Check for updates

Comprehensive benchmarking of density matrix functional approximations†‡

Mauricio Rodríguez-Mayorga,[ab] Eloy Ramos-Cordoba, [iD][ac] Mireia Via-Nadal,[a] Mario Piris[ad] and Eduard Matito [iD] *[ad]

The energy usually serves as a yardstick in assessing the performance of approximate methods in computational chemistry. After all, these methods are mostly used for the calculation of the electronic energy of chemical systems. However, computational methods should be also aimed at reproducing other properties, such strategy leading to more robust approximations with a wider range of applicability. In this study, we suggest a battery of ten tests with the aim to analyze density matrix functional approximations (DMFAs), including several properties that the exact functional should satisfy. The tests are performed on a model system with varying electron correlation, carrying a very small computational effort. Our results not only put forward a complete and exhaustive benchmark test for DMFAs, currently lacking, but also reveal serious deficiencies of existing approximations that lead to important clues in the construction of more robust DMFAs.

1 Introduction

Density matrix functional theory (DMFT) is among the computational methods that have experienced a most important advance in the last years. Its foundations are more than forty years old[1] but the most important progress in the field has occurred in the last twenty years.[2,3] Namely, the use of the natural orbital representation of the first-order reduced density matrix has brought many[4–22] density matrix functional approximations (DMFAs) within the context of what is known in the literature as natural orbital functional theory. Some of these functionals provide very accurate energies, sometimes competing with high-level electronic structure methods.[23–27]

Despite the success, the account of dynamic correlation still poses a great challenge for DMFAs[2,14,23] and calls for means to separate dynamic and nondynamic correlation within DMFT.[22,28,29] The development of more accurate DMFAs also depends on appropriate benchmark tests and, to our knowledge, only a recent paper addresses the validation of most DMFAs in the literature, comparing their performance in the energy calculation of few-electron systems with different electron correlation.[23]

Since functionals are mostly used to calculate the electronic energy of chemical systems, it does not strike as a surprise that the energy usually serves as a yardstick in benchmarking DMFAs. However, it is becoming commonly accepted that energy functionals should be also aimed at reproducing other properties[30] in order to construct more robust approximations with a wider range of applicability. In this line, some of us have recently tested the spin structure of several DMFAs[27,31] using the local spin[32] as a benchmarking tool.

In the present study, we suggest a battery of ten tests to analyze DMFAs, including several properties that the exact functional should satisfy. We submit fifteen functionals to this series of constrictive tests using a model system with varying electron correlation, the two-electron harmonium atom,[33] carrying a very small computational effort.

Although the construction of computational methods that can tackle both one- and two-electron systems is relevant in other frameworks such as density functional theory,[34] it is believed that they do not pose a serious challenge in DMFT. The reason behind this idea is the existence of a quasi-exact closed-shell two-electron expression of the second-order reduced density matrix in terms of natural orbitals.[35] However, there are several facts that go against this idea. First of all, most DMFAs do not reduce to the quasi-exact expression for two-electron systems and, therefore, their calibration is justified. Second, the latter expression depends on some phase factors (*vide infra*) that change importantly under certain circumstances,[36–39] including the strong-correlation regime of

[a] Kimika Fakultatea, Euskal Herriko Unibertsitatea UPV/EHU, and Donostia International Physics Center (DIPC), P.K. 1072, 20080 Donostia, Euskadi, Spain. E-mail: ematito@gmail.com
[b] Institut de Química Computacional i Catàlisi (IQCC) i Departament de Química, University of Girona, 17071 Girona, Catalonia, Spain
[c] Department of Chemistry, University of California Berkeley, 94720, Berkeley, CA, USA
[d] IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Euskadi, Spain

† In memoriam of Jose Ignacio Eguiazabal, researcher and Director of Polymer Science and Technology Department (UPV-EHU).
‡ Electronic supplementary information (ESI) available. See DOI: 10.1039/c7cp03349d

the system studied in this work. Finally, whereas energetic differences between the quasi-exact result and the exact one are often negligible, there are several properties studied in this paper that show non-negligible differences for the quasi-exact functional.

All in all, we shall see that the set of tests suggested in this work poses a great challenge for DMFAs and reveals various defects of the approximations that were hindered by a reasonable performance in energy benchmark tests. The current strategy can be easily extended to a larger number of electrons, thus setting new challenges for the few DMFAs that reduce to the quasi-exact expression. We are confident that DMFT developers will benefit from the results obtained in this paper and will use this test set as a means to construct more robust approximations.

2 Methodology

The second-order reduced density matrix (2-RDM),[40,41]

$$\rho_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1', \mathbf{x}_2') = N(N-1) \int d\mathbf{x}_3 \cdots \int d\mathbf{x}_N$$
$$\times \Psi^*(\mathbf{x}_1', \mathbf{x}_2', \mathbf{x}_3, \dots, \mathbf{x}_N) \Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N), \tag{1}$$

where we have adopted McWeeny's normalization[41] and $\mathbf{x} = (\mathbf{r}, \sigma)$, is the simplest function in terms of which the explicit expression of the electronic energy of a physical system is known.[42] Hence, approximations to the 2-RDM in terms of simpler quantities provide estimates of the energy that, in principle, should reduce the associated computational cost. Most quantum mechanical calculations employ orbital basis sets, and it is thus costumary to express the 2-RDM in a given orbital basis. In this paper, we adopt the basis of natural orbitals,

$$^2D_{ij,kl}^{\sigma\sigma'} = \left\langle \Psi \left| a_{i\sigma}^\dagger a_{j\sigma'}^\dagger a_{l\sigma'} a_{k\sigma} \right| \Psi \right\rangle, \tag{2}$$

where $a_{i\sigma}^\dagger$ ($a_{i\sigma}$) is the creation (annihilation) operator acting over natural orbital i with spin σ; hereafter we will refer to $^2D_{ij,kl}^{\sigma\sigma'}$ as the two-density matrix (2-DM). The spinless 2-RDM is a twelve-variable function, whereas the 2-DM is a four-index tensor of dimension M^4. In the present study we focus on 2-DM approximations built from natural occupation numbers (ONs), $\{n_i\}_{i=1}^M$, where M is the size of the basis set.[1–3] The approximate 2-DMs here studied are built upon the simplification of the 2-DM being a sparse matrix with only three types of non-zero elements: $^2D_{ij,ij}^{\sigma\sigma}$ and $^2D_{ij,ij}^{\sigma\sigma'}$, $^2D_{ij,ji}^{\sigma\sigma}$ and $^2D_{ij,ji}^{\sigma\sigma'}$, and $^2D_{ii,jj}^{\sigma\sigma'}$. The opposite-spin elements are actually sufficient to express the exact 2-DM of a two-electron closed-shell system (see eqn (12)). Each ON-based 2-DM approximation actually provides a DMFA. Among the DMFAs, the simplest one is the single-determinant (SD) approximation, whose expression reads[40]

$$^2D_{ij,kl}^{\mathrm{SD},\alpha\beta} = n_i^\alpha n_j^\beta \delta_{ik}\delta_{jl} \tag{3}$$

for the opposite-spin elements and

$$^2D_{ij,kl}^{\mathrm{SD},\alpha\alpha} = n_i^\alpha n_j^\alpha (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) \tag{4}$$

for the like-spin ones. Upon optimization of natural orbitals and ONs, the SD approximation produces the Hartree–Fock energy. In this work, we optimize neither the orbitals nor the occupations (*vide infra*) and, in order to avoid confusion with the Hartree–Fock method (which does not employ fractional occupancies), we have preferred to call this approximation SD. We will consider 15 DMFAs, all of which are *JKL*-only functionals, *i.e.*, functionals that only need Coulomb, exchange and time-inversion two-electron integrals.[43] We have classified them in two groups: those that only modify the exchange part of the functional and those that modify both the exchange and Coulomb parts. The latter group corresponds to the functionals developed by one of us and known as Piris natural orbital functionals (PNOFs). The first group of DMFAs uses eqn (3) and

$$^2D_{ij,kl}^{X,\alpha\alpha} = n_i^\alpha n_j^\alpha \delta_{ik}\delta_{jl} - f_X(n_i^\alpha, n_j^\alpha)\delta_{il}\delta_{jk}, \tag{5}$$

the expression of $f_X(n_i, n_j)$ determining the functional.[44] These functionals are *JK*-only functionals but, in practice, they use the simple SD approximation for the terms involving two-electron Coulomb integrals (*J*) and, therefore, they will be referred as *K*-functionals hereafter. Their expressions are collected in Table 1.

PNOFs actually correspond to approximations to the two-particle cumulant matrix ($^2\Gamma$),[11] which is defined as the

Table 1 $f(n_i, n_j)$ functions (see eqn (5)) that define the *K*-functionals. In a two-electron closed-shell system $F_L = 1$

DMFA	$f(n_i, n_j)$		Parameters	Ref.
SD	$n_i n_j$			40
MBB^a	$(n_i n_j)^{1/2}$			4 and 5
BBC2^b	n_i	$i = j$		10
	$-(n_i n_j)^{1/2}$	$i \neq j \wedge i \in (F_L; \infty) \wedge j \in (F_L; \infty)$		
	$n_i n_j$	$i \neq j \wedge i \in [1; F_L] \wedge j \in [1; F_L]$		
	$(n_i n_j)^{1/2}$	Otherwise		
CA^c	$[n_i(1-n_i)n_j(1-n_j)]^{1/2} + n_i n_j$			7
CGA^d	$\dfrac{n_i n_j + [n_i(2-n_i)n_j(2-n_j)]^{1/2}}{2}$			9
ML^e	$n_i n_j \dfrac{a_0 + a_1 n_i n_j}{1 + b_1 n_i n_j}$		$a_0 = 126.3101$ $a_1 = 2213.33$ $b_1 = 2338.64$	13
MLSIC^f	$n_i n_j \dfrac{a_0 + a_1 n_i n_j}{1 + b_1 n_i n_j}$ $\quad i \neq j$ $n_i n_j \quad i = j$		$a_0 = 1298.78$ $a_1 = 35114.4$ $b_1 = 36412.2$	13
GU^g	$(n_i n_j)^{1/2}$ $\quad i \neq j$ $n_i n_j \quad i = j$			6
POWER^h	$(n_i n_j)^\alpha$			15–17

^{a} Introduced independently by Müller and by Buijse and Baerends.[5,8] ^{b} The BBC2 functional coincides with BBC1[10] for a two-electron closed-shell system. ^{c} Csányi and Arias functional. ^{d} Csányi, Goedecker and Arias functional. ^{e} Marques and Lathiotakis functional. ^{f} Marques and Lathiotakis functional corrected for self-interaction. ^{g} Goedecker and Umrigar functional. ^{h} The α parameter of the POWER functional is fitted for each ω, $\alpha(\omega)$, in order to reproduce the exact V_{ee} value (see the ESI for further details).

difference between the exact 2-DM and the SD approximation,[45]

$$^2D_{ij,kl}^{\sigma\sigma'} = {}^2D_{ij,kl}^{\mathrm{SD},\sigma\sigma'} + {}^2\Gamma_{ij,kl}^{\sigma\sigma'}. \tag{6}$$

The different PNOF expressions are constructed in terms of the auxiliary Δ and Π matrices according to the following recipe:

$$^2\Gamma_{ij,kl}^{\mathrm{PNOF}n,\alpha\beta} = -\Delta_{ij}^{\alpha\beta}\delta_{ik}\delta_{jl} + \Pi_{ik}\delta_{ij}\delta_{kl}, \tag{7}$$

$$^2\Gamma_{ij,kl}^{\mathrm{PNOF}n,\alpha\alpha} = -\Delta_{ij}^{\alpha\alpha}(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}). \tag{8}$$

Table 2 collects Δ and Π matrices for all PNOFs. In all PNOFs $\Delta_{ij}^{\alpha\alpha} = \Delta_{ij}^{\alpha\beta} \equiv \Delta_{ij}$ excepting PNOF3 that takes $\Delta_{ij}^{\alpha\alpha} = 0$. In this work, approximate 2-DMs are constructed from full configuration interaction (FCI) ONs and, therefore, the original PNOF5 and PNOF6,[20,21] which impose perfect-pairing constraints in the ONs, cannot be employed. Alternatively, in

this paper we employ the extended versions of PNOF5[46] and PNOF6[47] that are free of these restrictions for a closed-shell two-electron system. Hereafter, the names PNOF5 and PNOF6 refer to the extended versions of these DMFAs. Unlike the original PNOF6, the extended version of PNOF6 can be actually calculated in three different ways, depending on the definition of S_γ, which will be called down (d), up (u) and average (h),

$$S_\gamma^{\mathrm{d}} = \sum_{i=1}^{F_{\mathrm{L}}} \gamma_i, \quad S_\gamma^{\mathrm{u}} = \sum_{i>F_{\mathrm{L}}}^{M} \gamma_i, \quad S_\gamma^{\mathrm{h}} = \frac{S^{\mathrm{d}} + S^{\mathrm{u}}}{2}, \tag{9}$$

where

$$\gamma_i = n_i(1 - n_i) + \kappa_i^2 - \kappa_i \sum_{i=j}^{F_{\mathrm{L}}} \kappa_j, \tag{10}$$

Table 2 Δ and Π non-zero matrix elements. The diagonal elements coincide for all functionals: $\Delta_{ii} = n_i^2$ and $\Pi_{ii} = n_i$. $S_{\mathrm{F}} = \sum_{i=1}^{F_{\mathrm{L}}} h_i$, $T_{ij} = n_i n_j - \Delta_{ij}$, $h_i = 1 - n_i$, and S_γ^{x} and γ_i are defined in eqn (9) and (10), respectively. Ω_g is the subspace containing orbital g, which is below the Fermi level, and several orbitals above the Fermi level. In a two-electron closed-shell system $F_{\mathrm{L}} = 1$

	Δ_{ij}	Π_{ij}	Cases ($i \neq j$)	Ref.
PNOF2	$h_i h_j$	$\sqrt{n_i n_j} + \sqrt{h_i h_j} + T_{ij}$	$i \wedge j \in [1,F_{\mathrm{L}}]$	12
	$n_j h_i \left(\frac{1-S_{\mathrm{F}}}{S_{\mathrm{F}}}\right)$	$\sqrt{n_i n_j} - \sqrt{n_j h_i} + T_{ij}$	$i \in [1,F_{\mathrm{L}}] \wedge j \in (F_{\mathrm{L}},M)$	
	$n_i h_j \left(\frac{1-S_{\mathrm{F}}}{S_{\mathrm{F}}}\right)$	$\sqrt{n_i n_j} - \sqrt{n_i h_j} + T_{ij}$	$j \in [1,F_{\mathrm{L}}] \wedge i \in (F_{\mathrm{L}},M)$	
	$n_i n_j$	T_{ij}	$i \wedge j \in (F_{\mathrm{L}},M)$	
PNOF3	$h_i h_j$	$n_i n_j - \sqrt{n_i n_j}$	$i \wedge j \in [1,F_{\mathrm{L}}]$	18
	$n_j h_i \left(\frac{1-S_{\mathrm{F}}}{S_{\mathrm{F}}}\right)$	$n_i n_j - \sqrt{n_i n_j} - \sqrt{n_j h_i}$	$i \in [1,F_{\mathrm{L}}] \wedge j \in (F_{\mathrm{L}},M)$	
	$n_i h_j \left(\frac{1-S_{\mathrm{F}}}{S_{\mathrm{F}}}\right)$	$n_i n_j - \sqrt{n_i n_j} - \sqrt{n_i h_j}$	$j \in [1,F_{\mathrm{L}}] \wedge i \in (F_{\mathrm{L}},M)$	
	$n_i n_j$	$n_i n_j + \sqrt{n_i n_j}$	$i \wedge j \in (F_{\mathrm{L}},M)$	
PNOF4	$h_i h_j$	$-\sqrt{h_i h_j}$	$i \wedge j \in [1,F_{\mathrm{L}}]$	
	$n_j h_i \left(\frac{1-S_{\mathrm{F}}}{S_{\mathrm{F}}}\right)$	$-\sqrt{\left(\frac{h_i n_j}{S_{\mathrm{F}}}\right)\left(n_i - n_j + \frac{h_i n_j}{S_{\mathrm{F}}}\right)}$	$i \in [1,F_{\mathrm{L}}] \wedge j \in (F_{\mathrm{L}},M)$	
	$n_i h_j \left(\frac{1-S_{\mathrm{F}}}{S_{\mathrm{F}}}\right)$	$-\sqrt{\left(\frac{h_j n_i}{S_{\mathrm{F}}}\right)\left(n_j - n_i + \frac{h_j n_i}{S_{\mathrm{F}}}\right)}$	$j \in [1,F_{\mathrm{L}}] \wedge i \in (F_{\mathrm{L}},M)$	
	$n_i n_j$	$\sqrt{n_i n_j}$	$i \wedge j \in (F_{\mathrm{L}},M)$	
PNOF5	$n_i n_j$	$-\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i = g \vee j = g)$	20
	$n_i n_j$	$\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i \wedge j \in (F_{\mathrm{L}},M])$	
PNOF6x	$\mathrm{e}^{-2S_{\mathrm{F}}} h_i h_j$	$-\mathrm{e}^{-S_{\mathrm{F}}}\sqrt{h_i h_j}$	$i \wedge j \in [1,F_{\mathrm{L}}]$	21
	$\gamma_i \gamma_j / S_\gamma^{\mathrm{x}}$	$-\sqrt{\left(n_i h_j + \frac{\gamma_i \gamma_j}{S_\gamma^{\mathrm{x}}}\right)\left(n_j h_i + \frac{\gamma_i \gamma_j}{S_\gamma^{\mathrm{x}}}\right)}$	$i \in [1,F_{\mathrm{L}}] \wedge j \in (F_{\mathrm{L}},M)$	
x = d, u, h	$\gamma_i \gamma_j / S_\gamma^{\mathrm{x}}$	$-\sqrt{\left(n_i h_j + \frac{\gamma_i \gamma_j}{S_\gamma^{\mathrm{x}}}\right)\left(n_j h_i + \frac{\gamma_i \gamma_j}{S_\gamma^{\mathrm{x}}}\right)}$	$j \in [1,F_{\mathrm{L}}] \wedge i \in (F_{\mathrm{L}},M)$	
	$\mathrm{e}^{-2S_{\mathrm{F}}} n_i n_j$	$\mathrm{e}^{-S_{\mathrm{F}}}\sqrt{n_i n_j}$	$i \wedge j \in (F_{\mathrm{L}},M)$	
PNOF7	$n_i n_j$	$-\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i = g \vee j = g)$	22
	$n_i n_j$	$\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i \wedge j \in (F_{\mathrm{L}},M])$	
	0	$-\sqrt{n_i h_i n_j h_j}$	$(i \vee j) \in [1,F_{\mathrm{L}}] \wedge ((i \in \Omega_g \wedge j \notin \Omega_g) \vee (j \in \Omega_g \wedge i \notin \Omega_g))$	
	0	$\sqrt{n_i h_i n_j h_j}$	$(i \wedge j) \in (F_{\mathrm{L}},\infty) \wedge ((i \in \Omega_g \wedge j \notin \Omega_g) \vee (j \in \Omega_g \wedge i \notin \Omega_g))$	

and

$$\kappa_i = \begin{cases} (1 - n_i)\mathrm{e}^{-S_F} & i \in [1, F_L] \\ n_i \mathrm{e}^{-S_F} & i \in (F_L, M] \end{cases} \qquad (11)$$

F_L being the last occupied orbital below the Fermi level.

For two-electron closed-shell systems, the 2-DM in terms of ONs is known up to a phase factor, Φ_{ij},[35] the following being the only non-zero elements:

$${}^2D_{ii,jj}^{\alpha\beta} = \Phi_{ij}\sqrt{n_i^{\alpha} n_j^{\beta}} \quad (i \neq j), \qquad (12)$$

where $\Phi_{ij} = \pm 1$, depending on the nature of orbitals i and j. The most convenient way to choose the phase factors is to split the set of orbitals into two groups: the orbitals above and below the Fermi level, and choose $\Phi_{ij} = 1$ for i and j belonging to the same group and $\Phi_{ij} = -1$ otherwise.[36] Let us call this approximation the fixed-phases (FP) approximation. FP is very accurate for most two-electron systems[3,36] with only a few exceptions[37,39,48] that Giesbertz *et al.* attribute to long-range Coulomb interactions.[38] For two-electron systems, PNOF4, PNOF5 and PNOF7 reduce to FP with the mentioned phase factors, *i.e.*, eqn (12) with $\Phi_{1i} = -1(i \neq 1)$ and $\Phi_{ij} = 1$ for all other cases. Hence, hereafter we will only discuss PNOF4 results. Conversely, PNOF2, PNOF3 and PNOF6 expressions were not defined for two-electron systems and, therefore, the expression given in Table 2 does not reduce to FP.[49] In this work we have decided to study these expressions in a two-electron model.

Not all approximate 2-DMs correspond to an N-particle fermionic wavefunction. The set of 2-DMs that satisfy this condition are called N-representable 2-DMs. Non N-representable 2-DMs might lead to spurious results such as non-variational energies.[19,42] The set of conditions that guarantees the N-representability of the 2-DM is known[50] but its calculation involves higher-order density matrices. There are three conditions that only require the 2-DM for its calculation, the P-, G- and Q-conditions,[51–53] which concern the positive semidefinite character of P, Q and G matrices,[44]

$$P_{ij,kl}^{\sigma\sigma'} = \left\langle \Psi \left| a_{i\sigma}^{\dagger} a_{j\sigma'}^{\dagger} a_{l\sigma'} a_{k\sigma} \right| \Psi \right\rangle, \qquad (13)$$

$$Q_{ij,kl}^{\sigma\sigma'} = \left\langle \Psi \left| a_{i\sigma} a_{j\sigma'} a_{l\sigma'}^{\dagger} a_{k\sigma}^{\dagger} \right| \Psi \right\rangle, \qquad (14)$$

$$G_{ij,kl}^{\sigma\sigma'} = \left\langle \Psi \left| a_{i\sigma} a_{j\sigma'}^{\dagger} a_{l\sigma'}^{\dagger} a_{k\sigma} \right| \Psi \right\rangle. \qquad (15)$$

In order to test these conditions, one must build these matrices and check the sign of the corresponding eigenvalues. Notice that the P matrix coincides with the 2-DM and, therefore, the P condition is equivalent to the non-negativity condition of the geminal occupancies.

The pair density, $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$, is the diagonal part of the 2-RDM (*i.e.*, eqn (1) when $\mathbf{r}_1 = \mathbf{r}_1'$ and $\mathbf{r}_2 = \mathbf{r}_2'$) upon integration over spin and it is the only part of the 2-RDM needed to calculate the electron–electron repulsion energy (V_{ee}). Although $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ is a simpler function than the 2-RDM, it depends on six variables and it is difficult to analyze. Fortunately, there is no need of the

full knowledge of $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ to compute V_{ee}. The calculation of the electronic repulsion only requires the radial intracule[54] density[41]

$$I(r_{12}) = r_{12}^2 \int \mathrm{d}\mathbf{r}_1' \mathrm{d}\mathbf{r}_2' \mathrm{d}\Omega_{12} \rho_2(\mathbf{r}_1', \mathbf{r}_2') \delta(\mathbf{r}_{12} - \mathbf{r}_1' + \mathbf{r}_2'), \qquad (16)$$

where $\mathrm{d}\Omega_{12} = \sin\theta_{12}\mathrm{d}\theta_{12}\mathrm{d}\phi_{12}$. The radial intracule density is a one dimensional function that provides a graphical means to analyze $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ at different interelectronic separations and a simple expression to calculate V_{ee},

$$V_{ee} = \int_0^\infty \mathrm{d}r_{12} \frac{I(r_{12})}{r_{12}}. \qquad (17)$$

$\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ also enters the expression of the so-called delocalization index (DI),[55,56] which is a measure of covariance between the electron population of two regions, A and B,

$$\delta(A, B) = -2 \int_A \int_B \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 [\rho_2(\mathbf{r}_1, \mathbf{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)], \qquad (18)$$

where $\rho(\mathbf{r})$ is the electron density. The DI has been used in the past to calibrate the performance of several approximations.[57–64] In the present work, the DMFAs use the exact natural orbitals and occupancies and, therefore, the second term in the r.h.s. of eqn (18) is identical in both the exact calculation and the DMFA. Hence, the DI difference actually measures the difference between the exact and the DMFA number of electron pairs (one in region A and another in B).

3 Computational details

We will test several DMFAs in the two-electron harmonium atom—a model system with the following Hamiltonian,[33,65,66]

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{2}\omega^2 r_1^2 + \frac{1}{2}\omega^2 r_2^2 + \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|}, \qquad (19)$$

where the ω parameter is the confinement strength and tunes the electron correlation in a continuous manner: low-ω values correspond to a strong-correlation regime whereas weakly correlated systems are produced at large ω. The harmonium atom has been widely used for benchmarking and developing functionals[23,64,67–77] due to the availability of benchmark results.[48,78–80] The harmonium atom is one of the most difficult systems for computational methods[64,68,76,77] and, therefore, is a formidable test-bed for DMFAs. In addition to harmonium, there are other model systems that pose a great challenge for computational methods, such as the Hubbard model[81] or uniform gases of electrons trapped in rings, spheres and geometrical objects of higher dimensions.[82–85]

FCI calculations were performed for the ground-state singlet two-electron harmonium atom using 20 values of the ω parameter: 0.03, 0.033, 0.036, 0.0365373, 0.04, 0.05, 0.06, 0.08, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 1, 2, 5, 10, 100 and 1000.[86] We used a modified version of the code developed by Knowles and Handy[87,88] and a variationally optimized even-tempered basis set of seven S, P, D and F Gaussian functions, which form a total of 112 basis functions.[86] The exact 2-DM and the ONs were

calculated from the FCI expansion coefficients using the DMN[89] in-house code. The radial intracule density was computed using RHO2_OPS[90] code which uses the algorithm proposed by Cioslowski and Liu.[91] The calculation of the DI was performed with the in-house RHO_OPS[92] and ESI-3D[57,93,94] codes.

In the current study, exact (within the given basis set) natural orbitals and occupancies are used in fifteen DMFAs (BBC2, CA, CGA, GU, MBB, POWER, ML, MLSIC, PNOF2, PNOF3, PNOF4, three PNOF6 definitions and SD) to evaluate their performance in a series of tests. Namely, we have used the expressions given in Tables 1 and 2 with FCI ONs to generate the corresponding approximate 2-DM, which are subsequently analyzed using ten different tests:

(i) calculation of the 2-DM trace, (ii) cumulative absolute error (CAE) for the diagonal elements, *i.e.*,

$$\text{CAE}_\text{D}\left[{}^2\mathbf{D}^X\right] = \sum_{\substack{ij \\ \sigma\sigma'}} \left| {}^2D_{ij,ij}^{X,\sigma\sigma'} - {}^2D_{ij,ij}^{\sigma\sigma'} \right|, \quad (20)$$

(iii) CAE for all the elements of the 2-DM,

$$\text{CAE}\left[{}^2\mathbf{D}^X\right] = \sum_{\substack{ijkl \\ \sigma\sigma'}} \left| {}^2D_{ij,kl}^{X,\sigma\sigma'} - {}^2D_{ij,kl}^{\sigma\sigma'} \right|, \quad (21)$$

(iv) the correct antisymmetry of the 2-DM, *i.e.*,

$$\text{Err}_\text{A}\left[{}^2\mathbf{D}^{X,\sigma\sigma'}\right] = \sum_{ijkl,\sigma} \left| {}^2D_{ij,kl}^{X,\sigma\sigma} + {}^2D_{ij,lk}^{X,\sigma\sigma} + {}^2D_{ji,lk}^{X,\sigma\sigma} + {}^2D_{ji,kl}^{X,\sigma\sigma} \right|, \quad (22)$$

(v) P, Q, and G N-representability conditions, eqn (13)–(15), (vi) the DI between the two symmetric regions generated by a bisecting plane passing through the center of mass, (vii) the average interelectronic distance and (viii) its variance, (ix) the interelectronic repulsion, V_ee, and (x) the radial intracule density profile.

4 Results

4.1 The diagonal elements: sum rule and cumulative absolute error

The plot in Fig. 1 shows the 2-DM trace errors for several DMFAs (the exact trace equals two in McWeeny's normalization[41]). BBC2, CA, CGA, MBB, PNOF2 and PNOF4 have not been included in this plot because they satisfy the sum rule. The larger the ω value, the less important the correlation effects in the harmonium atom. Indeed, for large values of ω all approximations perform very well because correlation effects are negligible. However, when correlation increases, SD produces very poor results. GU and MLSIC have the same diagonal elements ($f(n_i,n_i) = n_i^2$) and, therefore, give exactly the same trace as the SD approximation for the two-electron case. PNOF3 coincides with SD because the former only modifies the opposite-spin elements in the cumulant construction and the opposite-spin cumulant of PNOF3 (like the exact one) does not contribute to the sum rule. ML is based on a Padé approximant including some fitted parameters that result in wrong
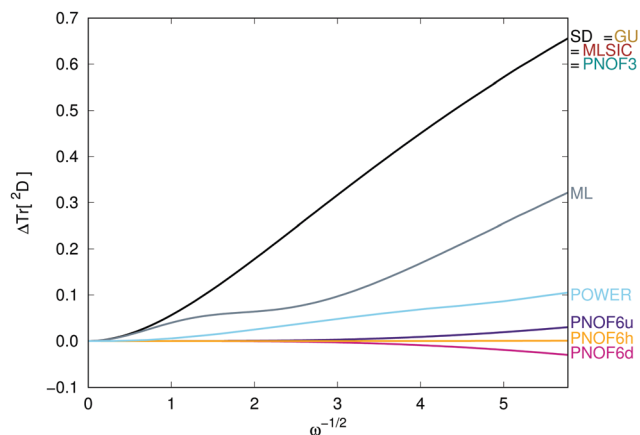

Fig. 1 2-DM trace error against $\omega^{-1/2}$. BBC2, CA, CGA, MBB, PNOF2 and PNOF4 have not been included because they satisfy the sum rule.

trace numbers. The POWER functional presents non-negligible errors in the trace as the correlation increases. PNOF6 shows small trace deviations, the three PNOF6 versions (PNOF6u, PNOF6h and PNOF6d) differing on the value of S_γ (actually, the S_γ definition, eqn (9), is responsible for the violation of the sum rule). The S_γ of PNOF6h provides the smallest error.

To get further insight about the error committed in the 2-DM diagonal elements, we have analyzed the diagonal CAE, eqn (20), as a function of $\omega^{-1/2}$ (Fig. 2). PNOF2 and PNOF4 approximations have not been included in Fig. 2 because the error produced by these approximations is lower than 10^{-4}. GU and MLSIC produce exactly the same error as the SD approximation because the self-interaction correction enforced in these approximations results in ${}^2D_{ij,ij}$ terms equal to those produced by SD. ML, which showed better trace numbers than SD for all ω values, presents a larger diagonal CAE indicating important error cancellation in the calculation of the trace. PNOF3 shows smaller diagonal CAE than SD, which necessarily come from the opposite-spin diagonal components, because the same-spin components of the 2-DM are identical in both
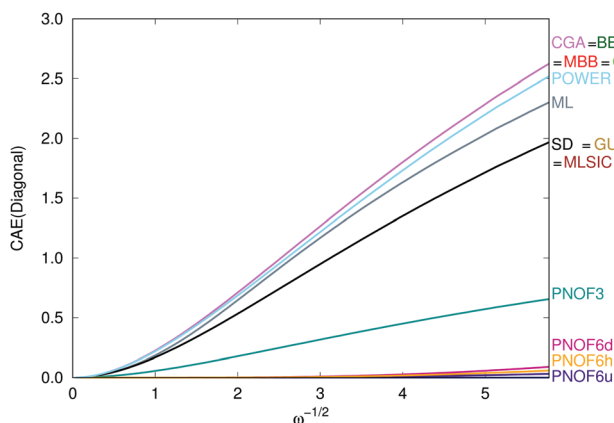

Fig. 2 Cumulative absolute error (CAE) for the diagonal elements of the 2-DM against $\omega^{-1/2}$ (eqn (20)). PNOF2 and PNOF4 approximations have not been included because they present errors lower than 10^{-4}.

approximations. All PNOF6 versions provide very small CAE, the best among the three definitions being PNOF6u (hence PNOF6h best trace numbers were due to error cancellation). BBC2, CA, CGA, MBB and POWER show larger diagonal CAE than SD due to the unphysical $^2D_{ii,ii}^{\sigma\sigma}$ elements that provide a correct total trace but contribute to important self-interaction errors. In fact, any K-functional studied in this work produces the same diagonal elements than the SD approximation (*i.e.*, $f(n_i,n_i) = n_i^2$) if we remove the unphysical elements $^2D_{ii,ii}^{\sigma\sigma}$ that are included in some DMFAs.

4.2 Cumulative absolute error

The total CAE (eqn (21)) is plotted against $\omega^{-1/2}$ in Fig. 3. All fifteen DMFAs provide CAE when correlation increases. A troublesome result is that all K-functionals perform worse than the SD approximation for all ω values. Most of the approximations show a monotonic increase of the error excepting MLSIC, probably due to the parameterization of this approximation. BBC2 and MBB values coincide for all ω and present the largest total CAE. Their 2-DMs do fully not coincide but for a closed-shell two-electron system these two DMFAs only differ in the phase of some unphysical $^2D_{ii,ii}^{\sigma\sigma}$ elements, which obviously contribute to the same CAE. All PNOFs perform better than SD indicating that the cumulant correction of PNOFs improves in the right direction. The most recently developed approximations, PNOF4 to PNOF7, show the best agreement with the exact 2-DM, giving only a small total CAE. Among the three PNOF6, PNOF6u performs marginally better than the rest. PNOF4 is actually exact in a wide range of ω values, only deviating at the high-correlation regime. In the high-correlation regime ($\omega \leq 0.1$) several phases of PNOF4 do not coincide with the exact ones, preventing PNOF4 from reproducing the exact elements of the 2-DM.

4.3 Antisymmetry

The electronic wavefunction must be antisymmetric due to the fermionic character of electrons. The 2-DM preserves the antisymmetric nature inherited from the wavefunction and, therefore, deviations from the antisymmetry condition, eqn (22), can be also regarded as violations of a necessary N-representability
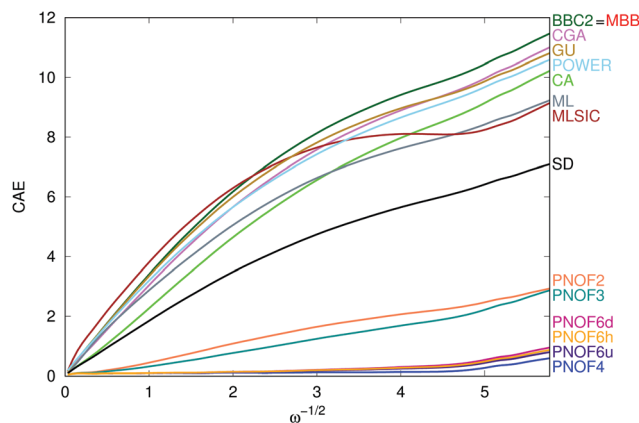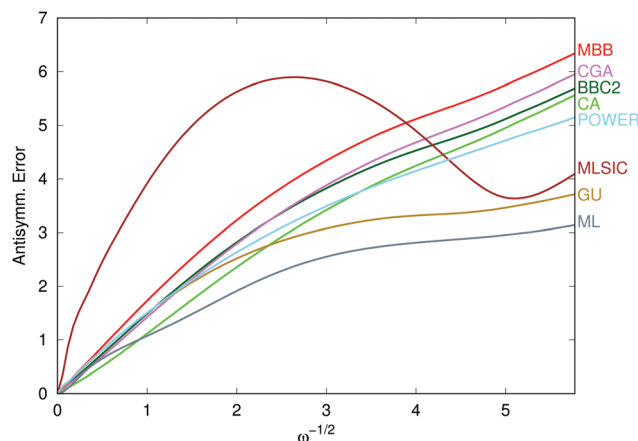

Fig. 4 Antisymmetry error of the 2-DMs (eqn (22)) against $\omega^{-1/2}$. PNOFs and SD have not been included because they satisfy the antisymmetry condition.

condition. Functionals that do not satisfy this condition fail to correctly treat the fermionic nature of electrons.

PNOFs were built in order to satisfy the correct antisymmetry of the 2-DM by constructing approximations from an inherently antisymmetric cumulant structure. SD is also antisymmetric by construction. Fig. 4 shows eqn (22) against $\omega^{-1/2}$ for the other DMFAs. The best K-functional is ML, which error is almost half the error of MBB in the high-correlation regime. The DMFAs that deviate most from the antisymmetry condition are BBC2, CA, CGA, and MBB with errors growing as $\omega^{-1/2}$. Conversely, the only DMFA that does not show a monotonic increase of the error with correlation is MLSIC, once again, putting forward the parameterized nature of this DMFA. A self-interaction correction applied to MBB produces the GU approximation,[8] resulting in smaller antisymmetry errors. One can easily prove that among all K-functionals that one could devise, the only one that satisfies the antisymmetric condition is SD. This result evinces the need for designing functionals that, at least, include J and K components, beyond the SD approximation.

4.4 N-Representability

In order to check the deviation from the N-representability conditions, we have computed the eigenvalues of matrices P, Q and G (eqn (13)–(15)) and summed all negative ones. In Fig. 5 we have plotted the result of the sum against $\omega^{-1/2}$.

The 2-DMs of PNOFs give rise to non-negative basis-set-independent eigenvalues associated to P-, Q- and G-conditions (see Appendix I for the basis-set dependent and independent eigenvalues). However, one cannot anticipate the conditions that the basis-set-dependent eigenvalues might impose in the functional structure and only PNOF2, PNOF4 and SD satisfy the N-representability conditions studied in this work. PNOF3 and PNOF6 perform very well even in the high-correlation regime with the only exception of PNOF3 that shows significant deviations in the G-condition when correlation increases, in line with previous findings.[19] PNOF6 shows small negative
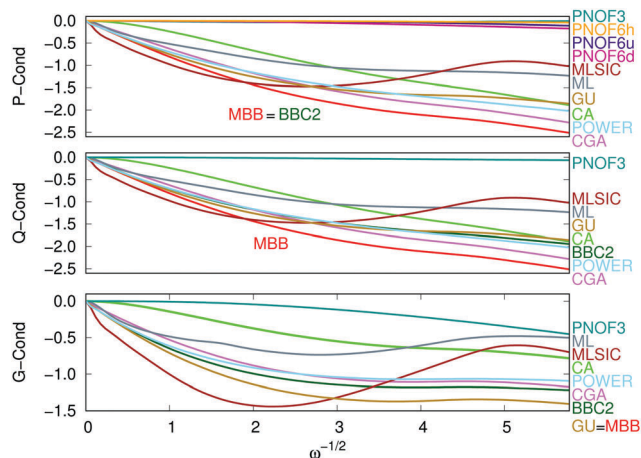

Fig. 3 Cumulative absolute error (CAE) for the whole 2-DM against $\omega^{-1/2}$ (eqn (21)).

Fig. 5 Sum of all negative eigenvalues of P, Q and G matrices against $\omega^{-1/2}$. PNOF2, PNOF4 and SD have not been included because they satisfy the N-representability conditions studied.

eigenvalues of the P matrix that change depending on the definition of S_γ. In fact, if the wavefunction would correspond to a perfect-pairing situation (in which natural orbitals are coupled by pairs, each pair occupancy summing exactly to one electron) the P-condition would be satisfied for a two-electron closed-shell system regardless the definition of S_γ. All K-functionals show significant deviations from P-, Q- and G-conditions that rapidly increase with electron correlation. The largest errors are presented by MBB, but BBC2, CA, CGA, GU and POWER also present non-negligible errors. ML and MLSIC approximations show non-monotonic increase of the errors due to their parameterized nature.

4.5 Delocalization index

The difference between the approximate DI and the exact one is plotted in Fig. 6. In the present case, where the density is computed from exact ONs for both methods (see eqn (18)), the latter quantity also corresponds to the difference between the exact and the approximate number of electron pairs between
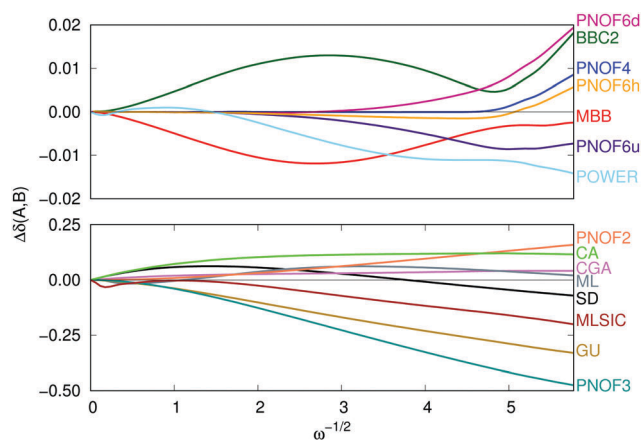
two regions. Upon increase of electron correlation effects, the number of electron pairs between regions is expected to decrease, as found by other calculations of the DI in molecules under DMFAs (see the ESI‡ for the exact values).[57,60,95–97]

In general, the gross number of pairs is pretty well described by BCC2, MBB, PNOF4, PNOF6 and POWER (that present errors below 3%), whereas CA, GU, MLSIC, PNOF2, PNOF3 and SD present errors ranging between 20% to 75%. Interestingly, GU, MBB, MLSIC, PNOF3 and PNOF6u systematically underestimate the DI (*i.e.*, they overestimate the number of electron pairs) whereas BBC2, CA, CGA, PNOF2 and PNOF6d always overestimate it. In line with the results obtained in the sum rule and N-representability tests, PNOF6h provides better results than PNOF6u and PNOF6d.

4.6 Interelectronic distance

In this section we analyze four quantities related to the interelectronic distance: (i) the mean value, $\langle r_{12} \rangle$, (ii) the variance, $\sigma^2 = \langle r_{12}^2 \rangle - \langle r_{12} \rangle^2$, (iii) the interelectronic density distribution through radial intracule density profiles, and (iv) the V_{ee}. All these quantities can be calculated from the radial intracule density, eqn (16).

4.6.1 $\langle r_{12} \rangle$ and σ^2. The analysis of the $\langle r_{12} \rangle$ and the σ^2 reveals important aspects of the effects of electron correlation in DMFAs. Usually large $\langle r_{12} \rangle$ values go with smaller V_{ee}, but some exceptions exist.[98] Hence, large (small) $\Delta\langle r_{12} \rangle$ are common in methods that overestimate (underestimate) electron correlation, whereas the variance of the probability distribution measures the spread of the interelectronic distribution. Fig. 7 shows that all DMFAs deviate from the exact $\langle r_{12} \rangle$ as correlation increases. BBC2($\approx$MBB), PNOF3, PNOF6d and PNOF6h overestimate correlation effects, whereas CA, CGA, GU, ML, MLSIC, POWER, PNOF2, PNOF4, PNOF6u and SD underestimate electron correlation at all ω values. This is actually the only test (together with V_{ee}, *vide infra*) where SD performs clearly and systematically worse than the other DMFAs. The error of PNOF4 is small and only due to the choice of phase factors. PNOF6h performs somewhat better than the other two PNOF6 versions.

Fig. 8 plots the difference between the approximate variance, computed from DMFAs, and the exact one against ω^{-2}.
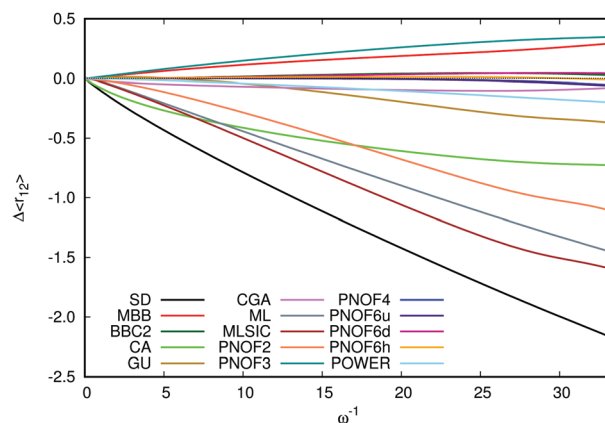

Fig. 6 Error in the DI against $\omega^{-1/2}$. (top) DMFAs with errors below 3% and (bottom) DMFAs with errors above 3%.

Fig. 7 Error in $\langle r_{12} \rangle$ against ω^{-1}.

Only PNOF4 and PNOF6 show a good agreement with both the exact $\langle r_{12} \rangle$ and σ^2. Interestingly, all DMFAs either underestimate or overestimate the distribution spread around the average value, the SD approximation showing the largest overestimation and MBB presenting the sharpest distribution among all DMFAs.

4.6.2 Radial intracule density. In Fig. 9 we have plotted the difference between the DMFA and the exact radial intracule density for three values of the ω parameter which cover low- ($\omega = 1000$), medium- ($\omega = 0.5$) and high-correlation ($\omega = 0.03$) regimes. These profiles allow a range-separation analysis of electron correlation, the typical profile of the method that lacks electron correlation being negative at short distances and positive at large distances.[99] The negative and the positive regions compensate for all DMFAs that satisfy the sum rule. For $\omega = 1000$ we observe that all DMFAs produce exact results except CA, CGA, ML, MLSIC and SD. One should keep in mind that there is residual correlation even at the limit of large ω,[48,78] however, thus far this is the only analyzed property that some DMFAs fail to reproduce at the weak correlation limit. CA and SD actually coincide because CA recovers the SD expression at this limit. At the medium-correlation regime, no functional is exact except PNOF4 and PNOF6, whereas at the high-correlation regime even PNOF4 and PNOF6 show some deviations. PNOF4 presents the smallest error while PNOF6h shows the best performance among PNOF6 versions. GU, ML, MLSIC, PNOF3 and SD exhibit larger negative values at short and medium ranges that are not compensated by positive ones at larger separations because these functionals do not satisfy the sum rule.

We may classify the functionals according to their profile. The only functionals that are not included in this classification are PNOF4 and PNOF6, which show the smallest errors, and MLSIC which presents the largest errors. In Fig. 9, for the medium-correlation regime (see the l.h.s. $\omega = 0.5$ plot), there is a first type of profile including CA, ML, PNOF2 and SD, which consists in the typical profile of methods that underestimate electron correlation. Namely, CA and SD underestimate short-range correlation, whereas ML and PNOF2 underestimate mid-range correlation. These four DMFAs also underestimated

importantly the value of $\langle r_{12} \rangle$. The r.h.s. of the $\omega = 0.5$ plot, including BBC2, CGA, GU, MBB, PNOF3 and POWER, shows an unusual intracule density profile with overestimation of short-range correlation and underestimation of mid- and long-range correlation. The latter group of functionals actually provided quite accurate $\langle r_{12} \rangle$ for small and medium-correlation regimes and, at high correlation, either overestimated $\langle r_{12} \rangle$ or presented values that are not much smaller than the exact one (see Fig. 7). The profile of the intracule density at $\omega = 0.03$ can be also used to classify the functionals. We first find a group of DMFAs (including BBC2, CA, CGA, MBB and POWER), which presents two maxima: one at short-range and the other at long-range (see $\omega = 0.03$ r.h.s plot in Fig. 9). A second group of DMFAs includes GU and PNOF3 that underestimate the interelectronic separation at all ranges except at very short range (see also $\omega = 0.03$ r.h.s plot in Fig. 9). The last group of DMFAs (ML, MLSIC, PNOF2 and SD) show large underestimation of short- and medium-range correlation (see $\omega = 0.03$ l.h.s plot in Fig. 9).

Although most DMFAs show similar profiles at different electron correlation regimes, the values of r_{12} at which they underestimate/overestimate the interelectronic separation changes with ω. Hence, if one would use these functionals (at least those that preserve the profile with electron correlation) in a range-separation scheme[100,101] the attenuating parameter[102] should depend on ω.[103,104] An inspection of the intersection values at different ω puts forward that $\Delta I(r_{12}) = 0$ occurs at values of r_{12} that change with $\omega^{-1/2}$. Taking the latter point as the point at which short- and long-range separation functions coincide would ensure that errors are kept at different correlation regimes. If we choose the error function, $\text{erf}(\mu r_{12})$, as the range-separation function it is easy to prove that the attenuating parameter, μ, should be proportional to $\omega^{1/2}$, in line with the well-known fact that the attenuating parameter should change with electron correlation.[103–106]

The radial intracule density of harmonium atom at $\omega = 0.03$ presents negative probabilities in the short-range region (Fig. S2 in the ESI‡). Hence, we are prompted to attribute the overestimation of short-range correlation at the high-correlation regime in a number of DMFAs (BBC2, CA, CGA, GU, MBB, PNOF3, PNOF6 and POWER) to the unphysical behavior of the associated pair density, which can be traced back to the violation of the P condition. The only functional that actually shows overestimation of short-range correlation at $\omega = 0.03$ and it is not due to negative radial intracule density values is PNOF4. Despite their negative radial intracule density values, the $\langle r_{12} \rangle$ values computed with BBC2, CA, CGA, GU, MBB, POWER, PNOF3 and PNOF6 functionals are not among the worst ones.

4.6.3 Electron repulsion energy. Thus far, all the tests considered in this work did not measure the ability of DMFAs to reproduce the electronic energy. In this section, we analyze the performance of the DMFAs in reproducing the V_{ee}, the only fraction of the energy that is actually approximated in DMFT. See previous publications of our group for a similar analysis of DMFAs in other systems.[23,27]

Fig. 10 shows the relative error in the V_{ee} against $\omega^{-1/2}$. We have not included the POWER functional because it was
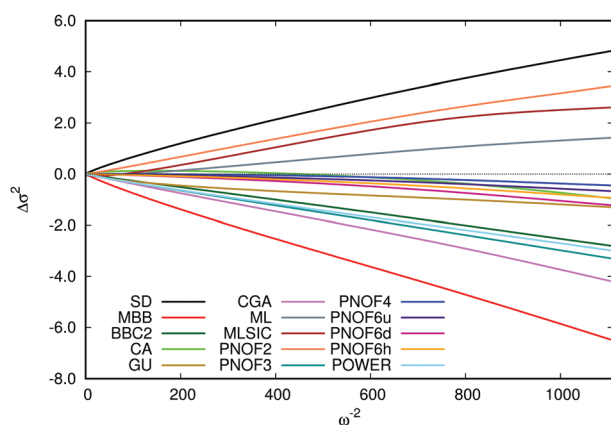

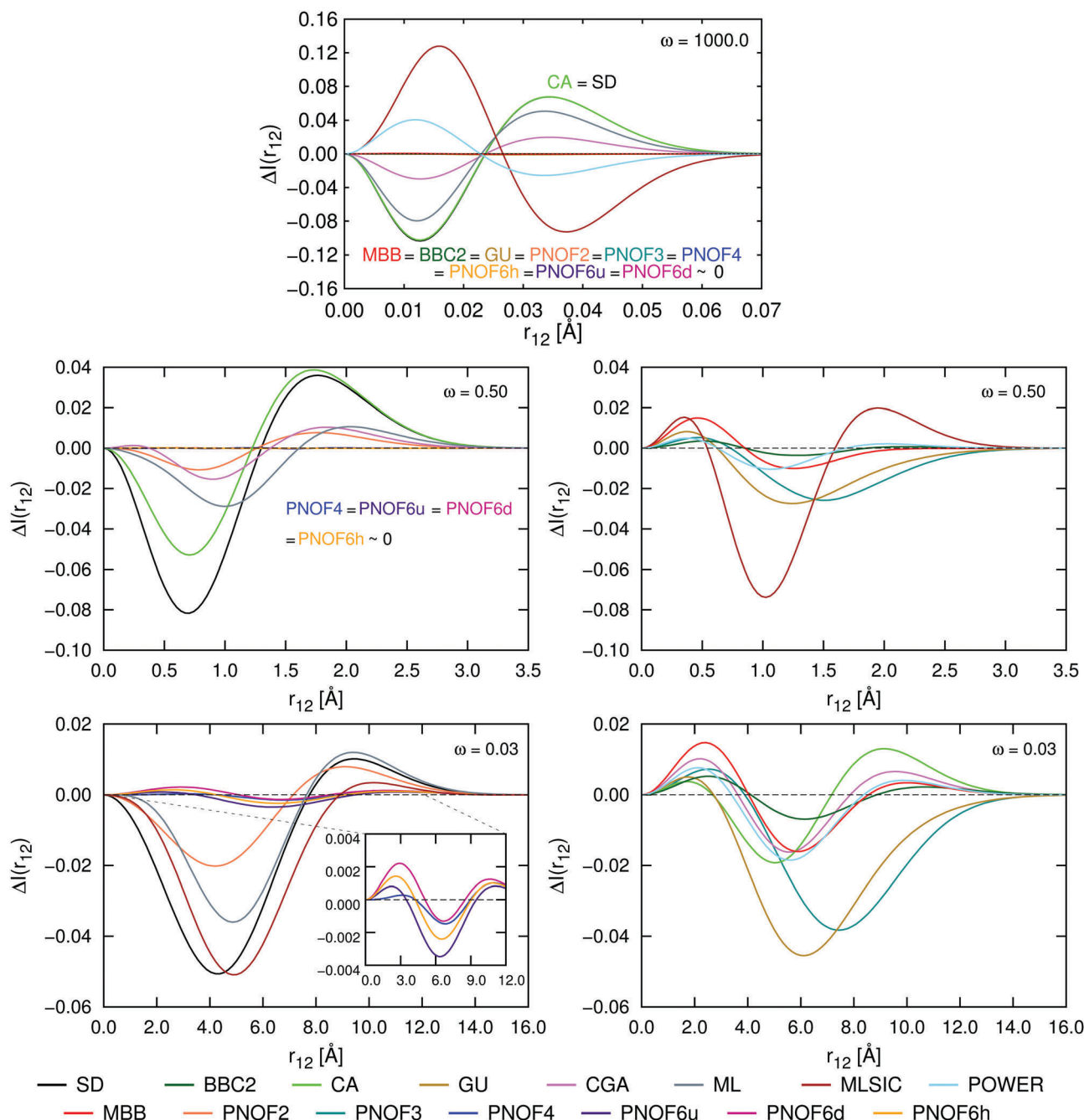**Fig. 8** Error in σ^2 against ω^{-2}.

Fig. 9 Difference between the exact and the approximate radial intracule density (eqn (16)) for three values of ω (1000, 0.5 and 0.03).

optimized to reproduce the exact V_{ee}, yielding errors below 10^{-5} a.u. (see the ESI‡ for further details). CA, GU, ML, PNOF2, PNOF4, PNOF6u and PNOF6h underestimate correlation energy, whereas BBC2, CGA, MBB, MLSIC and PNOF6d overestimate it. SD performs very poorly with a relative correlation error that grows linearly with $\omega^{-1/2}$. Interestingly, despite the wrong behavior found in the previous tests, all other DMFAs perform better. Namely, BBC2, CA, CGA, PNOF4 and PNOF6 present errors below 10%. PNOF4 is virtually exact for all values of ω, whereas PNOF6h and PNOF6d provide very accurate estimates and only show some minor deviations at the

high-correlation regime. This fact puts forward the need for tests not based in the energy to reveal some inherent important problems in DMFAs.

5 Discussion and conclusions

Despite its simplicity, the two-electron harmonium atom has proven an excellent model for benchmarking.[64,107] Even though two-electron systems should not pose a great challenge for DMFAs, the present paper has unveiled many problems and
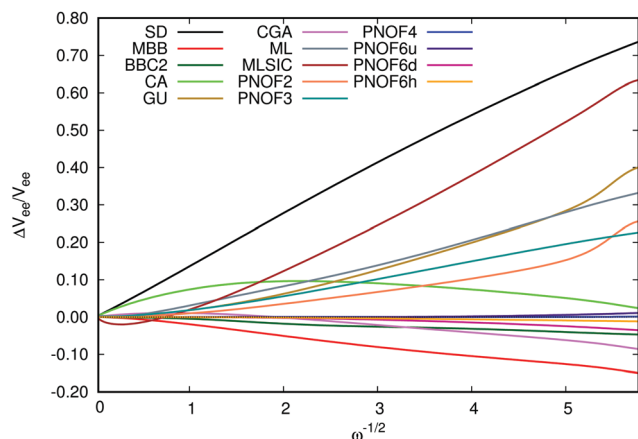
Fig. 10 Relative error in V_{ee} against $\omega^{-1/2}$.

strengths of current approximations. In the following we summarize these results and put forward various suggestions for the development of DMFAs.

The study of the diagonal and the antisymmetry of the 2-DM reveals the limited accuracy of K-functionals. The only K-functional that can satisfy the antisymmetry requirements of the 2-DM is the SD approximation, which, actually, performs better than (or equal to) any other K-functional in the calculation of the trace and the assessment of the CAE diagonal elements. These results evince that the construction of a DMFA needs to consider, at least, both J and K terms.

Results obtained with ML and MLSIC functionals, which present errors oscillating with the confinement strength in many tests, suggest caution in using fitted parameters for the construction of DMFAs.

The DI test warns against the use of the SD approximation in strongly correlated systems. The widely used MBB approximation remains a good approximation in any correlation regime[57] but we suggest the usage of PNOF6h or PNOF7 as they provide more accurate results. However, the performance of DMFAs in the DI for molecules is an open question that needs to be addressed.

As expected, the results of PNOF4 (equivalent to PNOF5, PNOF7 and FP) are significantly better than those obtained with other DMFAs with the only exception of PNOF6. PNOF4 test results reveal that the phase dilemma[36] reminds an open problem, which affects the energy in a lesser extent than other properties. Namely, at the high-correlation regime, non-negligible errors arise in the off-diagonal elements of the 2-DM that result in inaccurate DI, interelectronic distances, variances, and radial intracule densities.

The study of radial intracule densities unveils that many DMFAs present negative probabilities at short interelectronic distances, in connection with the violation of the P condition. On the other hand, numerical inspection indicates that most DMFAs present $\Delta I(r_{12}) = 0$ for r_{12} values that are proportional to $\omega^{-1/2}$. In this sense, it seems natural to choose the latter point as the crossing point between the attenuating functions that separate short- and long-range regions in range-separated

functionals. Assuming that the attenuating functions are error functions depending on an attenuating parameter,[100,101] μ, it is easy to show that μ should be proportional to $\omega^{1/2}$. This fact can be exploited in the construction of new range-separation methods.

Let us notice that many functionals performed reasonably well in the calculation of the exact electronic energy but produce important errors in the calculation of other properties, supporting the claim that functional development should consider other properties besides the electronic energy.[30,31] In the calibration of DMFAs one can use properties such as the intracule pair density, which are particularly challenging for DMFAs, or other properties, such as the expected value of the interelectronic separation and its variance, which are convenient because they are easy to compute.

Finally, we can draw the conclusion that a DMFA should attain as many N-representable properties as possible because the best-performing functionals are those that satisfy most of these conditions. Some of these N-representable conditions can be imposed in the construction of the functional.

The results of this work suggest the construction of a DMFA following some simple and somewhat expected rules: (i) consider both J and K energy components beyond the single-determinant approximation, (ii) impose the known N-representability conditions, (iii) refrain from using empirical parameterization; and calibrate the functionals using (iv) the energy and other properties, (v) a model with tunable electron correlation to consider various correlation regimes.

In the present study we have suggested a battery of ten tests to analyze DMFAs, including several properties that the exact functional should satisfy. The tests are performed on a two-electron model system with varying electron correlation and carrying a very small computational effort. The test can be easily extended to a larger number of electrons, thus setting new challenges for DMFAs. We are confident that DMFA developers will benefit from the results obtained in this paper and will use this test set as a means to construct more robust approximations.

Conflicts of interest

There are no conflicts to declare.

Appendix I: N-representability conditions

The analysis of P, Q and G N-representability conditions, involves the diagonalization of the P, Q and G matrices (eqn (13)–(15)) for each spin case ($\alpha\alpha$, $\beta\beta$, $\alpha\beta$ and $\beta\alpha$). Some of these matrices render themselves to an analytic diagonalization, producing eigenvalues that can be used to constrain the 2-DM elements that enter DMFAs; other matrices cannot be diagonalized and they only impose conditions that can be checked *a posteriori*. In this section we present the eigenvalues

and the constraints associated to the P, Q and G conditions produced by K-functionals and PNOFs.

K-Functionals

These DMFAs assume that the opposite-spin cumulant terms are zero. Hence, the eigenvalues associated to the opposite-spin P, Q and G matrices are identical for all methods:

• P-condition (opposite-spin):

$$\bar{P}_{ij}^{\sigma\sigma'} = n_i^\sigma n_j^{\sigma'} \geq 0 \tag{23}$$

• G-condition (opposite-spin):

$$\bar{G}_{ij}^{\sigma\sigma'} = n_i^\sigma \left(1 - n_j^{\sigma'}\right) \geq 0 \tag{24}$$

• Q-condition (opposite-spin):

$$\bar{Q}_{ij}^{\sigma\sigma'} = \left(1 - n_i^\sigma\right)\left(1 - n_j^{\sigma'}\right) \geq 0 \tag{25}$$

These eigenvalues are necessarily non-negative and, therefore, the opposite-spin matrices do not impose additional constraints in these functionals. Conversely, the same-spin matrices depend on the definition of $f(n_i,n_j)$ and thus produce *ad hoc* eigenvalues. These eigenvalues are collected in the following matrices:

• P-condition (same-spin):

$$\bar{P}_{ij}^{\sigma\sigma} = \begin{cases} n_i^2 - f(n_i, n_i) & \text{if } i = j \\ n_i n_j + f(n_i, n_j) & \text{if } i < j \\ n_i n_j - f(n_j, n_i) & \text{if } i > j \end{cases} \tag{26}$$

• Q-condition (same-spin):

$$\bar{Q}_{ij}^{\sigma\sigma} = \begin{cases} n_i^2 - f(n_i, n_i) & \text{if } i = j \\ n_i n_j - f(n_i, n_j) & \text{if } i < j \\ 2(1 - n_i - n_j) + n_i n_j + f(n_i, n_j) & \text{if } i > j \end{cases} \tag{27}$$

• G-condition (same-spin):

$$\bar{G}_{ij}^{\sigma\sigma} = \begin{cases} \bar{\gamma}_i & \text{if } i = j \\ n_i - f(n_i, n_j) & \text{if } i < j \\ n_j - f(n_i, n_j) & \text{if } i > j \end{cases} \tag{28}$$

where $\bar{\gamma}_i$ are the (basis-set dependent) eigenvalues of the following matrix:

$$\gamma_{ij} = \begin{cases} n_i(n_i + 1) - f(n_i, n_i) & \text{if } i = j \\ n_i n_j & \text{if } i \neq j \end{cases} \tag{29}$$

where we have assumed that all occupancies n_i refer to σ-spin natural orbitals (*i.e.*, $n_i \equiv n_i^\sigma$). The latter matrices dimension is M, the size of the basis set. In the case of P and Q conditions, K-functionals produce two sets (one for each spin case: $\alpha\alpha$ and $\beta\beta$) of M^2 basis-set-independent eigenvalues, giving rise to some conditions that one can impose in the corresponding functional. The G condition, on the other hand, produces $M(M-1)$ basis-set independent eigenvalues (and some corresponding

conditions on $f(n_i,n_j)$) and M basis-set-dependent eigenvalues that can only be checked *a posteriori*.

PNOF

In the case of PNOF, the matrices to diagonalize are given in terms of Δ and Π. The conditions will involve like- and opposite-spin components of the 2-DM because, unlike K-functionals, PNOFs construct all the spin-components of the cumulant. We collect below the eigenvalues of these matrices:

• P-condition (same-spin):

$$\bar{P}_{ij}^{\sigma\sigma} = \begin{cases} 2(n_i n_j - \Delta_{ij}) & \text{if } i > j \\ 0 & \text{if } i \leq j \end{cases} \tag{30}$$

• Q-condition (same-spin):

$$\bar{Q}_{ij}^{\sigma\sigma} = \begin{cases} 2(h_i h_j - \Delta_{ij}) & \text{if } i > j \\ 0 & \text{if } i \leq j \end{cases} \tag{31}$$

• G-condition (same-spin):

$$\bar{G}_{ij}^{\sigma\sigma} = \begin{cases} n_i h_j + \Delta_{ij} & \text{if } i \neq j \\ \bar{\gamma}_i & \text{if } i = j \end{cases} \tag{32}$$

where $h_i = (1 - n_i)$ and $\bar{\gamma}_i$ are the (basis-set dependent) eigenvalues of the following matrix:

$$\gamma_{ij} = \begin{cases} n_i & \text{if } i = j \\ n_i n_j - \Delta_{ij} & \text{if } i \neq j \end{cases} \tag{33}$$

where $n_i = n_i^\sigma$ for eqn (30)–(33).

The opposite-spin components of P, Q and G matrices produce the following eigenvalues:

• P-condition (opposite-spin):

$$\bar{P}_{ij}^{\sigma\sigma'} = \begin{cases} n_i n_j - \Delta_{ij} & \text{if } i \neq j \\ \bar{\gamma}_i & \text{if } i = j \end{cases} \tag{34}$$

where $\bar{\gamma}_i$ are the (basis-set dependent) eigenvalues of the following matrix:

$$\gamma_{ij} = \begin{cases} n_i & \text{if } i = j \\ \Pi_{ij} & \text{if } i \neq j \end{cases} \tag{35}$$

• Q-condition (opposite-spin):

$$\bar{Q}_{ij}^{\sigma\sigma'} = \begin{cases} h_i h_j - \Delta_{ij} & \text{if } i \neq j \\ \bar{\gamma}_i & \text{if } i = j \end{cases} \tag{36}$$

$\bar{\gamma}_i$ are the (basis-set dependent) eigenvalues of the following matrix:

$$\gamma_{ij} = \begin{cases} h_i & \text{if } i = j \\ \Pi_{ij} & \text{if } i \neq j \end{cases} \tag{37}$$

- *G*-condition (opposite-spin):

$$\bar{G}_{ij}^{\sigma\sigma'} = \begin{cases} \dfrac{n_i h_j + \Delta_{ij} + \Delta_{ji} + n_j h_i + \sqrt{(n_i h_j - n_j h_i)^2 + 4\Pi_{ij}^2}}{2} & \text{if } i < j \\[3mm] \dfrac{n_i h_j + \Delta_{ij} + \Delta_{ji} + n_j h_i - \sqrt{(n_i h_j - n_j h_i)^2 + 4\Pi_{ij}^2}}{2} & \text{if } i > j \\[3mm] n_i & \text{if } i = j \end{cases} \tag{38}$$

where $n_i = n_i^\sigma$ for eqn (34)–(38), which are only valid for a closed-shell restricted system.

Acknowledgements

References

1 T. Gilbert, *Phys. Rev. B: Solid State*, 1975, **12**, 2111.
2 M. Piris and J. Ugalde, *Int. J. Quantum Chem.*, 2014, **114**, 1169–1175.
3 K. Pernal and K. J. H. Giesbertz, *Top. Curr. Chem.*, 2015, **368**, 125.
4 A. M. K. Müller, *Phys. Lett.*, 1984, **105A**, 446–452.
5 M. A. Buijse, PhD thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1991.
6 S. Goedecker and C. J. Umrigar, *Phys. Rev. Lett.*, 1998, **81**, 866–869.
7 G. Csányi and T. A. Arias, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **61**, 7348.
8 M. A. Buijse and E. J. Baerends, *Mol. Phys.*, 2002, **100**, 401–421.
9 G. Csányi, S. Goedecker and T. A. Arias, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2002, **65**, 032510.
10 O. Gritsenko, K. Pernal and E. J. Baerends, *J. Chem. Phys.*, 2005, **122**, 204102.
11 M. Piris, *Int. J. Quantum Chem.*, 2006, **106**, 1093–1104.
12 M. Piris, X. Lopez and J. M. Ugalde, *J. Chem. Phys.*, 2007, **126**, 214103.
13 M. A. L. Marques and N. N. Lathiotakis, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2008, **77**, 032509.
14 D. R. Rohr, K. Pernal, O. V. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2008, **129**, 164105.
15 J. Cioslowski and K. Pernal, *J. Chem. Phys.*, 1999, **111**, 3396–3400.
16 J. Cioslowski and K. Pernal, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2000, **61**, 034503.
17 S. Sharma, J. K. Dewhurst, N. N. Lathiotakis and E. K. Gross, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 201103.
18 M. Piris, J. M. Matxain, X. Lopez and J. M. Ugalde, *J. Chem. Phys.*, 2010, **132**, 031103.
19 M. Piris, J. M. Matxain, X. Lopez and J. M. Ugalde, *J. Chem. Phys.*, 2010, **133**, 111101.
20 M. Piris, X. Lopez, F. Ruipérez, J. M. Matxain and J. M. Ugalde, *J. Chem. Phys.*, 2011, **134**, 164102.
21 M. Piris, *J. Chem. Phys.*, 2014, **141**, 044107.
22 M. Piris, *Phys. Rev. Lett.*, 2017, **119**, 063002.
23 J. Cioslowski, M. Piris and E. Matito, *J. Chem. Phys.*, 2015, **143**, 214101.
24 F. Ruipérez, M. Piris, J. Ugalde and J. Matxain, *Phys. Chem. Chem. Phys.*, 2013, **15**, 2055–2062.
25 X. Lopez, F. Ruipérez, M. Piris, J. M. Matxain, E. Matito and J. M. Ugalde, *J. Chem. Theory Comput.*, 2012, **8**, 2646–2652.
26 X. Lopez, M. Piris, M. Nakano and B. Champagne, *J. Phys. B: At., Mol. Opt. Phys.*, 2013, **47**, 015101.
27 E. Ramos-Cordoba, X. Lopez, M. Piris and E. Matito, *J. Chem. Phys.*, 2015, **143**, 164112.
28 E. Ramos-Cordoba, P. Salvador and E. Matito, *Phys. Chem. Chem. Phys.*, 2016, **18**, 24015–24023.
29 E. Ramos-Cordoba and E. Matito, *J. Chem. Theory Comput.*, 2017, **13**, 2705.
30 M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew and K. A. Lyssenko, *Science*, 2017, **355**, 49–52.
31 E. Ramos-Cordoba, P. Salvador, M. Piris and E. Matito, *J. Chem. Phys.*, 2014, **141**, 234101.
32 E. Ramos-Cordoba, E. Matito, I. Mayer and P. Salvador, *J. Chem. Theory Comput.*, 2012, **8**, 1270–1279.
33 N. R. Kestner and O. Sinanoglu, *Phys. Rev.*, 1962, **128**, 2687.
34 E. Matito, D. Casanova, X. Lopez and J. M. Ugalde, *Theor. Chim. Acta*, 2016, **135**, 226.
35 P.-O. Löwdin and H. Shull, *Phys. Rev.*, 1956, **101**, 1730–1739.
36 K. Pernal and J. Cioslowski, *J. Chem. Phys.*, 2004, **120**, 5987–5992.
37 J. Cioslowski and K. Pernal, *Chem. Phys. Lett.*, 2006, **430**, 188–190.
38 K. J. H. Giesbertz and R. van Leeuwen, *J. Chem. Phys.*, 2013, **139**, 104110.
39 X. W. Sheng, Ł. M. Mentel, O. V. Gritsenko and E. J. Baerends, *J. Chem. Phys.*, 2013, **138**, 164105.
40 P.-O. Löwdin, *Phys. Rev.*, 1955, **97**, 1474–1489.
41 R. McWeeny, *Rev. Mod. Phys.*, 1960, **32**, 335–369.
42 A. J. Coleman and V. I. Yukalov, *Reduced density matrices: Coulson's challenge*, Springer Verlag, Berlin, 2000, vol. 72.
43 M. Piris, *J. Math. Chem.*, 1999, **25**, 47–54.
44 J. M. Herbert and J. E. Harriman, *J. Chem. Phys.*, 2003, **118**, 10835–10846.
45 W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.*, 1999, **110**, 2800–2809.
46 M. Piris, J. Matxain and X. Lopez, *J. Chem. Phys.*, 2013, **139**, 234109.
47 M. Piris and N. H. March, *J. Phys. Chem. A*, 2015, **119**, 10190–10194.
48 J. Cioslowski and K. Pernal, *J. Chem. Phys.*, 2000, **113**, 8434.
49 M. Piris, *Int. J. Quantum Chem.*, 2013, **113**, 620–630.
50 D. A. Mazziotti, *Phys. Rev. Lett.*, 2012, **108**, 263002.

51 A. J. Coleman, *Rev. Mod. Phys.*, 1963, **35**, 668–687.

52 C. Garrod and J. K. Percus, *J. Math. Phys.*, 1964, **5**, 1756–1776.

53 F. Weinhold and E. B. Wilson Jr, *J. Chem. Phys.*, 1967, **47**, 2298–2311.

54 A. S. Eddington, *Fundamental theory*, Cambridge University Press, Cambridge, 1946.

55 R. F. W. Bader and M. E. Stephens, *Chem. Phys. Lett.*, 1974, **26**, 445.

56 X. Fradera, M. A. Austen and R. F. W. Bader, *J. Phys. Chem. A*, 1999, **103**, 304–314.

57 E. Matito, M. Solà, P. Salvador and M. Duran, *Faraday Discuss.*, 2007, **135**, 325–345.

58 F. Feixas, E. Matito, M. Duran, M. Solà and B. Silvi, *J. Chem. Theory Comput.*, 2010, **6**, 2736–2742.

59 F. Feixas, J. Vandenbussche, P. Bultinck, E. Matito and M. Solà, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20690–20703.

60 M. García-Revilla, E. Francisco, A. Costales and A. M. Pendás, *J. Phys. Chem. A*, 2012, **116**, 1237–1250.

61 F. Feixas, M. Solà, J. M. Barroso, J. M. Ugalde and E. Matito, *J. Chem. Theory Comput.*, 2014, **10**, 3055–3065.

62 F. Feixas, M. Rodrguez-Mayorga, E. Matito and M. Solà, *Comput. Theor. Chem.*, 2015, **1053**, 173–179.

63 M. Rodríguez-Mayorga, E. Ramos-Cordoba, P. Salvador, M. Solà and E. Matito, *Mol. Phys.*, 2016, **114**, 1345.

64 M. Rodríguez-Mayorga, E. Ramos-Cordoba, F. Feixas and E. Matito, *Phys. Chem. Chem. Phys.*, 2017, **19**, 4522.

65 E. Santos, *An. R. Soc. Esp. Fis. Quim.*, 1968, **64**, 117.

66 M. Taut, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1993, **48**, 3561.

67 P. Hessler, J. Park and K. Burke, *Phys. Rev. Lett.*, 1999, **82**, 378.

68 P. M. Laufer and J. B. Krieger, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1986, **33**, 1480–1491.

69 S. Kais, D. R. Hersbach, N. C. Handy, C. W. Murray and G. J. Laming, *J. Chem. Phys.*, 1993, **99**, 417.

70 C. Filippi, C. J. Umrigar and M. Taut, *J. Chem. Phys.*, 1994, **100**, 1290.

71 C.-J. Huang and C. J. Umrigar, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1997, **56**, 290.

72 M. Taut, A. Ernst and H. Eschrig, *J. Phys. B: At., Mol. Opt. Phys.*, 1998, **31**, 2689.

73 Z. Qian and V. Sahni, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1998, **57**, 2527.

74 S. Ivanov, K. Burke and M. Levy, *J. Chem. Phys.*, 1999, **110**, 10262.

75 E. V. Ludeña, V. Karasiev, A. Artemiev and D. Gómez, in *Functional N-representability in density Matrix and Density Functional Theory: An illustration for Hooke's Atom*, ed. J. Cioslowski, Kluwer Academic/Plenum Publishers, New York, 2000, ch. 10.

76 W. M. Zhu and S. B. Trickey, *J. Chem. Phys.*, 2006, **125**, 094317.

77 J. Cioslowski and E. Matito, *J. Chem. Theory Comput.*, 2011, **7**, 915.

78 J. Cioslowski and E. Matito, *J. Chem. Phys.*, 2011, **134**, 116101.

79 J. Cioslowski, K. Strasburger and E. Matito, *J. Chem. Phys.*, 2012, **136**, 194112.

80 J. Cioslowski, K. Strasburger and E. Matito, *J. Chem. Phys.*, 2014, **141**, 044128.

81 I. Mitxelena, M. Piris and M. Rodríguez-Mayorga, *J. Phys.: Condens. Matter*, 2017, in press, http://iopscience.iop.org/article/10.1088/1361-648X/aa80ca.

82 P.-F. Loos and P. M. Gill, *Phys. Rev. Lett.*, 2009, **103**, 123008.

83 P.-F. Loos and P. M. Gill, *Chem. Phys. Lett.*, 2010, **500**, 1.

84 P.-F. Loos and P. M. Gill, *Phys. Rev. Lett.*, 2012, **108**, 083002.

85 P.-F. Loos and P. M. Gill, *J. Chem. Phys.*, 2013, **138**, 164124.

86 E. Matito, J. Cioslowski and S. F. Vyboishchikov, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6712.

87 P. J. Knowles and N. C. Handy, *Chem. Phys. Lett.*, 1984, **111**, 315–321.

88 P. J. Knowles and N. C. Handy, *Comput. Phys. Commun.*, 1989, **54**, 75.

89 E. Matito and F. Feixas, *DMn program*, University of Girona (Spain) and University of Szczecin (Poland), 2009.

90 M. Rodríguez-Mayorga, *RHO2-OPS: 2-DM Operations*, Institute of Computational Chemistry and Catalysis, University of Girona, Catalonia, Spain, 2016.

91 J. Cioslowski and G. Liu, *J. Chem. Phys.*, 1996, **105**, 4151–4158.

92 M. Rodríguez-Mayorga, *RHO-OPS: Density Operations*, Institute of Computational Chemistry and Catalysis, University of Girona, Catalonia, Spain, 2015.

93 E. Matito, *ESI-3D: Electron Sharing Indices Program for 3D Molecular Space Partitioning*, Institute of Computational Chemistry and Catalysis, University of Girona, Catalonia, Spain, 2015.

94 E. Matito, M. Duran and M. Solà, *J. Chem. Phys.*, 2005, **122**, 014109.

95 Y. G. Wang, C. Matta and N. H. Werstiuk, *J. Comput. Chem.*, 2003, **24**, 1720–1729.

96 F. Feixas, J. Jiménez-Halla, E. Matito, J. Poater and M. Solà, *J. Chem. Theory Comput.*, 2010, **6**, 1118–1130.

97 I. Ruiz, E. Matito, F. J. Holgun-Gallego, E. Francisco, Á. M. Pendás and T. Rocha-Rinza, *Theor. Chem. Acc.*, 2016, **135**, 209.

98 J. M. Mercero, M. Rodríguez-Mayorga, E. Matito, X. Lopez and J. M. Ugalde, *Can. J. Chem.*, 2017, **94**, 998–1001.

99 C. A. Coulson and A. H. Neilson, *Proc. Phys. Soc., London*, 1961, **78**, 831.

100 A. Savin, *Int. J. Quantum Chem.*, 1988, **34**, 59–69.

101 A. Savin, *Density functional methods in chemistry*, Springer, 1991, pp. 213–230.

102 P. M. Gill, R. D. Adamson and J. A. Pople, *Mol. Phys.*, 1996, **88**, 1005–1009.

103 T. M. Henderson, A. F. Izmaylov, G. E. Scuseria and A. Savin, *J. Chem. Theory Comput.*, 2008, **4**, 1254–1262.

104 B. G. Janesko, T. M. Henderson and G. E. Scuseria, *Phys. Chem. Chem. Phys.*, 2009, **11**, 443–454.

105 J. Jaramillo, G. E. Scuseria and M. Ernzerhof, *J. Chem. Phys.*, 2003, **118**, 1068–1073.

106 T. M. Henderson, A. F. Izmaylov, G. E. Scuseria and A. Savin, *J. Chem. Phys.*, 2007, **127**, 221103.

107 M. Rodriguez-Mayorga, MSc thesis, Univ. Girona, Girona, 2013.