Computer Physics Communications 182 (2011) 1235–1244

Contents lists available at ScienceDirect

Computer Physics Communications

www.elsevier.com/locate/cpc

A primal–dual semidefinite programming algorithm tailored to the variational determination of the two-body density matrix

Brecht Verstichel [a,*], Helen van Aggelen [b], Dimitri Van Neck [a], Patrick Bultinck [b], Stijn De Baerdemacker [c]

[a] Ghent University, Center for Molecular Modeling, Technologiepark 903, 9052 Zwijnaarde, Belgium
[b] Ghent University, Department of Inorganic and Physical Chemistry, Krijgslaan 281 (S3), B-9000 Gent, Belgium
[c] Department of Physics and Astronomy, Proeftuinstraat 86, 9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 17 December 2010
Received in revised form 4 February 2011
Accepted 15 February 2011
Available online 26 February 2011

Keywords:
Density matrix
Variational
Semidefinite programming

ABSTRACT

The quantum many-body problem can be rephrased as a variational determination of the two-body reduced density matrix, subject to a set of N-representability constraints. The mathematical problem has the form of a semidefinite program. We adapt a standard primal–dual interior point algorithm in order to exploit the specific structure of the physical problem. In particular the matrix-vector product can be calculated very efficiently. We have applied the proposed algorithm to a pairing-type Hamiltonian and studied the computational aspects of the method. The standard N-representability conditions perform very well for this problem.

1. Introduction

It was realized in the 1950s [1,2] that the energy of a quantum many-body system can be expressed in terms of the two-body reduced density matrix (2DM), when only one- and two-body interactions are present. This insight led to the idea of variationally determining the 2DM by minimizing the energy, henceforth referred to as the v2DM method. Once the 2DM is known, all other physical properties that can be expressed as one- or two-body operators can be extracted. In this way the 2DM effectively replaces the wave function and we have "quantum mechanics without wave functions" [3]. Early attempts, however, produced unrealistic results [4] and it was soon realized [5] that non-trivial constraints are needed to ensure that the 2DM is derivable from a physical wave function. These constraints were called N-representability conditions by Coleman [6], and Garrod and Percus [7] derived two such conditions, the so-called Q and G conditions, which can be expressed as matrix-positivity constraints. With these constraints there were some attempts, some of which quite successful, to solve this problem numerically in the 1970s [8–11]. However the method was soon abandoned because of the computational cost. Interest in the subject was renewed at the beginning of this century, when Nakata et al. [12] and Mazziotti [13] realized that the v2DM problem can be formulated as a semidefinite program (SDP) for which general-purpose primal–dual SDP solvers can be used [14], and they calculated the ground-state properties of small atoms and molecules. Primal–dual interior point methods are the "Rolls Royce" of SDP algorithms, having several appealing features, but they require a lot of storage and are computationally expensive. These early calculations were therefore limited to small systems (minimal basis set). Mazziotti [15] then developed an algorithm that transforms the SDP into a non-linear optimization program solved by a gradient-only method. This reduced the cost of the storage and the basic floating point operations, but at the cost of these nice convergence properties of the interior point methods. In this paper we adapt a standard primal–dual interior point algorithm [16] to the specific case of v2DM, in an attempt to retain the nice convergence properties, while reducing the storage and computational cost. In Section 2 we present an introduction to the theory of N-representability, v2DM and some mathematical properties of the constraints. In Section 3 we discuss the representation of the problem as a primal–dual semidefinite program, and introduce the method we use to solve it. Then we apply the algorithm to a BCS (Bardeen–Cooper–Shrieffer) [17] or pairing-type Hamiltonian in Section 4 and present the physical results and computational aspects. A summary is provided in Section 5.

2. Variational density matrix determination

When only two-body interactions are present, the Hamiltonian of a physical system can be written as:

$$\hat{H} = \sum_{\alpha\gamma} t_{\alpha\gamma} a_\alpha^\dagger a_\gamma + \frac{1}{4} \sum_{\alpha\beta\gamma\delta} V_{\alpha\beta;\gamma\delta} a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma, \tag{1}$$

* Corresponding author.
E-mail address: brecht.verstichel@ugent.be (B. Verstichel).

0010-4655/$ – see front matter
doi:10.1016/j.cpc.2011.02.005

using second quantized notation where $a_\alpha^\dagger$ (a_α) creates (annihilates) a fermion in a single-particle (sp) state α [18]. The expectation value of the energy in an arbitrary N-particle state $|\Psi^N\rangle$ can be expressed in terms of the 2DM only,

$$E(\Gamma) = \operatorname{Tr}\Gamma H^{(2)} = \sum_{\alpha<\beta;\,\gamma<\delta} \Gamma_{\alpha\beta;\gamma\delta} H^{(2)}_{\alpha\beta;\gamma\delta}, \tag{2}$$

with the 2DM defined as:

$$\Gamma_{\alpha\beta;\gamma\delta} = \langle\Psi^N|a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma|\Psi^N\rangle, \tag{3}$$

and the reduced two-particle Hamiltonian,

$$H^{(2)}_{\alpha\beta;\gamma\delta} = \frac{1}{N-1}(\delta_{\alpha\gamma}t_{\beta\delta} - \delta_{\alpha\delta}t_{\beta\gamma} - \delta_{\beta\gamma}t_{\alpha\delta} + \delta_{\beta\delta}t_{\alpha\gamma}) + V_{\alpha\beta;\gamma\delta}. \tag{4}$$

The idea of v2DM is to determine the ground-state energy and other two- or one-body properties by minimizing the energy (2) using the 2DM as a variable. The 2DM is a much more compact object than the wave function because one keeps the dimension of two-particle (tp) space, no matter how many particles are involved. The problem is that there is no straightforward way to know whether an arbitrary matrix in tp-space Γ is derivable from a physical wave function as in Eq. (3). Actually, it is sufficient that Γ is derivable from an ensemble of N-particle wave functions, and this is called the N-representability problem [6]. Some obvious necessary N-representability constraints are apparent from the definition (3):

$$\text{trace condition} \quad \operatorname{Tr}\Gamma = \sum_{\alpha<\beta}\Gamma_{\alpha\beta;\alpha\beta} = \frac{N(N-1)}{2}, \tag{5}$$

$$\text{antisymmetry} \quad \Gamma_{\alpha\beta;\gamma\delta} = -\Gamma_{\beta\alpha;\gamma\delta} = -\Gamma_{\alpha\beta;\delta\gamma} = \Gamma_{\beta\alpha;\delta\gamma}, \tag{6}$$

$$\text{Hermiticity} \quad \Gamma_{\alpha\beta;\gamma\delta} = \Gamma_{\gamma\delta;\alpha\beta}, \tag{7}$$

but it turns out that there are many non-trivial constraints needed to ensure that a 2DM is physical.

2.1. N-representability

The necessary and sufficient conditions for N-representability are formally known [19]. A tp-matrix is N-representable if and only if, for every two-body Hamiltonian $\hat{H}_\nu$, the following inequality is satisfied:

$$\operatorname{Tr} H_\nu^{(2)}\Gamma \geqslant E_0(H_\nu), \tag{8}$$

where $E_0(H_\nu)$ is the exact N-particle ground-state energy corresponding to the Hamiltonian. This is hardly a practical approach, as one needs to know the ground-state energy of every two-body Hamiltonian. Therefore one resorts to certain classes of Hamiltonians for which a lower bound to the ground-state energy is known. A Hamiltonian class that is used as necessary constraint is

$$\langle\Psi^N|B^\dagger B|\Psi^N\rangle \geqslant 0, \tag{9}$$

which leads to positivity conditions of linear matrix maps of the 2DM. If we want (9) to be restricted to tp-space there are three possible forms of the operator $B^\dagger$, leading to three conditions on the density matrix:

a. $B^\dagger = \sum_{\alpha\beta} p_{\alpha\beta} a_\alpha^\dagger a_\beta^\dagger$ leads to the trivial $\mathcal{P}$-condition:

$$\mathcal{P}(\Gamma) = \Gamma \succcurlyeq 0, \tag{10}$$

which imposes positive semidefiniteness on the 2DM.

b. $B^\dagger = \sum_{\alpha\beta} q_{\alpha\beta} a_\alpha a_\beta$ leads to the $\mathcal{Q}$-condition:

$$\mathcal{Q}(\Gamma) \succcurlyeq 0, \tag{11}$$

where the linear matrix map $\mathcal{Q}$ is defined as

$$\begin{aligned}\mathcal{Q}(\Gamma)_{\alpha\beta;\gamma\delta} &= \langle\Psi^N|a_\alpha a_\beta a_\delta^\dagger a_\gamma^\dagger|\Psi^N\rangle \\ &= \Gamma_{\alpha\beta;\gamma\delta} + (\delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\delta})\frac{\bar{\bar{\Gamma}}}{N(N-1)} \\ &\quad - \delta_{\alpha\gamma}\rho_{\beta\delta} + \delta_{\alpha\delta}\rho_{\beta\gamma} + \delta_{\beta\gamma}\rho_{\alpha\delta} - \delta_{\beta\delta}\rho_{\alpha\gamma},\end{aligned} \tag{12}$$

with

$$\rho_{\alpha\gamma} = \frac{1}{N-1}\bar{\Gamma}_{\alpha\gamma} = \frac{1}{N-1}\sum_\beta \Gamma_{\alpha\beta;\gamma\beta}, \tag{13}$$

the one-body reduced density matrix (1DM), and with

$$\bar{\bar{\Gamma}} = \sum_{\alpha\beta}\Gamma_{\alpha\beta;\alpha\beta}, \tag{14}$$

the unrestricted trace of the 2DM.

c. $B^\dagger = \sum_{\alpha\beta} g_{\alpha\beta} a_\alpha^\dagger a_\beta$ which leads to the $\mathcal{G}$-condition:

$$\mathcal{G}(\Gamma) \succcurlyeq 0, \tag{15}$$

with the linear matrix map $\mathcal{G}$ defined as

$$\begin{aligned}\mathcal{G}(\Gamma)_{\alpha\beta;\gamma\delta} &= \langle\Psi^N|a_\alpha^\dagger a_\beta a_\delta^\dagger a_\gamma|\Psi^N\rangle \\ &= \delta_{\beta\delta}\rho_{\alpha\gamma} - \Gamma_{\alpha\delta;\gamma\beta}.\end{aligned} \tag{16}$$

Another Hamiltonian class for which a lower bound to the ground-state energy is known gives rise to the so-called three-index conditions:

$$\langle\Psi^N|\{B^\dagger, B\}|\Psi^N\rangle \geqslant 0. \tag{17}$$

In this article we will use two conditions that come from Eq. (17).

d. $B^\dagger = \sum_{\alpha\beta\gamma} t_{\alpha\beta\gamma} a_\alpha^\dagger a_\beta^\dagger a_\gamma^\dagger$ leads to the $\mathcal{T}_1$-condition:

$$\mathcal{T}_1(\Gamma) \succcurlyeq 0, \tag{18}$$

with the linear matrix map $\mathcal{T}_1$ defined as

$$\begin{aligned}&\mathcal{T}_1(\Gamma)_{\alpha\beta\gamma;\delta\epsilon\zeta} \\ &= \langle\Psi^N|a_\alpha^\dagger a_\beta^\dagger a_\gamma^\dagger a_\zeta a_\epsilon a_\delta + a_\alpha a_\beta a_\gamma a_\zeta^\dagger a_\epsilon^\dagger a_\delta^\dagger|\Psi^N\rangle \\ &= (\delta_{\gamma\zeta}\delta_{\beta\epsilon}\delta_{\alpha\delta} - \delta_{\gamma\epsilon}\delta_{\alpha\delta}\delta_{\beta\zeta} + \delta_{\alpha\zeta}\delta_{\gamma\epsilon}\delta_{\beta\delta} - \delta_{\gamma\zeta}\delta_{\alpha\epsilon}\delta_{\beta\delta} \\ &\quad + \delta_{\beta\zeta}\delta_{\alpha\epsilon}\delta_{\gamma\delta} - \delta_{\alpha\zeta}\delta_{\beta\epsilon}\delta_{\gamma\delta})\frac{\bar{\bar{\Gamma}}}{N(N-1)} \\ &\quad - (\delta_{\gamma\zeta}\delta_{\beta\epsilon} - \delta_{\beta\zeta}\delta_{\gamma\epsilon})\rho_{\alpha\delta} + (\delta_{\gamma\zeta}\delta_{\alpha\epsilon} - \delta_{\alpha\zeta}\delta_{\gamma\epsilon})\rho_{\beta\delta} \\ &\quad - (\delta_{\beta\zeta}\delta_{\alpha\epsilon} - \delta_{\alpha\zeta}\delta_{\beta\epsilon})\rho_{\gamma\delta} + (\delta_{\gamma\zeta}\delta_{\beta\delta} - \delta_{\beta\zeta}\delta_{\gamma\delta})\rho_{\alpha\epsilon} \\ &\quad - (\delta_{\gamma\zeta}\delta_{\alpha\delta} - \delta_{\alpha\zeta}\delta_{\gamma\delta})\rho_{\epsilon\beta} + (\delta_{\beta\zeta}\delta_{\alpha\delta} - \delta_{\alpha\zeta}\delta_{\beta\delta})\rho_{\gamma\epsilon} \\ &\quad - (\delta_{\beta\delta}\delta_{\gamma\epsilon} - \delta_{\beta\epsilon}\delta_{\gamma\delta})\rho_{\alpha\zeta} + (\delta_{\gamma\epsilon}\delta_{\alpha\delta} - \delta_{\alpha\epsilon}\delta_{\gamma\delta})\rho_{\beta\zeta} \\ &\quad - (\delta_{\beta\epsilon}\delta_{\alpha\delta} - \delta_{\alpha\epsilon}\delta_{\beta\delta})\rho_{\gamma\zeta} + \delta_{\gamma\zeta}\Gamma_{\alpha\beta;\delta\epsilon} - \delta_{\beta\zeta}\Gamma_{\alpha\gamma;\delta\epsilon} \\ &\quad + \delta_{\alpha\zeta}\Gamma_{\beta\gamma;\delta\epsilon} - \delta_{\gamma\epsilon}\Gamma_{\alpha\beta;\delta\zeta} + \delta_{\beta\epsilon}\Gamma_{\alpha\gamma;\delta\zeta} - \delta_{\alpha\epsilon}\Gamma_{\beta\gamma;\delta\zeta} \\ &\quad + \delta_{\gamma\delta}\Gamma_{\alpha\beta;\epsilon\zeta} - \delta_{\beta\delta}\Gamma_{\alpha\gamma;\epsilon\zeta} + \delta_{\alpha\delta}\Gamma_{\beta\gamma;\epsilon\zeta}.\end{aligned} \tag{19}$$

e. $B^\dagger = \sum_{\alpha\beta\gamma} t_{\alpha\beta\gamma} a_\alpha^\dagger a_\beta^\dagger a_\gamma$ leads to the $\mathcal{T}_2$-condition:

$$\mathcal{T}_2(\Gamma) \succcurlyeq 0, \tag{20}$$

with the linear matrix map $\mathcal{T}_2$ defined as

$$\begin{aligned}\mathcal{T}_2(\Gamma)_{\alpha\beta\gamma;\delta\epsilon\zeta} &= \langle \Psi^N | a_\alpha^\dagger a_\beta^\dagger a_\gamma a_\zeta^\dagger a_\epsilon a_\delta + a_\gamma^\dagger a_\beta a_\alpha a_\delta^\dagger a_\epsilon^\dagger a_\zeta | \Psi^N \rangle \\ &= (\delta_{\alpha\delta}\delta_{\beta\epsilon} - \delta_{\alpha\epsilon}\delta_{\beta\delta})\rho_{\gamma\zeta} + \delta_{\gamma\zeta}\Gamma_{\alpha\beta;\delta\epsilon} - \delta_{\alpha\delta}\Gamma_{\gamma\epsilon;\zeta\beta} \\ &\quad + \delta_{\beta\delta}\Gamma_{\gamma\epsilon;\zeta\alpha} + \delta_{\alpha\epsilon}\Gamma_{\gamma\delta;\zeta\beta} - \delta_{\beta\epsilon}\Gamma_{\gamma\delta;\zeta\alpha}. \end{aligned} \tag{21}$$

The optimization problem that we have to solve can be summarized as:

$$\min_\Gamma \operatorname{Tr} \Gamma H^{(2)}, \tag{22}$$

under the condition that

$$\operatorname{Tr} \Gamma = \frac{N(N-1)}{2}, \tag{23}$$

$$\mathcal{L}(\Gamma) \succcurlyeq 0 \quad \forall \mathcal{L} \in \{\mathcal{P}, \mathcal{Q}, \mathcal{G}, \mathcal{T}_1, \mathcal{T}_2\}. \tag{24}$$

2.2. Hermitian adjoint maps

For the following it is useful to introduce the Hermitian adjoints of matrix maps introduced in the previous section. The Hermitian adjoint maps are defined through:

$$\operatorname{Tr} \mathcal{L}_i(\Gamma) A = \operatorname{Tr} \mathcal{L}_i^\dagger(A)\Gamma, \tag{25}$$

in which A is a matrix of the same dimension as the image of the map $\mathcal{L}_i$ in question (*e.g.* a three-particle matrix for a $\mathcal{T}_1$ map, *etc.*), and the traces sum over the appropriate indices. The $\mathcal{P}$ and $\mathcal{Q}$ maps are Hermitian, so they are identical to their Hermitian adjoints. For the other maps however this is not the case. Using Eq. (25) the Hermitian adjoint of the $\mathcal{G}$ map can be shown to have the form:

$$\begin{aligned}\mathcal{G}^\dagger(A)_{\alpha\beta;\gamma\delta} &= \frac{1}{N-1}[\delta_{\beta\delta}\bar{A}_{\alpha\gamma} - \delta_{\alpha\delta}\bar{A}_{\beta\gamma} - \delta_{\beta\gamma}\bar{A}_{\alpha\delta} + \delta_{\alpha\gamma}\bar{A}_{\beta\delta}] \\ &\quad - A_{\alpha\delta;\gamma\beta} + A_{\beta\delta;\gamma\alpha} + A_{\alpha\gamma;\delta\beta} - A_{\beta\gamma;\delta\alpha}, \end{aligned} \tag{26}$$

in which a particle–hole matrix A is mapped onto tp-matrix space and

$$\bar{A}_{\alpha\gamma} = \sum_\lambda A_{\alpha\lambda;\gamma\lambda}. \tag{27}$$

The $\mathcal{T}_1$-operator maps a tp-matrix onto a three-particle matrix, so its Hermitian adjoint has to map a three-particle matrix A onto tp-space. Solving Eq. (25) with $\mathcal{L} = \mathcal{T}_1$ one finds that:

$$\begin{aligned}\mathcal{T}_1^\dagger(A)_{\alpha\beta;\gamma\delta} &= \frac{2}{N(N-1)}(\delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma})\operatorname{Tr} A + \bar{A}_{\alpha\beta;\gamma\delta} \\ &\quad - \frac{1}{2(N-1)}[\delta_{\beta\delta}\bar{\bar{A}}_{\alpha\gamma} - \delta_{\alpha\delta}\bar{\bar{A}}_{\beta\gamma} - \delta_{\beta\gamma}\bar{\bar{A}}_{\alpha\delta} + \delta_{\alpha\gamma}\bar{\bar{A}}_{\beta\delta}], \end{aligned} \tag{28}$$

with

$$\bar{A}_{\alpha\beta;\gamma\delta} = \sum_\lambda A_{\alpha\beta\lambda;\gamma\delta\lambda}, \tag{29}$$

$$\bar{\bar{A}}_{\alpha\gamma} = \sum_{\lambda\kappa} A_{\alpha\lambda\kappa;\gamma\lambda\kappa}. \tag{30}$$

In the same way one can derive for $\mathcal{L} = \mathcal{T}_2$ that

$$\begin{aligned}\mathcal{T}_2^\dagger(A)_{\alpha\beta;\gamma\delta} &= \frac{1}{2(N-1)}[\delta_{\beta\delta}\tilde{\bar{A}}_{\alpha\gamma} - \delta_{\alpha\delta}\tilde{\bar{A}}_{\beta\gamma} - \delta_{\beta\gamma}\tilde{\bar{A}}_{\alpha\delta} + \delta_{\alpha\gamma}\tilde{\bar{A}}_{\beta\delta}] + \bar{A}_{\alpha\beta;\gamma\delta} \\ &\quad - [\tilde{A}_{\delta\alpha;\beta\gamma} - \tilde{A}_{\delta\beta;\alpha\gamma} - \tilde{A}_{\gamma\alpha;\beta\delta} + \tilde{A}_{\gamma\beta;\alpha\delta}], \end{aligned} \tag{31}$$

with this time A a matrix on two-particle–one-hole space and

$$\tilde{\bar{A}}_{\alpha\gamma} = \sum_{\lambda\kappa} A_{\lambda\kappa\alpha;\lambda\kappa\gamma}, \tag{32}$$

$$\bar{A}_{\alpha\beta;\gamma\delta} = \sum_\lambda A_{\alpha\beta\lambda;\gamma\delta\lambda}, \tag{33}$$

$$\tilde{A}_{\alpha\beta;\gamma\delta} = \sum_\lambda A_{\lambda\alpha\beta;\lambda\gamma\delta}. \tag{34}$$

3. Primal–dual semidefinite program

The variational method described in the previous section can be formulated as a primal–dual semidefinite program. A general 2DM, describing an N-particle system can be expanded in an arbitrary orthogonal basis $\{f^i\}$ of traceless matrix space as

$$\Gamma = \frac{N(N-1)}{M(M-1)}\mathbb{1}_{\text{tp}} + \sum_i \gamma_i f^i, \tag{35}$$

with M the dimension of single-particle (sp) space, and the unit matrix on tp-space defined as

$$(\mathbb{1}_{\text{tp}})_{\alpha\beta;\gamma\delta} = \delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma}. \tag{36}$$

The energy of the system can be written as a function of the γ's as

$$\operatorname{Tr} \Gamma H^{(2)} = \frac{N(N-1)}{M(M-1)} \operatorname{Tr} H^{(2)} + \sum_i \gamma_i \operatorname{Tr} H^{(2)} f^i. \tag{37}$$

Because the necessary N-representability conditions can be written as linear homogeneous matrix maps of Γ, we can also write them as a function of the γ's:

$$\mathcal{L}(\Gamma) = \frac{N(N-1)}{M(M-1)}\mathcal{L}(\mathbb{1}_{\text{tp}}) + \sum_i \gamma_i \mathcal{L}(f^i) \succcurlyeq 0. \tag{38}$$

If we now consider the direct sum of the linear spaces associated with the maps and define the block matrices:

$$u^0 = \frac{N(N-1)}{M(M-1)} \bigoplus_k \mathcal{L}_k(\mathbb{1}_{\text{tp}}) \quad \text{and} \quad u^i = \bigoplus_k \mathcal{L}_k(f^i), \tag{39}$$

then we can formulate v2DM as a standard dual semidefinite program [14]:

$$\min_\gamma \gamma^T h \quad \text{on condition that} \quad Z = u^0 + \sum_i \gamma_i u^i \succcurlyeq 0, \tag{40}$$

in which $h^i = \operatorname{Tr} H^{(2)} f^i$. The primal problem corresponding to (40) optimizes the matrix variable X, the problem being defined as:

$$\max_X (-\operatorname{Tr} X u^0) \quad \text{on condition that} \quad \operatorname{Tr} X u^i = h^i \quad \text{and} \quad X \succcurlyeq 0. \tag{41}$$

X will be a block matrix because the u-matrices are block matrices. The primal–dual gap η is defined as the difference between the primal and the dual cost function for a certain primal–dual point (X, Z):

$$\eta = \gamma^T h + \operatorname{Tr} u^0 X = \sum_i \gamma_i \operatorname{Tr} X u^i + \operatorname{Tr} X u^0 = \operatorname{Tr} XZ \geqslant 0, \tag{42}$$

as X and Z are positive semidefinite matrices. We can see that the smallest value of η will be reached when both the primal and the dual problem are optimal. It can be proven that if the primal and the dual problem are both strictly feasible, then the primal–dual

gap vanishes at their solution [14]. This means that the primal–dual gap can be used as a convergence criterion for the algorithm. Even better, at any point during the optimization, the error on the current value is limited from above by the primal–dual gap. Note that in our previous implementation [20], a dual-only algorithm was used. (The assignation of a problem as "primal" or "dual" is largely matter of convention. In [21], *e.g.*, the opposite convention is used.) The properties of the present primal–dual method can lead to a serious reduction in computation time since we can stop the algorithm at a prescribed error estimate.

3.1. Equations of motion

There are several known methods to solve a semidefinite program. In this paper a path-following interior point method is used. The central path is defined as the set of primal–dual points for which

$$XZ = \frac{\eta}{n}\mathbb{1}_{\text{sup}}, \tag{43}$$

with n the total dimension of the X and Z matrices and $\mathbb{1}_{\text{sup}}$ the direct sum of the unity matrices on the different constraint spaces:

$$\mathbb{1}_{\text{sup}} = \bigoplus_k \mathbb{1}_k. \tag{44}$$

In the path-following algorithm [16] we try to follow the central path, reducing the primal–dual gap along the way. Consider a primal–dual point (X, Z) on the central path with primal–dual gap η. We want to know what is the primal–dual point on the central path with primal–dual gap scaled down with a factor ν. Rephrasing, we are looking for the (Δ_X, Δ_Z) that solve:

$$(X + \Delta_X)(Z + \Delta_Z) = \frac{\nu\eta}{n}\mathbb{1}_{\text{sup}}. \tag{45}$$

There are several ways to symmetrize these equations. Using the method proposed by [16], two equivalent equations (called the dual and the primal) are obtained, *i.e.* one has to solve the equations

$$\text{(dual):}\quad \Delta_X + D^{-1}\Delta_Z D^{-1} = \frac{\nu\eta}{n}Z^{-1} - X, \tag{46}$$

$$\text{(primal):}\quad \Delta_Z + D\Delta_X D = \frac{\nu\eta}{n}X^{-1} - Z, \tag{47}$$

and under the condition that:

$$\operatorname{Tr}\Delta_X u^i = 0 \quad\text{and}\quad \Delta_Z = \sum_i (\delta\gamma)_i u^i, \tag{48}$$

and with

$$D(X, Z) = X^{-\frac{1}{2}}\left(X^{\frac{1}{2}} Z X^{\frac{1}{2}}\right)^{\frac{1}{2}} X^{-\frac{1}{2}}. \tag{49}$$

3.1.1. Solution to the dual equation

In order to obtain the primal–dual direction (Δ_X, Δ_Z), the dual equation (46) is first projected onto the space spanned by the non-orthogonal basis $\{u^i\}$ (which we will call $\mathcal{U}$-space). With B denoting the right-hand side of (46) and making use of Eq. (48) we obtain:

$$\sum_j \underbrace{\left(\operatorname{Tr} D^{-1}u^j D^{-1}u^i\right)}_{\mathcal{H}^D_{ij}} \Delta\gamma_j = \operatorname{Tr} B u^i, \tag{50}$$

which can be seen to be a symmetrical, positive-definite linear system and as such can be solved iteratively using the linear conjugate gradient method. This can be done without explicit construction of the dual Hessian matrix $\mathcal{H}^D$ or any reference to the non-orthogonal basis set $\{u^i\}$. This is because $\mathcal{H}^D$ can be seen as a map from traceless tp-matrix space onto itself, by using the Hermitian adjoints of the linear maps $\mathcal{L}$. Consider an arbitrary traceless tp-matrix:

$$\epsilon = \sum_j \epsilon_j f^j. \tag{51}$$

Using (25) and the fact that the $\mathcal{L}$'s are linear and homogeneous we obtain that the image of ϵ under the dual Hessian map can be written as:

$$\mathcal{H}^D\epsilon = \hat{P}_{\text{Tr}}\left[\sum_k \mathcal{L}_k^\dagger\left(D_k^{-1}\mathcal{L}_k(\epsilon)D_k^{-1}\right)\right], \tag{52}$$

in which the D_k are the blocks of the D matrix corresponding to the different constraints $\mathcal{L}_k$, and $\hat{P}_{\text{Tr}}$ stands for the projection operator onto traceless tp-matrix space:

$$\hat{P}_{\text{Tr}}(A) = A - \frac{2\operatorname{Tr} A}{M(M-1)}\mathbb{1}_{\text{tp}}. \tag{53}$$

3.1.2. Solution to the primal equation

The solution of the primal equation (47) is obtained in the same manner, by projecting this equation onto $\mathcal{C}$-space, the orthogonal complement of $\mathcal{U}$-space. With B denoting the right-hand side of Eq. (47) and making use of Eq. (48) one gets:

$$\sum_j \underbrace{\left(\operatorname{Tr} Dc^j Dc^i\right)}_{\mathcal{H}^P_{ij}} \delta x_j = \operatorname{Tr} Bc^i, \tag{54}$$

where we have used

$$\Delta_X = \sum_i \delta x_i c^i. \tag{55}$$

This is again a symmetrical positive-definite system of linear equations that can be solved iteratively using the linear conjugate gradient method. As with the dual equation it can be solved without explicit construction of the Hessian matrix $\mathcal{H}^P$, or any reference to the basis set $\{c^i\}$, because $\mathcal{H}^P$ can be seen as a map from $\mathcal{C}$-space onto itself. For an arbitrary matrix in $\mathcal{C}$-space:

$$\epsilon = \sum_i \epsilon_i c^i, \tag{56}$$

the image of ϵ under the primal Hessian map is

$$\mathcal{H}^P\epsilon = \hat{P}_{\mathcal{C}}[D\epsilon D] \tag{57}$$

in which $\hat{P}_{\mathcal{C}}$ is the projection onto $\mathcal{C}$-space. This projection can be executed quickly by using the inverse of the overlap matrix of the $\mathcal{U}$-space basis vectors. Suppose we have an arbitrary block matrix A of the same dimension as X and Z. First we project it onto the space spanned by the basis $\{u^0, u^i\} = \{u^\alpha\}$. The projected matrix A' reads as:

$$A' = \sum_{\alpha\beta} \operatorname{Tr}\left[Au^\alpha\right]\left(\mathcal{S}^{-1}\right)_{\alpha\beta} u^\beta, \tag{58}$$

where the overlap matrix $\mathcal{S}$ appears because of the non-orthogonality of the basis. Due to the special properties of the linear matrix maps $\mathcal{L}$ that determine the basis matrices u^α, the inverse overlap matrix can also be considered as a map from tp-space onto itself (see Appendices A and B for the actual analytic expression of this map). The projected matrix A' can now be written in block-matrix form as:

$$A' = \bigoplus_l \mathcal{L}_l\left[\mathcal{S}^{-1}\left(\sum_k \mathcal{L}_k^\dagger(A_k)\right)\right]. \tag{59}$$

To project A onto $\mathcal{U}$-space we still have to remove the component along the u^0-matrix:

$$\hat{P}_{\mathcal{U}} A = A' - \left(\frac{\operatorname{Tr} u^0 A'}{\operatorname{Tr} u^0 u^0}\right) u^0. \tag{60}$$

Since $\mathcal{C}$-space is the orthogonal complement of the $\mathcal{U}$-space, the desired projection of A onto the $\mathcal{C}$-space is simply given by

$$\hat{P}_{\mathcal{C}} A = A - \hat{P}_{\mathcal{U}} A. \tag{61}$$

3.2. Outline of the algorithm

In this section a short outline of the algorithm will be presented. The first step is to initialize the primal–dual variables, after which they are directed towards the central path. Then the actual minimization of the primal–dual gap takes place, which is done in a predictor–corrector loop.

3.2.1. Initialization

We need a feasible primal–dual starting point. An initial feasible dual point $Z^{(0)}$, *i.e.* a matrix that satisfies the inequality (40), is easily found by setting

$$Z^{(0)} = u^0, \tag{62}$$

which corresponds to setting all the γ_i's equal to zero. A feasible primal starting point will have to satisfy Eq. (41). To construct such a point we take a completely random matrix X and project it onto a matrix X' for which

$$\operatorname{Tr} X' u^i = h^i. \tag{63}$$

This is again achieved using the inverse overlap matrix of the $\{u^\alpha\}$ basis,

$$X' = X - \underbrace{\sum_{\alpha\beta} \left(\operatorname{Tr} X u^\alpha - h^\alpha\right) \mathcal{S}^{-1}_{\alpha\beta} u^\beta}_{X^\perp}. \tag{64}$$

The last term on the right-hand side can be computed as:

$$X^\perp = \bigoplus_l \mathcal{L}_l \left[\mathcal{S}^{-1}\left(\sum_k \mathcal{L}_k^\dagger(X_k) - H^{(2)}\right)\right]. \tag{65}$$

At this point, X' satisfies the equality constraint (63), and one just has to add u^0, with a positive scaling factor that is large enough to ensure positive semidefiniteness:

$$X^{(0)} = X' + \alpha u^0 \succcurlyeq 0. \tag{66}$$

3.2.2. Centering run

Before the actual program can be started, a couple of centering steps have to be taken, which is done by solving Eqs. (46) and (47) with $\nu = 1$. The purpose is to go sufficiently near the central path, without bothering about the primal–dual gap. In a first step, Eq. (50) which has the smallest dimension, is solved using the conjugate gradient method, and the dual solution Δ_Z is obtained. The primal solution Δ_X then follows from the dual equation (46) by substitution. For these initial centering steps, both linear systems are so well conditioned that hardly any iterations are needed for convergence. As a measure for the distance from the center we use the potential [14]:

$$\Phi(X, Z) = -\ln\det X - \ln\det Z, \tag{67}$$

which is minimal (for points with the same primal–dual gap $\eta = \operatorname{Tr} XZ$) on the central path for which Eq. (43) is satisfied:

$$\Phi\left(X^c, Z^c\right) = -n \ln \frac{\eta}{n}. \tag{68}$$

When the potential difference (which is always positive):

$$\begin{aligned}\Psi(X, Z) &= \Phi(X, Z) - \Phi\left(X^c, Z^c\right) \\ &= n \ln \operatorname{Tr} XZ - n \ln n - \ln\det X - \ln\det Z,\end{aligned} \tag{69}$$

is sufficiently small, the centering run is stopped.

3.2.3. Predictor–corrector run

In this part of the program the primal–dual gap is minimized by alternating predictor and corrector steps. A predictor step tries to reduce the primal–dual gap by solving Eqs. (46) and (47) with $\nu = 0$. This is done in exactly the same way as for the centering run, by first solving (50) for Δ_Z, then substituting into (46) to obtain an approximate primal step Δ_X. The final primal step Δ_X is obtained by solving (54) using the conjugate gradient method with the approximate Δ_X as a starting point. Note that when the primal–dual gap decreases, the condition number of the primal and dual Hessian matrices increases and more iterations are needed before convergence is reached. One can adjust the convergence criteria of the primal and dual conjugate gradient loops, in order to minimize the combined number of iterations.

At this point we have a predictor direction (Δ_X, Δ_Z). The logarithmic potential $\phi(\alpha) = \Psi(X + \alpha\Delta_X, Z + \alpha\Delta_Z)$ in the predictor direction (see Eq. (69)) can be simply evaluated for any value of α by precomputing the eigenvalues λ_i^X of $X^{-\frac{1}{2}}\Delta_X X^{-\frac{1}{2}}$ and λ_i^Z of $Z^{-\frac{1}{2}}\Delta_Z Z^{-\frac{1}{2}}$. One then has

$$\begin{aligned}\phi(\alpha) = \Psi(X, Z) &+ \ln\left[1 + \alpha(c_X + c_Z)\right] \\ &- \sum_i \ln\left(1 + \alpha\lambda_i^X\right) - \sum_i \ln\left(1 + \alpha\lambda_i^Z\right),\end{aligned} \tag{70}$$

where

$$c_Z = \frac{1}{\eta} \operatorname{Tr} X\Delta_Z \quad \text{and} \quad c_X = \frac{1}{\eta} \operatorname{Tr} Z\Delta_X. \tag{71}$$

With a standard bisection method one can now compute the step size α corresponding to the maximal deviation from the central path we want to allow.

After the predictor step, a corrector step is taken, which is equivalent to the centering step described previously (see Section 3.2.2). The alternation of predictor and corrector steps continues until the primal–dual gap is smaller then the desired value.

4. Application to the BCS Hamiltonian

4.1. The BCS Hamiltonian

The algorithm introduced in Section 3 is applied to the BCS Hamiltonian [17]. The BCS Hamiltonian is an interesting system that models the competition between a single-particle operator and a schematic pairing interaction:

$$\hat{H} = \sum_{i\sigma} \epsilon_i a_{i\sigma}^\dagger a_{i\sigma} - g \sum_{ij} a_{i\uparrow}^\dagger a_{i\downarrow}^\dagger a_{j\downarrow} a_{j\uparrow}. \tag{72}$$

Here the single-particle levels are denoted with an index $i = 1, \ldots, M$, and the up (down) spin as $\sigma = \uparrow(\downarrow)$. When the pairing strength g is small compared to the single-particle level spacing, the energy is minimized by filling up the single-particle orbitals up to the Fermi level. With increasing g however, it becomes advantageous to form pairs, *i.e.* it is energetically favorable to maximize the ground-state occupation of the fermion pair state $\sum_i a_{i\uparrow}^\dagger a_{i\downarrow}^\dagger$. This problem is hard to solve using standard perturbative methods

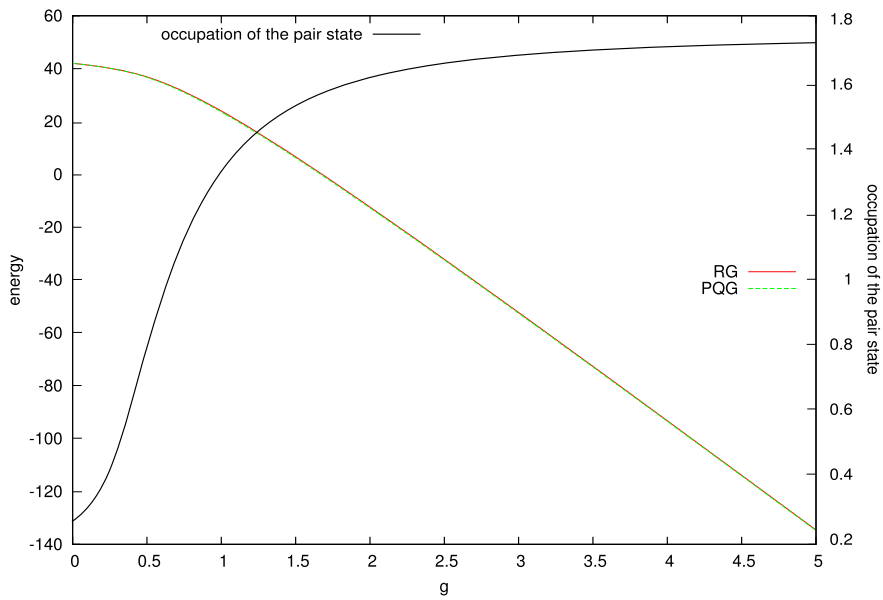

Fig. 1. The ground-state energy as calculated by v2DM(PQG) and the Richardson–Gaudin equations (RG), together with the pair occupation in the ground state by v2DM(PQG), as a function of the pairing interaction strength g.

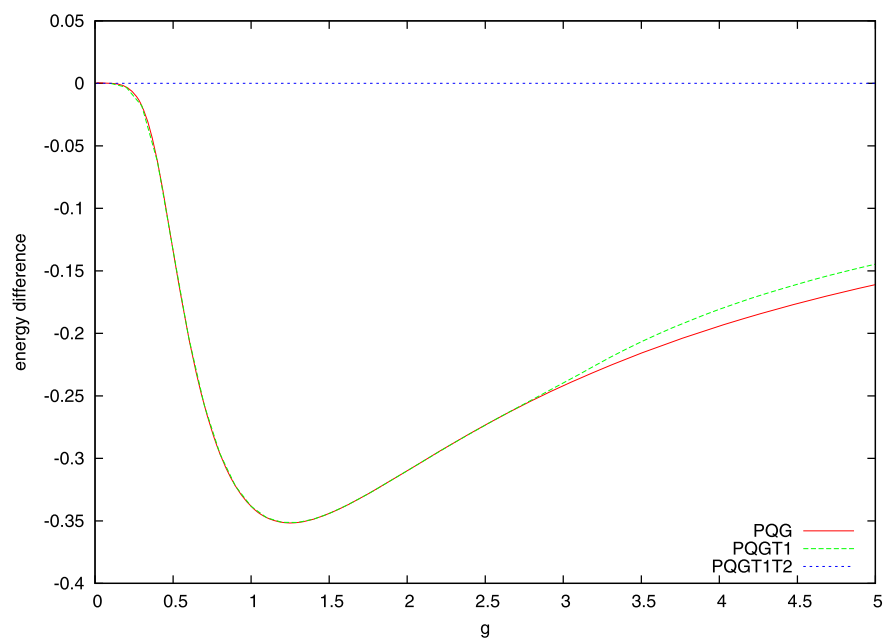

Fig. 2. The difference between the ground-state energy calculated by v2DM with various constraints, and the exact solution, as a function of pairing strength g.

as these tend to break down when pairs are formed. An exact solution based on the Bethe-ansatz exists for this problem, however, and involves solving a system of non-linear equations [22]. These equations are notoriously difficult to solve because, for certain critical values of g, the equations become singular. Several approaches have been suggested for solving these equations [23,24]. In this paper we follow the approach recently proposed by De Baerdemacker [25]. The exact ground-state energies as a function of g are compared to the v2DM results calculated within the present formalism.

4.2. Results

We have studied the Hamiltonian equation (72) with $M = 12$ doubly degenerate equidistant single-particle levels and $N = 12$ fermions, and g ranging from 0 to 5 in steps of 0.01. v2DM calculations were performed with respectively PQG, $PQGT_1$ and $PQGT_1T_2$ constraints. The resulting ground-state energy is compared to the exact solution in Fig. 1. For all values of g the agreement is already remarkably good at the PQG level. To appreciate how the result improves when constraints are added the difference between the various v2DM results and the exact solution is plotted in Fig. 2. Note that the difference is always negative, since v2DM provides a variational lower bound. As one observes, all approximations describe exactly the non-interacting small-g limit. When g becomes larger, there is competition between different types of ground states and the performance of PQG gets worse up to $g \approx 1.4$. For larger g the PQG result becomes better again. In fact, we checked (by omitting the single-particle piece) that also the $g \to \infty$ limit becomes exact for PQG, which is a peculiarity of the schematic pairing force. The $PQGT_1$ results show that the T_1 constraint only becomes active around $g = 2.5$, and ensures faster convergence to

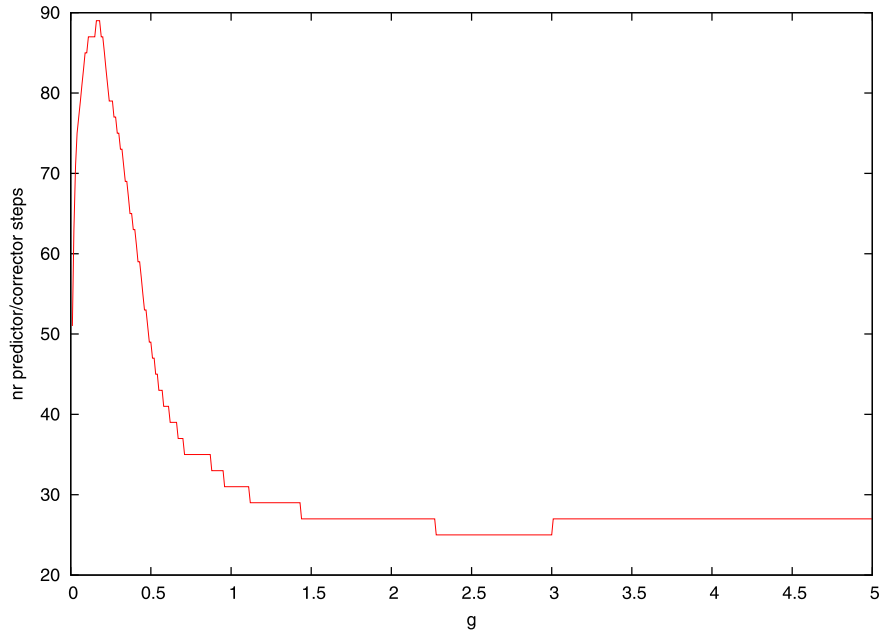

Fig. 3. Number of predictor and corrector steps needed for convergence, as a function of the pairing strength g in v2DM(PQG).

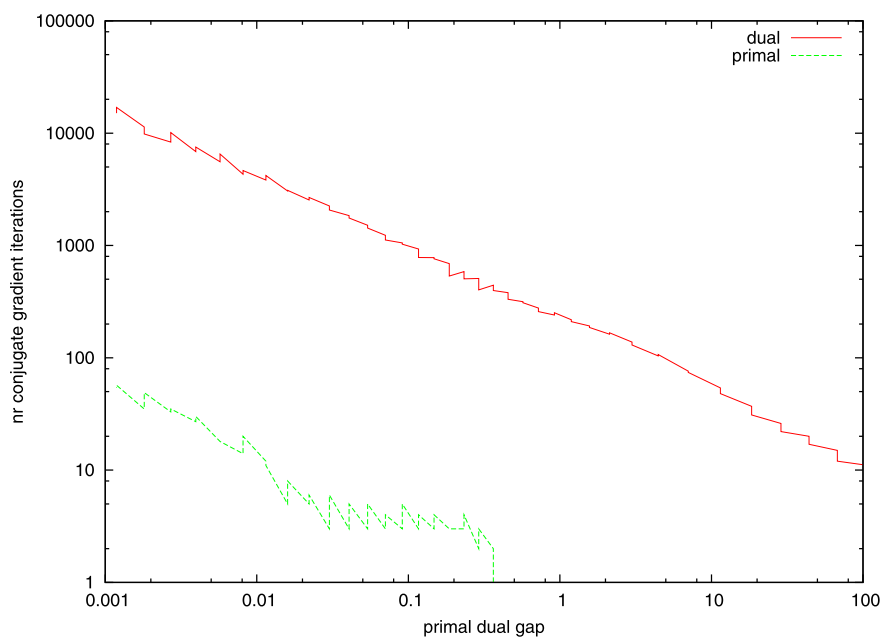

Fig. 4. Number of primal and dual conjugate gradient iterations needed for convergence, as a function of the primal–dual gap η for $g = 0.25$ in v2DM(PQG).

the exact $g \to \infty$ limit. Somewhat surprisingly, adding the T_2 condition is sufficient for obtaining the exact solution at all values of g.

4.3. Computational performance

Some of the computational aspects of the algorithm are worth pointing out. It is interesting to see *e.g.* that depending on the pairing interaction parameter g, the convergence properties of the algorithm change. In Fig. 3 the joint number of predictor and corrector steps needed for convergence, is plotted as a function of g. One observes a sharp peak at fairly small g, just when the perturbative regime is left and the structure of the ground state changes. For $g = 0.25$, which is at the position of the peak in Fig. 3, we have plotted in Fig. 4 the number of conjugate gradient iterations needed for convergence, of both the dual and the primal linear system, as a function of the primal–dual gap η. As expected, the number of iterations for the dual problem increases with decreasing primal–dual gap, as the linear system grows ill-conditioned. The primal conjugate gradient loop only becomes active for small values of η. This signals that the numerical stability becomes too small to generate a high quality approximation for Δ_X using the Δ_Z obtained in the dual conjugate gradient loop. Anyway, the needed number of primal iteration remains insignificant compared to the dual ones, for all values of η. The situation at $g = 0.25$ is the worst case. For larger values of g, where the number of predictor–corrector steps is smaller and approximately constant (see Fig. 3), the number of conjugate gradient iterations is also drastically reduced. A typical behaviour is plotted in Fig. 5 for $g = 4$. The calculations have been performed on a single-core Toshiba lap-

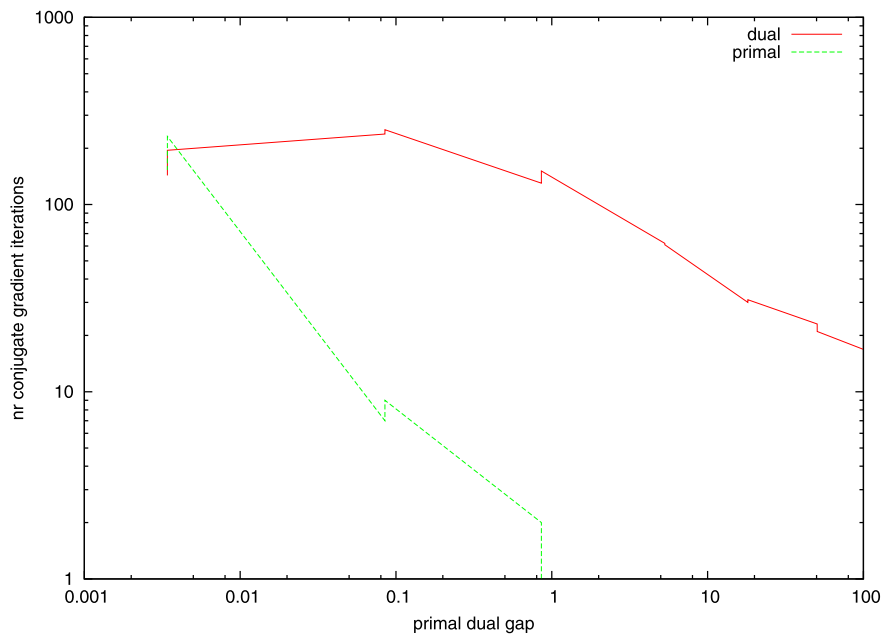

Fig. 5. Number of primal and dual conjugate gradient iterations needed for convergence, as a function of the primal–dual gap η for $g = 4$ in v2DM(PQG).

top with a 2.26 GHz Intel Centrino processor. In the region of g where convergence is slowest, a PQG calculation took up to half an hour whereas for other g values the calculation ends in less than a minute. Apart from spin, no symmetry was taken into account.

5. Summary and discussion

Interacting quantum many-particle systems lie at the heart of most issues in condensed matter, molecular/atomic and nuclear physics. Their analysis may be rephrased as the problem of minimizing the energy, expressed as a linear function of a two-body density matrix, subject to the N-representability constraint that the 2DM can be derived from a physical N-particle system. By working solely with the 2DM, rather than with the N-particle wave function itself, the problem of the exponentially exploding dimension of N-particle Hilbert space with increasing N is circumvented. The complexity of the problem is shifted, however, to the characterization of the N-particle representable 2DMs. In practice, a limited set of necessary but not sufficient conditions for N-representability are imposed during the minimization, resulting in a strict lower bound to the energy, which converges to the exact energy when more and more N-representability conditions are imposed.

Commonly used N-representability conditions impose the positive semidefiniteness of a set of linear matrix functionals of the 2DM. In this way the quantum many-body problem is converted into a well established field of optimization techniques called semidefinite programming. Standard packages for SDP, however, fail to take into account properties of the physical problem that can be exploited.

Using specific mathematical properties of the constraints for the v2DM problem, we have adapted a standard primal–dual interior point method to be computationally cheaper, both in storage as in floating point operations. We make extensive use of the algebra of linear matrix maps to calculate efficiently some intermediate quantities. During the Newton minimization procedure, a new direction in 2DM space is found iteratively using the conjugate gradient algorithm, thereby exploiting the fact that the product of the Hessian with a 2DM is considerably cheaper for the physical problem at hand than in a general situation.

As an example we have applied the algorithm to a BCS-type Hamiltonian. We found that the standard constraints work very well for this kind of problem. The computational performance of the method was analyzed, and it was shown that the convergence behaviour is dependent on the value of the pairing strength parameter. As in our previous algorithm [20] the method slows down near the solution, because the matrices involved become ill conditioned. The present primal–dual algorithm allows to control this since the primal–dual gap provides an upper bound to the remaining error. Therefore the algorithm can be stopped when the required accuracy is reached, saving many unnecessary iterations.

Acknowledgements

We gratefully acknowledge financial support from FWO-Flanders and the Research Council of Ghent University. We would like to thank Paul W. Ayers for his useful suggestions. B.V., H.V.A., P.B. and D.V.N. are Members of the QCMM Alliance Ghent–Brussels.

Appendix A. Calculation of the overlap-matrix map

The overlap matrix of the non-orthogonal basis set $\{u^\alpha\}$ is defined as:

$$\mathcal{S}_{\alpha\beta} = \mathrm{Tr}\, u^\alpha u^\beta. \tag{A.1}$$

Using the Hermitian adjoints of the linear maps $\mathcal{L}$ we can rewrite this as:

$$\mathcal{S}_{\alpha\beta} = \sum_k \mathrm{Tr}\big[\mathcal{L}_k^\dagger\big(\mathcal{L}_k\big(f^\alpha\big)\big) f^\beta\big], \tag{A.2}$$

in which $\{f^\alpha\}$ *is* an orthogonal basis of tp-matrix space. This means that the overlap matrix can be seen as a linear map from tp-space onto itself, whose action onto a tp-matrix Γ is:

$$\mathcal{S}(\Gamma) = \sum_k \mathcal{L}_k^\dagger\big(\mathcal{L}_k(\Gamma)\big). \tag{A.3}$$

It turns out that this map can be written as a generalized $\mathcal{Q}$ map, which is defined as:

$$\mathcal{Q}(a,b,c)(\Gamma)_{\alpha\beta;\gamma\delta} = a\Gamma_{\alpha\beta;\gamma\delta} + b(\delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma})\bar{\bar{\Gamma}} - c(\delta_{\alpha\gamma}\bar{\Gamma}_{\beta\delta} - \delta_{\beta\gamma}\bar{\Gamma}_{\alpha\delta} - \delta_{\alpha\delta}\bar{\Gamma}_{\beta\gamma} + \delta_{\beta\delta}\bar{\Gamma}_{\alpha\gamma}). \tag{A.4}$$

This is like a $\mathcal{Q}$-map (12) but with general coefficients (a,b,c). The proof is somewhat tedious and proceeds by considering every $\mathcal{L}_k$ separately.

A.1. $\mathcal{P}^2$

It is trivial to see that $\mathcal{P}^2(\Gamma) = \Gamma$ and that this is a generalized $\mathcal{Q}$ map with coefficients

$$a = 1, \qquad b = 0, \qquad c = 0. \tag{A.5}$$

A.2. $\mathcal{Q}^2$

To reexpress $\mathcal{Q}^2$ we first calculate the various pieces,

$$\bar{\mathcal{Q}}(\Gamma)_{\alpha\gamma} = \left[\frac{M-N-1}{N(N-1)}\right]\delta_{\alpha\gamma}\bar{\bar{\Gamma}} - \left[\frac{M-N-1}{N-1}\right]\bar{\Gamma}_{\alpha\gamma}, \tag{A.6}$$

$$\bar{\bar{\mathcal{Q}}}(\Gamma) = \left[\frac{(M-N)(M-N-1)}{N(N-1)}\right]\bar{\bar{\Gamma}}. \tag{A.7}$$

Substitute into Eq. (12) leads once again to a generalized $\mathcal{Q}$ map with coefficients:

$$a = 1, \qquad b = \frac{4N^2 + 2N - 4NM + M^2 - M}{N^2(N-1)^2}, \qquad c = \frac{2N - M}{(N-1)^2}. \tag{A.8}$$

A.3. $\mathcal{G}^\dagger\mathcal{G}$

With the same strategy one finds on the basis of Eq. (26) and

$$\bar{\mathcal{G}}(\Gamma)_{\alpha\gamma} = \frac{M-1}{N-1}\bar{\Gamma}_{\alpha\gamma}, \tag{A.9}$$

that substituting into (26) leads to another generalized $\mathcal{Q}$ map with coefficients:

$$a = 4, \qquad b = 0, \qquad c = \frac{2N - M - 2}{(N-1)^2}. \tag{A.10}$$

A.4. $\mathcal{T}_1^\dagger\mathcal{T}_1$

The needed terms are now:

$$\bar{\mathcal{T}}_1(\Gamma)_{\alpha\beta;\gamma\delta} = (M-4)\Gamma_{\alpha\beta;\gamma\delta} + \left[\frac{M-N-2}{N(N-1)}\right](\delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma})\bar{\bar{\Gamma}} \tag{A.11}$$

$$- \left[\frac{M-N-2}{N-1}\right]\hat{A}[\delta_{\alpha\gamma}\bar{\Gamma}_{\beta\delta} - \delta_{\beta\gamma}\bar{\Gamma}_{\alpha\delta} - \delta_{\alpha\delta}\bar{\Gamma}_{\beta\gamma} + \delta_{\beta\delta}\bar{\Gamma}_{\alpha\gamma}], \tag{A.12}$$

$$\bar{\bar{\mathcal{T}}}_1(\Gamma)_{\alpha\gamma} = \left[\frac{(M-N-2)(M-N-1)}{N(N-1)}\right]\delta_{\alpha\gamma}\bar{\bar{\Gamma}} - \left[\frac{(M-3)(M-2N)}{N-1}\right]\bar{\Gamma}_{\alpha\gamma}, \tag{A.13}$$

$$\bar{\bar{\bar{\mathcal{T}}}}_1(\Gamma) = \left[\frac{(M-2)(M(M-1) - 3N(M-N))}{N(N-1)}\right]\bar{\bar{\Gamma}}, \tag{A.14}$$

and substitution into Eq. (28) leads to the coefficients:

$$a = M - 4,$$
$$b = \frac{M^3 - 6M^2N - 3M^2 + 12MN^2 + 12MN}{3N^2(N-1)^2} + \frac{2M - 18N^2 - 6N^3}{3N^2(N-1)^2},$$
$$c = -\frac{M^2 + 2N^2 - 4MN - M + 8N - 4}{2(N-1)^2}.$$

A.5. $\mathcal{T}_2^\dagger\mathcal{T}_2$

Finally, needed for the calculation of the last map are:

$$\bar{\mathcal{T}}_2(\Gamma)_{\alpha\beta;\gamma\delta} = \frac{\bar{\bar{\Gamma}}}{N-1}(\delta_{\alpha\gamma}\delta_{\beta\delta} - \delta_{\alpha\delta}\delta_{\beta\gamma}) + M\Gamma - [\delta_{\alpha\gamma}\bar{\Gamma}_{\beta\delta} - \delta_{\beta\gamma}\bar{\Gamma}_{\alpha\delta} - \delta_{\alpha\delta}\bar{\Gamma}_{\beta\gamma} + \delta_{\beta\delta}\bar{\Gamma}_{\alpha\gamma}], \tag{A.15, A.16}$$

$$\tilde{\mathcal{T}}_2(\Gamma)_{\alpha\beta;\gamma\delta} = \frac{M-N}{N-1}\bar{\Gamma}_{\beta\delta}\delta_{\alpha\gamma} + \delta_{\beta\delta}\bar{\Gamma}_{\alpha\gamma} - (M-2)\Gamma_{\alpha\delta;\gamma\beta}, \tag{A.17}$$

$$\tilde{\bar{\mathcal{T}}}_2(\Gamma)_{\alpha\gamma} = \left[\frac{M(M-N) - (N-1)(M-2)}{N-1}\right]\bar{\Gamma}_{\alpha\gamma} + \delta_{\alpha\gamma}\bar{\bar{\Gamma}}, \tag{A.18}$$

which, when substituted into Eq. (31) gives the following coefficients:

$$a = 5M - 8, \qquad b = \frac{2}{N-1}, \qquad c = \frac{2N^2 + (M-2)(4N-3) - M^2}{2(N-1)^2}. \tag{A.19}$$

The overlap-matrix map is just the sum of the various terms obtained, and hence also a generalized $\mathcal{Q}$ map with rather complex coefficients.

Appendix B. Inverse of generalized $\mathcal{Q}$ map

The inverse of a generalized $\mathcal{Q}$ map can be shown to be another generalized $\mathcal{Q}$ map. Consider for brevity the notation:

$$\mathcal{Q}(a,b,c)(\Gamma) = Q, \tag{B.1}$$

then applying partial trace operations on Eq. (A.4) leads to:

$$\bar{\bar{\Gamma}} = \frac{\bar{\bar{Q}}}{a + M(M-1)b - 2(M-1)c}, \tag{B.2}$$

$$\bar{\Gamma}_{\alpha\gamma} = \frac{1}{a - c(M-2)} \times \left[\bar{Q}_{\alpha\gamma} - \frac{b(M-1) - c}{a + M(M-1)b - 2(M-1)c}\delta_{\alpha\gamma}\bar{\bar{Q}}\right]. \tag{B.3}$$

Upon substitution into Eq. (A.4) and solving for Γ one obtains,

$$\Gamma = \mathcal{Q}^{-1}(a,b,c)(Q) = \mathcal{Q}(a',b',c')(Q), \tag{B.4}$$

where

$$a' = \frac{1}{a}, \tag{B.5}$$

$$b' = \frac{ba + bcM - 2c^2}{a[c(M-2) - a][a + bM(M-1) - 2c(M-1)]}, \tag{B.6}$$

$$c' = \frac{c}{a[c(M-2) - a]}. \tag{B.7}$$

These are important relations since they allow to evaluate the action of the inverse overlap matrix on a tp-matrix as fast as a $\mathcal{Q}$ map, *i.e.* at a computational cost which is negligible compared to the other matrix manipulations.

References

[1] K. Husimi, Proc. Phys. Math. Soc. Japan 22 (1940) 264.
[2] P. Löwdin, Phys. Rev. 97 (1955) 6.
[3] A.J. Coleman, V.I. Yukalov, Reduced Density Matrices: Coulson's Challange, Springer-Verlag, New York, 2000.
[4] J. Mayer, Phys. Rev. 100 (1955) 6.
[5] R.H. Tredgold, Phys. Rev. 105 (1957) 5.
[6] A.J. Coleman, Rev. Mod. Phys. 35 (1963) 3.
[7] C. Garrod, J.K. Percus, J. Math. Phys. 5 (1964) 12.
[8] C. Garrod, M.A. Fusco, Int. J. Quantum Chem. 10 (1976) 495.
[9] C. Garrod, M.V. Mihailovic, M. Rosina, J. Math. Phys. 16 (1975) 868.
[10] M. Rosina, C. Garrod, J. Comp. Phys. 18 (1975) 300.
[11] M. Mihailovic, M. Rosina, Nucl. Phys. A 237 (1974) 221.
[12] M. Nakata, H. Nakatsuji, M. Ehara, M. Fukuda, K. Nakata, K. Fujisawa, J. Chem. Phys. 114 (2001) 19.
[13] D.A. Mazziotti, Phys. Rev. A 65 (2002) 062511.
[14] L. Vandenberghe, S. Boyd, SIAM Rev. 38 (1996) 1.
[15] D.A. Mazziotti, Phys. Rev. Lett. 93 (2004) 213001.
[16] J.F. Sturm, Primal–Dual Interior Point Approach to Semidefinite Programming, Tinbergen Institute Research Series, Amsterdam Thesis Publishers, 1997.
[17] J. Bardeen, L.N. Cooper, J.R. Schrieffer, Theory of Superconductivity 108 (1957).
[18] W.H. Dickhoff, D. Van Neck, Many-Body Theory Exposed! World Scientific, 2008.
[19] P.W. Ayers, Phys. Rev. A 74 (2006) 042502.
[20] B. Verstichel, H. van Aggelen, D. Van Neck, P.W. Ayers, P. Bultinck, Phys. Rev. A 80 (2009) 032508.
[21] E. Cancès, G. Stoltz, M. Lewin, J. Chem. Phys. 125 (2006) 064101.
[22] R.W. Richardson, N. Sherman, Nucl. Phys. 52 (1964) 221.
[23] R.W. Richardson, Phys. Rev. 141 (1966) 949.
[24] S. Rombouts, D. Van Neck, J. Dukelsky, Phys. Rev. C 69 (2004) 061303(R).
[25] S. De Baerdemacker, arXiv:1011.3691v1 [nucl-th], 2010, J. Phys. Conf. Ser., in press.