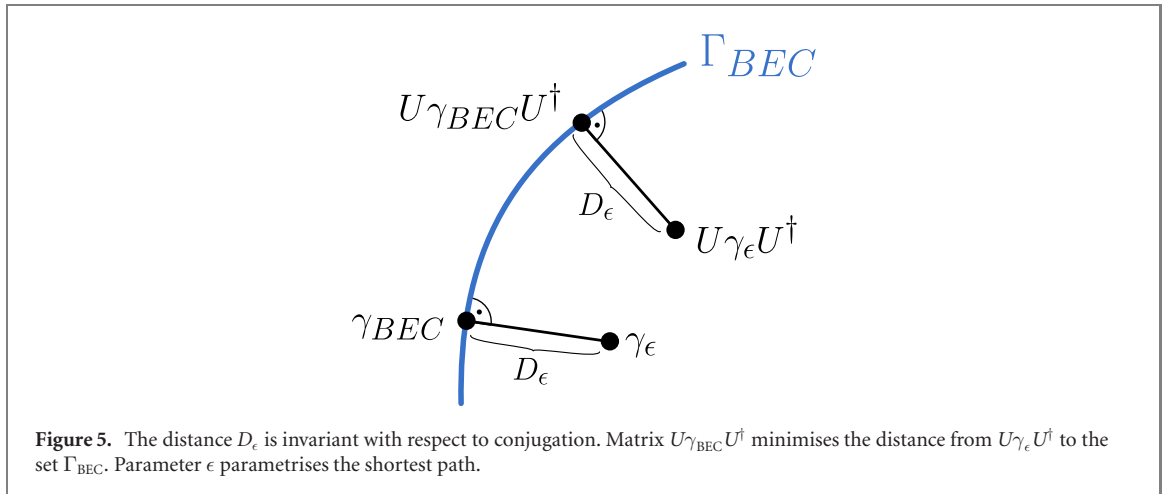
where constants  $c_1, c_2$  depend on  $N, d$ , parameters  $\{\alpha_i\}_{i=1}^{d-1}$  as well as the precise form of  $\hat{W}$  and the unitary  $U$ . The expressions for  $c_1$  and  $c_2$  can be viewed as the following combinations of the two-body integrals  $c_1 = 2\sum_{k=1}^{d-1} \alpha_k \mathfrak{R}I_k$ ,  $c_2 = \sum_{k=1}^{d-1} \alpha_k^2 I_{k,k} + 2\sum_{k<l=1}^{d-1} \alpha_k \alpha_l \mathfrak{R}I_{k,l}$ , where

$$I_k := \frac{1}{\sqrt{2N(N-1)}} \langle N_{(1)} | \hat{U}^\dagger \hat{W} \hat{U} b_1^2 (b_{k+1}^\dagger)^2 | N_{(1)} \rangle,$$

$$I_{k,l} := \frac{1}{2N(N-1)} \langle N_{(1)} | b_{l+1}^2 (b_1^\dagger)^2 \hat{U}^\dagger \hat{W} \hat{U} b_1^2 (b_{k+1}^\dagger)^2 | N_{(1)} \rangle.$$

The two-body integrals can be obtained as closed-form expressions if  $d = 2$ , while for  $d > 2$  they are computable as convergent series. Importantly, in the following sections we argue that the coefficient  $c_1$  is generically nonzero. Using equation (17) this yields the following upper bound for the functional in terms of the distance  $D_\epsilon < \beta_d$

$$\tilde{\mathcal{F}}(U\gamma_\epsilon U^\dagger) \leq \left(1 - \frac{D_\epsilon}{\beta_d}\right) \tilde{\mathcal{F}}(U\gamma_{\text{BEC}}U^\dagger) + c_1\sigma\sqrt{\frac{D_\epsilon}{\beta_d} \left(1 - \frac{D_\epsilon}{\beta_d}\right)} + c_2\frac{D_\epsilon}{\beta_d}. \quad (19)$$



The final key observation is that the upper bound in equation (19) is exact for  $\epsilon = 0$ . Hence, while approaching the completely condensed state by taking  $\epsilon \ll 1$  (and hence  $D_\epsilon \ll 1$ ), the gradient of the upper bound with respect to the distance  $D_\epsilon$  is also an upper bound for the gradient of the exact functional (see figure 6). Furthermore, the binary parameter  $\sigma$  can be chosen so that the gradient of the expression (19) is repulsive for small  $D_\epsilon$  and diverges to  $-\infty$  as  $-\frac{|c_1|}{2\sqrt{\beta_d}} \frac{1}{\sqrt{D_\epsilon}}$  in the leading order, a result anticipated in earlier works [18, 23].

## 6. Calculating the coefficient $c_1$

In this section we provide arguments for the crucial fact that thanks to the specific form of our proposed ansatz state the coefficient  $c_1$  in the bound (19) is generically non-zero.

Although our proof of the bound (19) is elementary and requires only minimal assumptions about the Hamiltonian, it becomes heavy with calculations in its most general instance. However, most of the key arguments appear already when considering a dimer,  $d = 2$ . Hence, it will be most instructive to gradually increase the generality of our considerations, beginning with the Bose–Hubbard dimer, then moving to a dimer with an arbitrary pair interaction and finally formulating the proof in its full generality.

### 6.1. The Bose–Hubbard dimer

The interaction term for the Bose–Hubbard Hamiltonian on  $d$  sites reads

$$\hat{W}_{\text{BH}}^{(d)} = V \sum_{i=1}^d \hat{n}_i^2$$

with  $V \in \mathbb{R}$  being a real parameter.  $V > 0$  corresponds to an on-site repulsion, while  $V < 0$  corresponds to an on-site attraction. The most difficult part of the proof of bound (19) relies on computing  $U^\dagger \hat{W}_{\text{BH}} U$ . The general strategy is to express the interaction term in terms of generators of the unitary group and then use some well-known facts concerning conjugation of the generators by unitary matrices (in particular, the correspondence with rotations in  $(d^2 - 1)$ -dimensions [34]). Let us next show how this strategy is realized in the case of the Hubbard dimer where  $\hat{W}_{\text{BH}} = V(\hat{n}_1^2 + \hat{n}_2^2)$ . Our first goal is to express  $\hat{W}_{\text{BH}}$  in terms of the

$SU(2)$  generators

$$\hat{X} := b_1^\dagger b_2 + b_2^\dagger b_1, \quad \hat{Y} := i \left( -b_1^\dagger b_2 + b_2^\dagger b_1 \right), \quad \hat{Z} := b_1^\dagger b_1 - b_2^\dagger b_2 = \hat{n}_1 - \hat{n}_2. \quad (20)$$

In the following, we will often arrange the above generators into a vector  $\hat{\vec{R}} = (\hat{R}_1, \hat{R}_2, \hat{R}_3)$ , where  $\hat{R}_1 = \hat{X}$ ,  $\hat{R}_2 = \hat{Y}$ ,  $\hat{R}_3 = \hat{Z}$ . In the above uniform notation, the generators satisfy the standard  $SU(2)$  commutation relations, i.e.  $[\hat{R}_i, \hat{R}_j] = 2i\epsilon_{ijk}\hat{R}_k$ , where  $\epsilon_{ijk}$  is the Levi-Civita symbol (not to be confused with the real parameter  $\epsilon$ ).

With the above notation established, the Bose–Hubbard interaction term can be written as

$$\hat{W}_{\text{BH}}^{(2)} = \frac{V}{2} \left( \hat{N}^2 + \hat{Z}^2 \right),$$

where  $\hat{N}$  is the total particle number operator. For  $d > 2$ , we analogously define  $\hat{Z}_{ij} = \hat{n}_i - \hat{n}_j$  to obtain

$$\hat{W}_{\text{BH}}^{(d)} = \frac{V}{d} \left( \hat{N}^2 + \sum_{i < j} \hat{Z}_{ij}^2 \right).$$

Let us next take a slight detour and revisit some algebraic properties of conjugation in  $d = 2$ . By definition, any unitary operator for  $d = 2$  can be written in terms of generators (20) as  $\hat{U} = e^{i\alpha\vec{\tau}\cdot\hat{\vec{R}}}$ , where  $\vec{\tau}$  is a vector in three dimensions of unit length. Recall the following key formula for the conjugation in  $d = 2$

$$e^{-i\alpha\vec{\tau}\cdot\hat{\vec{R}}}\hat{R}_j e^{i\alpha\vec{\tau}\cdot\hat{\vec{R}}} = \sum_{i=1}^3 C_{ij}(\alpha, \vec{\tau})\hat{R}_i. \quad (21)$$

Matrix  $[C_{ij}](\alpha, \vec{\tau})$  is real and is well-known to describe a rotation by angle  $2\alpha$  about the axis  $\vec{\tau}$  [34]. The exact expressions for its entries can be found e.g. in [35].

In particular, for the Bose–Hubbard dimer, we have

$$\langle \Psi_{\epsilon, \sigma} | \hat{U}^\dagger \hat{W}_{\text{BH}}^{(2)} \hat{U} | \Psi_{\epsilon, \sigma} \rangle = \frac{V}{2} \left( N^2 + \langle \Phi_{\epsilon, \sigma}^U | \Phi_{\epsilon, \sigma}^U \rangle \right),$$

where  $|\Phi_{\epsilon, \sigma}^U\rangle := \hat{U}^\dagger \hat{Z} \hat{U} | \Psi_{\epsilon, \sigma} \rangle$ . Formula (21) applied for  $j = 3$  gives

$$\begin{aligned} \langle \Phi_{\epsilon, \sigma}^U | \Phi_{\epsilon, \sigma}^U \rangle &= \sum_{i,j=1}^3 C_{i,3} C_{j,3} \langle \Psi_{\epsilon, \sigma} | \hat{R}_j \hat{R}_i | \Psi_{\epsilon, \sigma} \rangle \\ &= \sum_{i=1}^3 C_{i,3}^2 \langle \Psi_{\epsilon, \sigma} | \hat{R}_i^2 | \Psi_{\epsilon, \sigma} \rangle. \end{aligned}$$

The last line of the above equation follows from the fact that  $\langle \Psi_{\epsilon, \sigma} | \hat{R}_j \hat{R}_i | \Psi_{\epsilon, \sigma} \rangle = 0$  whenever  $j = 3$  and  $i \neq 3$  or  $i = 3$  and  $i \neq 3$ . This is because  $|\Psi_{\epsilon, \sigma}\rangle$  is a superposition of  $|N_1\rangle$  and  $b_0^2 (b_1^\dagger)^2 |N_1\rangle$ , so there have to be two different annihilations in  $\hat{R}_j \hat{R}_i$  to get a nonzero expectation value. Moreover,  $|\Psi_{\epsilon, \sigma}\rangle$  is a real vector, thus  $\langle \Psi_{\epsilon, \sigma} | (\hat{X}\hat{Y} + \hat{Y}\hat{X}) | \Psi_{\epsilon, \sigma} \rangle = 0$ . Next, we check the following facts by a straightforward calculation

$$\begin{aligned} \hat{X} | \Psi_{\epsilon, \sigma} \rangle &= \mathcal{N}_\epsilon \left( \left( \sqrt{N} + \sigma\epsilon\sqrt{2(N-1)} \right) |N-1, 1\rangle + \sigma\epsilon\sqrt{3(N-2)} |N-3, 3\rangle \right), \\ \hat{Y} | \Psi_{\epsilon, \sigma} \rangle &= i\mathcal{N}_\epsilon \left( \left( \sqrt{N} - \sigma\epsilon\sqrt{2(N-1)} \right) |N-1, 1\rangle + \sigma\epsilon\sqrt{3(N-2)} |N-3, 3\rangle \right), \\ \hat{Z} | \Psi_{\epsilon, \sigma} \rangle &= \mathcal{N}_\epsilon \left( N |N, 0\rangle + \sigma\epsilon(N-4) |N-2, 2\rangle \right), \end{aligned} \quad (22)$$

where we used the shorthand notation  $|K, L\rangle$  for the normalised state  $(b_1^\dagger)^K (b_2^\dagger)^L |\text{vac}\rangle$ . The above calculation is key for our proof. Namely, computing expectation values of  $\hat{R}_i^2$  generates terms proportional to  $\epsilon$  due to the fact that squares of the form  $(\sqrt{N} \pm \sigma\epsilon\sqrt{2(N-1)})^2$  appear on the way. Collecting all

contributions proportional to  $\epsilon$  and  $\epsilon^2$ , we find the following constants from equation (18).

$$\begin{aligned} \tilde{\mathcal{F}}(U\gamma_{\text{BEC}}U^\dagger) &= \frac{NV}{2} (N(1 + C_{3,3}^2) + C_{1,3}^2 + C_{2,3}^2), \\ c_1 &= V\sqrt{2N(N-1)} (C_{1,3}^2 - C_{2,3}^2), \\ c_2 &= \frac{V}{2} (N^2 + (5N-8)(C_{1,3}^2 + C_{2,3}^2) + (N-4)^2 C_{3,3}^2). \end{aligned}$$

Using the axis-angle representation of the rotation matrix  $[C_{i,j}](\vec{r}, \alpha)$  [35], we obtain explicitly

$$c_1 = V\sqrt{2N(N-1)} ((r_1^2 - r_2^2)(4r_3^2 \sin^4(\alpha) - \sin^2(2\alpha)) + 8r_1r_2r_3 \sin^2(\alpha) \sin(2\alpha)).$$

### 6.2. The dimer with an arbitrary interaction

The most general interaction from equation (25) for  $d = 2$  can always be recast into a combination of  $\hat{R}_i\hat{R}_j$  with the complex coefficients  $\Omega_{ij} = \omega_{ij} + i\tau_{ij}$ . Taking into account that the interaction  $\hat{W}$  is Hermitian, we have  $\hat{W} = \sum_{i \leq j} \hat{W}_{ij}$ , where

$$\hat{W}_{ij} = \omega_{ij} (\hat{R}_i\hat{R}_j + \hat{R}_j\hat{R}_i) + i\tau_{ij} (\hat{R}_i\hat{R}_j - \hat{R}_j\hat{R}_i).$$

In particular, if  $\Omega_{3,3}$  is the only nonzero coefficient, we recover the Bose–Hubbard interaction (up to a constant term).

Applying the conjugation formula (21) to such a most general  $\hat{W}$ , we obtain

$$\hat{U}^\dagger \hat{W} \hat{U} = \sum_{k \leq l} (\tilde{\omega}_{k,l} (\hat{R}_k\hat{R}_l + \hat{R}_l\hat{R}_k) + i\tilde{\tau}_{k,l} (\hat{R}_k\hat{R}_l - \hat{R}_l\hat{R}_k)), \tag{23}$$

where

$$\begin{aligned} \tilde{\omega}_{k,l} &= \sum_{i \leq j} \omega_{ij} (C_{k,i}C_{l,j} + C_{k,j}C_{l,i}), \\ \tilde{\tau}_{k,l} &= \sum_{i \leq j} \tau_{ij} (C_{k,i}C_{l,j} - C_{k,j}C_{l,i}). \end{aligned}$$

The rest of the proof proceeds in a very similar way to the Bose–Hubbard case. Namely, we notice that  $\langle \Psi_{\epsilon,\sigma} | (\hat{R}_k\hat{R}_l + \hat{R}_l\hat{R}_k) | \Psi_{\epsilon,\sigma} \rangle = 0$  whenever  $k \neq l$ . Similarly,  $\langle \Psi_{\epsilon,\sigma} | (\hat{R}_k\hat{R}_3 - \hat{R}_3\hat{R}_k) | \Psi_{\epsilon,\sigma} \rangle = 0$  whenever  $k \neq 3$ . Moreover, the only terms proportional to  $\epsilon$  (i.e. contributing to  $c_1$  from equation (18)) come from components of equation (23) with  $k = l \neq 3$ . Thus, using equation (22) we obtain

$$c_1 = 2\sqrt{2N(N-1)} (\tilde{\omega}_{1,1} - \tilde{\omega}_{2,2}).$$

Closed forms of the remaining coefficients from the bound (18) read

$$\begin{aligned} \tilde{\mathcal{F}}(U\gamma_{\text{BEC}}U^\dagger) &= N(N\tilde{\omega}_{3,3} + \tilde{\omega}_{2,2} + \tilde{\omega}_{1,1} - 4\tilde{\tau}_{1,2}) \\ c_2 &= (5N-8)(\tilde{\omega}_{1,1} + \tilde{\omega}_{2,2}) + \tilde{\omega}_{3,3}(N-4)^2 - 2\tilde{\tau}_{1,2}(N-4). \end{aligned}$$

### 6.3. The general case

Our previous considerations concerning an interacting dimer generalise fairly straightforwardly to  $d > 2$ . Firstly, we have more generators of the single-particle unitary operations. Their forms are completely analogous to the dimer case, namely for every pair  $1 \leq i < j \leq d$  we define the following operators forming the  $SU(d)$  algebra

$$\hat{X}_{ij} := b_i^\dagger b_j + b_j^\dagger b_i, \quad \hat{Y}_{ij} := i(-b_i^\dagger b_j + b_j^\dagger b_i), \quad \hat{Z}_{ij} := b_i^\dagger b_i - b_j^\dagger b_j = \hat{n}_i - \hat{n}_j. \tag{24}$$

Thus, we have  $3d(d-1)/2$  generators. However, the number of linearly independent generators is  $d^2 - 1$  due to the relations  $\hat{Z}_{ij} + \hat{Z}_{jk} = \hat{Z}_{ik}$ . We will also arrange all generators in a linear order and use a single greek symbol to enumerate the generators, so that the set of linearly independent generators can be written in a uniform way as  $\{\hat{R}_\mu\}_{\mu=1}^{d^2-1}$ . Thus, in full analogy to the dimer case, we can write a general pair interaction as  $\hat{W} = \sum_{\mu \leq \nu} \hat{W}_{\mu,\nu}$ , where

$$\hat{W}_{\mu,\nu} = \omega_{\mu,\nu} (\hat{R}_\mu\hat{R}_\nu + \hat{R}_\nu\hat{R}_\mu) + i\tau_{\mu,\nu} (\hat{R}_\mu\hat{R}_\nu - \hat{R}_\nu\hat{R}_\mu) \tag{25}$$

with  $\omega_{\mu,\nu}$  and  $\tau_{\mu,\nu}$  being some given real constants.

Conjugating  $\hat{R}_\mu$  by a single-particle unitary of the form  $\hat{U} = e^{i\hat{A}}$  with  $\hat{A} = \sum_\mu a_\mu \hat{R}_\mu$  results with the transformation

$$e^{-i\hat{A}} \hat{R}_\nu e^{i\hat{A}} = \sum_{\mu=1}^{d^2-1} C_{\mu,\nu}(A) \hat{R}_\mu, \quad (26)$$

where coefficients  $C_{\mu,\nu}(A)$  are real and can be computed by taking the exponent of the so-called adjoint action  $ad_A(R_\mu) := [A, R_\mu]$ . The coefficient matrix  $[C_{\mu,\nu}(A)]$  describes a rotation in  $d^2 - 1$  dimensions [34]. Thus, we get

$$\hat{U}^\dagger \hat{W} \hat{U} = \sum_{\gamma \leq \delta} (\tilde{\omega}_{\gamma,\delta} (\hat{R}_\gamma \hat{R}_\delta + \hat{R}_\delta \hat{R}_\gamma) + i\tilde{\tau}_{\gamma,\delta} (\hat{R}_\gamma \hat{R}_\delta - \hat{R}_\delta \hat{R}_\gamma)), \quad (27)$$

where

$$\begin{aligned} \tilde{\omega}_{\gamma,\delta} &= \sum_{\mu \leq \nu} \omega_{\mu,\nu} (C_{\gamma,\mu} C_{\delta,\nu} + C_{\gamma,\nu} C_{\delta,\mu}), \\ \tilde{\tau}_{\gamma,\delta} &= \sum_{\mu \leq \nu} \tau_{\mu,\nu} (C_{\gamma,\mu} C_{\delta,\nu} - C_{\gamma,\nu} C_{\delta,\mu}). \end{aligned}$$

Let us remark that the coefficients in equations (26) and (27) depend only on the Lie algebra element  $A$ , not on the particular representation. In particular, the same equations hold for the fermionic systems described in in section 7.

The rest of the proof relies on keeping track of contributions proportional to  $\epsilon$  and  $\epsilon^2$  in expressions of the type  $\langle \Psi_{\epsilon,\sigma} | \hat{R}_\gamma \hat{R}_\delta | \Psi_{\epsilon,\sigma} \rangle$ . The calculation is tedious, but straightforward, thus we move it to appendix A.1. For the sake of completeness, below we write down the explicit expression for  $c_1$  from equation (18). As it turns out, the only contributions to  $c_1$  come from expressions  $\langle \Psi_{\epsilon,\sigma} | \hat{R}_\gamma^2 | \Psi_{\epsilon,\sigma} \rangle$ , where  $\hat{R}_\gamma = \hat{X}_{1,k}$  or  $\hat{R}_\gamma = \hat{Y}_{1,k}$  for some  $k$ . Thus, the result is completely analogous to the formula describing  $c_1$  in the dimer case. Denote by  $\omega_k^{(x)}$  the coefficient  $\tilde{\omega}_{\gamma,\gamma}$  in equation (27) which multiplies  $\hat{X}_{1,k}^2$ . Similarly, define  $\omega_k^{(y)}$  as the coefficient  $\tilde{\omega}_{\gamma,\gamma}$  in equation (27) which multiplies  $\hat{Y}_{1,k}^2$ . Then, as explained in appendix A.1

$$c_1 = 4\sqrt{2N(N-1)} \sum_{k=1}^{d-1} \alpha_k (\omega_k^{(x)} - \omega_k^{(y)}), \quad (28)$$

where  $\{\alpha_k\}_{k=1}^{d-1}$  are parameters of the ansatz state (15).

## 7. Repulsive gradient in the fermionic RDMFT for the singlet sector

In this section we show that an analogous repulsive gradient force is present in systems of interacting spin-1/2 fermions for Hamiltonians preserving the total spin which we assume to be zero by selecting the singlet sector. The fermions occupy  $d$  sites/orbitals with the corresponding annihilation operators  $\{a_{i\uparrow}, a_{i\downarrow}\}_{i=1}^d$ . We start by defining the auxiliary operators  $f_{i,j}$  which will be helpful in carrying the calculations in an almost complete analogy to the bosonic case.

$$f_{i,j} := \sum_{\sigma \in \{\uparrow, \downarrow\}} a_{j\sigma}^\dagger a_{i\sigma}, \quad 1 \leq i \leq j \leq d.$$

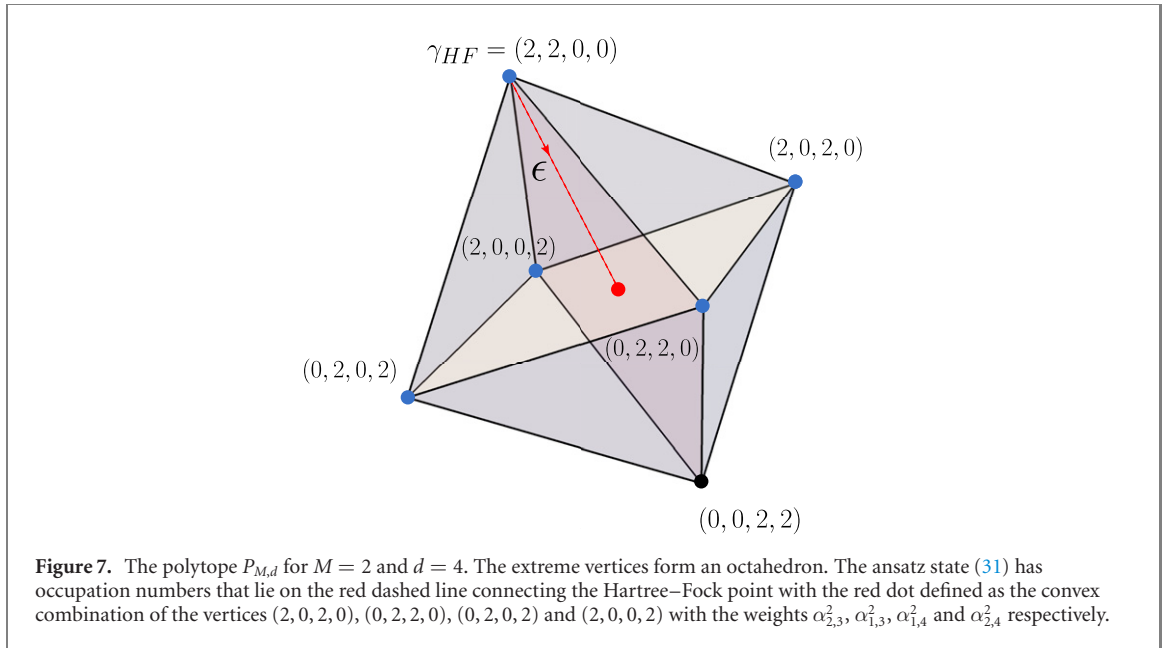
The one-particle reduced density matrix is given by

$$\gamma(|\Psi\rangle)_{ij} = \sum_{\sigma} \langle \Psi | a_{i\sigma}^\dagger a_{j\sigma} | \Psi \rangle = \langle \Psi | f_{i,j}^\dagger | \Psi \rangle.$$

The single-particle operators form an  $SU(d)$  algebra and are represented on the singlet sector as

$$\hat{X}_{i,j}^f := f_{i,j} + f_{i,j}^\dagger, \quad \hat{Y}_{i,j}^f := i(f_{i,j} - f_{i,j}^\dagger), \quad \hat{Z}_{i,j}^f := [f_{i,j}^\dagger, f_{i,j}] = \sum_{\sigma} (\hat{n}_{i\sigma} - \hat{n}_{j\sigma}), \quad (29)$$

where  $i < j$ . As before, for any single-particle unitary  $\hat{U}$  the one-particle reduced density matrices transform equivariantly according to the formula  $\gamma(\hat{U}|\Psi\rangle) = U\gamma(|\Psi\rangle)U^\dagger$ . We also consider the *spectral polytope*  $P_{M,d}$  which is the polytope formed by the spectra of the one-particle reduced density matrices of the singlet



states of  $N = 2M$  fermions occupying  $d$  orbitals (or sites). Importantly, the polytope  $P_{M,d}$  has a tractable description [37]—it is defined only by the Pauli constraints

$$0 \leq n_i \leq 2, \quad \sum_{i=1}^d n_i = 2M,$$

where  $n_i = n_{i\uparrow} + n_{i\downarrow}$  is the number of fermions occupying the  $i$ th orbital. Clearly,  $P_{M,d}$  is not a simplex when  $d > 3$  or  $M > 1$ . It is a convex hull of  $\binom{d}{M}$  vertices whose occupation numbers  $n_i \in \{0, 2\}$  and sum up to  $2M$  (see figure 7).

For  $1 \leq i_1 < i_2 < \dots < i_M \leq d$  we define the corresponding singlet Hartree–Fock state as

$$|i_{1\downarrow}, i_{1\uparrow}, \dots, i_{M\downarrow}, i_{M\uparrow}\rangle := a_{i_{1\downarrow}}^\dagger a_{i_{1\uparrow}}^\dagger \dots a_{i_{M\downarrow}}^\dagger a_{i_{M\uparrow}}^\dagger |\text{vac}\rangle.$$

Any point from the polytope  $P_{M,d}$  can be written as the spectrum of the one-particle reduced density matrix of the following combination of the Hartree–Fock states

$$|\Psi\rangle = \sum_{1 \leq i_1 < i_2 < \dots < i_M \leq d} \alpha_{i_1, \dots, i_M} |i_{1\downarrow}, i_{1\uparrow}, \dots, i_{M\downarrow}, i_{M\uparrow}\rangle.$$

The one-particle reduced density matrix corresponding to the above state is easily seen to be diagonal and to correspond to the convex combination of the polytope’s vertices with the coefficients  $|\alpha_{i_1, \dots, i_M}|^2$ .

Any Hamiltonian conserving the total spin of  $2M$  fermions occupying  $d$  orbitals when truncated to the singlet sector can be written as

$$\hat{H} = \sum_{i < j} \left( t_{ij} \hat{X}_{ij}^f + y_{ij} \hat{Y}_{ij}^f + v_{ij} \hat{Z}_{ij}^f \right) + \hat{W} + \mu \hat{N}, \quad (30)$$

where the interaction term  $\hat{W}$  is an arbitrary linear combination of products of operators (29) with real coefficients (i.e. a real polynomial in  $\hat{X}_{ij}^f$ ,  $\hat{Y}_{ij}^f$  and  $\hat{Z}_{ij}^f$ ). In particular, the Fermi–Hubbard interaction is obtained as  $\hat{W}_{\text{FH}} = \sum_{i < j} V_{ij} \left( \hat{Z}_{ij}^f \right)^2$ ,  $V_{ij} \in \mathbb{R}$ . For the attractive interaction ( $V_{ij} < 0$ ) the Hamiltonian (30) with the Bose–Hubbard interaction and only real hopping terms ( $y_{ij} = 0$ ) extended to the entire Fock space is known to have a unique singlet ground state [36]. For the repulsive interaction, the ground state is a singlet state only at the half-filling ( $N = d$ ) and only when the graph defined by the hopping geometry is bipartite with the two disjoint sets of vertices having the same cardinality. This is satisfied for instance for the chain-geometry. In what follows, we will assume that  $\hat{W}$  is a pair interaction, i.e. it only consists of terms of degree 2 in the operators  $\hat{X}_{ij}^f$ ,  $\hat{Y}_{ij}^f$  and  $\hat{Z}_{ij}^f$ .

We consider the following variational ansatz state.

$$|\Psi_{\epsilon,\sigma}^f\rangle := \frac{1}{\sqrt{1+\epsilon^2}} \left( 1 + \sigma\epsilon \sum_{i=1}^M \sum_{j=M+1}^d \alpha_{ij} a_{j\downarrow}^\dagger a_{i\uparrow}^\dagger a_{i\uparrow} a_{j\downarrow} \right) |1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow\rangle, \quad (31)$$

where  $\sigma \in \{-1, 1\}$  and  $\alpha_{ij}$  are real parameters such that  $\sum_{i=1}^M \sum_{j=M+1}^d \alpha_{ij}^2 = 1$ . The one-particle reduced density matrix of the above ansatz state is diagonal and its entries  $\gamma(|\Psi_{\epsilon,\sigma}^f\rangle)_{k,k} = \langle \Psi_{\epsilon,\sigma} | \hat{n}_{k\downarrow} + \hat{n}_{k\uparrow} | \Psi_{\epsilon,\sigma} \rangle$  read

$$\begin{aligned} \gamma(|\Psi_{\epsilon,\sigma}^f\rangle)_{k,k} &= \frac{2}{1+\epsilon^2} \left( 1 + \epsilon^2 \sum_{j=M+1}^d \alpha_{k,j}^2 \right), \quad 1 \leq k \leq M, \\ \gamma(|\Psi_{\epsilon,\sigma}^f\rangle)_{l,l} &= \frac{2\epsilon^2}{1+\epsilon^2} \sum_{i=1}^M \alpha_{i,l}^2, \quad M < l \leq d. \end{aligned}$$

Clearly, for  $\epsilon = 0$  the ansatz state is a Hartree–Fock state and  $\gamma(|\Psi_{0,\sigma}^f\rangle) = \text{diag}(2, \dots, 2, 0, \dots, 0) =: \gamma_{\text{HF}}$ . If  $\epsilon > 0$ , the parameters  $\alpha_{ij}$  swiwe through all possible directions coming out of the vertex  $\gamma_{\text{HF}}$  and going inside the polytope  $P_{M,d}$  (see figure 7). Consequently, for a fixed choice of the coefficients  $\alpha_{ij}$  and for any single-particle unitary  $\hat{U}$  the parameter  $\epsilon$  parametrizes the shortest path from the point  $U\gamma(|\Psi_{\epsilon,\sigma}^f\rangle)U^\dagger$  to the set of the Hartree–Fock states

$$\Gamma_{\text{HF}} := \{U\gamma_{\text{HF}}U^\dagger \mid U - d \times d \text{ unitary}\}.$$

Alternatively, one may think of the ansatz state  $\hat{U}|\Psi_{\epsilon,\sigma}^f\rangle$  as having the form (31) where the annihilation operators  $\{a_{i\uparrow}, a_{i\downarrow}\}_{i=1}^d$  are replaced by the natural orbital basis annihilators  $\{\tilde{a}_{i\uparrow}, \tilde{a}_{i\downarrow}\}_{i=1}^d$  ordered decreasingly in terms of their corresponding natural occupation numbers. The relevant distance measure is the Hilbert–Schmidt distance from equation (16). In particular, for any single-particle unitary  $\hat{U}$  we have

$$D_\epsilon := \min_{\tilde{\gamma} \in \Gamma_{\text{HF}}} D(\tilde{\gamma}, \gamma(\hat{U}|\Psi_{\epsilon,\sigma}^f\rangle)) = D(\gamma_{\text{HF}}, \gamma(|\Psi_{\epsilon,\sigma}^f\rangle)) = \delta_d \frac{\epsilon^2}{1+\epsilon^2},$$

where  $\delta_d = 2 \left( \sum_{i=1}^M \left( \sum_{j=M+1}^d \alpha_{ij}^2 \right)^2 + \sum_{j=M+1}^d \left( \sum_{i=1}^M \alpha_{ij}^2 \right)^2 \right)^{1/2}$ . By evaluating the sum of the  $d - M$  lowest natural occupation numbers of the state  $\hat{U}|\Psi_{\epsilon,\sigma}^f\rangle$ , we obtain  $N - N_{\text{HF}} = \sum_{i=M+1}^d \gamma(|\Psi_{\epsilon,\sigma}^f\rangle)_{i,i}$ . By the number  $N_{\text{HF}}$  we understand the total number of fermions occupying the  $M$  most occupied natural orbitals. Using the fact that the squares of the coefficients  $\alpha_{ij}$  sum up to one, we get that  $N - N_{\text{HF}} = 2\epsilon^2/(1+\epsilon^2)$  and thus the distance  $D_\epsilon$  has the interpretation as

$$D_\epsilon = \frac{1}{2} \delta_d (N - N_{\text{HF}}).$$

Our goal here is to show that the corresponding fermionic RDMFT functional is upper bounded in the vicinity of the set  $\Gamma_{\text{HF}}$  by

$$\tilde{\mathcal{F}}(U\gamma(|\Psi_{\epsilon,\sigma}^f\rangle)U^\dagger) \leq \left( 1 - \frac{D_\epsilon}{\delta_d} \right) \tilde{\mathcal{F}}(U\gamma_{\text{HF}}U^\dagger) + c_1^f \sigma \sqrt{\frac{D_\epsilon}{\delta_d} \left( 1 - \frac{D_\epsilon}{\delta_d} \right)} + c_2^f \frac{D_\epsilon}{\delta_d}. \quad (32)$$

By the reasoning presented in section 5 this implies that the gradient of the RDMFT functional with respect to the distance  $D_\epsilon$  diverges repulsively as  $-\frac{|c_1^f|}{2\sqrt{\delta_d}} \frac{1}{\sqrt{D_\epsilon}}$  in the vicinity of the Hartree–Fock point  $U\gamma_{\text{HF}}U^\dagger$ . To this end, we show that the following holds

$$\langle \Psi_{\epsilon,\sigma}^f | \hat{U}^\dagger \hat{W} \hat{U} | \Psi_{\epsilon,\sigma}^f \rangle = \frac{\tilde{\mathcal{F}}(U\gamma_{\text{HF}}U^\dagger) + c_1^f \sigma \epsilon + c_2^f \epsilon^2}{1+\epsilon^2} \quad (33)$$

for any pair-interaction  $\hat{W}$ . As before, the parameters  $c_1^f$  and  $c_2^f$  are linear combinations of the appropriate two-body integrals. For the most general interaction given by equation (25) the proof of equation (33) is done using the conjugation formula (27). The details are deferred to appendix A.2. As it turns out, the only contributions to  $c_1^f$  in the lhs of equation (33) come from the expressions  $\langle \Psi_{\epsilon,\sigma} | \hat{R}_\gamma^2 | \Psi_{\epsilon,\sigma} \rangle$ , where  $\hat{R}_\gamma = \hat{X}_{k,l}^f$  or  $\hat{R}_\gamma = \hat{Y}_{k,l}^f$  for  $1 \leq k \leq M$  and  $l > M$ . Thus, the result is completely analogous to the formula describing



$c_1$  in the bosonic case. Denote by  $\omega_{k,l}^{(x)}$  the coefficient  $\tilde{\omega}_{\gamma,\gamma}$  in equation (27) which multiplies  $(\hat{X}_{k,l}^f)^2$ . Similarly, define  $\omega_{k,l}^{(y)}$  as the coefficient  $\tilde{\omega}_{\gamma,\gamma}$  in equation (27) which multiplies  $(\hat{Y}_{k,l}^f)^2$ . Then,

$$c_1^f = 8 \sum_{k=1}^M \sum_{l=M+1}^d \alpha_{k,l} \left( \omega_{k,l}^{(x)} - \omega_{k,l}^{(y)} \right), \quad (34)$$

where  $\alpha_{k,l}$  are the parameters of the ansatz state (31). Thus, the coefficient  $c_1^f$  is generically nonzero.

### 7.1. The form of the approximate functional

For any 1RDM  $\gamma$  the approximate functional is computed from the ansatz state (31) by using the one-particle basis of the natural orbitals corresponding to the decreasingly ordered spectrum (i.e. the natural occupation numbers) of  $\gamma$ . Denote by  $2 \geq \nu_1 \geq \nu_2 \geq \dots \geq \nu_d \geq 0$  the decreasingly ordered natural occupation numbers of  $\gamma$ . Then,  $\gamma$  is a  $d \times d$  matrix  $\gamma = \sum_{i=1}^d \nu_i |\phi_i\rangle\langle\phi_i|$  and we define the operators  $\{\tilde{a}_{i\downarrow}, \tilde{a}_{i\uparrow}\}_{i=1}^d$  so that they refer to the annihilation operators corresponding to the natural orbitals  $\{|\phi_i\rangle\}_{i=1}^d$ . According to the equation (31), after slight modifications, the corresponding ansatz state reads

$$|\Psi_\epsilon^f(\gamma)\rangle := \frac{1}{\sqrt{1+\epsilon^2}} \left( 1 + \epsilon \sum_{i=1}^M \sum_{j=M+1}^d s_{ij} \sqrt{\alpha_{ij}} \tilde{a}_{j\downarrow}^\dagger \tilde{a}_{j\uparrow}^\dagger \tilde{a}_{i\uparrow} \tilde{a}_{i\downarrow} \right) |\text{HF}(\gamma)\rangle, \quad (35)$$

where  $|\text{HF}(\gamma)\rangle := |\phi_{1\downarrow}, \phi_{1\uparrow}, \dots, \phi_{M\downarrow}, \phi_{M\uparrow}\rangle$  and  $\alpha_{ij} \geq 0$ ,  $\sum_{i=1}^M \sum_{j=M+1}^d \alpha_{ij} = 1$  and  $s_{ij} \in \{-1, 1\}$ . The coefficients  $\{\alpha_{ij}\}$  and  $\epsilon$  are determined by demanding the 1RDM of  $|\Psi_{\epsilon,\sigma}^f(\gamma)\rangle$  to have the natural occupation numbers  $\{\nu_i\}_{i=1}^d$ . This leads to

$$\epsilon = \left( \frac{2}{N - N_{\text{HF}}} - 1 \right)^{-1/2}, \quad N_{\text{HF}} := \sum_{i=1}^M \nu_i$$

and to the following linear equations for the  $\alpha_{ij}$ 's:

$$\sum_{j=M+1}^d \alpha_{kj} = 1 - \frac{2 - \nu_k}{\sum_{i=M+1}^d \nu_i}, \quad 1 \leq k \leq M, \quad \sum_{i=1}^M \alpha_{i,l} = \frac{\nu_l}{\sum_{i=M+1}^d \nu_i}, \quad M < l \leq d \quad (36)$$

on top of the normalization. When  $d > 3$ , the solutions of the equation (36) together with the normalization and the non-negativity conditions form a convex set of a positive dimension as the number of unknowns exceeds the number of equations. Let us denote the space of  $\{\alpha_{ij}\}$  satisfying the equation (36) and the conditions  $\alpha_{ij} \geq 0$  by  $\Delta(\bar{\nu})$ . The set  $\Delta(\bar{\nu})$  is a simplex of dimension  $(d-M)(M-1) - M$ . The approximate functional is computed as the minimum of the expectation value  $\langle \Psi_\epsilon^f(\gamma) | \hat{W} | \Psi_\epsilon^f(\gamma) \rangle$  over all solutions from the simplex  $\Delta(\bar{\nu})$ . This ultimately boils down to the following expression

$$G(\gamma) = \min_{\{\alpha_{ij}\} \in \Delta(\bar{\nu})} \left\{ \frac{1}{1+\epsilon^2} \left( I_0 + 2\epsilon \sum_{i=1}^M \sum_{j=M+1}^d s_{ij} \sqrt{\alpha_{ij}} \mathfrak{R} \mathcal{I}_{(i,j)} + \epsilon^2 \sum_{i,k=1}^M \sum_{j,l=M+1}^d s_{ij} s_{kl} \sqrt{\alpha_{ij} \alpha_{kl}} \mathcal{I}_{(i,j),(k,l)} \right) \right\}, \quad (37)$$

where the two-fermion integrals read

$$I_0(\gamma) := \langle \text{HF}(\gamma) | \hat{W} | \text{HF}(\gamma) \rangle, \quad \mathcal{I}_{(i,j)}(\gamma) = \langle \text{HF}(\gamma) | \hat{W} \tilde{a}_{j\downarrow}^\dagger \tilde{a}_{j\uparrow}^\dagger \tilde{a}_{i\uparrow} \tilde{a}_{i\downarrow} | \text{HF}(\gamma) \rangle,$$

$$\mathcal{I}_{(i,j),(k,l)}(\gamma) = \langle \text{HF}(\gamma) | \tilde{a}_{k\downarrow}^\dagger \tilde{a}_{k\uparrow}^\dagger \tilde{a}_{l\uparrow} \tilde{a}_{l\downarrow} \hat{W} \tilde{a}_{j\downarrow}^\dagger \tilde{a}_{j\uparrow}^\dagger \tilde{a}_{i\uparrow} \tilde{a}_{i\downarrow} | \text{HF}(\gamma) \rangle$$

and

$$s_{ij} = -\text{sgn}(\mathfrak{R} \mathcal{I}_{(i,j)}(\gamma)).$$

Thus, the computation of  $G(\gamma)$  requires performing a convex minimization over the set  $\Delta(\bar{\nu})$ . Importantly, the region of applicability of the approximate functional is given by the condition  $N_{\text{HF}} \geq N - 2$ . This is always satisfied when  $N \leq 4$ , however when  $N > 4$  the approximate functional is defined only on a subset of the set of all the  $N$ -representable 1RDMs. Thus, we expect the approximate functional to perform best for low particle numbers.

For the dimer ( $d = 2$ ), there is no convex optimization required and the functional (37) simplifies to

$$G(\gamma) = \frac{1}{1 + \epsilon^2} (I_0 - 2\epsilon |\Re I_{2,2}| + \epsilon^2 I_{(2,2),(2,2)}). \quad (38)$$

In fact, this is the exact RDMFT functional whenever  $\gamma \neq \text{diag}(1, 1)$ . This is because any singlet state of two spin-1/2 fermions with  $d = 2$  whose one-body reduced density matrix is not the identity matrix necessarily has the form (35) up to a relative phase, when written in the natural orbital basis. The expression (38) is equivalent to the expressions for the exact density functional studied in other works [18, 19, 43]. Another simplification arises when  $N = 2$ , regardless of  $d$ . Then, the coefficients  $\{\alpha_{1,j}\}_{j=2}^d$  follow immediately from the equation (36) as

$$\alpha_{1,j} = \frac{\nu_j}{\sum_{i=M+1}^d \nu_i}.$$

## 7.2. Electron transfer in the Fermi–Hubbard model

The repulsive Fermi–Hubbard chain is a good testing ground for our proposed approximate functional. The ground state of  $N$  electrons on the Fermi–Hubbard chain of the length  $d = 2N$  (half-filling) is known to be a singlet state [36]. The Hamiltonian reads

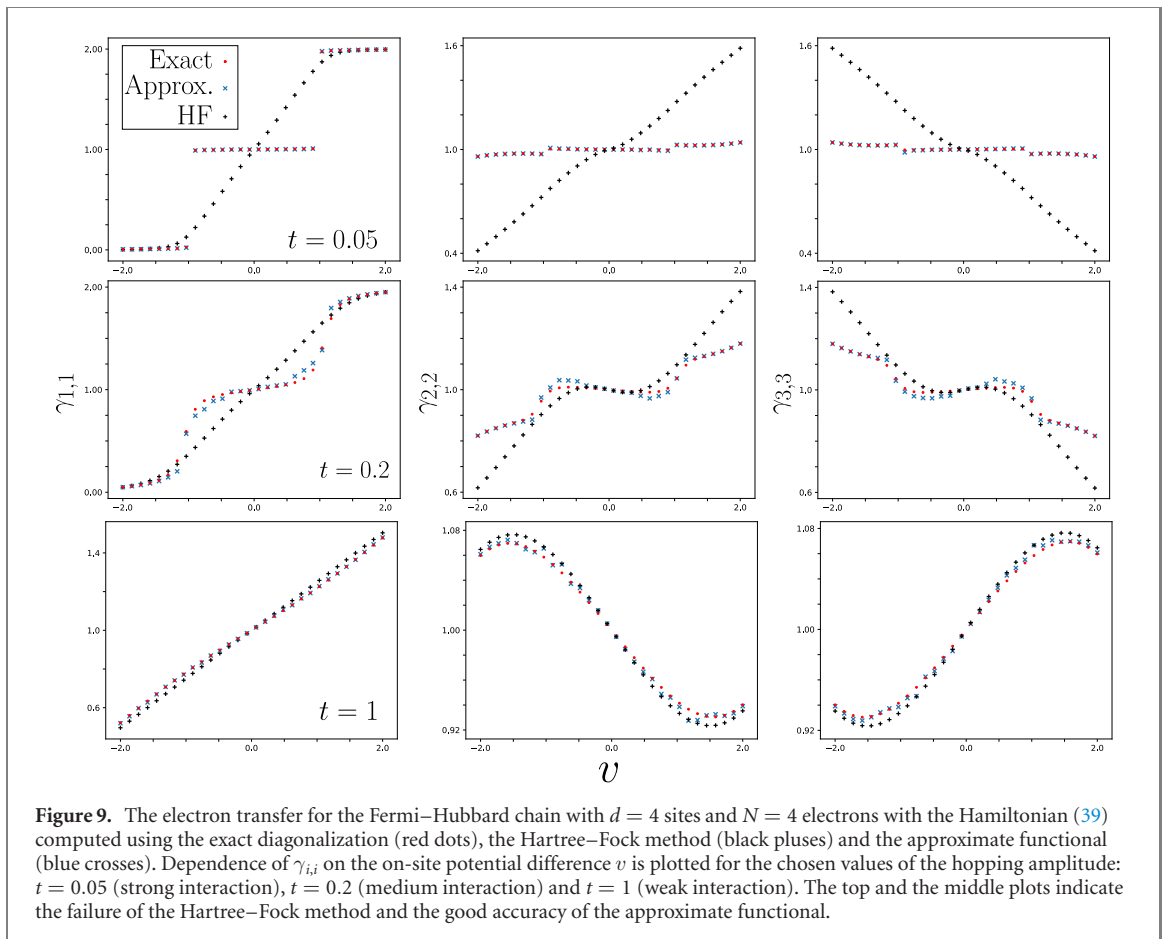
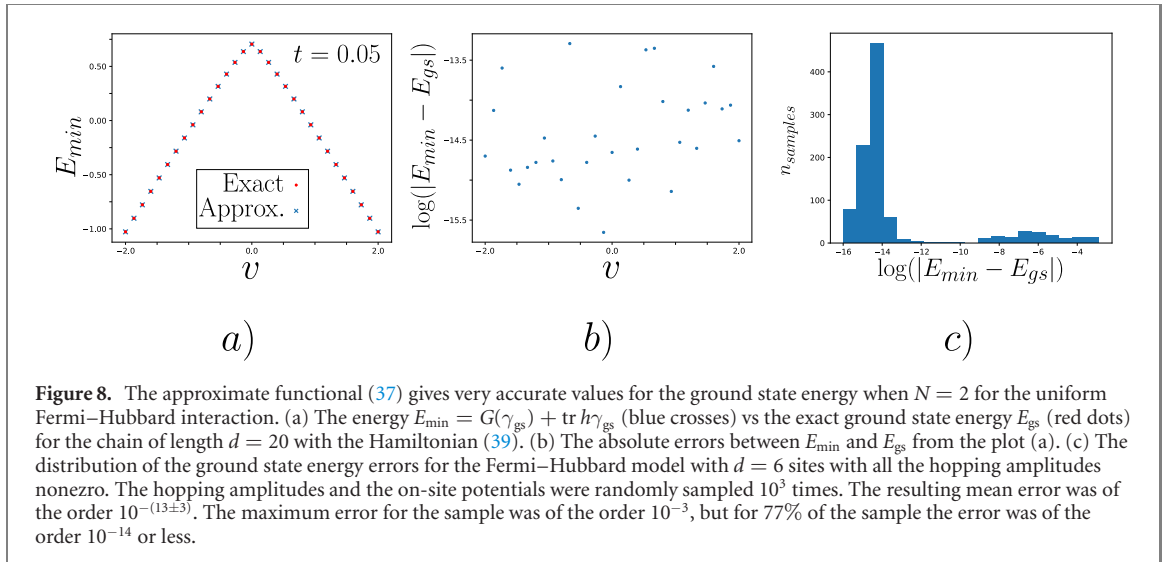
$$\hat{H}_{\text{HF}} = \sum_{i=1}^{d-1} t_{i,i+1} \sum_{\sigma \in \{\uparrow, \downarrow\}} \left( a_{i\sigma}^\dagger a_{(i+1)\sigma} + a_{(i+1)\sigma}^\dagger a_{i\sigma} \right) + \sum_{i=1}^d \sum_{\sigma \in \{\uparrow, \downarrow\}} v_i n_{i\sigma} + \sum_{i=1}^d V_i n_{i\uparrow} n_{i\downarrow}$$

with  $V_i > 0$  for all  $i$ . In the calculations to follow, we will assume an open chain with (i) uniform hopping terms, i.e.  $t_{i,i+1} \equiv -t$  with  $t \geq 0$  for all (i), (ii) uniform interaction strength, i.e.  $V_i \equiv V$  for all (i) and (iii) the on-site potentials  $v_i$  that change by even increments from  $v_1 = 0$  to  $v_d = v$ , i.e.  $v_i = v(i-1)/(d-1)$ . Then, up to a polynomial term in the total number of electrons, the Hamiltonian  $\hat{H}_{\text{HF}}$  is easily seen to have the form

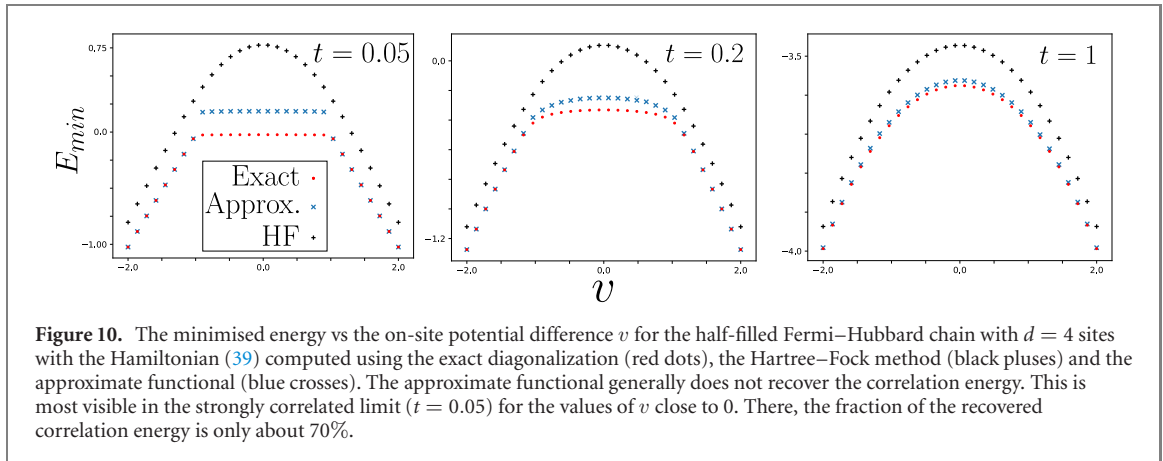
$$\hat{H}_{\text{HF}} = -t \sum_{i=1}^{d-1} \hat{X}_{i,i+1}^f + \frac{v}{2} \sum_{i=1}^{d-1} \frac{i(i-d)}{d-1} \hat{Z}_{i,i+1}^f + \frac{1}{2d} V \sum_{i < j=1}^d \left( \hat{Z}_{ij}^f \right)^2. \quad (39)$$

The exact RDMFT functional in the dimer case ( $d = 2$ ) has been studied in [19] where its performance has been compared with other approximate density functionals. In particular, the authors of [19] studied the electron transfer problem which asks about how the diagonal entries of the 1RDM of the ground state for the Hamiltonian (39) depend on the on-site potential difference  $v$  between the endpoints of the chain. It has been shown that the Müller and Power approximate density functionals [29, 30] give qualitatively incorrect results for the electron transfer, where the failure is most visible in the strong interaction limit ( $t \gg V$ ). Other works show that approximate molecular density functionals generally fail to give the qualitatively correct description of the electron transfer and point out that this is an important challenge to improve the performance of the approximate functionals in this area [32, 33]. In the following part of this section, we show that our proposed approximate density functional does indeed perform well in the electron transfer calculations of small Hubbard chains. In particular, we show that our functional gives accurate values of the ground state energy for the two-electron singlet systems and that it describes the electron transfer with a remarkable accuracy for the four-electron chain with four sites.

By applying the functional (37) to the two-electron ( $N = 2$ ) Fermi–Hubbard chain, we observed numerically that for the uniform interaction and for all the sampled values of the parameters  $t \in [0, 1]$  and  $v \in [-2, 2]$  of the Hamiltonian (39), the approximate functional  $G$  agreed with the exact functional within the numerical accuracy. In other words, the ground state energy error  $|G(\gamma_{\text{gs}}) + \text{tr } h \gamma_{\text{gs}} - E_{\text{gs}}|$  was of the mean order  $10^{-14.5}$  (see figures 8(a) and (b)) which is the same as the error of the  $LU$  diagonalization algorithm which was used in the calculation of  $E_{\text{gs}}$ . Thus, the approximate functional coincides with the exact functional for the Hubbard chain with the Hamiltonian (39) within the numerical accuracy range. We also tested other geometries of the Hubbard model. In particular, for the Fermi–Hubbard model on  $d = 6$  sites with a uniform repulsive interaction we sampled  $10^3$  values of the hopping and the on-site potentials. There, the hopping and the on-site potentials were allowed to be arbitrary and they did not correspond to the chain geometry. The above parameters were sampled from the flat distributions  $t_{i,j} \in [-1, 1]$  and  $v_i \in [-2, 2]$ . For each set of the sampled parameters a minimization of the approximate density functional  $G(\gamma)$  has been performed. This showed the mean energy error of the order  $10^{-(13 \pm 3)}$  (see figure 8(c)). This number is to be compared with the error of the  $LU$  diagonalization algorithm which was used in the calculation of  $E_{\text{gs}}$  which was of the order  $10^{-15}$  or with the typical difference between the lowest and the highest eigenvalue of the random Hamiltonian which was of the order  $10^1$ .



The above good results for the ground state energy of the two-electron systems are due to the fact that the excitations of the tested systems exhibit localization in the natural orbital basis making the superposition of doubly-excited Slater determinants in the ansatz state (31) an accurate guess for the ground state of the system. The ground states of systems of  $N > 2$  electrons generally involve higher excitations and thus the correlation energy is expected to be recovered by our proposed functional only to a small extent. This can be seen already in the half-filled chain of length four ( $N = d = 4$ ) studied in the remaining part of this section. However, as the numerical results show, the doubly-excited ansatz is enough to recover the electron transfer in the ground state of this system.



For  $N = d = 4$  equation (36) leads to the following expressions for the coefficients  $\alpha_{1,3}, \alpha_{1,4}, \alpha_{2,3}, \alpha_{2,4}$  of the ansatz state (35).

$$\frac{2 - \nu_2 - \nu_4}{\nu_3 + \nu_4} \leq \alpha_{1,3} \leq \min \left\{ \frac{\nu_3}{\nu_3 + \nu_4}, \frac{2 - \nu_2}{\nu_3 + \nu_4} \right\}, \quad \alpha_{1,4} = \frac{2 - \nu_2}{\nu_3 + \nu_4} - \alpha_{1,3},$$

$$\alpha_{2,3} = \frac{\nu_3}{\nu_3 + \nu_4} - \alpha_{1,3}, \quad \alpha_{2,4} = \alpha_{1,3} - \frac{2 - \nu_2 - \nu_4}{\nu_3 + \nu_4}.$$

Thus, the minimization in equation (37) is done simply over the interval containing the coefficient  $\alpha_{1,3}$ . The results for the ground state energy and the electron transfer are plotted in figures 9 and 10, where the interaction strength has been fixed to  $V = 1$  due to the fact that the physically meaningful quantity is the ratio  $t/V$ . As we mentioned before, our approximate functional gives a qualitatively correct electron transfer with a remarkably good accuracy, especially in the limit of the strong interaction. On the other hand, the correlation energy is recovered only to a small extent as can be seen in figure 10.

Technical note: the numerical minimization of  $G(\gamma)$  over the set of 1RDMs has been done using the adaptive Nelder–Mead method [38, 39] where the convergence threshold was the difference in function's arguments between iterations which was set to  $10^{-8}$ . Typically, the convergence was attained after about  $10^4$  iterations.

## 8. Summary and discussion

Using the ansatz states (15) and (31) we have constructed universal upper bounds (19) and (32) for the bosonic and fermionic RDMFT functionals in a finite region around of the set of fully condensed (respectively the Hartree–Fock) states for any pair interaction. The upper bounds depend on the (generalised) depletion parameter which is the number of non-condensed bosons,  $N - N_{\text{BEC}}$ , or the number of fermions which do not occupy the  $N/2$  highest occupied natural orbitals,  $N - N_{\text{HF}}$ . Because the upper bound coincides with the exact RDMFT functional when  $N_{\text{BEC}} = N$  (all the bosons condensed) or  $N_{\text{HF}} = N$  (system in a Hartree–Fock state), the gradient of the upper bound is also an upper bound for the gradient of the exact functional with respect to the depletion parameter. The gradient is shown to diverge repulsively when  $N_{\text{BEC}} \rightarrow N$  ( $N_{\text{HF}} \rightarrow N$  respectively). This provides an alternative explanation for the existence of the quantum depletion effect which uses only the very fundamental properties of the quantum theory: the variational principle and the geometry of the set of one-particle reduced density matrices.

We also tested our proposed approximate functional by computing the electron transfer in small Hubbard chains involving  $N = 2$  and  $N = 4$  electrons. We showed that in these systems the functional performs remarkably better than some other widely used approximate functionals, and thus our results shed some light on the possible ways of improving the performance of the density matrix functional methods in the electron transfer computations. In particular, the ansatz states (15) and (31) take into account only double excitations in the natural orbitals. Such ansatz states are expected to work well only in small systems (like the ones tested in this paper) or in extremely inhomogeneous ones where the excitations are effectively localised in a small part of the natural orbital basis. Our functional seems to perform well only in the specific task of computing the electron transfer, as it generally does not recover the correlation energy in the strong interaction limit, as seen in figure 10. A possible way of constructing more accurate functionals would be to superpose Slater determinants with higher numbers of excitations, as one expects that the number of excitations grows linearly with the size of the system. Finally, another practical potential use for

our results could be a testing criterion for finding accurate density functionals—an accurate approximate density functional should reproduce the repulsive gradient.

Let us remark that one of the main obstructions in finding the exact RDMFT functional is that it requires a detailed knowledge of the geometry of the set of single-particle reduced density matrices. The description of this set has been long-recognised as one of the most important fundamental problems in quantum chemistry [44]. The solution of this problem in the late 2000s [37] provided new insights to the RDMFT and it may lead to describing the counterparts of the quantum depletion effect for more general systems of fermions or distinguishable particles using the generalised doubly-excited states [45]. Finally, note that throughout the paper we are implicitly assuming that the Levy's RDMFT functional  $\tilde{\mathcal{F}}$  is actually differentiable at points from the set  $\Gamma_{\text{BEC}} (\Gamma_{\text{HF}})$ . Earlier works studying the exact RDMFT functional for Hubbard dimers [18, 19] suggest that this is a reasonable assumption. However, we are not aware of any mathematically rigorous proof of the fact that the extended functional is differentiable at  $\Gamma_{\text{BEC}} (\Gamma_{\text{HF}})$  for any pair interaction  $\hat{W}$ . Perhaps so-called symplectic slice techniques used in [45] may shed some light on a solution of this problem.

## Acknowledgments

I would like to thank Christian Schilling for introducing me to the subject of the reduced density matrix functional theory and for the feedback at the early stages of the manuscript. I am also very grateful to Jonathan Robbins for providing careful feedback about the manuscript and for many useful discussions.

## Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

## Appendix A. Calculation of $c_1$ for $d \geq 2$

### A.1. The bosonic density functional

Recall the definition of our ansatz state

$$|\Psi_{\epsilon,\delta}\rangle = \mathcal{N}_\epsilon \left( |N_{(1)}\rangle + \sigma \epsilon \sum_{i=1}^{d-1} \alpha_i |(N-2)_{(1)}, 2_{(i+1)}\rangle \right), \quad (\text{A.1})$$

where  $\epsilon \geq 0$ ,  $\alpha_1^2 + \dots + \alpha_{d-1}^2 = 1$ ,  $\sigma = \pm 1$  and the normalization constant reads  $\mathcal{N}_\epsilon = 1/\sqrt{1+\epsilon^2}$ . The basis states are defined in terms of bosonic creation and annihilation operators

$$[b_i, b_j^\dagger] = \delta_{ij}, \quad [b_i^\dagger, b_j^\dagger] = [b_i, b_j] = 0, \quad \hat{n}_i := b_i^\dagger b_i$$

as

$$|(N_1)_{(1)}, (N_2)_{(2)}, \dots, (N_d)_{(d)}\rangle := \frac{1}{\sqrt{N_1! N_2! \dots N_d!}} (b_1^\dagger)^{N_1} (b_2^\dagger)^{N_2} \dots (b_d^\dagger)^{N_d} |\text{vac}\rangle.$$

We also have

$$\begin{aligned} b_i |(N_1)_{(1)}, \dots, (N_i)_{(i)}, \dots, (N_d)_{(d)}\rangle &= \sqrt{N_i} |(N_1)_{(1)}, \dots, (N_i - 1)_{(i)}, \dots, (N_d)_{(d)}\rangle, \\ b_i^\dagger |(N_1)_{(1)}, \dots, (N_i)_{(i)}, \dots, (N_d)_{(d)}\rangle &= \sqrt{N_i + 1} |(N_1)_{(1)}, \dots, (N_i + 1)_{(i)}, \dots, (N_d)_{(d)}\rangle. \end{aligned}$$

Recall also the definitions of generators of the single-particle unitaries.

$$\hat{X}_{ij} := b_i^\dagger b_j + b_j^\dagger b_i, \quad \hat{Y}_{ij} := i \left( -b_i^\dagger b_j + b_j^\dagger b_i \right), \quad \hat{Z}_{ij} := b_i^\dagger b_i - b_j^\dagger b_j = \hat{n}_i - \hat{n}_j. \quad (\text{A.2})$$

We aim to study in detail the expression  $\langle \Psi_{\epsilon,\delta} | \hat{U}^\dagger \hat{W} \hat{U} | \Psi_{\epsilon,\delta} \rangle$  using expansion equation (27), i.e.

$$\langle \Psi_{\epsilon,\delta} | \hat{U}^\dagger \hat{W} \hat{U} | \Psi_{\epsilon,\delta} \rangle = \sum_{\gamma \leq \delta} (\tilde{\omega}_{\gamma,\delta} \langle \Psi_{\epsilon,\delta} | (\hat{R}_\gamma \hat{R}_\delta + \hat{R}_\delta \hat{R}_\gamma) | \Psi_{\epsilon,\delta} \rangle + i \tilde{\tau}_{\gamma,\delta} \langle \Psi_{\epsilon,\delta} | (\hat{R}_\gamma \hat{R}_\delta - \hat{R}_\delta \hat{R}_\gamma) | \Psi_{\epsilon,\delta} \rangle). \quad (\text{A.3})$$

In particular, we want to extract all terms which yield the constant  $c_1$  from equation (18), i.e. terms which are proportional to  $\epsilon\mathcal{N}_\epsilon^2$  in equation (A.3). To this end, we focus on the expression  $(b_l^\dagger b_k \pm b_k^\dagger b_l) |\Psi_{\epsilon,\sigma}\rangle$ ,  $k \leq l$ , as its knowledge determines all the vectors  $\hat{R}_\mu |\Psi_{\epsilon,\delta}\rangle$  and thus all the expectation values in equation (A.3).

If  $k \neq 1$ , then  $(b_l^\dagger b_k \pm b_k^\dagger b_l) |N_{(1)}\rangle = 0$ . Furthermore, if  $k \neq 1$  and  $k \neq i+1$  and  $l \neq i+1$ ,  $(b_l^\dagger b_k \pm b_k^\dagger b_l) |(N-2)_{(1)}, 2_{(i+1)}\rangle = 0$ . Thus, the only nonzero contributions to  $(b_l^\dagger b_k \pm b_k^\dagger b_l) |\Psi_{\epsilon,\sigma}\rangle$  with  $k \neq 1$  come from situations when either  $k = i+1$  or  $l = i+1$  in the equation (A.1). Thus,

$$\begin{aligned} (b_l^\dagger b_k \pm b_k^\dagger b_l) |\Psi_{\epsilon,\sigma}\rangle &= \mathcal{N}_\epsilon \sigma \epsilon \left( \alpha_{k-1} b_l^\dagger b_k |(N-2)_{(1)}, 2_{(k)}\rangle \pm \alpha_{l-1} b_k^\dagger b_l |(N-2)_{(1)}, 2_{(l)}\rangle \right) \\ &= \mathcal{N}_\epsilon \sigma \epsilon \sqrt{2} (\alpha_{k-1} \pm \alpha_{l-1}) |(N-2)_{(1)}, 1_{(k)}, 1_{(l)}\rangle \quad \text{if } k > 1. \end{aligned} \quad (\text{A.4})$$

For  $k = 1$  and  $2 \leq l \leq d$ , we have

$$\begin{aligned} (b_l^\dagger b_1 \pm b_1^\dagger b_l) |N_{(1)}\rangle &= \sqrt{N} |(N-1)_{(1)}, 1_{(l)}\rangle, \\ (b_l^\dagger b_1 \pm b_1^\dagger b_l) |(N-2)_{(1)}, 2_{(i+1)}\rangle &= \sqrt{N-2} |(N-3)_{(1)}, 2_{(i+1)}, 1_{(l)}\rangle \quad \text{if } l \neq i+1, \\ (b_l^\dagger b_1 \pm b_1^\dagger b_l) |(N-2)_{(1)}, 2_{(l)}\rangle &= \sqrt{3(N-2)} |(N-3)_{(1)}, 3_{(l)}\rangle \pm \sqrt{2(N-1)} |(N-1)_{(1)}, 1_{(l)}\rangle. \end{aligned}$$

Putting the above results together, we obtain

$$\begin{aligned} (b_l^\dagger b_1 \pm b_1^\dagger b_l) |\Psi_{\epsilon,\sigma}\rangle &= \mathcal{N}_\epsilon \left( \left( \sqrt{N} \pm \sigma \epsilon \alpha_{l-1} \sqrt{2(N-1)} \right) |(N-1)_{(1)}, 1_{(l)}\rangle \right. \\ &\quad \left. + \sigma \epsilon \alpha_{l-1} \sqrt{3(N-2)} |(N-3)_{(1)}, 3_{(l)}\rangle \right. \\ &\quad \left. + \sigma \epsilon \sqrt{N-2} \sum_{i=1, i \neq l-1}^{d-1} \alpha_i |(N-3)_{(1)}, 2_{(i+1)}, 1_{(l)}\rangle \right) \end{aligned} \quad (\text{A.5})$$

For completeness, let us write down the action of  $\hat{Z}_{k,l}$

$$\begin{aligned} (\hat{n}_1 - \hat{n}_l) |\Psi_{\epsilon,\sigma}\rangle &= \mathcal{N}_\epsilon \left( N |N_{(1)}\rangle + \sigma \epsilon \alpha_{l-1} (N-4) |(N-2)_{(1)}, 2_{(l)}\rangle \right. \\ &\quad \left. + \sigma \epsilon (N-2) \sum_{i=1, i \neq l-1}^{d-1} \alpha_i |(N-2)_{(1)}, 2_{(i+1)}\rangle \right), \\ (\hat{n}_k - \hat{n}_l) |\Psi_{\epsilon,\sigma}\rangle &= 2\sigma \epsilon (\alpha_{k-1} - \alpha_{l-1}) \quad \text{if } k > 1. \end{aligned} \quad (\text{A.6})$$

By a visual inspection of equations (A.4), (A.5), and (A.6) one can see that

- $\langle \Psi_{\epsilon,\sigma} | \hat{Z}_{k,l} \hat{X}_{m,n} | \Psi_{\epsilon,\sigma} \rangle = \langle \Psi_{\epsilon,\sigma} | \hat{Z}_{k,l} \hat{Y}_{m,n} | \Psi_{\epsilon,\sigma} \rangle = 0$  for any  $k, l, m, n$  and the same holds for  $\hat{Y}_{m,n}$ .
- Terms  $\langle \Psi_{\epsilon,\sigma} | \hat{Z}_{k,l} \hat{Z}_{m,n} | \Psi_{\epsilon,\sigma} \rangle$  do not produce any contribution proportional to  $\epsilon\mathcal{N}_\epsilon^2$  for all  $k, l, m, n$ .
- Terms  $\langle \Psi_{\epsilon,\sigma} | \hat{X}_{k,l} \hat{Y}_{m,n} | \Psi_{\epsilon,\sigma} \rangle$  could potentially yield some contribution proportional to  $\epsilon\mathcal{N}_\epsilon^2$  if  $k = m = 1$  and  $n = l$ . However, the relevant term is  $i(\sqrt{N} + \sigma \epsilon \alpha_{l-1} \sqrt{2(N-1)}) (\sqrt{N} - \sigma \epsilon \alpha_{l-1} \sqrt{2(N-1)})$  where the interesting terms cancel out in the end.
- The only terms that yield expressions proportional to  $\epsilon\mathcal{N}_\epsilon^2$  are  $\langle \Psi_{\epsilon,\sigma} | \hat{X}_{1,l}^2 | \Psi_{\epsilon,\sigma} \rangle$  and  $\langle \Psi_{\epsilon,\sigma} | \hat{Y}_{1,l}^2 | \Psi_{\epsilon,\sigma} \rangle$  for  $1 < l \leq d$ . The relevant expressions read  $\pm \mathcal{N}_\epsilon^2 (\sqrt{N} \pm \sigma \epsilon \alpha_{l-1} \sqrt{2(N-1)})^2$  which contribute via  $\pm \sigma \epsilon \mathcal{N}_\epsilon^2 \alpha_{l-1} 2\sqrt{2N(N-1)}$ . This leads directly to equation (34).

## A.2. The fermionic density functional

Recall the ansatz singlet state

$$|\Psi_{\epsilon,\sigma}^f\rangle := \mathcal{N}_\epsilon \left( 1 + \sigma \epsilon \sum_{i=1}^M \sum_{j=M+1}^d \alpha_{i,j} a_{j\downarrow}^\dagger a_{i\uparrow}^\dagger a_{i\downarrow} a_{j\uparrow} \right) |1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow\rangle, \quad (\text{A.7})$$

where  $\mathcal{N}_\epsilon := \frac{1}{\sqrt{1+\epsilon^2}}$ ,  $\epsilon \geq 0$ ,  $\sigma \in \{-1, +1\}$  and  $\sum_{i=1}^M \sum_{j=M+1}^d \alpha_{ij}^2 = 1$ . We have also defined the single-particle operators of the  $SU(d)$ -algebra as

$$\hat{X}_{ij}^f := f_{ij} + f_{ij}^\dagger, \quad \hat{Y}_{ij}^f := i(f_{ij} - f_{ij}^\dagger), \quad \hat{Z}_{ij}^f := [f_{ij}^\dagger, f_{ij}] = \sum_\sigma (\hat{n}_{i\sigma} - \hat{n}_{j\sigma}) \tag{A.8}$$

via the auxiliary ‘ladder’ operators (in fact, these are the root operators of the  $SU(d)$ -algebra)

$$f_{ij} := \sum_{\sigma \in \{\uparrow, \downarrow\}} a_{j\sigma}^\dagger a_{i\sigma}, \quad 1 \leq i \leq j \leq d.$$

We aim to evaluate the expression  $\langle \Psi_{\epsilon, \delta}^f | \hat{U}^\dagger \hat{W} \hat{U} | \Psi_{\epsilon, \delta}^f \rangle$  using the expansion (27). In particular, we want to find the constant  $c_1^f$  in the rhs of equation (33).

One can verify in a straightforward way that the expressions of the form  $\langle \Psi_{\epsilon, \sigma}^f | \hat{Z}_{k,l}^f \hat{X}_{m,n}^f | \Psi_{\epsilon, \sigma}^f \rangle$  and  $\langle \Psi_{\epsilon, \sigma}^f | \hat{Z}_{k,l}^f \hat{Y}_{m,n}^f | \Psi_{\epsilon, \sigma}^f \rangle$  in equation (27) vanish thanks to the special form of the ansatz state. Moreover, the expressions  $\langle \Psi_{\epsilon, \sigma}^f | \hat{Z}_{k,l}^f \hat{Z}_{m,n}^f | \Psi_{\epsilon, \sigma}^f \rangle$  give only terms proportional to  $\mathcal{N}_\epsilon^2$  or to  $\epsilon^2 \mathcal{N}_\epsilon^2$ , thus they do not contribute to the  $c_1^f$ .

Thus, the only expressions which may give terms proportional to  $\epsilon \mathcal{N}_\epsilon^2$  in the equation (27) are of the forms  $\langle \Psi_{\epsilon, \sigma}^f | \hat{X}_{k,l}^f \hat{X}_{m,n}^f | \Psi_{\epsilon, \sigma}^f \rangle$ ,  $\langle \Psi_{\epsilon, \sigma}^f | \hat{X}_{k,l}^f \hat{Y}_{m,n}^f | \Psi_{\epsilon, \sigma}^f \rangle$ ,  $\langle \Psi_{\epsilon, \sigma}^f | \hat{Y}_{k,l}^f \hat{Y}_{m,n}^f | \Psi_{\epsilon, \sigma}^f \rangle$ . In order to evaluate these expressions, it is enough to find the vectors  $(f_{k,l} \pm f_{k,l}^\dagger) | \Psi_{\epsilon, \sigma}^f \rangle$  for  $k < l$ .

- Assume that  $k > M$ . Then,  $f_{k,l} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = 0$  and  $f_{k,l}^\dagger | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = 0$ . We also have

$$\begin{aligned} & (f_{k,l} \pm f_{k,l}^\dagger) a_{j\downarrow}^\dagger a_{j\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle \\ &= \delta_{kj} (| 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow, j\downarrow, l\uparrow \rangle - | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow, j\uparrow, l\downarrow \rangle) \\ & \pm \delta_{lj} (| 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow, k\downarrow, j\uparrow \rangle - | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow, k\uparrow, j\downarrow \rangle). \end{aligned}$$

- Assume that  $l \leq M$ . Then,  $f_{k,l} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = 0$  and  $f_{k,l}^\dagger | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = 0$ . We also have

$$\begin{aligned} & (f_{k,l} \pm f_{k,l}^\dagger) a_{j\downarrow}^\dagger a_{j\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle \\ &= \delta_{li} (| 1\downarrow, 1\uparrow, \dots, k\downarrow, \dots, i\uparrow, \dots, M\downarrow, M\uparrow \rangle - | 1\downarrow, 1\uparrow, \dots, k\uparrow, \dots, i\downarrow, \dots, M\downarrow, M\uparrow \rangle) \\ & \pm \delta_{ki} (| 1\downarrow, 1\uparrow, \dots, i\downarrow, \dots, l\uparrow, \dots, M\downarrow, M\uparrow \rangle - | 1\downarrow, 1\uparrow, \dots, i\uparrow, \dots, l\downarrow, \dots, M\downarrow, M\uparrow \rangle). \end{aligned}$$

- Assume that  $k \leq M$  and  $l > M$ . Then,  $f_{k,l}^\dagger | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = 0$  and

$$f_{k,l} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = | 1\downarrow, 1\uparrow, \dots, k\downarrow, \dots, M\downarrow, M\uparrow, l\uparrow \rangle - | 1\downarrow, 1\uparrow, \dots, k\uparrow, \dots, M\downarrow, M\uparrow, l\downarrow \rangle.$$

We also have

$$\begin{aligned} & f_{k,l}^\dagger a_{j\downarrow}^\dagger a_{j\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = \delta_{ik} \delta_{jl} (| 1\downarrow, 1\uparrow, \dots, i\downarrow, \dots, M\downarrow, M\uparrow, j\uparrow \rangle \\ & \quad - | 1\downarrow, 1\uparrow, \dots, i\uparrow, \dots, M\downarrow, M\uparrow, j\downarrow \rangle), \\ & f_{k,l} a_{j\downarrow}^\dagger a_{j\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow} | 1\downarrow, 1\uparrow, \dots, M\downarrow, M\uparrow \rangle = (1 - \delta_{ik})(1 - \delta_{jl}) a_{j\downarrow}^\dagger a_{j\uparrow}^\dagger a_{i\uparrow} a_{i\downarrow} (| 1\downarrow, 1\uparrow, \dots, k\downarrow, \dots, M\downarrow, M\uparrow, l\uparrow \rangle \\ & \quad - | 1\downarrow, 1\uparrow, \dots, k\uparrow, \dots, M\downarrow, M\uparrow, l\downarrow \rangle). \end{aligned}$$

Thus, by the same reasoning as the one presented in appendix A.1 the only contributions to  $\epsilon \mathcal{N}_\epsilon^2$  in the equation (27) come from the expectation values  $\langle \Psi_{\epsilon, \sigma}^f | (\hat{X}_{k,l})^2 | \Psi_{\epsilon, \sigma}^f \rangle$  and  $\langle \Psi_{\epsilon, \sigma}^f | (\hat{Y}_{k,l})^2 | \Psi_{\epsilon, \sigma}^f \rangle$  when  $k \leq M$  and  $l > M$ . Each of such contributions comes from the expressions  $2\mathcal{N}_\epsilon^2 (1 \pm \alpha_{k,l} \epsilon \sigma)^2$ .

### ORCID iDs

Tomasz Maciążek  <https://orcid.org/0000-0002-2364-1164>

### References

[1] Einstein A 1925 Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung *Sitzungsber. Phys. Math. Kl.* **1** 3

- [2] Bose S 1924 Plancks Gesetz und Lichtquantenhypothese *Z. Phys.* **26** 178
- [3] Anderson M H, Ensher J R, Matthews M R, Wieman C E and Cornell E A 1995 Observation of Bose–Einstein condensation in a dilute atomic vapor *Science* **269** 198
- [4] Davis K B, Mewes M-O, Andrews M R, van Druten N J, Durfee D S, Kurn D M and Ketterle W 1995 Bose–Einstein condensation in a gas of sodium atoms *Phys. Rev. Lett.* **75** 3969
- [5] Griffin A, Snoke D W and Stringari S (ed) 1995 *Bose–Einstein Condensation* (Cambridge: Cambridge University Press)
- [6] Bloch I, Hänsch T W and Esslinger T 1999 Atom laser with a cw output coupler *Phys. Rev. Lett.* **82** 3008
- [7] Bartenstein M, Altmeyer A, Riedl S, Jochim S, Chin C, Denschlag J H and Grimm R 2004 Crossover from a molecular Bose–Einstein condensate to a degenerate Fermi gas *Phys. Rev. Lett.* **92** 120401
- [8] Zwierlein M W, Stan C A, Schunck C H, Raupach S M F, Kerman A J and Ketterle W 2004 Condensation of pairs of fermionic atoms near a Feshbach resonance *Phys. Rev. Lett.* **92** 120403
- [9] Bourdel T, Khaykovich L, Cubizolles J, Zhang J, Chevy F, Teichmann M, Tarruell L, Kokkelmans S J J M F and Salomon C 2004 Experimental study of the BEC–BCS crossover region in lithium 6 *Phys. Rev. Lett.* **93** 050401
- [10] Bogoliubov N N 1947 On the theory of superfluidity *J. Phys. (U.S.S.R.)* **11** 23
- [11] Pitaevskii L P and Stringari S 2003 *Bose–Einstein Condensation* (Oxford: Clarendon)
- [12] Lopes R, Eigen C, Navon N, Clément D, Smith R P and Hadzibabic Z 2017 Quantum depletion of a homogeneous Bose–Einstein condensate *Phys. Rev. Lett.* **119** 190404
- [13] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas *Phys. Rev.* **136** B864
- [14] Lieb E H 1983 Density functionals for coulomb systems *Int. J. Quantum Chem.* **24** 243
- [15] Penrose O and Onsager L 1956 Bose–Einstein condensation and liquid helium *Phys. Rev.* **104** 756
- [16] Gilbert T L 1975 Hohenberg–Kohn theorem for nonlocal external potentials *Phys. Rev. B* **12** 2111
- [17] Levy M 1979 Universal variational functionals of electron densities, first-order density matrices, and natural spin–orbitals and solution of the  $v$ -representability problem *Proc. Natl Acad. Sci.* **76** 6062
- [18] Benavides-Riveros C L, Wolff J, Marques M A L and Schilling C 2020 Reduced density matrix functional theory for bosons *Phys. Rev. Lett.* **124** 180603
- [19] Cohen A J and Mori-Sanchez P 2016 Landscape of an exact energy functional *Phys. Rev. A* **93** 042511
- [20] Mori-Sánchez P and Cohen A J 2018 Exact density functional obtained via the Levy constrained search *J. Phys. Chem. Lett.* **9** 4910
- [21] Schilling C and Schilling R 2019 Diverging exchange force and form of the exact density matrix functional *Phys. Rev. Lett.* **122** 013001
- [22] Schilling C 2018 Communication: relating the pure and ensemble density matrix functional *J. Chem. Phys.* **149** 231102
- [23] Liebert J and Schilling C 2021 Functional theory for Bose–Einstein condensates *Phys. Rev. Res.* **3** 013282
- [24] Schmidt J, Fadel M and Benavides-Riveros C L 2021 Machine learning universal bosonic functionals *Phys. Rev. Res.* **3** L032063
- [25] Girardeau M D 1998 Comment on ‘Particle-number-conserving Bogoliubov method which demonstrates the validity of the time-dependent Gross–Pitaevskii equation for a highly condensed Bose gas’ *Phys. Rev. A* **58** 775
- [26] Seiringer R 2011 The excitation spectrum for weakly interacting bosons *Commun. Math. Phys.* **306** 565
- [27] Seiringer R 2014 Bose gases, Bose–Einstein condensation, and the Bogoliubov approximation *J. Math. Phys.* **55** 075209
- [28] Lieb E H 1981 Variational principle for many-fermion systems *Phys. Rev. Lett.* **46** 457
- [29] Müller A M K 1984 Explicit approximate relation between reduced two- and one-particle density matrices *Phys. Lett. A* **105** 446
- [30] Sharma S, Dewhurst J K, Lathiotakis N N and Gross E K U 2008 Reduced density matrix functional for many-electron systems *Phys. Rev. B* **78** 201103
- [31] Cioslowski J, Pernal K and Ziesche P 2002 Systematic construction of approximate one-matrix functionals for the electron–electron repulsion energy *J. Chem. Phys.* **117** 9560
- [32] Cohen A J, Mori-Sánchez P and Yang W 2012 Challenges for density functional theory *Chem. Rev.* **112** 289–320
- [33] Mori-Sánchez P and Cohen A J 2014 The derivative discontinuity of the exchange–correlation functional *Phys. Chem. Chem. Phys.* **16** 14378
- [34] Baker A 2003 *Matrix Groups: An Introduction to Lie Group Theory* (Berlin: Springer)
- [35] Varshalovich D A, Moskalev A N and Khersonskii V K 1988 *Quantum Theory of Angular Momentum* (Singapore: World Scientific)
- [36] Lieb E H 1989 Two theorems on the Hubbard model *Phys. Rev. Lett.* **62** 1201
- [37] Altunbulak M and Klyachko A 2008 The Pauli principle revisited *Commun. Math. Phys.* **282** 287–322
- [38] Nelder J A and Mead R 1965 A simplex method for function minimization *Comput. J.* **7** 308–13
- [39] Gao F and Han L 2012 Implementing the Nelder–Mead simplex algorithm with adaptive parameters *Comput. Optim. Appl.* **51** 259–77
- [40] Ruskai M B 1970  $N$ -representability problem: particle–hole equivalence *J. Math. Phys.* **11** 3218–24
- [41] Ruskai M B 2007 Connecting  $N$ -representability to Weyl’s problem: the one-particle density matrix for  $N = 3$  and  $R = 6$  *J. Phys. A: Math. Theor.* **40** F961–7
- [42] Borland R E and Dennis K 1972 The conditions on the one-matrix for three-body fermion wavefunctions with one-rank equal to six *J. Phys. B: At. Mol. Phys.* **5** 7–15
- [43] López-Sandoval R and Pastor G M 2002 Density-matrix functional theory of strongly correlated lattice fermions *Phys. Rev. B* **66** 155118
- [44] Coleman A J and Yukalov V I 2000 *Reduced Density Matrices: Coulson’s Challenge* (Berlin: Springer)
- [45] Maciążek T and Tsanov V 2017 Quantum marginals from pure doubly excited states *J. Phys. A: Math. Theor.* **50** 465304