The Journal of Chemical Physics

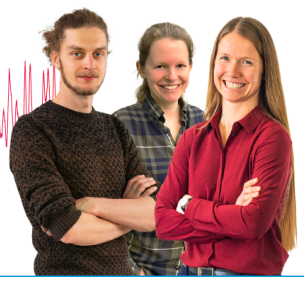
RESEARCH ARTICLE | AUGUST 25 2025

Deterministic optimization of Jastrow factors

Special Collection: Michele Parrinello Festschrift

Maria-Andreea Filip ; Evelin Martine Corvid Christlmaier ; J. Philip Haupt ; Daniel Kats ; Pablo López Ríos ; Ali Alavi

J. Chem. Phys. 163, 084107 (2025)
https://doi.org/10.1063/5.0284106



The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Deterministic optimization of Jastrow factors

Cite as: J. Chem. Phys. 163, 084107 (2025); doi: 10.1063/5.0284106
Submitted: 5 June 2025 • Accepted: 31 July 2025 •
Published Online: 25 August 2025

Maria-Andreea Filip,[1,2,a)] Evelin Martine Corvid Christlmaier,[1] J. Philip Haupt,[1] Daniel Kats,[1] Pablo López Ríos,[1] and Ali Alavi[1,2,b)]

AFFILIATIONS

[1] Max Planck Institute for Solid State Research, Heisenbergstr. 1, 70569 Stuttgart, Germany
[2] Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Note: This paper is part of the JCP Special Topic, Michele Parrinello Festschrift.
[a)] **Author to whom correspondence should be addressed:** maf63@cam.ac.uk
[b)] a.alavi@fkf.mpg.de

ABSTRACT

Highly flexible Jastrow factors have found significant use in stochastic electronic structure methods such as variational Monte Carlo (VMC) and diffusion Monte Carlo, as well as in quantum chemical transcorrelated (TC) approaches, which have recently seen great success in generating highly accurate electronic energies using moderately sized basis sets. In particular, for the latter, the intrinsic noise in the Jastrow factor due to its optimization by VMC can pose a problem, especially when targeting weak (non-covalent) interactions. In this paper, we propose a deterministic alternative to VMC Jastrow optimization, based on minimizing the "variance of the TC reference energy" in a standard basis set. Analytic expressions for the derivatives of the TC Hamiltonian matrix elements are derived and implemented. This approach can be used to optimize the parameters in the Jastrow functions, either from scratch or to refine an initial VMC-based guess, to produce noise-free Jastrows in a reproducible manner. Applied to the first row of atoms and molecules, the results show that the method yields Slater–Jastrow wavefunctions whose variances are almost as low as those obtained from standard VMC variance optimization, but whose energies are lower and comparable to those obtained from energy-minimization VMC. We propose that the method can be used both in the context of the transcorrelated method and in standard VMC as a new way to optimize Jastrow functions.

I. INTRODUCTION

Jastrow factors[1] are commonly used in various *ab initio* electronic structure methods, where the introduction of explicit particle–particle correlation has been found to lead to much improved accuracy and convergence properties compared to standard approaches. In variational formulations, the use of such wavefunctions leads to many-body integrals, the optimization of which can only be done using stochastic Monte Carlo techniques. This has led to the development of the family of variational Monte Carlo methods,[2–9] the output of which can be used as a trial wavefunction for the fixed-node diffusion Monte Carlo (DMC) procedure.[10–13]

There exists another class of explicitly correlated methods—the R12/F12 family of methods[14–30]—in which the quantum chemical wavefunction is *additively* augmented with short-ranged fixed-amplitude geminal-type functions, designed mainly to capture electron–electron cusp conditions. In these approaches, many-body integrals (beyond 4-electron integrals) are avoided, but the geminal functions themselves are typically not further optimized. These types of methods yield faster convergence to the basis-set limit of the underlying method (such as MP2 or CCSD) but are difficult to extend beyond the double-excitation level of coupled-cluster theory.[18–29]

Transcorrelation is another distinct class of explicitly correlated methods that share aspects of the previous two and that offer an interesting alternative. In this methodology, Jastrow factors are used to similarity transform the Hamiltonian itself.[31–41] This leads to a non-Hermitian Hamiltonian with 3-electron terms, approximate eigenvalues of which can be found by a variety of quantum chemical approaches,[41–51] provided they do not require a variational estimate of the energy. In general, this approach has been found to accelerate convergence to the basis-set limit and increase the single-reference character of the right eigenfunctions.[50,52,53] The use of very flexible Jastrow factors is a great advantage in this respect, although

they necessitate numerical quadrature of the required matrix elements. Thanks to normal-ordering approximations[42,43] such as the xTC approximation,[54] the TC Hamiltonian can be reduced to two-electron interactions, greatly decreasing the burden of additional TC integral calculations.

Both the DMC approach and many modern transcorrelated (TC) methods employ parameterized Jastrow factors,[31,55] which must be optimized. In both cases, this is usually done using VMC to minimize either the variational energy or the variance of the energy of the Slater–Jastrow wavefunction.[4,56] Stochastic optimization has proved quite effective in treating the often large number of parameters involved in this type of Jastrow factor, but it induces unavoidable noise in the final values. This is not an issue for the DMC energy, which is independent of the Jastrow factor, but may pose problems for TC methods, especially when small energy differences are being sought, which may be overwhelmed by the stochastic noise arising from the Jastrow factor.

An interesting question for the TC methodology is that, since it avoids many-electron integrals (beyond 3 electron terms) in the first place, can it be used to *deterministically* optimize the Jastrow parameters? Of course, due attention needs to be paid to the fact that, as a non-Hermitian problem, the TC energy itself cannot be used as a target function to optimize, and a target function based on a variance quantity must be used instead. In this work, we present such an approach, based on the minimization of the "variance of the TC reference energy" often employed in TC approaches.[53] We show that, when applied after VMC, this approach lowers the variance of energies obtained from different Jastrow optimization runs, with resulting transcorrelated energies maintaining similar accuracy to those obtained from pure VMC-optimized Jastrow factors. Furthermore, VMC energies themselves are shown to be lower than those obtained by stochastic minimization of the variance of the reference energy, while maintaining lower variances than those obtained by conventional VMC energy minimization.

II. TRANSCORRELATED METHODOLOGY

A. Similarity transform of the Hamiltonian

Rather than using the Slater–Jastrow form of the wavefunction,

$$|\Psi\rangle = e^{\tau}|\Phi\rangle, \tag{1}$$

as an ansatz to be directly optimized, in transcorrelated methods, one uses the Jastrow factor to similarity transform the Hamiltonian

$$\hat{H}_{\mathrm{TC}} = e^{-\tau}\hat{H}e^{\tau}, \tag{2}$$

which can then be used to calculate $|\Phi\rangle$ as the right-eigenfunction of $\hat{H}_{\mathrm{TC}}$. If the Jastrow function depends purely on the spatial coordinates of the particles, this Hamiltonian can be expressed as a Baker–Campbell–Hausdorff expansion, which terminates at the second nested commutator,

$$\hat{H}_{\mathrm{TC}} = \hat{H} + [\hat{H},\tau] + \frac{1}{2}[[\hat{H},\tau]\tau]. \tag{3}$$

Consider a Jastrow factor that contains no more than two-electron interactions,

$$\tau = \sum_{i<j} u(\mathbf{r}_i,\mathbf{r}_j), \tag{4}$$

where u is a symmetric correlation function. The TC Hamiltonian may then be rewritten as

$$\hat{H}_{\mathrm{TC}}[\tau] = \hat{H} - \hat{K}[\tau] - \hat{L}[\tau], \tag{5}$$

where we have emphasized the dependence of the additional terms on the Jastrow factor,

$$\begin{aligned}\hat{K}(\mathbf{r}_i,\mathbf{r}_j) = \frac{1}{2}\big[&\nabla_i^2 u(\mathbf{r}_i,\mathbf{r}_j) + \nabla_j^2 u(\mathbf{r}_i,\mathbf{r}_j)\\ &+ |\nabla_i u(\mathbf{r}_i,\mathbf{r}_j)|^2 + |\nabla_j u(\mathbf{r}_i,\mathbf{r}_j)|^2\big]\\ &+ \nabla_i u(\mathbf{r}_i,\mathbf{r}_j)\cdot\nabla_i + \nabla_j u(\mathbf{r}_i,\mathbf{r}_j)\cdot\nabla_j\end{aligned} \tag{6}$$

and

$$\begin{aligned}\hat{L}(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k) = &\nabla_i u(\mathbf{r}_i,\mathbf{r}_j)\cdot\nabla_i u(\mathbf{r}_i,\mathbf{r}_k)\\ &+ \nabla_j u(\mathbf{r}_j,\mathbf{r}_k)\cdot\nabla_j u(\mathbf{r}_j,\mathbf{r}_i)\\ &+ \nabla_k u(\mathbf{r}_k,\mathbf{r}_i)\cdot\nabla_k u(\mathbf{r}_k,\mathbf{r}_j).\end{aligned} \tag{7}$$

B. Stochastic Jastrow optimization

In most modern TC methods, $\hat{H}_{\mathrm{TC}}$ is obtained by optimizing a parameterized u with respect to some reference wavefunction $|\Phi\rangle$, which may be either the conventional Hartree–Fock determinant or a short configuration interaction (CI) expansion.[57] This is done using VMC and minimizing one of

- the variational energy of the Slater–Jastrow wavefunction,

$$\langle E\rangle = \frac{\langle\Phi|(e^{\tau})^{\dagger}\hat{H}e^{\tau}|\Phi\rangle}{\langle\Phi|(e^{\tau})^{\dagger}e^{\tau}|\Phi\rangle}, \tag{8}$$

- the variance of the energy of the Slater–Jastrow wavefunction,

$$\sigma^2 = \frac{\langle\Phi|(e^{\tau})^{\dagger}(\hat{H}-E)^2e^{\tau}|\Phi\rangle}{\langle\Phi|(e^{\tau})^{\dagger}e^{\tau}|\Phi\rangle}, \tag{9}$$

- the "variance of the reference energy,"[53,58]

$$\sigma_{\mathrm{ref}}^2 = \langle\Phi\|\hat{H}_{\mathrm{TC}} - E_{\mathrm{ref}}|^2|\Phi\rangle, \tag{10}$$

where $E_{\mathrm{ref}} = \langle\Phi|\hat{H}_{\mathrm{TC}}|\Phi\rangle$ is the (TC) reference energy.

The latter of these quantities was recently shown to provide the best compromise between value and spread of the VMC energy and has, therefore, been proposed as the standard approach for TC calculations.[53] In VMC, it can be computed as the sample variance of the local Slater–Jastrow energy over the reference distribution,

$$S_{\mathrm{ref}}^2 = \frac{1}{n_{\mathrm{opt}}-1}\sum_{n=1}^{n_{\mathrm{opt}}}\left|\frac{\hat{H}\Psi(\mathbf{R}_n)}{\Psi(\mathbf{R}_n)} - \widetilde{E}_{\mathrm{ref}}\right|^2, \tag{11}$$

where $\{\mathbf{R}_n\}$ are the electron distributions sampled from $|\Phi|^2$ and

$$\widetilde{E}_{\mathrm{ref}} = \frac{1}{n_{\mathrm{opt}}}\sum_{n=1}^{n_{\mathrm{opt}}}\frac{\hat{H}\Psi(\mathbf{R}_n)}{\Psi(\mathbf{R}_n)}. \tag{12}$$

Two types of parameterized Jastrow factors have been used recently in TC applications:

- Boys–Handy (BH) Jastrow factors[31] of the form

$$u(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_I)=\sum_{m,n,o}^{m+n+o\le 6} c_{mno}\bar{r}_{iI}^{m}\bar{r}_{jI}^{n}\bar{r}_{ij}^{o}, \tag{13}$$

where $\bar{r}$ is a scaled distance. This can take various forms, the most common of which is $\bar{r}=\frac{r}{1+r}$.

- Drummond–Towler–Needs (DTN) Jastrow factors[55,59] of the form

$$J(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_I)=\sum_{i<j}u(r_{ij})+\sum_{i,I}\chi(r_{iI})+\sum_{i<j,I}\zeta(r_{ij},r_{iI},r_{jI}), \tag{14}$$

with the electron–electron, electron–nucleus, and electron–electron–nucleus terms expressed as natural power expansions,

$$u(r_{ij})=t(r_{ij},L_u)\sum_{k}^{N_u}a_k r_{ij}^{k}, \tag{15}$$

$$\chi(r_{iI})=t(r_{iI},L_\chi)\sum_{k}^{N_\chi}b_k r_{iI}^{k}, \tag{16}$$

and

$$\zeta(r_{ij},r_{iI},r_{jI})=t(r_{iI},L_\zeta)t(r_{jI},L_\zeta)\times\sum_{klm}^{N_f}c_{klm}r_{ij}^{k}r_{iI}^{l}r_{jI}^{m}, \tag{17}$$

where $\{a_k\}$, $\{b_k\}$, and $\{c_{klm}\}$ are the linear coefficients, $t(r,L)=(1-r/L)^3\Theta(L-r)$, $\Theta(x)$ is the Heaviside function, and L_u, L_χ, and L_ζ are some appropriate cutoff radii.

As the nucleus positions are kept fixed, the terms of the DTN Jastrow can be combined as

$$\begin{aligned}\frac{1}{2}\sum_{i\neq j}^{N}(u_{ij}+\zeta_{ij})+\sum_{i}^{N}\chi_i&=\frac{1}{2}\sum_{i\neq j}^{N}(u_{ij}+\zeta_{ij})+\frac{1}{2}\left(\sum_{i}^{N}\chi_i+\sum_{j}^{N}\chi_j\right)\\&=\frac{1}{2}\sum_{i\neq j}^{N}\left(u_{ij}+\zeta_{ij}+\frac{1}{(N-1)}(\chi_i+\chi_j)\right),\end{aligned} \tag{18}$$

where we have used the notation $u_{ij}=u(r_{ij})$, $\chi_i=\sum_I\chi(r_{iI})$, and $\zeta_{ij}=\sum_I\zeta(r_{ij},r_{iI},r_{jI})$. Therefore, both Jastrow forms above can be rewritten in the general form of Eq. (4), with

$$u(\mathbf{r}_i,\mathbf{r}_j)=\sum_{l}f_l u_l(\mathbf{r}_i,\mathbf{r}_j). \tag{19}$$

For the BH Jastrow, l is a counter over the different (m,n,o) contributions, $f_l=c_{mno}$, and $u_l(\mathbf{r}_i,\mathbf{r}_j)=\bar{r}_{iI}^{m}\bar{r}_{jI}^{n}\bar{r}_{ij}^{o}$ for the corresponding (m,n,o). For the DTN Jastrow, f_l is one of a_k,b_k,c_{klm}, and u_l is the corresponding function, $t(r_{ij},L_u)r_{ij}^{k}$, $\sum_I\left(t(r_{iI},L_\chi)r_{iI}^{k}+t(r_{jI},L_\chi)r_{jI}^{k}\right)/(N-1)$, or $t(r_{iI},L_\zeta)\,t(r_{jI},L_\zeta)c_{klm}r_{ij}^{k}r_{iI}^{l}r_{jI}^{m}$. For convenience, we will assume the general form in Eq. (19) when deriving the gradients for the deterministic optimization of the Jastrow factors, although a similar derivation can be carried out without folding the one-electron terms of the Jastrow into the two-electron term.

C. Integral evaluation

Given a particular Jastrow, the resulting TC Hamiltonian may be written in spin–orbital notation as

$$\begin{aligned}\hat{H}_{\mathrm{TC}}=&\sum_{PQ}h_P^Q\hat{a}_P^\dagger\hat{a}_Q+\frac{1}{2}\sum_{PQRS}(V_{PR}^{QS}-K_{PR}^{QS})\hat{a}_P^\dagger\hat{a}_R^\dagger\hat{a}_S\hat{a}_Q\\&-\frac{1}{6}\sum_{PQRSTU}L_{PRT}^{QSU}\hat{a}_P^\dagger\hat{a}_R^\dagger\hat{a}_T^\dagger\hat{a}_U\hat{a}_S\hat{a}_Q,\end{aligned} \tag{20}$$

for a set of one-particle basis functions $\{\phi_P\}$, where $h_P^Q=-\frac{1}{2}\langle\phi_P|\nabla^2|\phi_Q\rangle-\sum_I Z_I\langle\phi_P|\mathbf{r}_I^{-1}|\phi_Q\rangle$, $V_{PR}^{QS}=\langle\phi_P\phi_R|\mathbf{r}_{12}^{-1}|\phi_Q\phi_S\rangle$, $K_{PR}^{QS}=\langle\phi_P\phi_R|\hat{K}|\phi_Q\phi_S\rangle$, and $L_{PRT}^{QSU}=\langle\phi_P\phi_R\phi_T|\hat{L}|\phi_Q\phi_S\phi_U\rangle$. One can compute the one-, two-, and three-body integrals of the Hamiltonian and use them in subsequent post-Hartree–Fock treatments. The three-body operator $\hat{L}$ is the bottleneck of this process, but it has been shown that excluding the explicit three-body contributions (L_{PRT}^{QSU} where all six indices are distinct) has little effect on the resulting energies.[43,49] The remaining three-body terms can be folded into the lower-order integrals, significantly reducing the cost of computing the Hamiltonian, in what is known as the xTC approximation.[54] This leads to a correction to the two-body term given by

$$\Delta U_{PR}^{QS}=-\sum_{TU}(L_{PRT}^{QSU}-L_{PRT}^{QUS}-L_{PRT}^{USQ})\gamma_U^T, \tag{21}$$

where $\gamma_U^T=\langle\Phi|\hat{a}_T^\dagger\hat{a}_U|\Phi\rangle$ are the elements of the density matrix of the reference wavefunction Φ. For a single-determinant reference, the one-body correction can then be computed as

$$\Delta h_P^Q=-\frac{1}{2}\sum_{RS}(\Delta U_{PR}^{QS}-\Delta U_{PR}^{SQ})\gamma_S^R, \tag{22}$$

and the zero-body correction needed to account for the normal ordering of the Hamiltonian as

$$\langle\Phi|L|\Phi\rangle=-\frac{1}{3}\sum_{PQ}\Delta h_P^Q\gamma_Q^P. \tag{23}$$

III. DETERMINISTIC JASTROW OPTIMIZATION

Access to the TC Hamiltonian integrals suggests a new path to Jastrow optimization that does not require a stochastic approach. If we consider the variance of the reference energy in Eq. (10), it may be rewritten as[53]

$$\sigma_{\mathrm{ref}}^2=\langle\Phi|(\hat{H}_{\mathrm{TC}}-E_{\mathrm{ref}})^\dagger(\hat{H}_{\mathrm{TC}}-E_{\mathrm{ref}})|\Phi\rangle, \tag{24}$$

$$=\sum_I\langle\Phi|(\hat{H}_{\mathrm{TC}}-E_{\mathrm{ref}})^\dagger|\Phi_I\rangle\langle\Phi_I|(\hat{H}_{\mathrm{TC}}-E_{\mathrm{ref}})|\Phi\rangle, \tag{25}$$

where $|\Phi_I\rangle$ are the Slater determinants spanning the complete problem Hilbert space. In the particular case where the reference wavefunction is the Hartree–Fock determinant $|\Phi_0\rangle$, Eq. (25) becomes

$$\sigma_{\mathrm{ref}}^2=\sum_{I\neq 0}\langle\Phi_I|\hat{H}_{\mathrm{TC}}|\Phi_0\rangle^2. \tag{26}$$

When $|\Phi_I\rangle$ only spans a finite basis set, Eq. (26) provides an approximation for the variance of the reference energy, which can be

easily computed using the TC Hamiltonian integrals. If the full TC Hamiltonian is used, the computation of this variance scales as $\mathcal{O}(N^3M^3)$, where N is the number of electrons and M is the number of basis functions, although in practice the contribution from triple excitations is found to be very small. By restricting the excited determinants to double excitations from the reference or by employing the xTC approximation, this cost can be reduced to $\mathcal{O}(N^2M^2)$. We propose using this approximate quantity as a cost function for Jastrow optimization and, in Sec. IV, investigate the requirements on the basis set to make this method viable, as well as its applicability in VMC and TC calculations.

A. Analytical gradients

An advantage of Eq. (26) as an optimization target is that analytical gradients can be easily computed using the same paradigm as for the matrix elements themselves. The gradients of the variance with respect to Jastrow parameters are given by

$$\frac{d\sigma^2_{\text{ref}}}{df_l} = \sum_{I\neq 0} 2\langle\Phi_I|\hat{H}_{\text{TC}}|\Phi_0\rangle \frac{d}{df_l}\langle\Phi_I|\hat{H}_{\text{TC}}|\Phi_0\rangle. \tag{27}$$

If we are interested in obtaining matrix elements of the form

$$\frac{\partial}{\partial f_l}\langle\Phi_I|H_{\text{TC}}|\Phi_0\rangle = -\langle\Phi_I|\frac{\partial(K+L)}{\partial f_l}|\Phi_0\rangle, \tag{28}$$

we are required to obtain the corresponding derivative integrals of the form $\langle PR|\frac{\partial K}{\partial f_l}|QS\rangle$ and $\langle PRT|\frac{\partial L}{\partial f_l}|QSU\rangle$.

Given the forms of $\hat{K}$ and $\hat{L}$ given in Eqs. (6) and (7), there are four types of terms that must be differentiated with respect to f_l,

$$K_1(\mathbf{r}_1,\mathbf{r}_2) = \nabla_1^2 u, \tag{29}$$

$$K_2(\mathbf{r}_1,\mathbf{r}_2) = (\nabla_1 u)^2, \tag{30}$$

$$K_3(\mathbf{r}_1,\mathbf{r}_2) = \nabla_1 u\cdot\nabla_1, \tag{31}$$

$$L_1(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = \nabla_1 u(\mathbf{r}_1,\mathbf{r}_2))\cdot\nabla_1 u(\mathbf{r}_1,\mathbf{r}_3). \tag{32}$$

In practice, the integrals of K_1 are computed by parts, so they take a form similar to K_3. However,

$$\begin{aligned}\langle PR|K_3|QS\rangle &= \iint \phi_P(\mathbf{r}_1)\phi_R(\mathbf{r}_2)\nabla_1 u\cdot\nabla_1[\phi_Q(\mathbf{r}_1)\phi_S(\mathbf{r}_2)]d\mathbf{r}_1 d\mathbf{r}_2\\ &= \iint \phi_P(\mathbf{r}_1)\phi_R(\mathbf{r}_2)\phi_Q(\mathbf{r}_2)\nabla_1 u\cdot\nabla_1\phi_S(\mathbf{r}_1)d\mathbf{r}_1 d\mathbf{r}_2,\end{aligned} \tag{33}$$

while

$$\begin{aligned}\langle PR|K_1|QS\rangle &= \iint \phi_P(\mathbf{r}_1)\phi_R(\mathbf{r}_2)(\nabla_1^2 u)\phi_Q(\mathbf{r}_1)\phi_S(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2\\ &= -\iint \nabla_1 u\cdot\nabla_1[\phi_P(\mathbf{r}_1)\phi_R(\mathbf{r}_2)\phi_Q(\mathbf{r}_1)\phi_S(\mathbf{r}_2)]d\mathbf{r}_1 d\mathbf{r}_2\\ &= -\iint \phi_R(\mathbf{r}_2)\phi_S(\mathbf{r}_2)\nabla_1 u\cdot[\phi_P(\mathbf{r}_1)\nabla_1\phi_Q(\mathbf{r}_1)\\ &\quad+\phi_Q(\mathbf{r}_1)\nabla_1\phi_P(\mathbf{r}_1)]d\mathbf{r}_1 d\mathbf{r}_2.\end{aligned} \tag{34}$$

The relevant derivatives of the Jastrow factor are, therefore,

$$\begin{aligned}\frac{\partial(\nabla_1 u)^2}{\partial f_l} &= 2\nabla_1 u\cdot\frac{\partial\nabla_1 u}{\partial f_l}\\ &= 2\nabla_1 u\cdot\frac{\partial}{\partial f_l}\nabla_1\sum_j f_j u_j = \nabla_1 u\cdot\nabla_1 u_l,\end{aligned} \tag{35}$$

$$\begin{aligned}\frac{\partial}{\partial f_l}(\nabla_1 u)\cdot\nabla_1 &= \frac{\partial}{\partial f_l}\left(\nabla_1\sum_j f_j u_j\right)\cdot\nabla_1\\ &= \frac{\partial}{\partial f_l}\sum_j f_j\nabla_1 u_j\cdot\nabla_1 = \nabla_1 u_l\cdot\nabla_1,\end{aligned} \tag{36}$$

$$\begin{aligned}\frac{\partial L_1}{\partial f_l} &= \frac{\partial}{\partial f_l}\left(\nabla_1\sum_j f_j u_j(\mathbf{r}_1,\mathbf{r}_2)\right)\cdot\left(\nabla_1\sum_k f_k u_k(\mathbf{r}_1,\mathbf{r}_3)\right)\\ &= \frac{\partial}{\partial f_l}\left(\sum_j f_j\nabla_1 u_j(\mathbf{r}_1,\mathbf{r}_2)\right)\cdot\left(\sum_k f_k\nabla_1 u_k(\mathbf{r}_1,\mathbf{r}_3)\right)\\ &= \nabla_1 u_l(\mathbf{r}_1,\mathbf{r}_2))\cdot\nabla_1 u(\mathbf{r}_1,\mathbf{r}_3)) + \nabla_1 u(\mathbf{r}_1,\mathbf{r}_2))\cdot\nabla_1 u_l(\mathbf{r}_1,\mathbf{r}_3)).\end{aligned} \tag{37}$$

Compared to an energy evaluation, the only new terms needed are $\nabla_1 u_l$ and $\nabla_2 u_l$. In practice, these terms are summed to obtain $\nabla_1 u$ and $\nabla_2 u$, so they are easily accessible.

In the TCHInt package[60] used for TC calculations, the K integrals are computed in two steps, the first being to compute the intermediate,

$$\begin{aligned}x_P^Q(\mathbf{r}_2) &= -\int\nabla_1 u\cdot[\phi_P(\mathbf{r}_1)\nabla_1\phi_Q(\mathbf{r}_1)+\phi_Q(\mathbf{r}_1)\nabla_1\phi_P(\mathbf{r}_1)]d\mathbf{r}_1\\ &\quad+\int\phi_P(\mathbf{r}_1)\nabla u^2\phi_Q(\mathbf{r}_1)d\mathbf{r}_1\\ &\quad+2\int\phi_P(\mathbf{r}_1)\nabla u\cdot\nabla\phi_Q(\mathbf{r}_1)d\mathbf{r}1\\ &= \int\nabla_1 u\cdot[\phi_P(\mathbf{r}_1)\nabla_1\phi_Q(\mathbf{r}_1)-\phi_Q(\mathbf{r}_1)\nabla_1\phi_P(\mathbf{r}_1)]d\mathbf{r}_1\\ &\quad+\int\phi_P(\mathbf{r}_1)\nabla u^2\phi_Q(\mathbf{r}_1)d\mathbf{r}_1,\end{aligned} \tag{38}$$

followed by

$$K_{PR}^{QS} = \frac{1}{2}\left[\int\phi_R(\mathbf{r}_2)x_P^Q(\mathbf{r}_2)\phi_S(\mathbf{r}_2)d\mathbf{r}_2+\int\phi_P(\mathbf{r}_2)x_R^S(\mathbf{r}_2)\phi_Q(\mathbf{r}_2)d\mathbf{r}_2\right]. \tag{39}$$

L is computed by first obtaining

$$\mathbf{V}_P^Q(\mathbf{r}) = \int\nabla u(\mathbf{r},\mathbf{r}_2)\phi_P(\mathbf{r}_2)\phi_Q(\mathbf{r}_2)d\mathbf{r}_2 \tag{40}$$

and then

$$\begin{aligned}L_{PRT}^{QSU} &= \int\phi_T(\mathbf{r})\phi_U(\mathbf{r})\mathbf{V}_P^Q(\mathbf{r})\cdot\mathbf{V}_R^S(\mathbf{r})\\ &\quad+\int\phi_R(\mathbf{r})\phi_S(\mathbf{r})\mathbf{V}_T^U(\mathbf{r})\cdot\mathbf{V}_P^Q(\mathbf{r})\\ &\quad+\int\phi_P(\mathbf{r})\phi_Q(\mathbf{r})\mathbf{V}_R^S(\mathbf{r})\cdot\mathbf{V}_T^U(\mathbf{r}).\end{aligned} \tag{41}$$

The same pattern may be followed to obtain the corresponding derivatives, computing first

$$\frac{dx_P^Q(\mathbf{r}_2)}{df_l} = -\frac{1}{2}\int \nabla_1 u_l \cdot \nabla_1(\phi_P(\mathbf{r}_1)\phi_Q(\mathbf{r}_1))d\mathbf{r}_1 + \int \phi_P(\mathbf{r}_1)\nabla_1 u_l \cdot \nabla_1\phi_Q(\mathbf{r}_1)d\mathbf{r}_1 + \int \phi_P(\mathbf{r}_1)\nabla_1 u_l \cdot \nabla_1 u\phi_Q(\mathbf{r}_1)d\mathbf{r}_1 = \frac{1}{2}\Big[\int \phi_P(\mathbf{r}_1)\nabla_1 u_l \cdot \nabla_1\phi_Q(\mathbf{r}_1)d\mathbf{r}_1 - \int \phi_Q(\mathbf{r}_1)\nabla_1 u_l \cdot \nabla_1\phi_P(\mathbf{r}_1)d\mathbf{r}_1\Big] + \int \phi_P(\mathbf{r}_1)\nabla_1 u_l.\nabla_1 u\phi_Q(\mathbf{r}_1)d\mathbf{r}_1 \quad (42)$$

and

$$\frac{d\mathbf{V}_P^Q(\mathbf{r}_2)}{df_l} = \int \nabla_1 u_l(\mathbf{r}_1,\mathbf{r}_2)\phi_P(\mathbf{r}_1)\phi_Q(\mathbf{r}_1)d\mathbf{r}_1. \quad (43)$$

These can then be used to obtain

$$\frac{dK_{PR}^{QS}}{df_l} = \int \phi_R(\mathbf{r}_2)\frac{dx_P^Q}{df_l}(\mathbf{r}_2)\phi_S(\mathbf{r}_2)d\mathbf{r}_2 + \int \phi_P(\mathbf{r}_2)\frac{dx_R^S}{df_l}(\mathbf{r}_2)\phi_Q(\mathbf{r}_2)d\mathbf{r}_2 \quad (44)$$

and

$$\frac{dL_{PRT}^{QSU}}{df_l} = +\int \phi_T(\mathbf{r})\phi_U(\mathbf{r})\frac{d\mathbf{V}_P^Q(\mathbf{r})}{df_l}\cdot\mathbf{V}_R^S(\mathbf{r}) + \int \phi_T(\mathbf{r})\phi_U(\mathbf{r})\mathbf{V}_P^Q(\mathbf{r})\cdot\frac{d\mathbf{V}_R^S(\mathbf{r})}{df_l} + \int \phi_R(\mathbf{r})\phi_S(\mathbf{r})\frac{d\mathbf{V}_T^U(\mathbf{r})}{df_l}\cdot\mathbf{V}_P^Q(\mathbf{r}) + \int \phi_R(\mathbf{r})\phi_S(\mathbf{r})\mathbf{V}_T^U(\mathbf{r})\cdot\frac{d\mathbf{V}_P^Q(\mathbf{r})}{df_l} + \int \phi_P(\mathbf{r})\phi_Q(\mathbf{r})\frac{d\mathbf{V}_R^S(\mathbf{r})}{df_l}\cdot\mathbf{V}_T^U(\mathbf{r}) + \int \phi_P(\mathbf{r})\phi_Q(\mathbf{r})\mathbf{V}_R^S(\mathbf{r})\cdot\frac{d\mathbf{V}_T^U(\mathbf{r})}{df_l}. \quad (45)$$

By using this intermediate structure, the cost of computing the K integrals scales as $\mathcal{O}(M^4N_{\text{grid}})$, while the J integrals scale as $\mathcal{O}(M^6N_{\text{grid}})$. The derivative integrals have the same scaling; however, since we only require those corresponding to triple excitations from the reference to compute Eq. (27), the scaling is reduced to $\mathcal{O}(N^2M^2N_{\text{grid}})$ and $\mathcal{O}(N^3M^3N_{\text{grid}})$, respectively.

B. xTC derivatives

In practice, rather than using the full TC Hamiltonian, we employ the xTC approximation, in which the two-body correction ΔU_{PR}^{QS} in Eq. (21) is computed from a series of intermediates shown in Table I, as[54]

$$\Delta U_{PR}^{QS} = -\mathscr{P}_{(RS)}^{(PQ)}\big(\rho_P^Q(i)A_R^S(i) + \mathbf{V}_P^Q(i)\cdot\mathbf{B}_R^S(i)\big). \quad (46)$$

The intermediates needed to compute the derivative of the two-body correction ΔU_{PR}^{QS} can be easily derived from the equations of the corresponding intermediates in the original xTC equations, as shown in Table I.

The full $\frac{d\Delta U_{PR}^{QS}}{df_l}$ term can then be obtained as

$$\frac{d\Delta U_{PR}^{QS}}{df_l} = -\mathscr{P}_{(PQ)}^{(RS)}\left(\rho_P^Q(i)\frac{dA_R^S(i)}{df_l} + \frac{d\mathbf{V}_P^Q(i)}{df_l}\cdot\mathbf{B}_R^S(i) + \mathbf{V}_P^Q(i)\cdot\frac{d\mathbf{B}_R^S(i)}{df_l}\right). \quad (47)$$

This is then used to compute the corresponding one- and zero-body corrections,

$$\frac{d\Delta h_P^Q}{df_l} = -\frac{1}{2}\left(\frac{d\Delta U_{PR}^{QS}}{df_l} - \frac{d\Delta U_{PR}^{SQ}}{df_l}\right)\gamma_S^R \quad (48)$$

and

$$\frac{d\langle\Phi_0|\hat{L}|\Phi_0\rangle}{df_l} = -\frac{1}{3}\frac{d\Delta h_P^Q}{df_l}\gamma_Q^P. \quad (49)$$

This reduces the overall cost of computing the integrals for optimization to $\mathcal{O}(N^2M^2N_{\text{grid}})$.

IV. RESULTS

All VMC calculations reported in this section have been carried out using the CASINO package.[9] Initial Hartree–Fock wavefunctions and integrals were obtained from PySCF,[61] while corresponding TC integral computation and deterministic Jastrow optimization

TABLE I. Derivatives with respect to f_l of various xTC intermediates, as given in Ref. 54.

ΔU_{PR}^{QS}	$d\Delta U_{PR}^{QS}/df_l$	ΔU_{PR}^{QS}	$d\Delta U_{PR}^{QS}/df_l$
$\rho_P^Q(i) = w_i\phi_P^*(i)\phi_Q(i)$		$A_R^S(i) = \tilde{V}_R^S(i) - \tilde{Z}_R^S(i)$	$\frac{dA_R^S(i)}{df_l} = \frac{d\tilde{V}_R^S(i)}{df_l} - \frac{d\tilde{Z}_R^S(i)}{df_l}$
$\mathbf{W}_i = \mathbf{V}_T^U(i)\gamma_U^T$	$\frac{d\mathbf{W}_i}{df_l} = \frac{d\mathbf{V}_T^U(i)}{df_l}\gamma_U^T$	$\mathbf{Y}_R^T(i) = \mathbf{V}_R^U(i)\gamma_U^T$	$\frac{d\mathbf{Y}_R^T(i)}{df_l} = \frac{d\mathbf{V}_R^U(i)}{df_l}\gamma_U^T$
$\tilde{V}_R^S(i) = \mathbf{W}_i\cdot\mathbf{V}_R^S(i)$	$\frac{d\tilde{V}_R^S(i)}{df_l} = \frac{d\mathbf{W}_i}{df_l}\cdot\mathbf{V}_R^S(i) + \mathbf{W}_i\cdot\frac{d\mathbf{V}_R^S(i)}{df_l}$	$\mathbf{Z}_R^S(i) = \rho_R^U\mathbf{X}_U^S(i) + \mathbf{Y}_R^T(i)\rho_T^S(i)$	$\frac{d\mathbf{Z}_R^S(i)}{df_l} = \rho_R^U\frac{d\mathbf{X}_U^S(i)}{df_l} + \frac{d\mathbf{Y}_R^T(i)}{df_l}\rho_T^S(i)$
$\mathbf{X}_U^S(i) = \mathbf{V}_T^S(i)\gamma_U^T$	$\frac{d\mathbf{X}_U^S(i)}{df_l} = \frac{d\mathbf{V}_T^S(i)}{df_l}\gamma_U^T$	$\tilde{W}(i) = \rho_T^U(i)\gamma_U^T$	
$\tilde{Z}_R^S(i) = \mathbf{V}_R^U(i)\cdot\mathbf{X}_U^S(i)$	$\frac{d\tilde{Z}_R^S(i)}{df_l} = \mathbf{V}_R^S(i)\cdot\frac{d\mathbf{X}_U^S(i)}{df_l} + \frac{d\mathbf{V}_R^S(i)}{df_l}\cdot\mathbf{X}_U^S(i)$	$\mathbf{B}_R^S(i) = \frac{1}{2}\tilde{W}(i)\mathbf{V}_R^S(i) - \mathbf{Z}_R^S(i)$	$\frac{d\mathbf{B}_R^S(i)}{df_l} = \frac{1}{2}\tilde{W}(i)\frac{\mathbf{V}_R^S(i)}{df_l} - \frac{d\mathbf{Z}_R^S(i)}{df_l}$

were carried out using the TCHInt library and its Python interface, PyTCHInt.[60] Integrals were computed using numerical quadrature over the direct product of atom-centered grids matching those used in PySCF. For the Jastrow optimization, grid level 1 was used to reduce integral cost, while grid level 2 was employed for the final integral computations, which could be carried out in under 5 minutes on one node for the largest system and basis set considered. To perform a CCSD(T) calculation with the xTC Hamiltonian, we used the ElemCo.jl package[62] to carry out a pseudocanonical λCCSD(T) calculation using biorthogonal orbitals,[63] which we will refer to as xTC-CCSD(T) for the remainder of this work. For the systems considered here, deterministic optimization of the Jastrow factor is the computational bottleneck. However, given the favorable scaling of the optimization ($\mathcal{O}(M^4)$) relative to CCSD(T) ($\mathcal{O}(M^7)$), and the fact that the integral computation process is very amenable to parallelization, we expect to see a crossover for larger systems.

A. Basis set requirements

The first important aspect to consider for this optimization scheme is the dependence of its performance on the basis set used. VMC optimized Jastrow factors are expected to be relatively independent of the basis set, in particular when the Hartree–Fock determinant is used as the reference wavefunction, as this is usually well described even in small basis sets. However, by introducing an incomplete basis in Eq. (26), this invariance is lost in the deterministic optimization scheme.

To demonstrate the effect of this, we consider the Li atom in basis sets from the cc-pVXZ and cc-pCVXZ (X = D, T, Q) families[64] and two simple BH Jastrow factors with a single optimizable parameter, corresponding to either an e–e or e–n term. The reference energies and their variances as a function of the Jastrow parameter, as computed using Eq. (26) or Eq. (11), are given in Fig. 1. It is clear that, while cc-pVXZ and cc-pCVXZ calculations lead to very similar VMC variances, this is not the case for the deterministic approach. For the cc-pVXZ basis sets, while the variance increases with basis set cardinality, it is significantly underestimated relative to VMC, while predicting a significantly shifted parameter value for the minimum. This effect is present in both the electron–electron and electron–nucleus Jastrow cases, although more pronounced in the former, and can be ascribed to basis set incompleteness errors. The cc-pCVXZ basis sets show significant improvement, with cc-pCVTZ results agreeing with the VMC results in both magnitude and the position of the minimum. The inclusion of additional orbitals with significant core character is, therefore, crucial for a correct estimate of the variance of the reference energy, and all deterministic

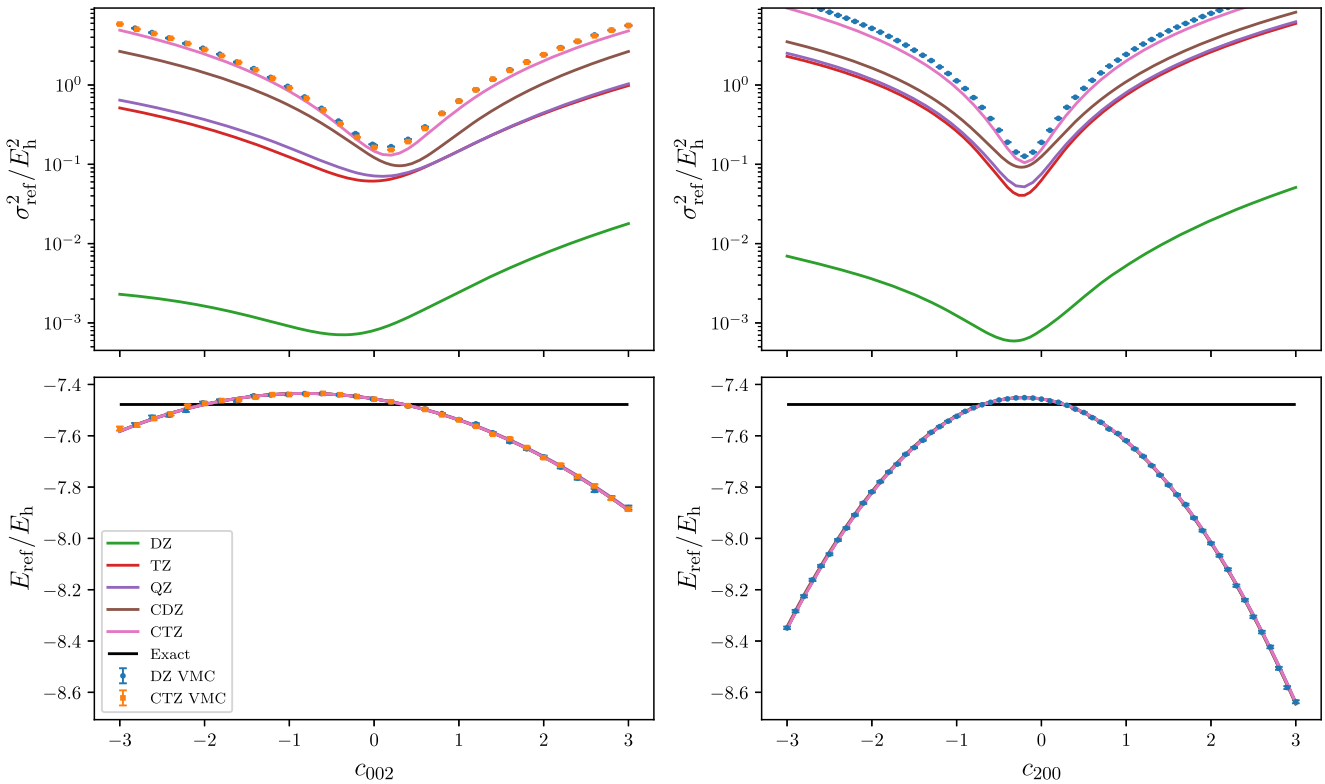


FIG. 1. Variance of the reference energy (top) and the reference energy itself (bottom) for the Li atom, as a function of the parameter controlling the first free electron–electron (left, $m = n = 0$, $o = 2$) and electron–nucleus [right, $m(n) = 2$, $n(m) = o = 0$] correlation term in a BH Jastrow, with all other optimizable parameters fixed at zero. The quantities are computed using different basis sets, both deterministically (solid lines) and with VMC (dotted lines). The cc-pCVTZ VMC curves are only reported for the left plots. Deterministic reference energies for all basis sets overlap to within the resolution of the plots.

Jastrow optimizations performed in the rest of this paper use basis sets of cc-pCVTZ quality. We note that, while the variance ranges over three orders of magnitude with different basis sets, the reference energies agree very well between basis sets. This suggests that while a relatively large basis set must be used to optimize the Jastrow factor, results are relatively transferable, and much smaller basis sets could be used for post-Hartree–Fock correlation treatments. We explore this hypothesis further for more complex systems and Jastrow factors in Secs. IV C and IV D.

B. Deterministic Jastrow optimization

For all results given below, the deterministic optimization of the linear parameters in the Jastrow factor is done by minimizing the variance computed according to Eq. (26), using the xTC approximation for the transcorrelated Hamiltonian integrals. The optimization is carried out using the L-BFGS algorithm.[65–69] Figure 2 shows an example of the convergence of the optimization algorithm with L-BFGS iteration for the O atom, comparing a pure deterministic optimization calculation to one initialized with VMC values for the Jastrow parameters. As expected, the latter starts much closer to the final result, and overall, they converge to very similar solutions for the Jastrow factor.

C. First row ionization potentials

We consider the first ionization potentials of atoms from the first row of the periodic table, Li–Ne. For each atom and ion, we consider two optimization routines, applied to a DTN Jastrow factor with $N_u = N_\chi = 4$ and $N_f = 2$: (a) we directly optimize the Jastrow factor deterministically; and (b) we initially do a VMC optimization, followed by a refinement of the Jastrow factor by deterministic optimization. We set the number of VMC configurations to target a standard deviation in E_{ref} of 0.1 mhartree. The deterministic optimization is carried out until the variance of the reference energy is converged to within 10^{-6} hartree^2, and we expect the observed standard deviations could be further reduced by employing a more stringent convergence criterion. In Fig. 3, we present ionization potential errors and standard deviations obtained at the reference energy level and using xTC-CCSD(T).[63] We find that reference energies obtained by either the (a) or (b) optimization routines generally agree with those from the purely stochastic optimization, although for some systems (Li, Al, and Ne), they exhibit lower errors. Significantly, the standard deviation of the reference energies decreases by more than an order of magnitude when deterministic optimization is used to refine the VMC-optimized Jastrow factors.

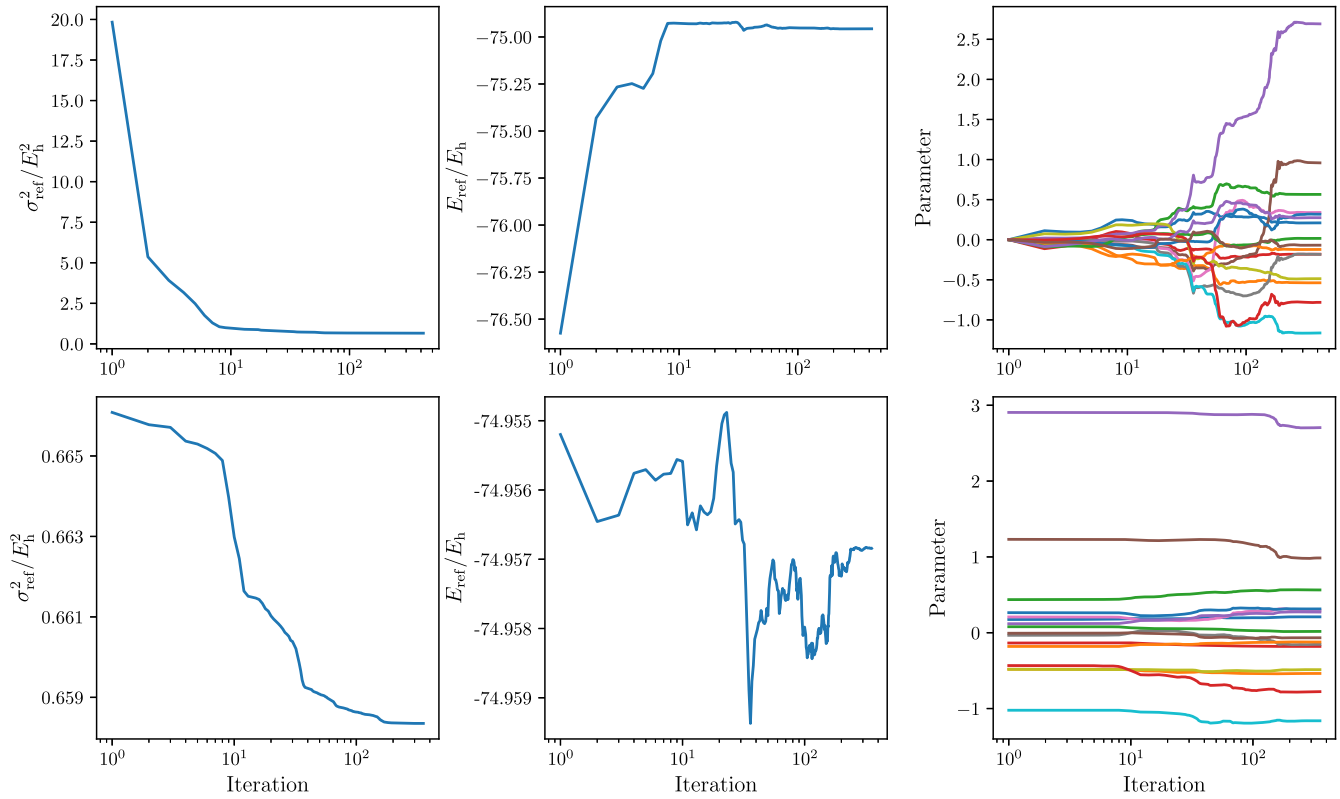


FIG. 2. Variation of the variance of the reference energy, the reference energy, and the various parameters in the Jastrow over the course of deterministic optimization for the oxygen atom in the cc-pCVTZ basis set. This particular Jastrow was of the DTN form, with 16 parameters in total (4 e–e terms, 4 e–n terms, and 8 e–e–n terms). The top panels correspond to the optimization from all-zero parameters, while the bottom panels correspond to the optimization from a VMC guess of the parameters. It is clear that, despite the two very different starting points, the Jastrow parameters converge to very similar values.

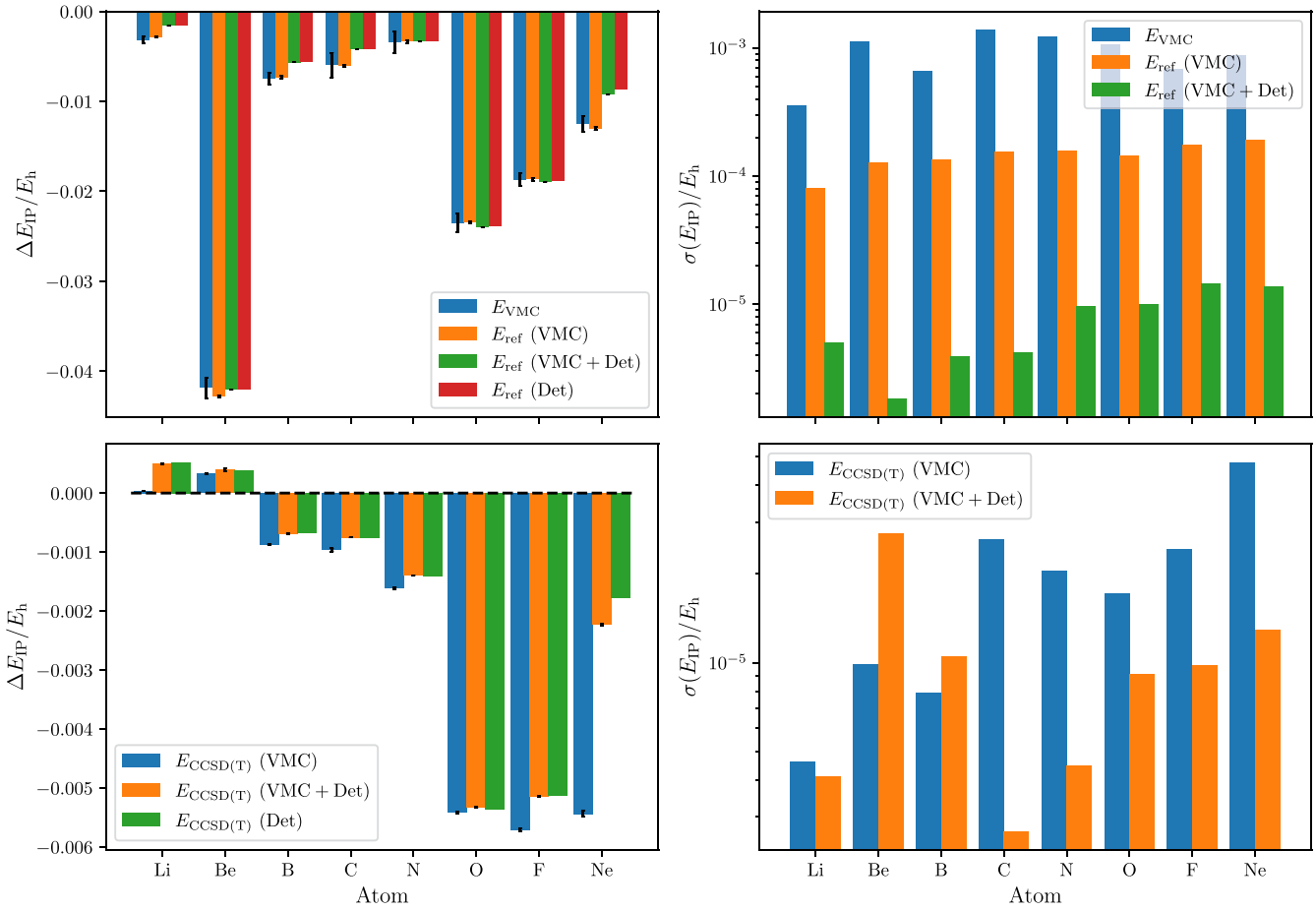


FIG. 3. Error in the IPs of first row elements as computed using VMC and reference energies (top left), the corresponding standard deviations (top right), the error in the IPs of first row elements as computed using xTC-CCSD(T) (bottom left), and the corresponding standard deviations (bottom right). Standard deviations are obtained from five independent VMC runs for each species. All results are obtained from the cc-pCVTZ basis set. The information in brackets denotes the optimization scheme employed.

At the xTC-CCSD(T) level, the decrease in variance is less pronounced but increases with system size. However, we once again observe a lowering of the error in the xTC-CCSD(T) IPs, which is as high as a factor of three for the Ne ionization.

The previous results are obtained using the cc-pCVTZ basis set[70,71] for both Jastrow optimization and the post-Hartree–Fock correlation treatment. We assess the transferability of the deterministically optimized Jastrow factors by using them to obtain reference and xTC-CCSD(T) energies in the cc-pVTZ basis set,[70] as shown in Fig. 4. In general, the results are comparable with those obtained from Jastrow factors optimized directly by VMC in the cc-pVTZ basis, with notable decreases in variance when employing deterministically optimized Jastrow factors. This suggests that post-Hartree–Fock calculations can be safely carried out in smaller basis sets than those used for Jastrow optimization, avoiding prohibitive scaling with basis set size.

D. First-row molecules

We consider first-row homonuclear diatomics N_2, O_2, and F_2 and heteronuclear diatomic CO and assess the performance of deterministically optimized Jastrows for the dissociation energy of these systems relative to pure VMC optimization. For more complex Jastrow factors, the latter has previously been shown to be sufficient to accurately describe these molecular properties using aug-cc-pVTZ basis sets.[70,72] As a proof of concept, we use here the same Jastrow parameterization from Sec. IV C, which has fewer terms than the one used in Ref. 54, resulting in less accurate results than those reported there. Total energies obtained by xTC-CCSD(T) are shown in Table II, with the corresponding atomization energies in Table III.

We carry out calculations in the cc-pVTZ, cc-pCVTZ, and aug-cc-pVTZ basis sets. Deterministic optimization of the Jastrow factor is always carried out in the cc-pCVTZ basis set, irrespective of which

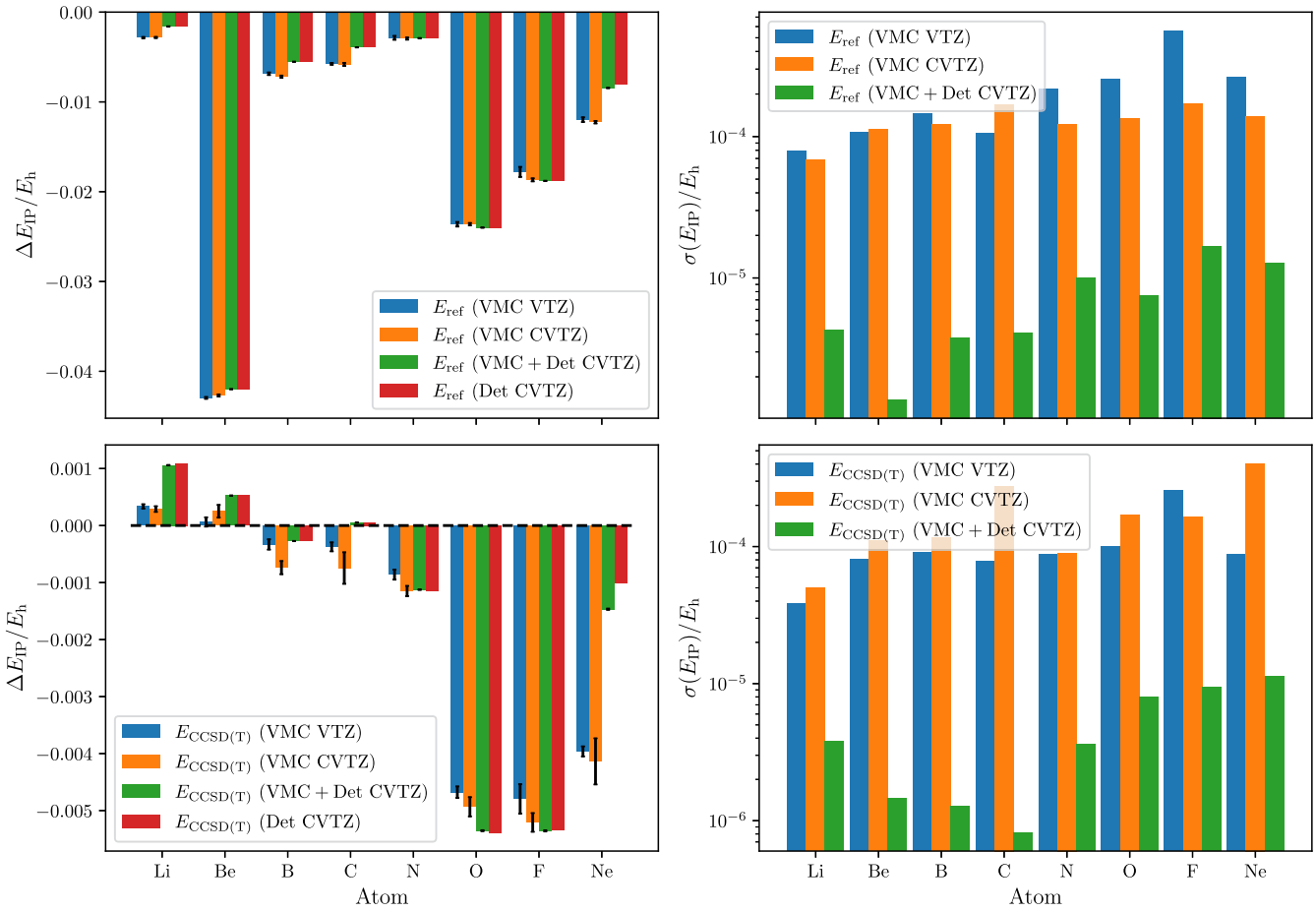


FIG. 4. Error in the IPs of first row elements as computed using reference energies (top left), the corresponding standard deviations (top right), the error in the IPs of first row elements as computed using TC-CCSD(T) (bottom left), and the corresponding standard deviations (bottom right). Standard deviations are obtained from five independent VMC runs for each species. Blue bars correspond to the results obtained from Jastrow factors optimized by VMC in the cc-pVTZ basis, while all others use the cc-pCVTZ basis for VMC and deterministic optimization. Reference and TC-CCSD(T) energies were computed in the cc-pVTZ basis. The information in brackets denotes the optimization scheme employed and the basis set used for final energy calculations.

TABLE II. Errors in the total energies of first row atoms, diatomics, and hydrides, computed using xTC-CCSD(T) with Jastrow factors optimized by VMC or by combined VMC and deterministic optimization, in Hartree. Total HEAT values for the energies are given in the second column. Calculations were carried out using cc-pVTZ (VTZ), cc-pCVTZ (CVTZ), and aug-cc-pVTZ (aVTZ) basis sets. For the VMC-only calculations, Jastrow optimization was carried out in the same basis set as the post–HF calculation. For VMC followed by deterministic optimization (VMC + det), both phases of Jastrow optimization were carried out in the cc-pCVTZ basis, regardless of the basis used for the post-HF calculation.

Molecule	Total energy (HEAT)	VMC (VTZ)	VMC (CVTZ)	VMC (aVTZ)	VMC + det (VTZ)	VMC + det (CVTZ)	VMC + det (aVTZ)
H	−0.500 022	0.000 124	⋯	0.000 109	0.000 126	⋯	− 0.000 02
C	−37.845 311	0.004 715	0.003 252	0.003 851	0.002 811	0.002 639	0.001 947
N	−54.589 743	0.008 351	0.006 626	0.006 508	0.006 510	0.005 674	0.004 698
O	−75.068 210	0.016 152	0.012 040	0.015 061	0.014 952	0.013 496	0.010 903
F	−99.735 148	0.023 856	0.018 668	0.023 393	0.023 495	0.020 955	0.018 150
N_2	−109.543 481	0.022 666	0.021 423	0.017 381	0.021 194	0.019 39	0.015 987
O_2	−150.328 884	0.036 044	0.033 140	0.026 535	0.030 875	0.026 545	0.022 136
F_2	−199.532 431	0.047 067	0.045 415	0.032 684	0.040 114	0.037 278	0.027 948
HF	−100.460 908	0.027 603	0.027 335	0.018 526	0.029 173	0.026 783	0.019 63
H_2O	−76.439 478	0.021 737	0.021 283	0.013 207	0.022 366	0.021 093	0.013 878
CO	−113.327 394	0.023 659	0.021 483	0.018 297	0.020 992	0.018 820	0.016 119

TABLE III. Errors in the atomization energies of first row molecules, computed using xTC-CCSD(T) with Jastrow factors optimized by VMC or by combined VMC and deterministic optimization, in Hartree. HEAT values for the atomization energies are given in the second column. Calculations were carried out using cc-pVTZ (VTZ), cc-pCVTZ (CVTZ), and aug-cc-pVTZ (aVTZ) basis sets. For the VMC-only calculations, Jastrow optimization was carried out in the same basis set as the post-HF calculation. For VMC followed by deterministic optimization (VMC + det), both phases of Jastrow optimization were carried out in the cc-pCVTZ basis, regardless of the basis used for the post-HF calculation.

Molecule	Atom. energy (HEAT)	VMC (VTZ)	VMC (CVTZ)	VMC + det (VTZ)	VMC + det (CVTZ)	VMC (aVTZ)	VMC + det (aVTZ)
N_2	0.363 995	−0.005 964	−0.008 171	−0.008 174	−0.008 042	−0.004 365	−0.006 591
O_2	0.192 464	−0.003 700	−0.003 018	−0.000 971	0.000 447	−0.002 455	−0.000 330
F_2	0.062 135	0.000 765	0.001 371	0.006 876	0.004 632	0.004 652	0.008 352
HF	0.225 638	−0.003 623	−0.003 818	−0.005 552	−0.005 702	0.000 251	−0.0 015
H_2O	0.871 246	−0.005 337	−0.005 874	−0.007 162	−0.007 345	−0.000 949	−0.003 015
CO	0.413 873	−0.002 792	−0.003 170	−0.003 229	−0.002 685	−0.002 406	−0.003 269

basis set is then used for the xTC-CCSD(T) calculation. For pure VMC optimization of the Jastrow factor, the target xTC-CCSD(T) basis set is employed. In all cases, deterministic optimization reduces the error in the total energies significantly, although the cancellation of errors sometimes proves less advantageous for the atomization energy.

We also consider two hydrides, HF and H_2O. The deterministic optimization routine employed in this paper generally requires core-like functions in the basis set used for Jastrow optimization. However, no such basis sets are available for the hydrogen atom, as it only has valence electrons. We find that in this case, using the cc-pVTZ basis set for hydrogen in the optimization does not introduce significant error relative to the energy obtained by VMC optimization. Therefore, for these molecules, we optimize the Jastrow factor using a cc-pCVTZ basis on the heavy atom and cc-pVTZ on hydrogen, with results also given in Tables II and III. We find that, in this case, deterministic optimization only reduces the total energy error in the core-valence basis set, and error cancellation in the atomization energy is generally poorer for the deterministically optimized Jastrow factors. Nevertheless, in most cases, results are of a similar quality to those obtained by pure VMC optimization of the Jastrow factors, and we expect that sufficiently flexible Jastrow parameterizations would be able to recover the sub-milliHartree accuracy of previously reported results in a more reliable, noise-free way.

E. Deterministic optimization for variational energies

Finally, we consider the quality of the variational energies of Slater–Jastrow wavefunctions obtained using deterministic optimization of the Jastrow factor, without further applying the transcorrelated methodology. Figure 5 shows the VMC energies E_{VMC} and their corresponding variances σ^2_{VMC} for first row atoms, computed by VMC minimization of either E_{VMC} or σ^2_{ref}, as well as

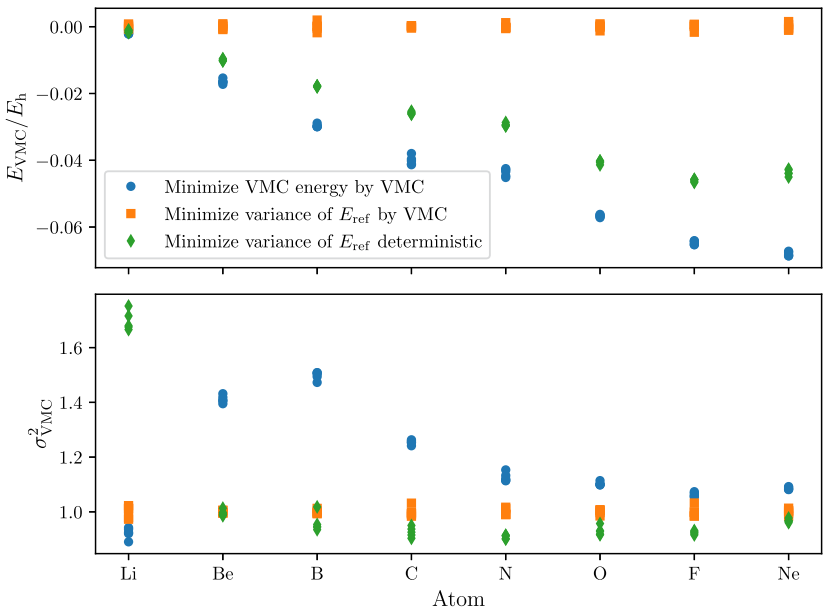


FIG. 5. VMC energies and variances obtained by variational energy minimization (blue), variance of the reference energy minimization (orange), and the proposed deterministic minimization protocol (green). Energies were shifted such that the mean of the E_{ref} variance minimization corresponds to 0. The variances were scaled such that the mean of the E_{ref} variance minimization corresponds to 1. All VMC calculations used the cc-pVTZ basis set for the Hartree–Fock distribution.

TABLE IV. Mean VMC energies and variances obtained by variational energy minimization, variance of the reference energy minimization, and the proposed deterministic minimization protocol. Energies are given in Hartree and variances in Hartree^2. All VMC calculations used the cc-pVTZ basis set for the Hartree–Fock distribution.

	Min. E_{VMC}		Min. σ^2_{ref}		Min. determ. σ^2_{ref}	
Atom	E_{VMC}	σ^2_{VMC}	E_{VMC}	σ^2_{VMC}	E_{VMC}	σ^2_{VMC}
Li	−7.4765	0.0132	−7.4748	0.0143	−7.4763	0.0243
Be	−14.6381	0.1074	−14.6216	0.0761	−14.6316	0.0758
B	−24.6172	0.4027	−24.5875	0.2689	−24.6054	0.2580
C	−37.8040	0.7121	−37.7639	0.5679	−37.7897	0.5261
N	−54.5461	1.3954	−54.5024	1.2382	−54.5317	1.1248
O	−75.0123	2.7475	−74.9563	2.4914	−74.9969	2.3180
F	−99.6715	5.0309	−99.6068	4.7323	−99.6529	4.3727
Ne	−128.8676	8.7779	−128.7995	8.0831	−128.8432	7.8219

those obtained by enhancing VMC minimization of σ^2_{ref} with deterministic optimization. Corresponding average values of E_{VMC} and σ^2_{VMC} are given in Table IV.

We find that the variational energies obtained by employing deterministic minimization are significantly lower than those from VMC optimization of σ^2_{ref}, approaching those obtained by direct variational energy minimization. However, σ^2_{VMC} remains much lower than the values obtained from energy minimization, suggesting this approach as a viable scheme to obtain VMC energies that are both variationally close to the ground state and have low variances, and thereby have utility beyond TC applications.

V. CONCLUSIONS

In this paper, we have presented a new deterministic Jastrow optimization scheme based on the minimization of the variance of the transcorrelated reference energy, as computed in a finite basis set. Core-like basis functions are crucial for the success of this approach, as in their absence, the computed variance is significantly underestimated, and the wrong parameterization is predicted for the minimum.

We benchmark this approach against the ionization potentials of first-row atoms, as well as the total energies and atomization energies of a representative set of second-row molecules. In general, results are of a similar standard to those obtained by VMC optimization. We find that, when used in conjunction with VMC, deterministic optimization significantly reduces the variance of the estimated reference energies, as well as the resulting xTC-CCSD(T) energies. For total molecular energies, it also results in lowered errors, although error cancellation in the atomization energies is slightly poorer in most cases, with the notable exception of the oxygen molecule.

The reduction of variance achievable through this deterministic optimization scheme opens up new directions for transcorrelated approaches, most notably the treatment of weak interactions, such as dispersion. In this case, the noise due to the VMC optimization would be sufficient to occlude any interaction, so we are interested in exploring the performance of deterministically optimized Jastrow factors in this setting in a future publication.

Recently, the use of effective core potentials (ECPs) in conjunction with transcorrelation has been proposed as a means to reduce the overhead of such calculations by reducing the required basis set size while maintaining and, in some cases, improving accuracy.[73] This reduction would be beneficial for deterministic optimization as well; however, we need to explore how this approach can be effectively combined with deterministic Jastrow optimization, particularly due to its effective removal of core orbitals from the wavefunction. In a similar direction, it may be possible to carry out the all-electron calculation but perform the Jastrow optimization using a frozen-core xTC approach to compute the variance. In the present work, we find that orbitals with significant density in the core region are required for the optimization, but the important contributions come from virtual orbitals, not so much from the occupied core orbitals, so ECP and frozen core approaches may be viable pathways to reducing the cost of deterministic Jastrow optimization.

Finally, this approach is promising for pure VMC applications as well, as it provides a better trade-off between VMC energy and variance than conventional VMC approaches based on direct energy or variance minimization, which tend to generate large values of whichever quantity is not being minimized. VMC optimization of the variance of the TC reference energy already provided an improvement in this case, but this is further enhanced by the use of the deterministic optimization scheme proposed in this work.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the funding from the Max Planck Society.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Maria-Andreea Filip: Conceptualization (lead); Data curation (lead); Investigation (lead); Methodology (equal); Software (lead); Writing – original draft (lead); Writing – review & editing (equal). **Evelin Martine Corvid Christlmaier**: Software (supporting); Writing – review & editing (equal). **J. Philip Haupt**: Software (supporting). **Daniel Kats**: Investigation (supporting); Software (supporting); Writing – review & editing (equal). **Pablo López Ríos**: Software (supporting); Writing – review & editing (equal). **Ali Alavi**: Conceptualization (supporting); Funding acquisition (lead); Methodology (equal); Project administration (lead); Resources (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] R. Jastrow, "Many-body problem with strong forces," Phys. Rev. **98**, 1479–1484 (1955).
[2] J. H. Bartlett, "Helium wave equation," Phys. Rev. **98**, 1067–1070 (1955).
[3] D. Ceperley, G. V. Chester, and M. H. Kalos, "Monte Carlo simulation of a many-fermion study," Phys. Rev. B **16**, 3081–3099 (1977).
[4] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, "Optimized trial wave functions for quantum Monte Carlo calculations," Phys. Rev. Lett. **60**, 1719–1722 (1988).
[5] K. E. Schmidt and J. W. Moskowitz, "Correlated Monte Carlo wave functions for the atoms He through Ne," J. Chem. Phys. **93**, 4172–4178 (1990).
[6] M. P. Nightingale and V. Melik-Alaverdian, "Optimization of ground- and excited-state wave functions and van der Waals clusters," Phys. Rev. Lett. **87**, 043401 (2001).
[7] J. Toulouse and C. J. Umrigar, "Optimization of quantum Monte Carlo wave functions by energy minimization," J. Chem. Phys. **126**, 084102 (2007).
[8] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, "Alleviation of the fermion-sign problem by optimization of many-body wave functions," Phys. Rev. Lett. **98**, 110201 (2007).
[9] R. J. Needs, M. D. Towler, N. D. Drummond, P. López Ríos, and J. R. Trail, "Variational and diffusion quantum Monte Carlo calculations with the CASINO code," J. Chem. Phys. **152**, 154106 (2020).
[10] J. B. Anderson, "A random-walk simulation of the Schrödinger equation: H^+_3," J. Chem. Phys. **63**, 1499 (1975).
[11] J. B. Anderson, "Quantum chemistry by random walk. H ^2P, H^+_3 D_{3h} $^1A'_1$, H_2 $^3\Sigma^+_u$, H_4 $^1\Sigma^+_g$, Be ^1S," J. Chem. Phys. **65**, 4121–4127 (1976).
[12] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," Phys. Rev. Lett. **45**, 566–569 (1980).
[13] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Rev. Mod. Phys. **73**, 33–83 (2001).
[14] E. A. Hylleraas, "Über den grundzustand des heliumatoms," Z. Phys. **48**, 469–494 (1928).
[15] E. A. Hylleraas, "Neue berechnung der energie des heliums im grundzustande, sowie des tiefsten terms von ortho-helium," Z. Phys. **54**, 347–366 (1929).
[16] E. A. Hylleraas, "Über den grundterm der zweielektronenprobleme von H^-, He, Li^+, Be^{++} usw.," Z. Phys. **65**, 209–225 (1930).
[17] W. Kutzelnigg, "r_{12}-Dependent terms in the wave function as closed sums of partial wave amplitudes for large I," Theoret. Chim. Acta **68**, 445–469 (1985).
[18] W. Klopper and W. Kutzelnigg, "Møller–Plesset calculations taking care of the correlation CUSP," Chem. Phys. Lett. **134**, 17–22 (1987).
[19] W. Klopper and W. Kutzelnigg, "Mp2-R12 calculations on the relative stability of carbocations," J. Phys. Chem. **94**, 5625–5630 (1990).
[20] W. Kutzelnigg and W. Klopper, "Wave functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. I. General theory," J. Chem. Phys. **94**, 1985–2001 (1991).
[21] J. Noga and W. Kutzelnigg, "Coupled cluster theory that takes care of the correlation CUSP by inclusion of linear terms in the interelectronic coordinates," J. Chem. Phys. **101**, 7738–7762 (1994).
[22] J. Noga, W. Klopper, and W. Kutzelnigg, "CC-R12: An explicitly correlated coupled-cluster theory," in *Recent Advances in Coupled-Cluster Methods* (World Scientific, 1997), Vol. 1–48, p. 1.
[23] S. Ten-no, "Initiation of explicitly correlated slater-type geminal theory," Chem. Phys. Lett. **398**, 56–61 (2004).
[24] E. F. Valeev, "Improving on the resolution of the identity in linear R12 ab initio theories," Chem. Phys. Lett. **395**, 190–195 (2004).
[25] S. Kedžuch, M. Milko, and J. Noga, "Alternative formulation of the matrix elements in MP2-R12 theory," Int. J. Quantum Chem. **105**, 929–936 (2005).
[26] T. B. Adler, G. Knizia, and H. J. Werner, "A simple and efficient CCSD(T)-F12 approximation," J. Chem. Phys. **127**, 221106 (2007).
[27] G. Knizia and H. J. Werner, "Explicitly correlated RMP2 for high-spin open-shell reference states," J. Chem. Phys. **128**, 154103 (2008).
[28] G. Knizia, T. B. Adler, and H. J. Werner, "Simplified CCSD(T)-F12 methods: Theory and benchmarks," J. Chem. Phys. **130**, 054104 (2009).
[29] C. Hättig, D. P. Tew, and A. Köhn, "Communications: Accurate and efficient approximations to explicitly correlated coupled-cluster singles and doubles, CCSD-F12," J. Chem. Phys. **132**, 231102 (2010).
[30] L. Kong, F. A. Bischoff, and E. F. Valeev, "Explicitly correlated R12/F12 methods for electronic structure," Chem. Rev. **112**, 75–107 (2012).
[31] S. F. Boys and N. C. Handy, "The determination of energies and wavefunctions with full electronic correlation," Proc. R. Soc. London, Ser. A **310**, 43–61 (1969).
[32] N. C. Handy, "Energies and expectation values for be by the transcorrelated method," J. Chem. Phys. **51**, 3205–3212 (1969).
[33] N. C. Handy, "On the minimization of the variance of the transcorrelated Hamiltonian," Mol. Phys. **21**, 817–828 (1971).
[34] M. Nooijen and R. J. Bartlett, "Elimination of Coulombic infinities through transformation of the Hamiltonian," J. Chem. Phys. **109**, 8232–8240 (1998).
[35] S. Ten-no, "Three-electron integral evaluation in the transcorrelated method using a frozen Gaussian geminal," Chem. Phys. Lett. **330**, 175–179 (2000).
[36] S. Ten-no, "A feasible transcorrelated method for treating electronic CUSPS using a frozen Gaussian geminal," Chem. Phys. Lett. **330**, 169–174 (2000).
[37] O. Hino, Y. Tanimura, and S. Ten-no, "Application of the transcorrelated Hamiltonian to the linearized coupled cluster singles and doubles model," Chem. Phys. Lett. **353**, 317–323 (2002).
[38] R. Sakuma and S. Tsuneyuki, "Electronic structure calculations of solids with a similarity-transformed Hamiltonian," J. Phys. Soc. Jpn. **75**, 103705 (2006).
[39] S. Tsuneyuki, "Transcorrelated method: Another possible way towards electronic structure calculation of solids," Prog. Theor. Phys. Suppl. **176**, 134–142 (2008).
[40] H. Luo and A. Alavi, "Combining the transcorrelated method with full configuration interaction quantum Monte Carlo: Application to the homogeneous electron gas," J. Chem. Theory Comput. **14**, 1403–1411 (2018).
[41] A. J. Cohen, H. Luo, K. Guther, W. Dobrautz, D. P. Tew, and A. Alavi, "Similarity transformation of the electronic Schrödinger equation via Jastrow factorization," J. Chem. Phys. **151**, 061101 (2019).
[42] K. Liao, T. Schraivogel, H. Luo, D. Kats, and A. Alavi, "Towards efficient and accurate ab initio solutions to periodic systems via transcorrelation and coupled cluster theory," Phys. Rev. Res. **3**, 033072 (2021).
[43] T. Schraivogel, A. J. Cohen, A. Alavi, and D. Kats, "Transcorrelated coupled cluster methods," J. Chem. Phys. **155**, 191101 (2021).
[44] A. Ammar, A. Scemama, and E. Giner, "Extension of selected configuration interaction for transcorrelated methods," J. Chem. Phys. **157**, 134107 (2022).
[45] A. Baiardi, M. Lesiuk, and M. Reiher, "Explicitly correlated electronic structure calculations with transcorrelated matrix product operators," J. Chem. Theory Comput. **18**, 4203–4217 (2022).
[46] A. Ammar, A. Scemama, and E. Giner, "Biorthonormal orbital optimization with a cheap core-electron-free three-body correlation factor for quantum Monte Carlo and transcorrelation," J. Chem. Theory Comput. **19**, 4883–4896 (2023).
[47] A. Ammar, A. Scemama, and E. Giner, "Transcorrelated selected configuration interaction in a bi-orthonormal basis and with a cheap three-body correlation factor," J. Chem. Phys. **159**, 114121 (2023).
[48] K. Liao, H. Zhai, E. M. C. Christlmaier, T. Schraivogel, P. L. Ríos, D. Kats, and A. Alavi, "Density matrix renormalization group for transcorrelated Hamiltonians: Ground and excited states in molecules," J. Chem. Theory Comput. **19**, 1734–1743 (2023).
[49] T. Schraivogel, E. M. C. Christlmaier, P. López Ríos, A. Alavi, and D. Kats, "Transcorrelated coupled cluster methods. II. Molecular systems," J. Chem. Phys. **158**, 214106 (2023).
[50] A. Ammar, A. Scemama, P.-F. Loos, and E. Giner, "Compactification of determinant expansions via transcorrelation," J. Chem. Phys. **161**, 084104 (2024).
[51] M. Mörchen, A. Baiardi, M. Lesiuk, and M. Reiher, "Non-iterative triples for transcorrelated coupled cluster theory," J. Chem. Theory Comput. **21**, 1588–1601 (2025).
[52] W. Dobrautz, H. Luo, and A. Alavi, "Compact numerical solutions to the two-dimensional repulsive Hubbard model obtained via nonunitary similarity transformations," Phys. Rev. B **99**, 075119 (2019).
[53] J. P. Haupt, S. M. Hosseini, P. López Ríos, W. Dobrautz, A. Cohen, and A. Alavi, "Optimizing Jastrow factors for the transcorrelated method," J. Chem. Phys. **158**, 224105 (2023).

[54]E. M. C. Christlmaier, T. Schraivogel, P. López Ríos, A. Alavi, and D. Kats, "xTC: An efficient treatment of three-body interactions in transcorrelated methods," J. Chem. Phys. **159**, 014113 (2023).
[55]N. D. Drummond, M. D. Towler, and R. J. Needs, "Jastrow correlation factor for atoms, molecules, and solids," Phys. Rev. B **70**, 235119 (2004).
[56]P. R. C. Kent, R. J. Needs, and G. Rajagopal, "Monte Carlo energy and variance-minimization techniques for optimizing many-body wave functions," Phys. Rev. B **59**, 12344–12351 (1999).
[57]J. P. Haupt, E. M. C. Christlmaier, P. L. Ríos, N. A. Bogdanov, D. Kats, and A. Alavi, "Transcorrelated methods for multireference problems," arXiv:2505.20187 [physics.chem-ph] (2025).
[58]The expression given here is a corrected version of Eq. (19) in Ref. 53, to agree with Eq. (20) of the same publication, which is the relevant VMC working equation.
[59]P. López Ríos, P. Seth, N. D. Drummond, and R. J. Needs, "Framework for constructing generic jastrow correlation factors," Phys. Rev. E **86**, 036703 (2012).
[60]Transcorrelated Hamiltonian integral library TCHInt to be released; available from the authors upon reasonable request.
[61]Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, "PySCF: The python-based simulations of chemistry framework," WIREs Comput. Mol. Sci. **8**, e1340 (2018).
[62]D. Kats, T. Schraivogel, J. Hauskrecht, C. Rickert, and F. Wu (2024). "ElemCo.jl: Julia program package for electron correlation methods," GitHub. https://github.com/fkfest/ElemCo.jl
[63]D. Kats, E. M. C. Christlmaier, T. Schraivogel, and A. Alavi, "Orbital optimisation in xTC transcorrelated methods," Faraday Discuss. **254**, 382–401 (2024).
[64]B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson, "Gaussian basis sets for use in correlated molecular calculations. VII. valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg," Theor. Chem. Acc. **128**, 69–82 (2011).
[65]C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 1. General considerations," IMA J. Appl. Math. **6**, 76–90 (1970).
[66]R. Fletcher, "A new approach to variable metric algorithms," Comput. J. **13**, 317–322 (1970).
[67]D. Goldfarb, "A family of variable-metric methods derived by variational means," Math. Comput. **24**, 23–26 (1970).
[68]D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," Math. Comput. **24**, 647–656 (1970).
[69]D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," Math. Program. **45**, 503–528 (1989).
[70]T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007–1023 (1989).
[71]D. E. Woon and T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon," J. Chem. Phys. **103**, 4572–4585 (1995).
[72]R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," J. Chem. Phys. **96**, 6796–6806 (1992).
[73]K. Simula, E. M. C. Christlmaier, M.-A. Filip, J. P. Haupt, D. Kats, P. Lopez-Rios, and A. Alavi, "Transcorrelated theory with pseudopotentials," J. Chem. Theory Comput. **21**, 5155–5170 (2025).