

Constrained-pairing mean-field theory. IV. Inclusion of corresponding pair constraints and connection to unrestricted Hartree–Fock theory

Takashi Tsuchimochi,¹ Thomas M. Henderson,² Gustavo E. Scuseria,^{2,a)} and Andreas Savin³

¹Department of Chemistry, Rice University, Houston, Texas 77005-1892, USA

²Department of Chemistry and Department of Physics and Astronomy, Rice University, Houston, Texas 77005-1892, USA

³Laboratoire de Chimie Théorique, CNRS, Université Pierre et Marie Curie, 4 Place Jussieu, F-75252 Paris, France

(Received 16 July 2010; accepted 26 August 2010; published online 5 October 2010)

Our previously developed constrained-pairing mean-field theory (CPMFT) is shown to map onto an unrestricted Hartree–Fock (UHF) type method if one imposes a corresponding pair constraint to the correlation problem that forces occupation numbers to occur in pairs adding to one. In this new version, CPMFT has all the advantages of standard independent particle models (orbitals and orbital energies, to mention a few), yet unlike UHF, it can dissociate polyatomic molecules to the correct ground-state restricted open-shell Hartree–Fock atoms or fragments. © 2010 American Institute of Physics. [doi:10.1063/1.3490478]

I. INTRODUCTION

In a recent series of papers,^{1–3} we have developed constrained-pairing mean-field theory (CPMFT), a method capable of describing static (strong) correlation in an accurate and efficient manner. The idea behind CPMFT is to make use of the pairing correlations (see below) that occur in a quasiparticle picture to describe static correlation in molecular systems. In CPMFT, we divide the natural orbitals into core, active, and virtual blocks. Each core orbital has unit occupation, each virtual orbital has zero occupation, and the active natural orbitals have fractional occupations n_i , where $0 < n_i < 1$. Static correlation is introduced by allowing electron pairs to have fractional occupations within an active space.

The use of a pairing interaction has many advantages. Unlike unrestricted Hartree–Fock (UHF), CPMFT has zero spin density everywhere for closed-shell systems. In the absence of static correlation, CPMFT reduces to restricted Hartree–Fock (RHF), while it dissociates polyatomic molecules to restricted open shell Hartree–Fock (ROHF) atoms or fragments. Essentially, the dissociation limit of CPMFT can be thought of as an ensemble solution. By reducing to RHF in the absence of strong correlation and ROHF at dissociation, CPMFT cleanly separates static from dynamic correlation, as previously shown in Ref. 3, where the CPMFT \mathbf{P} and \mathbf{K} density matrices were used to construct alternative densities to be used as inputs into traditional density functionals for the dynamical correlation energy. Remarkably, CPMFT accomplishes these feats at a mean field computational cost instead of the combinatorial blowup of complete active space (CASSCF) or full configuration interaction.

While CPMFT is clearly distinct from UHF, it shows

some unexpected connections. We can take advantage of these connections to simplify the formalism, make it more efficient, and establish interesting similarities. The purpose of this paper is to demonstrate these relations and the accompanying reformulation of CPMFT. Accordingly, we discuss this connection in Sec. III at some length, and show how we can use it to simplify the solution of the CPMFT equations. Section IV shows some numerical examples, and we provide conclusions in Sec. V. We include an Appendix that discusses some other formal properties of CPMFT. First, however, we provide a brief introduction to pairing correlations in Sec. II.

II. PAIRING CORRELATIONS AND THE QUASIPARTICLE PICTURE

Strong correlations in nuclear physics or superconductivity are often described as the formation of Cooper pairs. The theoretical machinery which does this is the Hartree–Fock–Bogoliubov (HFB) method.⁴ In HFB, we write the wave function $|\Phi_{\text{HFB}}\rangle$ as a single determinant of quasiparticles created by quasiparticle creation operators which are linear combinations of electron creation and annihilation operators. The quasiparticle wave function thus violates electron number conservation. Because the quasiparticle wave function is a single determinant, its associated density matrix \mathbf{R} is idempotent ($\mathbf{R}^2 = \mathbf{R}$) and Hermitian ($\mathbf{R} = \mathbf{R}^\dagger$). We have

$$\mathbf{R} = \begin{pmatrix} \gamma & \kappa \\ -\kappa^* & 1 - \gamma^* \end{pmatrix}. \quad (1)$$

Here, γ is the physical density matrix in the spinorbital basis; it is Hermitian but *not* idempotent. Information about pairing correlations is carried by the anomalous density matrix κ , which is antisymmetric by definition because $\kappa_{ij} = \langle a_i^\dagger a_j^\dagger \rangle$. We limit our discussion to the closed-shell case, in which case we have, for $\alpha\alpha$, $\alpha\beta$, $\beta\alpha$, and $\beta\beta$ blocks,

^{a)}Author to whom correspondence should be addressed. Electronic mail: guscus@rice.edu.

$$\boldsymbol{\gamma} = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix}, \quad (2a)$$

$$\boldsymbol{\kappa} = \begin{pmatrix} \mathbf{0} & \mathbf{K} \\ -\mathbf{K} & \mathbf{0} \end{pmatrix}, \quad (2b)$$

where \mathbf{P} is the closed-shell (spatial orbital) density matrix and \mathbf{K} is the (symmetric, positive semidefinite) closed-shell anomalous density matrix. We emphasize here that only the $\alpha\beta$ and $\beta\alpha$ blocks of $\boldsymbol{\kappa}$ are nonzero, so that we consider only singlet pairing.⁵ We should also mention that the notation here differs slightly from that used in Refs. 1–3, but does so in an attempt to make this manuscript self-contained and as clear as possible.

Idempotence of the quasiparticle density matrix \mathbf{R} yields two conditions on the electronic density matrix \mathbf{P} and the anomalous density matrix \mathbf{K} ,

$$\mathbf{PK} - \mathbf{KP} = \mathbf{0}, \quad (3a)$$

$$\mathbf{P} - \mathbf{P}^2 = \mathbf{K}^2. \quad (3b)$$

Physically, \mathbf{K}^2 is the “odd-electron distribution” of Yamaguchi,⁶ the “density of effectively unpaired electrons” of Staroverov and Davidson,⁷ and is related to Mayer’s “free valence index”⁸ once it is written in terms of the total density matrix $\boldsymbol{\gamma}^{\alpha\alpha} + \boldsymbol{\gamma}^{\beta\beta} = 2\mathbf{P}$. Essentially, \mathbf{K}^2 gauges the singlet diradical character of the system (or, for larger active spaces, the polyradical character) and is a local measure of electron entanglement.

The HFB energy is given as the expectation value of the Hamiltonian with respect to the HFB wave function,

$$E_{\text{HFB}} = \langle \Phi_{\text{HFB}} | H | \Phi_{\text{HFB}} \rangle, \quad (4a)$$

$$= 2h_{ij}P_{ij} + (2\langle ij|kl\rangle - \langle ij|lk\rangle)P_{ik}P_{jl} + \langle ij|kl\rangle K_{ij}K_{kl}, \quad (4b)$$

where summation over repeated indices here and throughout the manuscript is implied; h_{ij} are matrix elements of the one-electron part of the Hamiltonian and $\langle ij|kl\rangle$ are two-electron integrals in Dirac notation.

In order to determine the occupation numbers and natural orbitals, HFB variationally minimizes E_{HFB} subject to the constraint that the density matrix \mathbf{P} contains the correct number of particles,

$$\text{Tr}(\mathbf{P}) = N. \quad (5)$$

This condition is enforced by a chemical potential μ introduced as a Lagrange multiplier. The HFB formulation leads to equations similar to Hartree–Fock, which in the particular case of closed-shell systems are

$$\mathbf{R}_{\text{cs}} \mathcal{H}_{\text{HFB}} - \mathcal{H}_{\text{HFB}} \mathbf{R}_{\text{cs}} = \mathbf{0}, \quad (6)$$

where \mathbf{R}_{cs} is the closed-shell quasiparticle density matrix

$$\mathbf{R}_{\text{cs}} = \begin{pmatrix} \mathbf{P} & \mathbf{K} \\ \mathbf{K} & \mathbf{1} - \mathbf{P} \end{pmatrix}, \quad (7)$$

and \mathcal{H}_{HFB} is the double-Hamiltonian (DH) given by

$$\mathcal{H}_{\text{HFB}} = \begin{pmatrix} \mathbf{F}^{\text{cs}} + \mu N & \mathbf{\Delta} \\ \mathbf{\Delta} & -\mathbf{F}^{\text{cs}} - \mu N \end{pmatrix}. \quad (8)$$

Here \mathbf{F}^{cs} is the standard closed-shell Fock matrix and $\mathbf{\Delta}$ is known as the pairing Hamiltonian. These are given by

$$F_{ij}^{\text{cs}} = h_{ij} + (2\langle ik|jl\rangle - \langle ik|lj\rangle)P_{kl}, \quad (9a)$$

$$\Delta_{ij} = \langle ij|kl\rangle K_{kl}. \quad (9b)$$

The double-Hamiltonian \mathcal{H}_{HFB} is just the mean-field of the physical Hamiltonian with respect to the quasiparticle determinant.

Because the pairing energy of HFB [the term proportional to \mathbf{K}^2 in Eq. (4b)] is positive when the electron-electron interaction is repulsive, the variationally optimal solution is always $\mathbf{K} = \mathbf{0}$, and therefore, $\mathbf{\Delta} = \mathbf{0}$. In other words, HFB just returns the regular Hartree–Fock solution for Coulombic repulsive systems. In order to have HFB solutions with energies lower than Hartree–Fock, one needs a net attractive two-body interaction, as in the Bardeen–Cooper–Schrieffer picture of superconductivity (where it is provided by electron-phonon coupling) or in nuclear forces. In order to take advantage of the pairing picture for the conventional repulsive electron-electron interaction, and with the aim of describing strong correlations, CPMFT simply reverses the sign of the pairing energy. We thus have

$$E_{\text{CPMFT}} = 2h_{ij}P_{ij} + (2\langle ij|kl\rangle - \langle ij|lk\rangle)P_{ik}P_{jl} - \langle ij|kl\rangle K_{ij}K_{kl}. \quad (10)$$

The last term plays the role of a correlation energy—a correction to the closed-shell RHF-like energy expression—and will be referred to as such throughout this manuscript, but it is certainly not our previous definition of static correlation,² $E_{\text{CPMFT}} - E_{\text{RHF}}$, since \mathbf{P} is not \mathbf{P}_{RHF} .

In addition to changing the sign of the pairing energy, in CPMFT, we also restrict noninteger occupations to an active space, so that pairing only occurs between quasidegenerate orbitals. Changing the sign of the pairing term changes the sign of $\mathbf{\Delta}$ so that the double-Hamiltonian is

$$\mathcal{H}_{\text{CPMFT}} = \begin{pmatrix} \mathbf{F}^{\text{cs}} + \mu N & -\mathbf{\Delta} \\ -\mathbf{\Delta} & -\mathbf{F}^{\text{cs}} - \mu N \end{pmatrix}. \quad (11)$$

Otherwise, DH-CPMFT follows the same procedure as in HFB. However, changing the sign of the pairing energy and the pairing matrix severs the connection between the HFB wave function $|\Phi_{\text{HFB}}\rangle$ and the CPMFT energy. Note that what we have called simply CPMFT in Refs. 1–3 is here referred to as DH-CPMFT, whereas “CPMFT” here refers to the new formulation to be introduced below.

We can, indeed, view the CPMFT energy as the expectation value of a model Hamiltonian with respect to a particle-number violating determinant,

$$H_0 |\Phi\rangle = E_{\text{CPMFT}} |\Phi\rangle, \quad (12a)$$

$$H_0 = (F_{ij}^{\text{cs}} + h_{ij})a_i^\dagger a_j - \frac{1}{2}\Delta_{ij}a_i^\dagger a_j^\dagger - \frac{1}{2}\Delta_{ij}^* a_i a_j. \quad (12b)$$

This quadratic model Hamiltonian, however, is not the mean-field of the physical Hamiltonian with respect to a quasipar-

TABLE I. Summary of properties in UHF, HFB, and CPMFT for closed-shell systems. We show the correlation energy (i.e., the difference between the energy from the method and the closed-shell piece), the effective polarization, and the particle number fluctuations.

Method	E_c^a	Polarization	σ_N^2
UHF	$-v_{ij}^{kl}M_{ij}^kM_{kl}^j$	$\mathbf{M}=(\mathbf{A}-\mathbf{B})/2$	0
HFB	$v_{ij}^{kl}K^{ij}K_{kl}$	$\mathbf{K}= \mathbf{A}-\mathbf{B} /2$	$4 \text{Tr}(\mathbf{K}^2)$
CPMFT	$-v_{ij}^{kl}K^{ij}K_{kl}$	$\mathbf{K}= \mathbf{A}-\mathbf{B} /2$	0

^a $v_{ij}^{kl}=\langle ij|kl\rangle$ is a two-electron integral in Dirac notation.

ticle determinant. As previously noted,² we can interpret the CPMFT energy as a hybrid of Hartree–Fock and HFB where Hartree–Fock uses $2/r_{12}$ as the electron–electron repulsion operator and HFB uses $-1/r_{12}$.

Nevertheless, we have a fruitful alternative viewpoint, which is to envision the CPMFT energy expression of Eq. (10) as defining a model two-particle density matrix $\mathbf{\Gamma}$ such that the energy in the spin-orbital basis is

$$E_{\text{CPMFT}} = \text{Tr}(\mathbf{h}\boldsymbol{\gamma}) + \text{Tr}(\mathbf{v}\mathbf{\Gamma}_{\text{CPMFT}}), \quad (13)$$

where \mathbf{v} is the two-particle part of the Hamiltonian, and \mathbf{h} is the one-particle part. In terms of spin-orbitals, we have

$$(\mathbf{\Gamma}_{\text{CPMFT}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k - \frac{1}{2}\kappa_{ij}\kappa^{kl}, \quad (14)$$

with lower (upper) indices corresponding to bra (ket) indices. The first two terms in this model two-particle density matrix correspond to Hartree–Fock whereas the last term introduces static correlation via \mathbf{K} , which is a measure of nonidempotency for \mathbf{P} . This last term is an important quantity in the cumulant decomposition of density matrices,⁹ but in our work appears naturally from the idempotency of the quasi-particle density matrix. If we use this model two-particle density matrix to *define* expectation values of two-particle operators, then as shown in the Appendix, we find the important result that CPMFT has no particle number fluctuations. In making this choice, we are inevitably working with a density matrix functional and are effectively doing some form of a statistical ensemble theory. Table I collects results about the UHF two-particle density matrix, the CPMFT model two-particle density matrix, and the analogously defined HFB model two-particle density matrix. We derive these results in the Appendix.

Note that our model two-particle density matrix has appeared before in the literature as the corrected Hartree–Fock functional.¹⁰ Our model is, however, solved by diagonalization.⁵ More importantly, CPMFT restricts the noninteger occupation numbers to an active space only, and in the present work further enforces the corresponding pairs constraint, which we will now introduce.

III. CPMFT AND UHF

Consider the UHF treatment of a system where the number of spin-up and spin-down electrons is the same. The spin-up and spin-down density matrices $\boldsymbol{\gamma}^{\alpha\alpha}$ and $\boldsymbol{\gamma}^{\beta\beta}$ are both idempotent,

$$(\boldsymbol{\gamma}^{\alpha\alpha})^2 - \boldsymbol{\gamma}^{\alpha\alpha} = (\boldsymbol{\gamma}^{\beta\beta})^2 - \boldsymbol{\gamma}^{\beta\beta} = \mathbf{0}. \quad (15)$$

The charge density and spin magnetization (or polarization) density matrices are

$$\mathbf{P} = \frac{1}{2}(\boldsymbol{\gamma}^{\alpha\alpha} + \boldsymbol{\gamma}^{\beta\beta}), \quad (16a)$$

$$\mathbf{M} = \frac{1}{2}(\boldsymbol{\gamma}^{\alpha\alpha} - \boldsymbol{\gamma}^{\beta\beta}). \quad (16b)$$

Traditionally, the UHF energy¹¹ is expressed in terms of the $\boldsymbol{\gamma}^{\alpha\alpha}$ and $\boldsymbol{\gamma}^{\beta\beta}$ density matrices,

$$E_{\text{UHF}} = h_{ij}(\gamma_{ij}^{\alpha\alpha} + \gamma_{ij}^{\beta\beta}) + \frac{1}{2}\langle ij|kl\rangle(\gamma_{ik}^{\alpha\alpha} + \gamma_{ik}^{\beta\beta})(\gamma_{jl}^{\alpha\alpha} + \gamma_{jl}^{\beta\beta}) - \frac{1}{2}\langle ij|kl\rangle(\gamma_{il}^{\alpha\alpha}\gamma_{jk}^{\alpha\alpha} + \gamma_{il}^{\beta\beta}\gamma_{jk}^{\beta\beta}), \quad (17)$$

where we have put $\boldsymbol{\gamma}^{\alpha\alpha}$ and $\boldsymbol{\gamma}^{\beta\beta}$ in the same basis (say, the atomic orbital basis). Although it is almost never presented in this way, we can also write the UHF energy as a functional of \mathbf{P} and \mathbf{M} , which yields

$$E_{\text{UHF}}[\mathbf{P}, \mathbf{M}] = E_{\text{cs}}[\mathbf{P}] + E_c[\mathbf{M}], \quad (18a)$$

$$E_{\text{cs}}[\mathbf{P}] = 2h_{ij}P_{ij} + (2\langle ij|kl\rangle - \langle ij|lk\rangle)P_{ik}P_{jl}, \quad (18b)$$

$$E_c[\mathbf{M}] = -\langle ij|kl\rangle M_{il}M_{jk}. \quad (18c)$$

Here, E_{cs} indicates the usual RHF energy expression given in terms of the charge density matrix \mathbf{P} , while E_c carries the correlation energy in terms of the spin magnetization density matrix \mathbf{M} . An utterly unexpected result is that the closed-shell CPMFT energy expression of Eq. (10) is identical to the UHF energy expression of Eq. (18), except that the spin density matrix \mathbf{M} is replaced by the anomalous density matrix \mathbf{K} .¹² In cases in which UHF predicts static correlation by breaking symmetry (*i.e.* nonzero spin density),¹³ \mathbf{P} is not idempotent. Instead, it satisfies

$$\mathbf{P} - \mathbf{P}^2 = \frac{1}{2}(\boldsymbol{\gamma}^{\alpha\alpha} + \boldsymbol{\gamma}^{\beta\beta}) - \frac{1}{4}(\boldsymbol{\gamma}^{\alpha\alpha} + \boldsymbol{\gamma}^{\beta\beta})^2, \quad (19a)$$

$$= \frac{1}{4}(\boldsymbol{\gamma}^{\alpha\alpha} - \boldsymbol{\gamma}^{\beta\beta})^2, \quad (19b)$$

$$= \mathbf{M}^2. \quad (19c)$$

This is one consequence of the idempotence of $\boldsymbol{\gamma}^{\alpha\alpha}$ and $\boldsymbol{\gamma}^{\beta\beta}$. The second is

$$\mathbf{P}\mathbf{M} + \mathbf{M}\mathbf{P} = \mathbf{M}. \quad (20)$$

Note that the condition of Eq. (19) is the same as the CPMFT condition of Eq. (3b), again with \mathbf{M} taking the role of \mathbf{K} . Both the magnetization density matrix \mathbf{M} and the anomalous density matrix \mathbf{K} are Hermitian.

While CPMFT and UHF thus use the same energy expression (one with \mathbf{K} and the other with \mathbf{M}), \mathbf{K} and \mathbf{M} are not the same even though with the same density matrix \mathbf{P} , we have $\mathbf{K}^2 = \mathbf{M}^2$. There are also some other important differences. Both UHF and CPMFT impose an additional condition on these two matrices, which in UHF is given in Eq. (20) while in CPMFT is instead given in Eq. (3a). Additionally, \mathbf{K} is positive semidefinite while \mathbf{M} is traceless (and thus has both positive and negative eigenvalues). Finally, because in UHF we write \mathbf{P} as the half-sum of two idempotent ma-

trices, its eigenvalues occur in what is known as “corresponding pairs” n_i and $1-n_i$,^{14,15} a terminology that we here adopt.

That UHF has the corresponding pairs property has little to do with UHF per se. It originates simply from the observation¹⁶ that the eigenvalues of a matrix that is the half-sum of two idempotent matrices are 0, 1, $\frac{1}{2}$, or a corresponding pair $(n, 1-n)$. Similarly, the eigenvalues of a matrix written as the half-difference of two idempotent matrices are 0, $\pm\frac{1}{2}$, or a corresponding pair $(-n, n)$.¹⁷ Thus, for example, \mathbf{M} has eigenvalues adding to zero in pairs while \mathbf{P} has eigenvalues adding to one in pairs. Quite generally, any noninteger eigenvalues of the charge density matrix from a single determinant method will be either $\frac{1}{2}$ or occur in a corresponding pair. Eigenvalues of $\frac{1}{2}$ could be part of a corresponding pair (for entangled electrons) or may occur singly for open shells. We should be clear that while matrices written as the sum of two idempotent matrices exhibit the corresponding pairs property, the converse is not necessarily true; a matrix whose eigenvalues come in corresponding pairs may or may not be the sum of two idempotents.

Unlike UHF, the eigenvalues of \mathbf{P} in DH-CPMFT do not occur in corresponding pairs (except when the active space consists of two spatial orbitals). That said, the corresponding pairs property has some attractive features for CPMFT. Most important is that it eliminates overcorrelation between orbital pairs in different symmetries. This is ubiquitous for example in N_2 , where the variational principle drives occupancy into orbitals at low energies and one must introduce multiple chemical potentials to retain the correct total number of σ and π electrons. A corresponding pair constraint controls this unphysical “spilling” and has the inherent attractive feature of limiting strong correlations to be an affair between orbital pairs.

Previously, we had introduced the corresponding pairs feature within the DH-CPMFT framework using different chemical potentials (Lagrange multipliers) for different irreducible representations of the system. However, in the general case where no spatial symmetry is present, imposition of this constraint leads to one Lagrange multiplier per orbital pair and a rather complicated nonlinear optimization problem. A more satisfactory and much simpler approach, however, is to write the CPMFT density matrix as

$$\mathbf{P} = \frac{1}{2}(\mathbf{A} + \mathbf{B}), \quad (21)$$

where \mathbf{A} and \mathbf{B} are auxiliary density matrices, individually idempotent and Hermitian ($\mathbf{A}^2 = \mathbf{A} = \mathbf{A}^\dagger$ and similarly for \mathbf{B}). As with UHF, the decomposition above enforces the corresponding pairs condition automatically, and there is no need to enforce this condition via Lagrange multipliers. Eigenvalues of 0 or 1 in \mathbf{P} correspond to virtual or core orbitals, respectively, while paired eigenvalues correspond to active orbitals. Further, by choosing \mathbf{A} and \mathbf{B} to trace to half the number of electrons, we guarantee that \mathbf{P} does likewise, and we thus have no need of any chemical potential. By making this decomposition, in other words, we can avoid the Lagrange multipliers of the double-Hamiltonian approach entirely, and thus simplify the computation. Note that once we have converged solutions for \mathbf{A} and \mathbf{B} (and thus \mathbf{P} and

\mathbf{K}), we could, if desired, extract the Lagrange multipliers of the DH-CPMFT approach.

The critical mathematical difference between CPMFT as formulated in this manner and UHF is that in UHF, we get \mathbf{M} from the spin-up and spin-down density matrices, while in CPMFT, we get \mathbf{K} from the total density matrix alone (since \mathbf{K} satisfies the condition of Eq. (3b), commutes with \mathbf{P} , and is positive semi-definite). In other words, CPMFT with corresponding pairs defines \mathbf{P} from \mathbf{A} and \mathbf{B} as in Eq. (21), but differs from UHF in constructing

$$\mathbf{K} = \sqrt{\mathbf{P} - \mathbf{P}^2} = \frac{1}{2}\sqrt{(\mathbf{A} - \mathbf{B})^2} = \frac{1}{2}|\mathbf{A} - \mathbf{B}| \quad (22)$$

from auxiliary density matrices \mathbf{A} and \mathbf{B} while UHF builds \mathbf{P} and \mathbf{M} from $\gamma^{\alpha\alpha}$ and $\gamma^{\beta\beta}$, as shown in Eq. (16). Note in the last equation our definition of the absolute value of a matrix from the square root of the square. In practice, to calculate the absolute value of a matrix one needs to diagonalize it, flip the sign of the negative eigenvalues and transform back to the original basis. Both the square root and absolute value of a matrix are positive definite matrices and both have a convergent polynomial series expansion if the matrix is positive definite with eigenvalues between 0 and 1, as is the case here.

To make the comparison between CPMFT and UHF more concrete, consider the case where \mathbf{A} and \mathbf{B} are 2×2 matrices and let $\mathbf{M} = \frac{1}{2}(\mathbf{A} - \mathbf{B})$. Idempotency of \mathbf{A} and \mathbf{B} requires that in the natural orbital basis we have

$$\mathbf{A} = \begin{pmatrix} n & k \\ k & 1-n \end{pmatrix}, \quad (23a)$$

$$\mathbf{B} = \begin{pmatrix} n & -k \\ -k & 1-n \end{pmatrix}, \quad (23b)$$

$$\mathbf{P} = \begin{pmatrix} n & 0 \\ 0 & 1-n \end{pmatrix}, \quad (23c)$$

$$\mathbf{M} = \begin{pmatrix} 0 & k \\ k & 0 \end{pmatrix}, \quad (23d)$$

$$\mathbf{K} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}, \quad (23e)$$

$$k = \sqrt{n(1-n)}. \quad (23f)$$

When \mathbf{A} and \mathbf{B} are of larger dimension, then in the natural orbital basis they are block diagonal with 2×2 blocks of the form given above. This is essentially a consequence of Eq. (20), which in the natural basis becomes

$$(n_i + n_j)M_{ij} = M_{ij} = \frac{1}{2}(A_{ij} - B_{ij}), \quad (24)$$

the solutions to which are $M_{ij} = 0$ and $n_i + n_j = 1$. Because we also have $A_{ij} + B_{ij} = 2n_i\delta_{ij}$, we conclude that for $i \neq j$, we must either have $A_{ij} = B_{ij} = 0$ or $n_i + n_j = 1$ (in other words, the two eigenvalues form a corresponding pair). When the occupation numbers are degenerate, the natural orbitals are not uniquely defined and thus we can choose them such that \mathbf{A}

and \mathbf{B} still have this structure. In the core (virtual) space, $\mathbf{A}=\mathbf{B}=1$ ($\mathbf{A}=\mathbf{B}=\mathbf{0}$).

Before we continue to the working equations for CPMFT in this UHF-like framework, let us pause to make it explicit that CPMFT and UHF are different methods. While we have expressed the UHF energy as a density matrix functional, we could also write it as an expectation value,

$$E_{\text{UHF}} = \langle \Phi_{\text{UHF}} | H | \Phi_{\text{UHF}} \rangle, \quad (25)$$

with $|\Phi_{\text{UHF}}\rangle$ constrained to be a single determinant. This is not true of the CPMFT energy expression, and in fact there seems to be no wave function associated with CPMFT. This may seem somewhat surprising, in light of the intimate connection between CPMFT and HFB theory, in which there certainly is a wave function, albeit one which violates particle number conservation. As we have said, we lose the HFB wave function because we have by fiat changed the sign of the pairing energy. Additionally, unlike UHF, the spin density is zero everywhere for closed shells, even in the presence of static correlation.

One might wonder whether CPMFT is equivalent to projected UHF (PUHF). It is not. If one projects the UHF determinant onto a spin eigenfunction, one finds that the charge density matrix of the UHF determinant and the spin-projected state have the same eigenfunctions.¹⁵ Spin projection, in other words, changes only the occupation numbers of the charge density matrix, but not the natural orbitals. The fact that the UHF and CPMFT natural orbitals are different should lay to rest any concerns that CPMFT is just a projected UHF.

Another fundamental difference between CPMFT and UHF is the onset of the appearance of the solution with energy lower than RHF. As shown in our previous paper,² the CPMFT solution for a two-level model system appears when the RHF orbital energy gap reduces to

$$\varepsilon_2 - \varepsilon_1 < \frac{1}{2}\langle 11|11 \rangle + \frac{1}{2}\langle 22|22 \rangle + \langle 11|22 \rangle, \quad (26)$$

whereas the UHF Coulson-Fischer instability point is determined by

$$\varepsilon_2 - \varepsilon_1 < \langle 12|12 \rangle + \langle 11|22 \rangle. \quad (27)$$

Because all two-electron integrals in the equations above are positive, the CPMFT solution appears inevitably when the orbital gap closes and strong correlation is manifest, such as along a dissociation curve.

A. Working equations

Let us now return to the solution of the CPMFT equations in this UHF-like framework. For convenience, we repeat the energy expression here,

$$E_{\text{CPMFT}} = E_{\text{cs}} - \langle ij|kl \rangle K_{ij} K_{kl}. \quad (28)$$

We simply minimize the energy with respect to (idempotent) \mathbf{A} and \mathbf{B} matrices. The derivatives of E_{cs} in Eq. (28) with respect to \mathbf{A} and \mathbf{B} give the usual closed-shell Fock matrix obtained from \mathbf{P} . That is

$$\frac{\partial E_{\text{cs}}}{\partial A_{ij}} = \frac{\partial E_{\text{cs}}}{\partial B_{ij}} = \frac{1}{2} \frac{\partial E_{\text{cs}}}{\partial P_{ij}} = F_{ij}^{\text{cs}}. \quad (29)$$

The differences with UHF arise from differentiating the last term of the CPMFT energy, which we shall call E_c^{CPMFT} . Taking derivatives with respect to \mathbf{A} leads to an effective potential $\tilde{\Delta}$, given by

$$\tilde{\Delta}_{ij} = \frac{\partial E_c^{\text{CPMFT}}}{\partial A_{ij}} = \frac{\partial E_c^{\text{CPMFT}}}{\partial K_{kl}} \frac{\partial K_{kl}}{\partial A_{ij}} = -2\Delta_{kl} \frac{\partial K_{kl}}{\partial A_{ij}}. \quad (30)$$

This is essentially the same result that we get from differentiating E_c^{UHF} of Eq. (18),

$$\frac{\partial E_c^{\text{UHF}}}{\partial \gamma_{ij}^{\alpha\alpha}} = \frac{\partial E_c^{\text{UHF}}}{\partial M_{kl}} \frac{\partial M_{kl}}{\partial \gamma_{ij}^{\alpha\alpha}} = -2\Delta_{kl}^{\text{UHF}} \frac{\partial M_{kl}}{\partial \gamma_{ij}^{\alpha\alpha}}, \quad (31)$$

where

$$\Delta_{kl}^{\text{UHF}} = \langle km|nl \rangle M_{mn} = \langle kl|mn \rangle M_{mn} \quad (32)$$

looks just like Δ except we replace \mathbf{K} with \mathbf{M} . In UHF, however, we simply have

$$\frac{\partial M_{kl}}{\partial \gamma_{ij}^{\alpha\alpha}} = \frac{1}{2} \delta_{ik} \delta_{jl}, \quad (33)$$

while in CPMFT the derivative of \mathbf{K} with respect to \mathbf{A} is obtained by differentiating both sides of $\mathbf{K}^2 = \frac{1}{4}(\mathbf{A} - \mathbf{B})^2$. This gives

$$\frac{\partial K_{km}}{\partial A_{ij}} K_{ml} + K_{km} \frac{\partial K_{ml}}{\partial A_{ij}} = \frac{1}{2} (M_{jl} \delta_{ki} + M_{ki} \delta_{jl}). \quad (34)$$

In the natural orbital basis where \mathbf{K} is diagonal with eigenvalues K_i , we have

$$\frac{\partial K_{kl}}{\partial A_{ij}} = \frac{1}{2} \frac{M_{jl} \delta_{ki} + M_{ki} \delta_{jl}}{K_k + K_l}. \quad (35)$$

Thus, in the natural orbital basis the effective potential $\tilde{\Delta}$ is

$$\tilde{\Delta}_{ij} = -\frac{\Delta_{ij} M_{jl}}{K_i + K_l} - \frac{\Delta_{kj} M_{ki}}{K_k + K_j}. \quad (36)$$

Since

$$\frac{\partial \mathbf{K}}{\partial \mathbf{A}} = -\frac{\partial \mathbf{K}}{\partial \mathbf{B}}, \quad (37)$$

the equations we ultimately solve are $[\mathbf{F}^{\mathbf{A}}, \mathbf{A}] = \mathbf{0}$ and $[\mathbf{F}^{\mathbf{B}}, \mathbf{B}] = \mathbf{0}$, where $\mathbf{F}^{\mathbf{A}}$ and $\mathbf{F}^{\mathbf{B}}$ are effective Fock matrices given by

$$\mathbf{F}^{\mathbf{A}} = \mathbf{F}^{\text{cs}} + \tilde{\Delta}, \quad (38a)$$

$$\mathbf{F}^{\mathbf{B}} = \mathbf{F}^{\text{cs}} - \tilde{\Delta}. \quad (38b)$$

At first glance, the right-hand-side of Eq. (36) might appear to be divergent unless all K_i are nonzero. However, since forcing $\Delta_{ij}=0$ actually gives the condition $K_{ij}=0$, we simply set $\Delta_{ij}=0$ for the inactive-inactive (core and virtual) block where \mathbf{K} must be zero (because the occupation numbers are 0 or 1). Therefore, in Eq. (36), such divergent terms due to inactive orbitals are simply removed from the sum.

TABLE II. CPMFT energies of N_2 at $R=2.0$ Å. Also included are the number of diagonalization steps required, N_{diag} , and the number of SCF cycles required for convergence.

Scheme	Energy (a.u.)	N_{diag}	SCF cycles
DH-CPMFT(6,6) ^a	-108.799 017 62	118	32
DH-CPMFT(6,6) ^b	-108.797 154 42	121	34
CPMFT(6)	-108.797 154 42	12	12

^aSingle chemical potential.

^bCorresponding pairs enforced by multiple chemical potentials.

IV. RESULTS

We have implemented this version of CPMFT in the Gaussian suite of programs.¹⁸ Each calculation requires the specification of the number N_{act} of active natural orbitals. Due to the corresponding pairs constraint, the number of active electrons is always equal to N_{act} —in other words, we always work at half-filling. In order to obtain an appropriate initial guess for \mathbf{A} and \mathbf{B} , we mix the coefficients of the N_{act} orbitals closest to the Fermi level, just as one would do to break spatial symmetries in UHF. The natural orbital pairs closest to the Fermi energy correspond to those whose occupations are closest to half and half.

In single bond systems where we normally choose the active space to be two electrons in two orbitals, the corresponding pair constraint is automatically satisfied, and no difference is observed between the results using the present approach and those using our previous double-Hamiltonian approach (that is, diagonalization of the double-Hamiltonian constructed from \mathbf{F} and $\mathbf{\Delta}$). However, in DH-CPMFT, one must adjust the chemical potential μ at every iteration of the self-consistent field (SCF) procedure to control the number of electrons in the active space. Because we must adjust the chemical potential, we must diagonalize the double Hamiltonian of Eq. (11) several times in each SCF cycle, until the resulting density matrix has the proper trace. In contrast, the current approach requires no chemical potential, since we have $\text{Tr}(\mathbf{P})=1/2 \text{Tr}(\mathbf{A}+\mathbf{B})$. Because both \mathbf{A} and \mathbf{B} trace to the correct number of electrons, so too does \mathbf{P} . This is a significant operational advantage of the present implementation.

For systems with larger active spaces, the present approach differs from DH-CPMFT, although as mentioned above, we can impose the corresponding pairs constraint in DH-CPMFT in some special cases by including different chemical potentials for different irreducible representations. We illustrate this with the case of N_2 . Table II shows the total energy of N_2 at 2 Å. We use the cc-pVTZ basis set and choose six active orbitals and six active electrons. The current scheme gives a slightly higher energy than does DH-CPMFT with only one chemical potential, as one would expect since we have imposed an additional constraint on the system. Also as one would expect, it gives the same results as does DH-CPMFT with the corresponding pairs constraint enforced by additional Lagrange multipliers. However, removing the chemical potentials results in considerable computational savings. In Fig. 1, we show the N_2 dissociation curves from CPMFT in the double-Hamiltonian approach

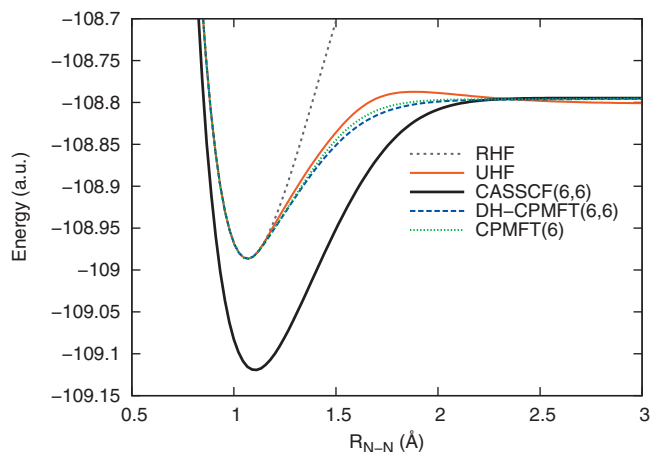


FIG. 1. Potential energy curves of N_2 calculated with the cc-pVTZ basis set.

and in the corresponding pairs framework. In this case, the corresponding pairs constraint has only a minor effect on the energy.

We have also performed a CPMFT calculation of the C_2 molecule with the 6-31G basis set. Near equilibrium, C_2 has significant static correlation due to near-degeneracy between the RHF occupied σ_{2s}^* and unoccupied σ_{2p_z} orbitals. As the molecule is stretched, however, the π_x , π_y , π_x^* , and π_y^* orbitals become degenerate, while the $\sigma_{2s}^*-\sigma_{2p_z}$ interaction becomes weak. We have therefore chosen our active space to be six electrons in six orbitals for this system. In Fig. 2 we show the total energy of C_2 as a function of bond length. The CASSCF energy includes all static correlation that results from these orbital interactions (plus some dynamical correlation). Without the corresponding pairs constraint, DH-CPMFT strongly overcorrelates nearly everywhere. Adding the corresponding pairs constraint significantly reduces this overcorrelation. Near equilibrium, it gives results between UHF and CASSCF. Unfortunately, it still overcorrelates as the molecule dissociates. This is due to electron “spilling” between σ_{2s}^* and σ_{2p_z} orbitals. As $R \rightarrow \infty$, only the π orbitals should be strongly correlated; including these σ orbitals in the active space at large internuclear separation allows them to correlate and lower the energy unphysically. If we remove two orbitals from the active space, we produce the curve

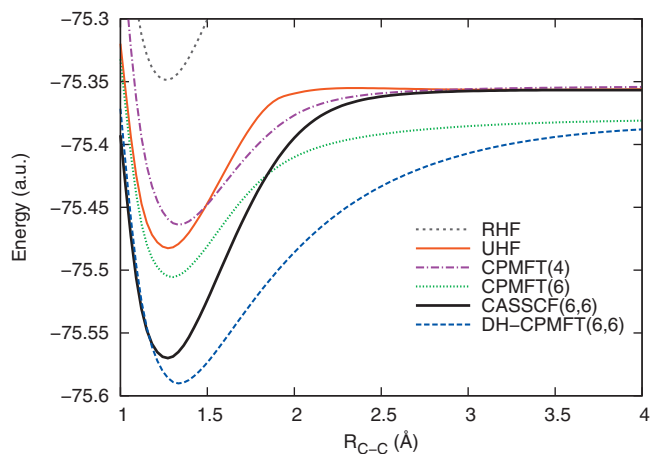


FIG. 2. Potential energy curves of C_2 calculated with the 6-31G basis set.

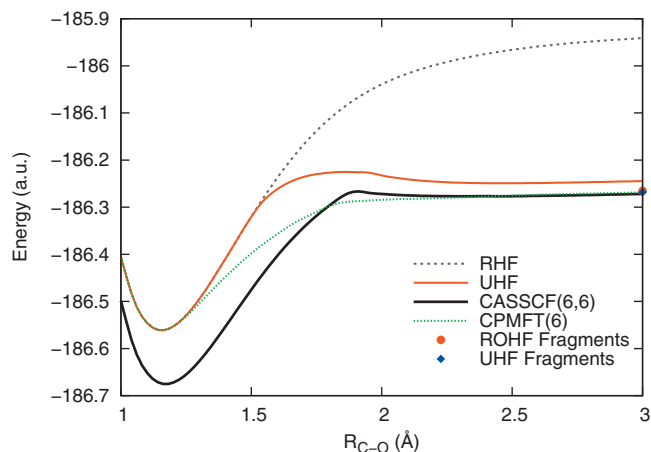


FIG. 3. Potential energy curves for the double dissociation of CO_2 calculated with the 3-21G basis set.

marked CPMFT(4). This goes to the correct dissociation limit, but undercorrelates at equilibrium where the active space should be larger. The correct solution for this molecule involves introducing renormalized one-body potentials in CPMFT(6) that eliminate the spilling at dissociation,² an approach that we will discuss in a forthcoming article. While going to the right dissociation limit is important, it is perhaps less critical than getting the correct behavior near equilibrium. Note that CPMFT(4) dissociates correctly to two ROHF carbon atoms, while UHF instead dissociates to two spin-contaminated UHF carbon atoms and CASSCF(6,6) has some dynamical correlation at dissociation.

Finally, we stress the differences between UHF and CPMFT by analyzing the dissociation of the CO_2 molecule. The ground state of CO_2 near equilibrium is a closed-shell singlet with no expected static correlation. Indeed both UHF and CPMFT reduce to the RHF solution near R_e . However, when the molecule is symmetrically stretched and the two oxygen atoms are simultaneously separated from the carbon atom, the correct dissociation limit corresponds to all three atoms in their triplet ground state. This situation cannot be handled by UHF. In CO_2 near R_e , there are six electrons associated with bond formation, three with spin-up and three with spin-down. At dissociation, UHF might assign two spin-up electrons to one oxygen atom and two spin-down electrons on the other, which puts both oxygen atoms in their triplet ground state. However, with only one electron of each spin remaining, the best UHF can do is to assign a singlet state to the carbon atom, which is clearly incorrect and not the lowest energy state. In simple words, UHF runs out of broken symmetry degrees of freedom (has only two) to model the dissociation of CO_2 (Fig. 3) and misses the correct dissociation limit by ~ 20 mhartree. The bumps in the dissociation curves correspond to crossings of different solutions to the respective SCF equations and we have plotted the lowest energy state at each R . Because spin states are treated in CPMFT through an “ensemble” representation, one that yields zero spin magnetization density everywhere, the CPMFT solution for this dissociation has two half spins up and two half spins down on each of the three atoms, leading to the correct energy corresponding to the sum of ROHF atomic

energies. Note that CPMFT(6) in Fig. 3 contains a one-body potential arising from an asymptotic constraint as explained in our previous publication.² We defer detailed discussion of the renormalization schemes used in CO_2 and applicable to C_2 within the current UHF-like context to a forthcoming publication.

V. CONCLUSIONS

We have developed a novel scheme for performing CPMFT calculations with occupation numbers occurring in corresponding pairs. In doing so, we eliminate all chemical potentials, and the effective Fock matrices \mathbf{F}^A and \mathbf{F}^B that are to be diagonalized are of half the dimension of the double Hamiltonian matrix in the previous DH-CPMFT scheme. Thus, the computational effort in our present implementation is greatly reduced over the previous formulation of CPMFT. The corresponding pairs constraint reduces the overcorrelation of C_2 near equilibrium, and has important consequences for the dissociation of heteronuclear systems. While the corresponding pair constraint could also be imposed in the DH-CPMFT framework by addition of one Lagrange multiplier per electron pair, the current approach imposes this constraint in a simpler black-box manner.

We have shown that this version of CPMFT is closely related to UHF theory. Unlike UHF, however, CPMFT incorporates static correlation by a different mechanism. The physical density matrix γ has identical spin-up and spin-down blocks, whereas the auxiliary \mathbf{A} and \mathbf{B} density matrices, in general, break symmetry. CPMFT can correctly dissociate polyatomic molecules into ROHF atoms or fragments, whereas UHF has problems with multiple entangled electrons at multiple centers, as shown for CO_2 above. In the present formulation, CPMFT becomes a density matrix functional that can be solved by diagonalization of effective Fock matrices providing orbitals and orbital energies. We wish to emphasize one more time that as we have demonstrated, a quasiparticle picture of strong correlations with the sign of the pairing interaction reversed yields an energy expression reminiscent of UHF.

Finally, we should note that in CPMFT different auxiliary \mathbf{A} and \mathbf{B} density matrices can lead to solutions with degenerate energies. The key quantities determining the energy in the model are \mathbf{P} and \mathbf{K} and there is a many-to-one mapping between \mathbf{A} and \mathbf{B} on the one hand and \mathbf{P} and \mathbf{K} on the other. At dissociation, for example, solutions where \mathbf{A} and \mathbf{B} orbitals are localized and delocalized (roughly corresponding to UHF and RHF orbitals) are degenerate. The existence of additional degenerate solutions in CPMFT (compared to UHF) can lead to convergence difficulties as the active space becomes large. Efficient ways of dealing with the additional degrees of freedom provided by the auxiliary \mathbf{A} and \mathbf{B} matrices are currently under investigation.

ACKNOWLEDGMENTS

This work was supported by NSF (Grant No. CHE-0807194), LANL (Subcontract No. 81277-001-10), and the Welch Foundation (Grant No. C-0036). A.S. acknowledges support from ANR (Grant No. 07-BLAN-0272). We thank

Carlos Jiménez–Hoyos and Jason Ellis for useful discussions and Dan Sorensen for pointing out Ref. 16.

APPENDIX: PROPERTIES OF THE CPMFT MODEL TWO-PARTICLE DENSITY MATRIX

The CPMFT model two-particle density matrix is

$$(\Gamma_{\text{CPMFT}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k - \frac{1}{2}\kappa_{ij}\kappa^{kl}, \quad (\text{A1})$$

where i, j, k , and l are spin-orbitals and γ and κ are the density matrix and anomalous density matrix in the spin-orbital basis (*i.e.*, they are of dimension $2N \times 2N$, where N is the size of the atomic orbital basis). In general, γ is Hermitian and κ is antisymmetric. When everything is real (which we take for simplicity; this does not affect our conclusions), the idempotent HFB quasiparticle density matrix is

$$\mathbf{R} = \begin{pmatrix} \gamma & \kappa \\ -\kappa & \mathbf{1} - \gamma \end{pmatrix}. \quad (\text{A2})$$

Idempotency tells us that

$$\gamma\kappa - \kappa\gamma = \mathbf{0}, \quad (\text{A3a})$$

$$\gamma^2 - \kappa^2 = \gamma. \quad (\text{A3b})$$

We recall that for closed shells,⁵

$$\gamma = \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} \end{pmatrix}, \quad (\text{A4a})$$

$$\kappa = \begin{pmatrix} \mathbf{0} & \mathbf{K} \\ -\mathbf{K} & \mathbf{0} \end{pmatrix}, \quad (\text{A4b})$$

$$\mathbf{0} = \mathbf{PK} - \mathbf{KP}, \quad (\text{A4c})$$

$$\mathbf{P} = \mathbf{P}^2 + \mathbf{K}^2. \quad (\text{A4d})$$

We can define an analogous model two-particle density matrix for HFB, for which all the conditions on κ , γ , \mathbf{K} , and \mathbf{P} are the same, but where

$$(\Gamma_{\text{HFB}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k + \frac{1}{2}\kappa_{ij}\kappa^{kl}. \quad (\text{A5})$$

Finally, the UHF two-particle density matrix is

$$(\Gamma_{\text{UHF}})_{ij}^{kl} = \frac{1}{2}\gamma_i^k\gamma_j^l - \frac{1}{2}\gamma_i^l\gamma_j^k, \quad (\text{A6})$$

where γ is idempotent. We have

$$\gamma = \begin{pmatrix} \gamma^{\alpha\alpha} & \mathbf{0} \\ \mathbf{0} & \gamma^{\beta\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{P} + \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{P} - \mathbf{M} \end{pmatrix}, \quad (\text{A7a})$$

$$\mathbf{P} = \mathbf{P}^2 + \mathbf{M}^2, \quad (\text{A7b})$$

$$\mathbf{M} = \mathbf{PM} + \mathbf{MP}. \quad (\text{A7c})$$

1. Partial trace of the two-particle density matrix

An important condition on the two-particle density matrix is that it traces to the one-particle density matrix. That is, we must have

$$\Gamma_{ij}^{il} = \frac{N-1}{2}\gamma_j^j. \quad (\text{A8})$$

We remind the reader that repeated indices are to be summed.

The partial trace condition is satisfied by the UHF two-matrix and the CPMFT model two-matrix, but not by the HFB model two-matrix,

$$\Gamma_{ij}^{il} = \frac{1}{2}(\gamma_i^j\gamma_j^i - \gamma_i^j\gamma_j^i \mp \kappa_{ij}\kappa^{il}), \quad (\text{A9a})$$

$$= \frac{1}{2}[N\gamma_j^j - (\gamma^2)_j^j \pm (\kappa^2)_j^j], \quad (\text{A9b})$$

$$= \frac{1}{2}[N\gamma_j^j - (\gamma + \kappa)_j^j \pm (\kappa^2)_j^j], \quad (\text{A9c})$$

$$= \frac{N-1}{2}\gamma_j^j - \frac{1}{2}[(\kappa^2)_j^j \mp (\kappa^2)_j^j]. \quad (\text{A9d})$$

Here, the top (bottom) sign in \pm and \mp corresponds to CP-MFT (HFB), and we have used antisymmetry of κ . Explicitly, we have

$$(\Gamma_{\text{CPMFT}})_{ij}^{il} = \frac{N-1}{2}\gamma_j^j, \quad (\text{A10a})$$

$$(\Gamma_{\text{HFB}})_{ij}^{il} = \frac{N-1}{2}\gamma_j^j - (\kappa^2)_j^j. \quad (\text{A10b})$$

Note that by N we mean the trace of the one-particle density matrix γ , which should be the number of particles in the system.

2. Particle number fluctuations

In order to work out particle number fluctuations, we need the expectation values of \hat{N} and \hat{N}^2 , with \hat{N} the number operator, given as

$$\hat{N} = \delta_{pq}a_p^\dagger a_q. \quad (\text{A11})$$

We have already noted that the expectation value of \hat{N} is just $\text{Tr}(\gamma)$. The expectation value of \hat{N}^2 requires the two-particle density matrix,

$$\langle \hat{N}^2 \rangle = \delta_{pq}\delta_{rs}\langle a_p^\dagger a_q a_r^\dagger a_s \rangle, \quad (\text{A12a})$$

$$= \delta_{pq}\delta_{rs}(-\langle a_p^\dagger a_r^\dagger a_q a_s \rangle + \delta_{qr}\langle a_p^\dagger a_s \rangle), \quad (\text{A12b})$$

$$= \delta_{pq}\delta_{rs}(2\Gamma_{pr}^{qs} + \delta_{qr}\gamma_p^s), \quad (\text{A12c})$$

$$= 2\Gamma_{pr}^{pr} + \gamma_p^p. \quad (\text{A12d})$$

If the two-particle density matrix obeys the partial trace condition, the particle number fluctuations are automatically zero. This is thus true of UHF and of CPMFT. However, HFB has particle number fluctuations,

$$\langle \hat{N}^2 \rangle_{\text{HFB}} = (N-1)\gamma_j^j - 2(\kappa^2)_j^j + \gamma_j^j = N^2 - 2\text{Tr}(\kappa)^2, \quad (\text{A13})$$

implying that

$$\sigma_N^2 = \langle \hat{N}^2 \rangle - \langle \hat{N} \rangle^2 = -2 \text{Tr}(\boldsymbol{\kappa}^2). \quad (\text{A14})$$

Note that this is positive, as it should be, since $-\boldsymbol{\kappa}^2 = \boldsymbol{\gamma} - \boldsymbol{\gamma}^2$ and occupation numbers are between zero and one, inclusive. In the closed-shell case, we have $\sigma_N^2 = 4 \text{Tr}(\mathbf{K}^2)$.

3. Spin contamination

Evaluating spin contamination is more complicated than evaluating particle number fluctuations, not least because we need an expression for $\langle \hat{S}^2 \rangle$ for a general two-particle density matrix Γ . We begin by noting that

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2, \quad (\text{A15a})$$

$$= \hat{S}_z + \hat{S}_z^2 + \hat{S}_- \hat{S}_+, \quad (\text{A15b})$$

where \hat{S}_\pm is the spin raising/lowering operator. We are interested here in the closed-shell case (i.e., $N_\alpha = N_\beta$ with a block diagonal $\boldsymbol{\gamma}$).

In the closed-shell case, the contribution to $\langle \hat{S}^2 \rangle$ from the first term is zero. We must evaluate the contribution from the next piece using our model two-particle density matrix. We have

$$\hat{S}_z^2 = \underbrace{\sum_i \hat{s}_z(i)^2}_{\hat{X}_z} + \underbrace{\sum_{i \neq j} \hat{s}_z(i) \hat{s}_z(j)}_{\hat{Y}_z} \quad (\text{A16})$$

The first (second) term is a one-particle (two-particle) operator. Note that we could also write

$$\hat{Y}_z = 2 \sum_{i > j} \hat{s}_z(i) \hat{s}_z(j), \quad (\text{A17})$$

which explains the factor of two that might otherwise appear to be missing below.

Evaluating the contribution to $\langle \hat{S}_z^2 \rangle$ from \hat{X}_z is straightforward, and we get just

$$\langle \hat{X}_z \rangle = \frac{1}{4}(N_\alpha + N_\beta) = \frac{1}{2} \text{Tr}(\mathbf{P}). \quad (\text{A18})$$

The nonzero matrix elements of \hat{Y}_z are

$$(Y_z)_{k_\alpha l_\alpha}^{i_\alpha j_\alpha} = \langle i_\alpha j_\alpha | \hat{Y}_z | k_\alpha l_\alpha \rangle = \frac{1}{2} \delta_k^i \delta_l^j, \quad (\text{A19a})$$

$$(Y_z)_{k_\alpha l_\beta}^{i_\alpha j_\beta} = \langle i_\alpha j_\beta | \hat{Y}_z | k_\alpha l_\beta \rangle = -\frac{1}{2} \delta_k^i \delta_l^j, \quad (\text{A19b})$$

$$(Y_z)_{k_\beta l_\alpha}^{i_\beta j_\alpha} = \langle i_\beta j_\alpha | \hat{Y}_z | k_\beta l_\alpha \rangle = -\frac{1}{2} \delta_k^i \delta_l^j, \quad (\text{A19c})$$

$$(Y_z)_{k_\beta l_\beta}^{i_\beta j_\beta} = \langle i_\beta j_\beta | \hat{Y}_z | k_\beta l_\beta \rangle = \frac{1}{2} \delta_k^i \delta_l^j. \quad (\text{A19d})$$

Here, we are working in an orthonormal basis set.

The relevant components of the CPMFT and HFB two-particle density matrices are

$$\Gamma_{i_\alpha l_\alpha}^{k_\alpha j_\alpha} = \frac{1}{2}(\gamma_{i_\alpha}^{k_\alpha} \gamma_{j_\alpha}^{l_\alpha} - \gamma_{i_\alpha}^{l_\alpha} \gamma_{j_\alpha}^{k_\alpha}), \quad (\text{A20a})$$

$$\Gamma_{i_\alpha l_\beta}^{k_\alpha j_\beta} = \frac{1}{2}(\gamma_{i_\alpha}^{k_\alpha} \gamma_{j_\beta}^{l_\beta} \mp \kappa_{i_\alpha j_\beta} \boldsymbol{\kappa}^{k_\alpha l_\beta}), \quad (\text{A20b})$$

$$\Gamma_{i_\beta l_\alpha}^{k_\beta j_\alpha} = \frac{1}{2}(\gamma_{i_\beta}^{k_\beta} \gamma_{j_\alpha}^{l_\alpha} \mp \kappa_{i_\beta j_\alpha} \boldsymbol{\kappa}^{k_\beta l_\alpha}), \quad (\text{A20c})$$

$$\Gamma_{i_\beta l_\beta}^{k_\beta j_\beta} = \frac{1}{2}(\gamma_{i_\beta}^{k_\beta} \gamma_{j_\beta}^{l_\beta} - \gamma_{i_\beta}^{l_\beta} \gamma_{j_\beta}^{k_\beta}), \quad (\text{A20d})$$

where the top (bottom) sign corresponds to CPMFT (HFB).

Contracting the density matrices with the matrix elements, we get

$$\langle \hat{Y}_z \rangle = \left(\frac{N_\alpha - N_\beta}{2} \right)^2 - \frac{1}{4} \text{Tr}(\boldsymbol{\gamma}_{\alpha\alpha}^2 + \boldsymbol{\gamma}_{\beta\beta}^2 \mp \boldsymbol{\kappa}_{\alpha\beta}^2 \mp \boldsymbol{\kappa}_{\beta\alpha}^2), \quad (\text{A21})$$

where we have used antisymmetry of $\boldsymbol{\kappa}$. Working in our closed-shell case, this reduces to

$$\langle \hat{Y}_z \rangle = -\frac{1}{2} \text{Tr}(\mathbf{P}^2 \mp \mathbf{K}^2). \quad (\text{A22})$$

In total, we find that $\langle \hat{S}_z^2 \rangle$ in CPMFT and HFB is given by

$$\langle \hat{S}_z^2 \rangle = \frac{1}{2} \text{Tr}(\mathbf{P} - \mathbf{P}^2 \pm \mathbf{K}^2), \quad (\text{A23a})$$

$$= \frac{1}{2} \text{Tr}(\mathbf{K}^2 \pm \mathbf{K}^2). \quad (\text{A23b})$$

Thus, we end up with

$$\langle \hat{S}_z^2 \rangle_{\text{HFB}} = 0, \quad (\text{A24a})$$

$$\langle \hat{S}_z^2 \rangle_{\text{CPMFT}} = \text{Tr}(\mathbf{K}^2). \quad (\text{A24b})$$

The contribution to $\langle \hat{S}^2 \rangle$ from $\hat{S}_- \hat{S}_+$ must also be evaluated using the model two-particle density matrix. Expanding this operator in terms of contributions from individual electrons, we have

$$\hat{S}_- \hat{S}_+ = \underbrace{\sum_i \hat{s}_-(i) \hat{s}_+(i)}_{\hat{X}} + \underbrace{\sum_{i \neq j} \hat{s}_-(i) \hat{s}_+(j)}_{\hat{Y}} \quad (\text{A25})$$

The first term is the one-particle operator \hat{X} , and the second is the two-particle operator \hat{Y} .

Since \hat{X} does nothing to down-spin electrons but annihilates up-spin electrons, we clearly have

$$\langle \hat{X} \rangle = N_\beta = \text{Tr}(\mathbf{P}). \quad (\text{A26})$$

To take the expectation value of \hat{Y} , it proves useful to symmetrize it so that it acts the same on the two electrons. Since operators acting on different electrons commute, we have

$$\hat{Y} = \sum_{i \neq j} \hat{s}_-(i) \hat{s}_+(j), \quad (\text{A27a})$$

$$= \frac{1}{2} \sum_{i \neq j} (\hat{s}_-(i) \hat{s}_+(j) + \hat{s}_+(i) \hat{s}_-(j)), \quad (\text{A27b})$$

$$= \sum_{i > j} (\hat{s}_-(i) \hat{s}_+(j) + \hat{s}_+(i) \hat{s}_-(j)). \quad (\text{A27c})$$

The only nonzero matrix elements of \hat{Y} are

$$Y_{k_\alpha l_\beta}^{i_\beta j_\alpha} = \langle i_\beta j_\alpha | \hat{Y} | k_\alpha l_\beta \rangle = \delta_k^i \delta_l^j, \quad (\text{A28a})$$

$$Y_{k\beta\alpha}^{i\alpha j\beta} = \langle i_{\alpha j\beta} | \hat{Y} | k_{\beta l\alpha} \rangle = \delta_k^i \delta_l^j. \quad (\text{A28b})$$

The relevant spin components of the CPMFT and HFB model two-particle density matrix are

$$\Gamma_{i\beta\alpha}^{k\beta l\alpha} = \frac{1}{2} (-\gamma_{i\alpha}^{l\alpha} \gamma_{j\beta}^{k\beta} \mp \kappa_{i\alpha j\beta} \kappa^{k\beta l\alpha}), \quad (\text{A29a})$$

$$\Gamma_{i\beta\alpha}^{k\alpha l\beta} = \frac{1}{2} (-\gamma_{i\beta}^{l\beta} \gamma_{j\alpha}^{k\alpha} \mp \kappa_{i\beta j\alpha} \kappa^{k\alpha l\beta}), \quad (\text{A29b})$$

where again CPMFT (HFB) corresponds to the top (bottom) sign.

Contracting the two-particle density matrix with the matrix elements gives us

$$\langle \hat{Y} \rangle = -\text{Tr}(\gamma_{\alpha\alpha} \gamma_{\beta\beta} \mp \kappa_{\alpha\beta} \kappa_{\alpha\beta}). \quad (\text{A30})$$

In the closed-shell case, using the results in Eq. (A4), this becomes

$$\langle \hat{Y} \rangle = -\text{Tr}(\mathbf{P}^2 \mp \mathbf{K}^2). \quad (\text{A31})$$

Then the expectation value of $\hat{S}_- \hat{S}_+$ is given by

$$\langle \hat{S}_- \hat{S}_+ \rangle = \text{Tr}(\mathbf{P} - \mathbf{P}^2 \pm \mathbf{K}^2), \quad (\text{A32a})$$

$$= \text{Tr}(\mathbf{K}^2 \pm \mathbf{K}^2). \quad (\text{A32b})$$

We therefore have

$$\langle \hat{S}_- \hat{S}_+ \rangle_{\text{HFB}} = 0, \quad (\text{A33a})$$

$$\langle \hat{S}_- \hat{S}_+ \rangle_{\text{CPMFT}} = 2 \text{Tr}(\mathbf{K}^2). \quad (\text{A33b})$$

Combining Eqs. (A24) and (A33) gives us the total spin contamination in HFB and in CPMFT,

$$\langle \hat{S}^2 \rangle_{\text{HFB}} = 0, \quad (\text{A34a})$$

$$\langle \hat{S}^2 \rangle_{\text{CPMFT}} = 3 \text{Tr}(\mathbf{K}^2). \quad (\text{A34b})$$

For UHF in cases in which there is strong correlation, we have the familiar formula

$$\langle \hat{S}^2 \rangle = s(s+1) + N_{\beta} - \text{Tr}(\gamma_{\alpha\alpha} \gamma_{\beta\beta}). \quad (\text{A35})$$

For the closed-shell case, using the results in Eq. (A7), we have

$$\begin{aligned} \langle \hat{S}^2 \rangle &= \text{Tr}[\mathbf{P} - (\mathbf{P} + \mathbf{M})(\mathbf{P} - \mathbf{M})] = \text{Tr}(\mathbf{P} - \mathbf{P}^2 + \mathbf{M}^2) \\ &= 2 \text{Tr}(\mathbf{M}^2). \end{aligned} \quad (\text{A36})$$

¹T. Tsuchimochi and G. E. Scuseria, *J. Chem. Phys.* **131**, 121102 (2009).

²G. E. Scuseria and T. Tsuchimochi, *J. Chem. Phys.* **131**, 164119 (2009).

³T. Tsuchimochi, G. E. Scuseria, and A. Savin, *J. Chem. Phys.* **132**, 024111 (2010).

⁴J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (MIT, Cambridge, MA, 1986).

⁵V. N. Staroverov and G. E. Scuseria, *J. Chem. Phys.* **117**, 11107 (2002).

⁶K. Takatsuka, T. Fueno, and K. Yamaguchi, *Theor. Chim. Acta* **48**, 175 (1978).

⁷V. N. Staroverov and E. R. Davidson, *Chem. Phys. Lett.* **330**, 161 (2000).

⁸I. Mayer, *Chem. Phys. Lett.* **440**, 357 (2007).

⁹W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **110**, 2800 (1999).

¹⁰G. Csányi and T. A. Arias, *Phys. Rev. B* **61**, 7348 (2000).

¹¹J. A. Pople and R. K. Nesbet, *J. Chem. Phys.* **22**, 571 (1954).

¹²Note that UHF subtracts $\langle ij|kl \rangle M_{ij} M_{jk}$ from the closed-shell energy while CPMFT subtracts $\langle ij|kl \rangle K_{ij} K_{kl}$. However, these are the same if the basis functions and the anomalous density matrix are real, as they generally are.

¹³H. Fukutome, *Int. J. Quantum Chem.* **20**, 955 (1981).

¹⁴A. T. Amos and G. G. Hall, *Proc. R. Soc. London* **A263**, 483 (1961).

¹⁵J. E. Harriman, *J. Chem. Phys.* **40**, 2827 (1964).

¹⁶V. Rabanovich, *Numer. Linear Algebra Appl.* **390**, 137 (2004).

¹⁷More precisely, the eigenvalues of $\alpha\mathbf{A} + \beta\mathbf{B}$ for idempotent \mathbf{A} and \mathbf{B} are 0, α , β , or a corresponding pair $(n, \alpha + \beta - n)$.

¹⁸M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN, Revision G.01, Gaussian, Inc., Wallingford, CT, 2007.