Communication: Generalized canonical purification for density matrix minimization

Lionel A. Truflandier, Rivo M. Dianzinga, and David R. Bowler

Citation: The Journal of Chemical Physics **144**, 091102 (2016); doi: 10.1063/1.4943213
View online: http://dx.doi.org/10.1063/1.4943213
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/144/9?ver=pdfcov
Published by the AIP Publishing

Communication: Generalized canonical purification for density matrix minimization

Lionel A. Truflandier,[1,a)] Rivo M. Dianzinga,[1] and David R. Bowler[2]
[1]*Institut des Sciences Moléculaires, Université Bordeaux, CNRS UMR 5255, 351 cours de la Libération, 33405 Talence cedex, France*
[2]*London Centre for Nanotechnology, UCL, 17-19 Gordon St., London WC1H 0AH, United Kingdom; Department of Physics and Astronomy, UCL, Gower St., London WC1E 6BT, United Kingdom; and International Centre for Materials Nanoarchitechtonics (MANA), National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan*

(Received 6 January 2016; accepted 22 February 2016; published online 3 March 2016)

A Lagrangian formulation for the constrained search for the N-representable one-particle density matrix based on the McWeeny idempotency error minimization is proposed, which converges systematically to the ground state. A closed form of the canonical purification is derived for which no *a posteriori* adjustment on the trace of the density matrix is needed. The relationship with comparable methods is discussed, showing their possible generalization through the *hole-particle* duality. The appealing simplicity of this *self-consistent* recursion relation along with its low computational complexity could prove useful as an alternative to diagonalization in solving dense and sparse matrix eigenvalue problems. © *2016 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4943213]

As suggested 60 years ago,[1] the idempotency property of the density matrix (DM) along with a minimization algorithm would be sufficient to solve for the electronic structure without relying on the time consuming step of calculating the eigenstates of the Hamiltonian matrix. The celebrated McWeeny purification formula[2] has inspired major advances in electronic structure theory based on (conjugate-gradient) DM minimization[3–8] (DMM) or DM polynomial expansion[9,10] (DMPE), where the DM is evaluated by the recursive application of projection polynomials (commonly referred to as *purification*). DMPE resolution includes the Chebyshev polynomial recursion,[9–15] the Newton-Schultz sign matrix iteration,[16–18] the trace-correcting[19] and the trace-resetting[20] purification (TCP and TRS, respectively), and the Palser and Manolopoulos canonical purification (PMCP).[21] They constitute, with sparse matrix algebra, the principal ingredient for efficient linear-scaling tight-binding (TB) and self-consistent field (SCF) theories.[22,23] Since all these methods were originally derived within the *grand canonical ensemble*,[24] for a given total number of states (M), none of them are expected to yield the correct number of occupied states (N) unless the chemical potential (μ) is known exactly. As a result, their implementation to the *canonical ensemble* involves heuristic considerations, where the value of μ[12] or the polynomial expansion[19] is adapted *a posteriori* to reach the correct value for N, which adds irremediably to the computational complexity. Despite the remarkable performances of the DMPE approaches for solving for sparse[6,25] and dense[26–28] DMs, it remains desirable to develop an approach that overcomes the use of the chemical potential while respecting the canonical requirement of constant-N.

In this letter, we derive a rigorous and variational constrained search for the one-particle density matrix which does not rely on *ad hoc* adjustments and respects the N-representability constraint throughout the minimization process. We shall start from the McWeeny unconstrained minimization of the error in the idempotency of the density matrix,[1] given by

$$\underset{D\to\mathcal{D}_\mu}{\text{minimize}}\ \Omega_{\text{McW}}\{D;(\mathcal{H},\mu)\}, \tag{1a}$$

$$\text{with}\ \ \Omega_{\text{McW}} = \text{Tr}\{(D^2-D)^2\}, \tag{1b}$$

where for a given fixed Hamiltonian[29] $\mathcal{H}$ and chemical potential μ, the density matrix $\mathcal{D}_\mu$ is the ground-state for that Hamiltonian and chemical potential. The initial guess (D_0) is generally constructed as a function $\mathcal{H}$, suitably scaled,

$$D_0 = \beta_1 I + \beta_2(\mu I - \mathcal{H}), \tag{2}$$

where β_1 and β_2 stand for preconditioning constants such that the eigenvalues of D_0 lie within a predefined range. The double-well shape of the McWeeny function with 3 stationary points: 2 minima at $x_p = 1$ and $x_{\bar{p}} = 0$ and 1 local maximum at $x_m = \frac{1}{2}$ (see Fig. 1(a), red curve), are important features in developing robust DMM algorithms. Finding the minimum of Ω_{McW} would be easily performed by stepwise gradient descent,[1] where the DM is updated at each iteration n,

$$D_{n+1} = D_n - \sigma_n \nabla\Omega_{\text{McW}}, \tag{3a}$$

$$\text{with}\ \ \nabla\Omega_{\text{McW}} = 2\left(2D_n^3 - 3D_n^2 + D_n\right), \tag{3b}$$

and $\sigma_n \geq 0$ represents the step length in the negative direction of the gradient. Considering an optimal fixed step length descent ($\sigma = 1/2$), on inserting Eq. (3b) into Eq. (3a), the McWeeny purification formula appears,

$$D_{n+1} = 3D_n^2 - 2D_n^3, \tag{4}$$

where the right-hand side of the equation above can be view as an auxiliary DM. For a well-conditioned D_0, i.e., $\lambda(D_0) \in [-\frac{1}{2}, \frac{3}{2}]$, repeated application of the recursion

a)Electronic mail: lionel.truflandier@u-bordeaux.fr

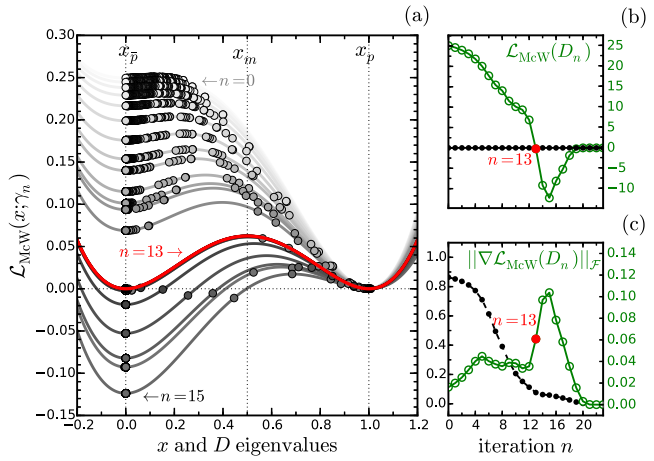

FIG. 1. (a) Convergence of the McWeeny Lagrangian and density matrix eigenvalues during the course of the minimization using a test Hamiltonian and an occupation factor $\theta = 0.10$. A grey scale is used to guide the eye during the processes of purification. Each curve is a plot of the function $\mathcal{L}_{\mathrm{McW}}(x;\gamma_n)$ computed at each iteration n. The red line corresponds to $\mathcal{L}_{\mathrm{McW}}(x;0) = \Omega_{\mathrm{McW}}$. (b) Convergence of $\mathcal{L}_{\mathrm{McW}}$ (green circles) and the trace conservation $\mathrm{Tr}\{D_n\} - N$ (black dots). (c) Convergence of $\|\nabla\mathcal{L}_{\mathrm{McW}}\|_{\mathcal{F}}$ (green circles) and $\|D_n\|_{\mathcal{F}} - N$ (black dots).

identity [Eq. (4)] naturally drives the eigenvalues of D_{n+1} towards 0 or 1. For basic TB Hamiltonians where the occupation factor ($\theta = N/M$) is close to 1/2 and μ can be determined by symmetry[21] or when the input DM is already strongly idempotent, the minimization principle (1a) is able, on its own, to deliver the correct N-representable ground-state DM ($\mathcal{D}$). Beyond these very specific cases, we have to enforce the objective function (1b) to keep N constant during the minimization. From Eq. (4), a sufficient condition would be to impose the trace of the auxiliary DM to give the correct number of occupied states. This leads us to solve a constrained optimization problem which can be formulated in terms of the McWeeny Lagrangian ($\mathcal{L}_{\mathrm{McW}}$) by

$$\underset{\substack{\{D\to\mathcal{D}|\mathrm{Tr}\{D\}=N\}\\ \gamma}}{\text{minimize}} \ \mathcal{L}_{\mathrm{McW}}\{D,\gamma;(\mathcal{H}),N\}, \tag{5a}$$

$$\text{with } \mathcal{L}_{\mathrm{McW}} = \Omega_{\mathrm{McW}} - \gamma\left(\mathrm{Tr}\{3D^2 - 2D^3\} - N\right), \tag{5b}$$

where γ is the constant-N Lagrange multiplier. The McWeeny Lagrangian can be minimized using

$$\nabla\mathcal{L}_{\mathrm{McW}} = \nabla\Omega_{\mathrm{McW}} - 6\gamma\left(D - D^2\right), \tag{6a}$$

$$\partial_\gamma\mathcal{L}_{\mathrm{McW}} = \mathrm{Tr}\{3D^2 - 2D^3\} - N. \tag{6b}$$

Taking trace Eq. (6a) we obtain the expression for γ,

$$\gamma = \frac{1}{3} - \frac{2}{3}c - \frac{1}{6}d, \tag{7a}$$

$$\text{with } c = \frac{\mathrm{Tr}\{D^2 - D^3\}}{\mathrm{Tr}\{D - D^2\}}, \tag{7b}$$

$$d = \frac{\mathrm{Tr}\{\nabla\mathcal{L}_{\mathrm{McW}}\}}{\mathrm{Tr}\{D - D^2\}}. \tag{7c}$$

Then, Eqs. (6a) and (7a) are updated at each iteration by requiring $\mathrm{Tr}\{\nabla\mathcal{L}_{\mathrm{McW}}\} = 0$, that is $d = 0$, for all D. As a result, given D_0 such that $\mathrm{Tr}\{D_0\} = N$ and $[\mathcal{H}, D_0] = 0$, from the fixed-step gradient descent minimization described above, we obtain a recursion formula,

$$D_{n+1} = D_n - \frac{1}{2}\nabla\mathcal{L}_{\mathrm{McW}}\{D_n;\gamma_n\}, \tag{8}$$

which guarantees $\mathrm{Tr}\{D_{n+1}\} = N$ and $[\mathcal{H}, D_{n+1}] = 0$, $\forall n$. Added to the preconditioning $\lambda(D_0) \in [0,1]$, the iterative process should approach the (one-particle) ground-state energy $\mathcal{E} = \mathrm{Tr}\{\mathcal{H}\mathcal{D}\}$ variationally. The parameter c [Eq. (7b)] is recognized as the unstable fixed point introduced in Ref. 21, where $c \in [0,1]$. As a result, the interval $[-\frac{1}{3}, \frac{1}{3}]$ constitutes the stable variational domain of γ.

The variation of the McWeeny Lagrangian function and the DM eigenvalues during the course of the minimization is presented in Fig. 1(a) for a test Hamiltonian with $N = 10$, $M = 100$, and a suitably conditioned initial guess (*vide infra*). The corresponding convergence profiles of $\mathcal{L}_{\mathrm{McW}}$ and $\|\nabla\mathcal{L}_{\mathrm{McW}}\|$ (green circles) are reported on Figs. 1(b) and 1(c), respectively, along with the trace conservation $\mathrm{Tr}\{D_n\} - N$ and the DM norm convergence $\|D_n\| - N$ (black dots). We may notice first that for $\gamma = 0$ (or $c = x_m = \frac{1}{2}$), $\mathcal{L}_{\mathrm{McW}}$ simplifies to Ω_{McW}. For intermediate states, $\gamma \in [-\frac{1}{3}, 0] \cup [0, \frac{1}{3}]$, the symmetry of Ω_{McW} is lost and the shape of $\mathcal{L}_{\mathrm{McW}}(x, \gamma_n)$ drives the eigenvalues in the *hole* (left) or in the *particle* (right) well. From the grey scale in Fig. 1(a), we observe how γ_n influences $\mathcal{L}_{\mathrm{McW}}$ (along the y-axis) at $x_{\bar{p}}$ and the abscissa of the second stationary point x_m which is free to move in $[x_{\bar{p}}, x_p]$. This yields to transform the *hole* well from a local ($n = 0$) to a global ($n = 15$) minimum (or conversely the *particule* well from a global to a local minimum). At the boundary values $\gamma = \{-\frac{1}{3}, \frac{1}{3}\}$, $x_{\bar{p}}$ and x_m merged to a saddle point in such a way that only one global minimum left at x_p. Notice that, for situations where $\gamma \notin [-\frac{1}{3}, \frac{1}{3}]$, the saddle point transforms to a maximum and runaway solutions may appear. Nevertheless, as long as D_0 is well conditioned, such kind of critical problem should not be encountered.

Figs. 1(b) and 1(c) highlight the minimization mechanism: (i) from iterate $n = 0$ to 12; $\gamma \to 0^+$, $\mathcal{L}_{\mathrm{McW}}$ follows the search direction and decreases monotonically. (ii) At iterate $n = 13$; $\gamma \simeq 0$, $\mathcal{L}_{\mathrm{McW}}$ is close to the target value but the gradient residual is nonzero. (iii) From $n = 14$ to 15; $\gamma < 0$, the search direction is inverted. (iv) At iterate $n = 16$, all the eigenvalues are trapped in their respective wells. (iii) From iterate $n = 17$ to 23, $\gamma \to 0^-$, we are in the McWeeny regime [Eq. (4)] and $\mathcal{L}_{\mathrm{McW}}$ eventually reaches the global minimum.

Taking advantage of the closure relation,

$$\bar{D} + D = I, \tag{9}$$

where $\bar{D}$ stands for the *hole* density matrix,[30] a more appealing form for the McWeeny canonical purification [Eq. (8)] can be derived by reformulating Eqs. (6a) and (7b) in terms of D and $\bar{D}$,

$$D_{n+1} = D_n + 2\left(D_n^2\bar{D}_n - \frac{\mathrm{Tr}\{D_n^2\bar{D}_n\}}{\mathrm{Tr}\{D_n\bar{D}_n\}}D_n\bar{D}_n\right). \tag{10}$$

Notice that since at convergence $D\bar{D} = 0$, $\mathrm{Tr}\{D\bar{D}\}$ must be chosen as the termination criterion in the recursion of Eq. (10) to avoid numerical instabilities when approaching the minima. The closed-form of this recurrence relation is remarkable: providing N and $\mathcal{H}$ used to build D_0 [Eq. (2)], we have

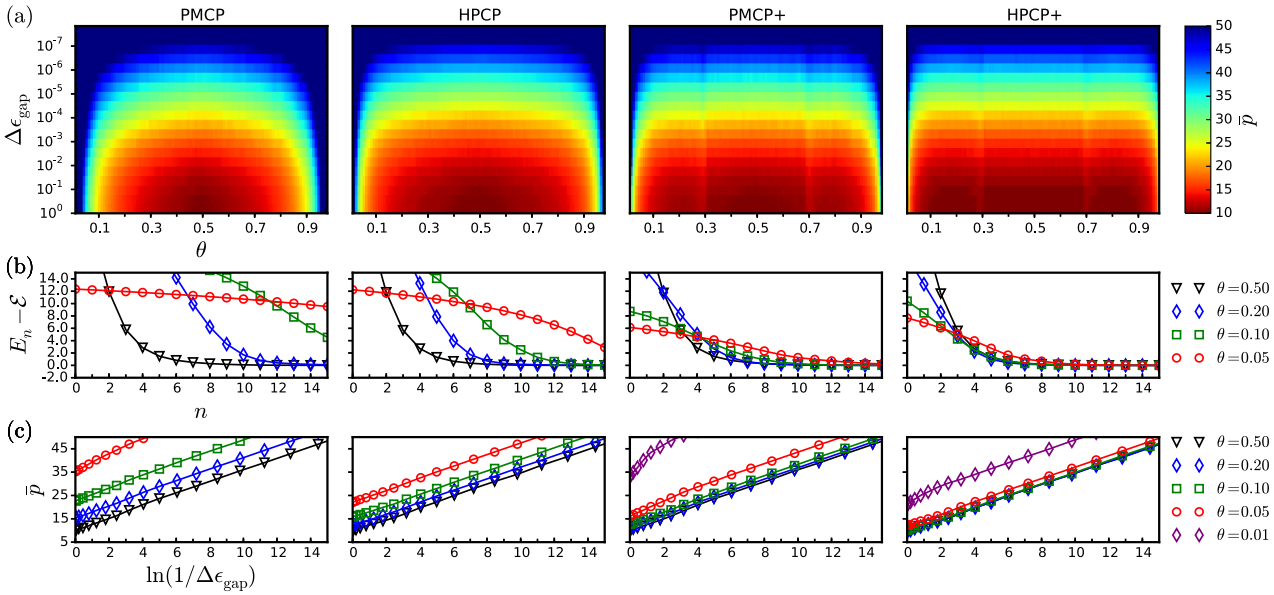

FIG. 2. (a) Color maps displaying the average number of purifications ($\bar{p}$) as the function of the filling factor (θ) and energy gap ($\Delta\epsilon_{\rm gap}$). Results obtained from the PMCP and HPCP methods using the initial guess of Eqs. (2)-(11) and (2)-(14) (notated PMCP+ and HPCP+). Each pixel on the maps corresponds to an average over 32 test Hamiltonians. (b) Energy convergence profiles with respect to the first 15 iterations for selected values of θ. (c) Average number of purifications as a function of $\ln(1/\Delta\epsilon_{\rm gap})$.

a self-consistent purification transformation which should converge to $\mathcal{D}$ without any support of heuristic adjustments. Indeed, Eq. (10) can also be derived from the PMCP relations by working on both D and $\bar{D}$ and enforcing relation (9) at each iteration (see the Appendix). Consequently, we can also demonstrate[31] that the *hole-particle* canonical purification (HPCP) of Eq. (10) converges quadratically on $\mathcal{D}$ as shown in Fig. 2(b).

To assess the efficiency and limitations of the HPCP, we have investigated the dependence of the number of purifications (p) on the occupation factor (θ) and the energy gap ($\Delta\epsilon_{\rm gap} = \epsilon_{N+1} - \epsilon_N$), defined by the higher-occupied (ϵ_N) and lower-unoccupied (ϵ_{N+1}) states. Similarly to the protocol of Niklasson,[15,19] sequences of $M \times M$ dense Hamiltonian matrices ($M = 100$) with vanishing off-diagonal elements were generated, having eigenvalues randomly distributed in the range $[-2.5, \epsilon_N] \cup [\epsilon_{N+1}, 2.5]$ for various $\Delta\epsilon_{\rm gap} \in [10^{-7}, 1.0]$. As a first test, results are compared to the PMCP,[21] along with the original initial guess [Eq. (2)], where $\beta_1 = \theta$ and $\beta_2 = \min\{\beta, \bar{\beta}\}$, with

$$\beta = \frac{\theta}{\widetilde{\mathcal{H}}_{\rm max} - \mu}, \quad \bar{\beta} = \frac{\bar{\theta}}{\mu - \widetilde{\mathcal{H}}_{\rm min}}, \quad \mu \simeq \widetilde{\mu} = \frac{{\rm Tr}\{\mathcal{H}\}}{M}, \tag{11}$$

and $\bar{\theta} = 1 - \theta = \bar{N}/M$, $\bar{N}$ being the number of unoccupied states. The lower and upper bounds of the Hamiltonian eigenspectrum ($\widetilde{\mathcal{H}}_{\rm min}$ and $\widetilde{\mathcal{H}}_{\rm max}$, respectively) were estimated from to the Geršgorin's disc theorem.[32] The preconditioning of D_0 given in Eq. (11) guarantees that the DM eigenvalues lie in the interval $[0, 1]$ and gives rise to the following additional constraints:

$$ {\rm Tr}\{D_0\} = N, \tag{12a}$$
$$ {\rm Tr}\{D_0\} > {\rm Tr}\{D_0^2\} > {\rm Tr}\{D_0^3\}, \tag{12b}$$
$$ {\rm Tr}\{D_0^3\} > 2{\rm Tr}\{D_0^2\} - {\rm Tr}\{D_0\}, \tag{12c}$$

which are also necessary and sufficient conditions for $c \in [0, 1]$ at the first iteration. Convergence was achieved with respect to the idempotency property, such that ${\rm Tr}\{D_n\bar{D}_n\} \leq 10^{-6}$ for all the calculations. Additional tests on the Frobenius norm[33] and the eigenvalues of the converged density matrix (D_∞) were performed, using

$$ \|D_\infty\|_{\mathcal{F}} - \sqrt{{\rm Tr}\{D_\infty\}} < 10^{-6}, \tag{13a}$$
$$ \|D_\infty\|_{\mathcal{F}} - N < 10^{-6}, \tag{13b}$$
$$ \|{\rm diag}\{D_\infty\} - {\rm diag}\{I_N, 0_{\bar{N}}\}\|_{\mathcal{F}} < 10^{-6}, \tag{13c}$$

which ensures that, at convergence, the representation of D_∞ is orthogonal, and D_∞ corresponds to $\mathcal{D}$.

The variation of the average number of purifications ($\bar{p}$) with respect to θ and $\Delta\epsilon_{\rm gap}$ is displayed in Fig. 2(a) using a color map for $\bar{p} \in [10, 50]$. For a given energy gap, the HPCP shows a net improvement over the PMCP approach regarding moderate low and high occupation factors. Nevertheless, as previously noted by Niklasson and Mazziotti,[19,30] the extreme values of θ remain pathological for the original canonical purification and to a lesser extent for the HPCP. One solution would be to break the symmetry of the McWeeny function by moving x_m towards x_p or $x_{\bar{p}}$ depending on the θ value. Basically, this requires a higher polynomial degree for $\Omega_{\rm McW}$, i.e., ${\rm Tr}\{(D^n - D)^2\}_{n>2}$, resulting in a higher computational complexity. Assuming optimal programming, we emphasize that the PMCP and HPCP involved only two matrix multiplications per iteration. As already proved in Ref. 21 and highlighted by the energy convergence profiles in Fig. 2(b), the PMCP and HPCP approach $\mathcal{E}$ monotonically.

The dependence of $\bar{p}$ on the band gap plotted in Fig. 2(c) confirms the early numerical experiments,[19,25] where $\bar{p}$ increases linearly with respect to $\ln(1/\Delta\epsilon_{\rm gap})$. The influence of θ is clearly apparent if we compare the minimum number of purifications as required for the wider band gap (y-axis

intercept), where for example, with $\theta = 0.5$, both canonical purifications reach the ideal value of about 10 purifications, whereas for $\theta = 0.05$, $\bar{p}_{\mathrm{HPCP}} = 23$ and $\bar{p}_{\mathrm{PMCP}} = 37$.

Let us consider how to improve the performance of the canonical purifications by working on the initial guess, regarding the *hole-particle* equivalence (or duality[30]). Instead of searching for D, we may choose to purify $\bar{D}$, which simply requires replacing D with $\bar{D}$ in relation (10). In that case, the initial hole density matrix, satisfying $\lambda(\bar{D}_0) \in [0,1]$, would be given by Eqs. (2) and (11), with $\beta_1 = \bar{\theta}$ and $\beta_2 = -\max\{\beta, \bar{\beta}\}$. Then, intuitively, the guess for the particle density matrix should be improved by using this additional information. Therefore, a more general preconditioning is proposed,

$$D_0^+ = \alpha D_0 + (1-\alpha)(I - \bar{D}_0), \tag{14}$$

where α can be viewed as a mixing coefficient.[34] Results obtained with this new preconditioning are plotted in Fig. 2 (notated PMCP+ and HPCP+). As evident from Fig. 2(a), the naive value of $\alpha = 0.5$ leads to a net improvement of the PMCP and HPCP performances over the range $0.3 < \theta < 0.7$, inside of which the number of purifications becomes independent of θ. Outside this interval, runaway solutions were encountered due to the ill-conditioning of c, where either of the constraints in Eq. (12b) or (12c) is violated. The solution to this problem is to perform a constrained search of α in Eq. (14), such that the first inequality of Eq. (12b) is respected, that is,

$$\underset{\substack{0\le\alpha\le 1\\ \delta>0}}{\mathrm{search}} \left\{ \mathrm{Tr}\{D_0^2\} = \begin{cases} N - \delta N, & \text{if } \theta < (1-\delta) \\ N - \delta\bar{N}, & \text{if } \theta > (1-\delta) \end{cases} \right\}, \tag{15}$$

which leads to solve a second-order polynomial equation in α, at the extra cost of only one matrix multiplication. Obviously, the parameter δ has to be carefully chosen such that the second equality of Eq. (12b) and condition (12c) are also respected. We found $\delta \simeq 2/3$ as the optimal value.[31] From Fig. 2, the benefits of this optimized preconditioning are clear when focussing within the range $[0.0, 0.3] \cup [0.7, 1.0]$, albeit with one or two extra purifications around the poles $\theta = \{0.3, 0.7\}$. These benefits are even clearer in Fig. 2(c), where we also show the plots of $\bar{p}$ as a function of $\ln(1/\Delta\epsilon_{\mathrm{gap}})$ for the test case $\theta = 0.01$. At the intercept, we find $\bar{p}_{\mathrm{PMCP}} \simeq 38$ compared to $\bar{p}_{\mathrm{HPCP}} \simeq 21$, showing the improvement bring by the hole-particle equivalence. We have also compared our method against the most efficient of the trace updating methods, TRS4,[20] and find that for non-pathological fillings, the two are comparable in efficiency. For the pathological cases, where TRS4 adjusts the polynomial, we found it more efficient, but at the expense of non-variational behaviour in the early iterations.

To conclude, we have shown how, by considering both electron and hole occupancies, the density matrix for a given system can be found efficiently while preserving N-representability. This opens the door to a more robust, stable ground state minimisation algorithm, with application to standard and linear scaling DFT approaches.

L.A.T. would like to acknowledge D. Hache for his unwavering support and midnight talks about how to move beads along a double-well potential.

APPENDIX: ALTERNATIVE DERIVATION OF THE HOLE-PARTICLE CANONICAL PURIFICATION

We demonstrate that by symmetrizing the Palser and Manolopoulos equations with respect to $\bar{D}$, the closed-form of Eq. (10) appears naturally. Let us start from Eq. (16) of Ref. 21,

$$\text{for } c_n \le \frac{1}{2}, \tag{A1a}$$

$$D_{n+1} = -\frac{1}{1-c_n} D_n^3 + \frac{1+c_n}{1-c_n} D_n^2 + \frac{1-2c_n}{1-c_n} D_n,$$

$$\text{for } c_n > \frac{1}{2},\ D_{n+1} = -\frac{1}{c_n} D_n^3 + \frac{1+c_n}{c_n} D_n^2, \tag{A1b}$$

with c_n given in Eq. (7b). We may search for purification relations *dual* to Eq. (A1), i.e., function of $\bar{D}$. We obtain

$$\text{for } \bar{c}_n \ge \frac{1}{2}, \tag{A2a}$$

$$\bar{D}_{n+1} = -\frac{1}{1-\bar{c}_n} \bar{D}_n^3 + \frac{1+\bar{c}_n}{1-\bar{c}_n} \bar{D}_n^2 + \frac{1-2\bar{c}_n}{1-\bar{c}_n} \bar{D}_n,$$

$$\text{for } \bar{c}_n < \frac{1}{2},\ \bar{D}_{n+1} = -\frac{1}{\bar{c}_n} \bar{D}_n^3 + \frac{1+\bar{c}_n}{\bar{c}_n} \bar{D}_n^2, \tag{A2b}$$

with $\bar{c}_n = 1 - c_n$. Instead of purifying either D or $\bar{D}$, we shall try to take advantage of the closure relation [Eq. (9)] in such a way that, if we choose to work within the subspace of occupied states, the purification of D [Eq. (A1)] is constrained to verify $D = I - \bar{D}$. By inserting this constraint in Eq. (A2), we obtain

$$\text{for } c_n \le \frac{1}{2},\quad D_{n+1} = I - \left(-\frac{1}{c_n}(I - D_n)^3 + \frac{2-c_n}{c_n}(I - D_n)^2 - \frac{1-2c_n}{c_n}(I - D_n) \right), \tag{A3a}$$

$$\text{for } c_n > \frac{1}{2},\quad D_{n+1} = I - \left(-\frac{1}{1-c_n}(I - D_n)^3 + \frac{2-c_n}{1-c_n}(I - D_n)^2 \right). \tag{A3b}$$

On multiplying Eqs. (A1a) and (A3a) by $(1-c_n)$ and c_n, respectively [or multiplying Eqs. (A1b) and (A3b) by c_n and $(1-c_n)$], and adding, we obtain

$$D_{n+1} = D_n + 2\left(D_n^2 \bar{D}_n - c_n D_n \bar{D}_n\right). \tag{A4a}$$

[1]R. McWeeny, Proc. R. Soc. A **235**, 496 (1956); **237**, 355 (1956); **241**, 239 (1957).
[2]R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[3]X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993).
[4]A. D. Daniels, J. M. Millam, and G. E. Scuseria, J. Chem. Phys. **107**, 425 (1997).

[5]J. M. Millam and G. E. Scuseria, J. Chem. Phys. **106**, 5569 (1997).
[6]A. D. Daniels and G. E. Scuseria, J. Chem. Phys. **110**, 1321 (1999).
[7]D. Bowler and M. Gillan, Comput. Phys. Commun. **120**, 95 (1999).
[8]M. Challacombe, J. Chem. Phys. **110**, 2332 (1999).
[9]S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994).
[10]S. Goedecker and M. Teter, Phys. Rev. B **51**, 9455 (1995).
[11]R. Baer and M. Head-Gordon, Phys. Rev. Lett. **79**, 3962 (1997).
[12]R. Baer and M. Head-Gordon, J. Chem. Phys. **107**, 10003 (1997).
[13]K. R. Bates, A. D. Daniels, and G. E. Scuseria, J. Chem. Phys. **109**, 3308 (1998).
[14]W. Liang, C. Saravanan, Y. Shao, R. Baer, A. T. Bell, and M. Head-Gordon, J. Chem. Phys. **119**, 4117 (2003).
[15]A. M. N. Niklasson, Phys. Rev. B **68**, 233104 (2003).
[16]K. Németh and G. E. Scuseria, J. Chem. Phys. **113**, 6035 (2000).
[17]G. Beylkin, N. Coult, and M. J. Mohlenkamp, J. Comput. Phys. **152**, 32 (1999).
[18]C. Kenney and A. Laub, SIAM J. Matrix Anal. Appl. **12**, 273 (1991).
[19]A. M. N. Niklasson, Phys. Rev. B **66**, 155115 (2002).
[20]A. M. N. Niklasson, C. J. Tymczak, and M. Challacombe, J. Chem. Phys. **118**, 8611 (2003).
[21]A. H. R. Palser and D. E. Manolopoulos, Phys. Rev. B **58**, 12704 (1998).
[22]D. R. Bowler and T. Miyazaki, Rep. Prog. Phys. **75**, 036503 (2012).
[23]S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[24]R. G. Parr and Y. Weitao, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York; Oxford, England, 1994).
[25]E. Rudberg and E. H. Rubensson, J. Phys.: Condens. Matter **23**, 075502 (2011).
[26]E. Chow, X. Liu, M. Smelyanskiy, and J. R. Hammond, J. Chem. Phys. **142**, 104103 (2015).
[27]M. J. Cawkwell, E. J. Sanville, S. M. Mniszewski, and A. M. N. Niklasson, J. Chem. Theory Comput. **8**, 4094 (2012).
[28]M. J. Cawkwell, M. A. Wood, A. M. N. Niklasson, and S. M. Mniszewski, J. Chem. Theory Comput. **10**, 5391 (2014).
[29]We will restrict our discussion to effective one-electron Hamiltonian operators expressed in a finite orthonormal Hilbert space.
[30]D. A. Mazziotti, Phys. Rev. E **68**, 066701 (2003).
[31]L. A. Truflandier, M. R. Dianzinga, and D. R. Bowler, "On grand canonical and canonical ensemble approaches for the one-particle density matrix minimization" (unpublished).
[32]S. Geršgorin, Proc. USSR Acad. Sci. **51**, 749 (1931); available at http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=im&paperid=5235&option_lang=eng.
[33]The Frobenius norm is defined by $\|D\|_{\mathcal{F}} = (\sum_{i,j} |D_{ij}|^2)^{1/2} = \sqrt{\mathrm{Tr}\{D^2\}}$. Notice that $\forall D$, such that $D^2 = D$, then $\|D\|_{\mathcal{F}} = \sqrt{\mathrm{Tr}\{D\}}$.
[34]In that case, it can be shown that $\lambda(D_0^+) \in [-\frac{1}{2}, \frac{3}{2}]$.