CHALMERS
UNIVERSITY OF TECHNOLOGY

Optimizing Jastrow factors for the transcorrelated method

Downloaded from: https://research.chalmers.se, 2024-10-12 04:52 UTC

Citation for the original published paper (version of record):
Haupt, J., Hosseini, S., López Ríos, P. et al (2023). Optimizing Jastrow factors for the transcorrelated method. Journal of Chemical Physics, 158(22). http://dx.doi.org/10.1063/5.0147877

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

The Journal of Chemical Physics

RESEARCH ARTICLE | JUNE 08 2023

Optimizing Jastrow factors for the transcorrelated method

J. Philip Haupt ; Seyed Mohammadreza Hosseini ; Pablo López Ríos ; Werner Dobrautz ; Aron Cohen ; Ali Alavi

Check for updates

J. Chem. Phys. 158, 224105 (2023)
https://doi.org/10.1063/5.0147877

View Online
Export Citation
CrossMark

04 July 2023 12:30:16

AIP Publishing

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Optimizing Jastrow factors for the transcorrelated method

Cite as: J. Chem. Phys. 158, 224105 (2023); doi: 10.1063/5.0147877
Submitted: 27 February 2023 • Accepted: 25 May 2023 •
Published Online: 8 June 2023

View Online Export Citation CrossMark

J. Philip Haupt,[1] Seyed Mohammadreza Hosseini,[1] Pablo López Ríos,[1,a)] Werner Dobrautz,[2] Aron Cohen,[3] and Ali Alavi[1,4,b)]

AFFILIATIONS

[1] Max-Planck Institute for Solid State Research, Heisenbergstr. 1, 70569 Stuttgart, Germany
[2] Department of Chemistry and Chemical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden
[3] DeepMind, 6 Pancras Square, London N1C 4AG, United Kingdom
[4] Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

[a)] **Author to whom correspondence should be addressed:** p.lopez.rios@fkf.mpg.de
[b)] **Electronic mail:** a.alavi@fkf.mpg.de

ABSTRACT

We investigate the optimization of flexible tailored real-space Jastrow factors for use in the transcorrelated (TC) method in combination with highly accurate quantum chemistry methods, such as initiator full configuration interaction quantum Monte Carlo (FCIQMC). Jastrow factors obtained by minimizing the variance of the TC reference energy are found to yield better, more consistent results than those obtained by minimizing the variational energy. We compute all-electron atomization energies for the challenging first-row molecules C_2, CN, N_2, and O_2 and find that the TC method yields chemically accurate results using only the cc-pVTZ basis set, roughly matching the accuracy of non-TC calculations with the much larger cc-pV5Z basis set. We also investigate an approximation in which pure three-body excitations are neglected from the TC-FCIQMC dynamics, saving storage and computational costs, and show that it affects relative energies negligibly. Our results demonstrate that the combination of tailored real-space Jastrow factors with the multi-configurational TC-FCIQMC method provides a route to obtaining chemical accuracy using modest basis sets, obviating the need for basis-set extrapolation and composite techniques.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0147877

04 July 2023 12:30:16

I. INTRODUCTION

The numerical study of the many-body electronic Schrödinger equation is plagued by a number of difficulties that originate from the singular nature of the Coulomb potential, which at particle coalescence points induces a non-smooth behavior (cusps) in the wave function.[1,2] This gives rise to slow convergence of quantum chemical methods with respect to the size of the basis sets used. One cure for this problem is to introduce explicit dependence in the wave function on electron–electron and electron–nucleus distances.[3] This allows us to analytically encode the non-smooth behavior into the wave function ansatz, leaving only a relatively smooth wave function for further treatment. Within the so-called Jastrow ansatz,[4] the electronic wave function is written as

$$\Psi = e^{J}\Phi, \tag{1}$$

where $J = J(\mathbf{r}_1, \ldots, \mathbf{r}_{N_e})$ is a symmetric correlation factor that contains optimizable parameters and depends on the positions of the N_e electrons, while Φ is an antisymmetric function, which can be taken to be a single Slater determinant, such as the Hartree–Fock (HF) state, or a multi-configurational function, such as a configurationinteraction (CI) wave function. With J taking care of the majority of the non-analytic behavior, Φ can be expected to be a relatively smooth function, accurately expressible using small basis sets.

The presence of J complicates the many-body integrals required to solve the Schrödinger equation, which can usually only be handled by the real-space variational and diffusion quantum Monte Carlo (VMC and DMC) methods.[5,6] The transcorrelated (TC) method of Boys and Handy[7–9] provides a framework for the treatment of such an ansatz in second quantization, which requires only up to three-body integrals to be evaluated but involves a non-Hermitian

© Author(s) 2023

effective Hamiltonian. In the early work on the TC method, the non-Hermiticity of the formalism was found to be highly problematic and led to a waning of interest in the whole methodology.

There was a revival of interest in the TC method in the late 1990s with Nooijen and Bartlett[10] and Ten-no,[11–13] who also successfully combined the TC method with quantum chemical methods, such as Møller–Plesset perturbation theory and the linearized coupled cluster method. Tsuneyuki, Umezawa, and co-workers developed a Slater–Jastrow treatment of the TC Hamiltonian using the VMC method[14] and applied transcorrelation to solid-state systems.[15–18]

With the development of the R12/F12 class of explicitly correlated methods,[19–27] the interest in the TC method in quantum chemistry once again declined. A second revival of the TC method has recently been spurred by the realization that the non-Hermiticity of the TC Hamiltonian is unproblematic[28] for methods such as full configuration-interaction quantum Monte Carlo (FCIQMC)[29–31] while providing substantial improvements in basis-set convergence[32] and in the compactness of CI expansions, for example for strongly correlated two-dimensional Hubbard models.[33] Coupled-cluster methods have also been developed to treat the TC Hamiltonian, both in its full form and via accurate approximations for the efficient treatment of three-body interactions.[34–37] A consistent finding in these works has been that transcorrelation improves not only the basis-set convergence, as expected, but also the effective level of theory of the underlying correlation method, providing a major motivation to develop this methodology with the ultimate aim of studying strongly correlated *ab initio* systems.

The TC method amounts to a non-unitary transformation of the Hamiltonian of the system, and methods to solve such equations are, in general, not guaranteed to converge to the exact total energy from above as the complete basis set limit is approached. The choice of Jastrow factor proves to be critical in the TC method since a poor choice can lead to highly non-variational TC energies and poor error cancellation in energy differences. In this paper, we present a method for the optimization of Jastrow factors catering to the TC approach, which we demonstrate in combination with the FCIQMC method.

TC-FCIQMC is a stochastic eigensolver for the TC Hamiltonian and allows for large CI solutions to be obtained for its eigenvectors. This methodology has been used in the past to study first-row atoms[32] using Boys–Handy-type Jastrow factors[38] pre-optimized for use in the VMC method,[39] the binding curve of the beryllium dimer using a Boys–Handy Jastrow with an exponential kernel optimized with VMC variance minimization,[40] and ultra-cold atomic systems with contact interactions.[41,42] With the present method, new highly flexible Jastrow factors tailored to each system can be obtained for large molecules, allowing for the methodology to be extended, both within the context of TC-FCIQMC and in other approaches, such as TC density matrix renormalization group,[43,44] TC selected-CI approaches,[45] and TC coupled-cluster theory,[34,35] or for reducing the resources required in a quantum computing setting.[46,47] In our tests, we found the atomization energies of various challenging first-row molecules to be chemically accurate using the moderately sized standard cc-pVTZ basis set.[48]

The rest of this paper is structured as follows: In Sec. II, we give an overview of the broader theoretical framework used in our calculations. The details of our proposed optimization methodology are given in Sec. III along with data for various first-row atoms and molecules to support our choices. In Sec. IV, we analyze the accuracy of the TC method in calculations of the atomization energies of these molecules, which we compare with their non-TC counterparts, and we present our conclusions in Sec. V. Hartree atomic units ($\hbar = |e| = m_e = 4\pi\epsilon_0 = 1$) are used throughout unless stated otherwise.

II. METHODOLOGICAL FRAMEWORK

Substituting the Jastrow ansatz of Eq. (1) into the Schrödinger equation, $\hat{H}\Psi = E\Psi$, we obtain the similarity-transformed Schrödinger equation,

$$\hat{H}_{\mathrm{TC}}\Phi = E\Phi, \tag{2}$$

where the transcorrelated Hamiltonian is

$$\hat{H}_{\mathrm{TC}} = e^{-J}\hat{H}e^{J} \tag{3}$$

and Φ is the right-eigenvector of $\hat{H}_{\mathrm{TC}}$. We approach solving the similarity-transformed Schrödinger equation in two stages: first, we obtain a suitable J using VMC-based optimization, and then, we obtain Φ as a CI wave function expanded in a standard quantum chemical basis set using the TC-FCIQMC method.

The TC-FCIQMC method is based on the imaginary-time Schrödinger equation with the TC Hamiltonian,

$$-\frac{\partial\Phi}{\partial\tau} = \hat{H}_{\mathrm{TC}}\Phi, \tag{4}$$

whose long-τ solution is the ground-state wave function that satisfies Eq. (2). As in the non-TC FCIQMC algorithm,[29] Φ is expressed as a general linear combination of Slater determinants,

$$|\Phi\rangle = \sum_I c_I|D_I\rangle, \tag{5}$$

where the c_I coefficients are sampled by walkers in the FCIQMC simulation. Once the ground state has been reached, the total energy can be evaluated by projection onto the HF determinant,

$$E_{\mathrm{proj}} = \frac{\langle D_{\mathrm{HF}}|\hat{H}_{\mathrm{TC}}|\Phi\rangle}{\langle D_{\mathrm{HF}}|\Phi\rangle}, \tag{6}$$

which is averaged over time steps at finite walker numbers to obtain a statistically meaningful result.

As has been shown by Luo and Alavi,[28] the non-Hermitian nature of $\hat{H}_{\mathrm{TC}}$ does not hamper the convergence of the FCIQMC simulation, and the tools developed for that method apply to $\hat{H}_{\mathrm{TC}}$ as they do to $\hat{H}$; this includes the initiator approximation,[30] which we use in all of the FCIQMC calculations reported in the present work. Furthermore, Dobrautz *et al.*[33] showed for the strongly correlated 2D Hubbard model that $\hat{H}_{\mathrm{TC}}$ has a significantly more compact ground-state right eigenvector than $\hat{H}$ does, implying a reduction in the initiator error, faster convergence with walker number, and improved size consistency of the results.

The TC Hamiltonian can be exactly evaluated using the Baker–Campbell–Hausdorff expansion, which for Jastrow factors

04 July 2023 12:30:16

© Author(s) 2023

dependent only on electronic positions truncates exactly at second order,

$$\hat{H}_{\mathrm{TC}} = e^{-J}\hat{H}e^{J} = \hat{H} + [\hat{H}, J] + \frac{1}{2}[[\hat{H}, J], J]. \tag{7}$$

For a Jastrow factor containing up to two electron contributions,

$$J = \sum_{i<j} u(\mathbf{r}_i, \mathbf{r}_j), \tag{8}$$

the explicit form of $\hat{H}_{\mathrm{TC}}$ is

$$\begin{aligned}\hat{H}_{\mathrm{TC}} &= \hat{H} - \sum_i \left(\frac{1}{2}\nabla_i^2 J + (\nabla_i J)\cdot\nabla_i + \frac{1}{2}(\nabla_i J)^2\right)\\ &= \hat{H} - \sum_{i<j}\hat{K}(\mathbf{r}_i, \mathbf{r}_j) - \sum_{i<j<k}\hat{L}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k),\end{aligned} \tag{9}$$

where the additional two- and three-body terms are

$$\begin{aligned}\hat{K}(\mathbf{r}_i, \mathbf{r}_j) &= \frac{1}{2}\big[\nabla_i^2 u(\mathbf{r}_i, \mathbf{r}_j) + \nabla_j^2 u(\mathbf{r}_i, \mathbf{r}_j)\\ &\quad + |\nabla_i u(\mathbf{r}_i, \mathbf{r}_j)|^2 + |\nabla_j u(\mathbf{r}_j, \mathbf{r}_i)|^2\big]\\ &\quad + \nabla_i u(\mathbf{r}_i, \mathbf{r}_j)\cdot\nabla_i + \nabla_j u(\mathbf{r}_i, \mathbf{r}_j)\cdot\nabla_j,\\ \hat{L}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) &= \nabla_i u(\mathbf{r}_i, \mathbf{r}_j)\cdot\nabla_i u(\mathbf{r}_i, \mathbf{r}_k)\\ &\quad + \nabla_j u(\mathbf{r}_j, \mathbf{r}_i)\cdot\nabla_j u(\mathbf{r}_j, \mathbf{r}_k)\\ &\quad + \nabla_k u(\mathbf{r}_k, \mathbf{r}_i)\cdot\nabla_k u(\mathbf{r}_k, \mathbf{r}_j).\end{aligned} \tag{10}$$

Using this first-quantized Hamiltonian, we can construct a second-quantized Hamiltonian for a given set of orthonormal spatial orbitals $\{\phi_1, \ldots, \phi_{n_{\mathrm{orb}}}\}$, with corresponding fermionic spin-$\frac{1}{2}$ creation (annihilation) operators $a_{p\sigma}^\dagger$ ($a_{p\sigma}$),

$$\begin{aligned}\hat{H}_{\mathrm{TC}} &= \sum_{pq\sigma} h_q^p a_{p\sigma}^\dagger a_{q\sigma} + \frac{1}{2}\sum_{pqrs}(V_{rs}^{pq} - K_{rs}^{pq})\sum_{\sigma\tau} a_{p\sigma}^\dagger a_{q\tau}^\dagger a_{s\tau} a_{r\sigma}\\ &\quad - \frac{1}{6}\sum_{pqrstu} L_{stu}^{pqr}\sum_{\sigma\tau\lambda} a_{p\sigma}^\dagger a_{q\tau}^\dagger a_{r\lambda}^\dagger a_{u\lambda} a_{t\tau} a_{s\sigma},\end{aligned} \tag{11}$$

where $h_q^p = -\frac{1}{2}\langle\phi_p|\nabla^2|\phi_q\rangle$ and $V_{rs}^{pq} = \langle\phi_p\phi_q|r_{12}^{-1}|\phi_r\phi_s\rangle$ are the one- and two-body terms of the original Hamiltonian, and $K_{rs}^{pq} = \langle\phi_p\phi_q|\hat{K}|\phi_r\phi_s\rangle$ and $L_{stu}^{pqr} = \langle\phi_p\phi_q\phi_r|\hat{L}|\phi_s\phi_t\phi_u\rangle$ are the corresponding terms arising from the similarity transformation.[32] Note that the three-body operator $\hat{L}$ is Hermitian, and for real orbitals, the L_{stu}^{pqr} tensor has 48-fold symmetry, a useful property to reduce the storage requirement for these integrals.

By construction, Jastrow factors can be used to impose local Kato cusp conditions[1] based on the relative spin of electron pairs, but they are ill-suited to describing more rigorous electronic-state dependent cusp conditions.[2] However, in the present work, we use spin-independent Jastrow factors for simplicity, which we constrain to obey the opposite-spin electron–electron cusp condition since it is physically more important than the parallel-spin cusp condition. The use of spin-dependent Jastrow factors would require replacing the spatial-orbital indices above with spin-orbital indices, resulting in an order of magnitude more three-body integrals to be computed and stored. Notwithstanding this increase in memory requirements, spin-dependent Jastrow factors may offer other advantages, such as faster basis set convergence,[24] and will be investigated in future work.

III. OPTIMIZATION METHODOLOGY

In this section, we present our methodological choices for the optimization of Jastrow factors to be used in TC methods. To illustrate the effect of these choices, we compute the ground-state energies of the all-electron C, N, and O atoms and the C_2, CN, N_2, and O_2 molecules at their equilibrium geometries,[49–51] listed in Table I, with non-TC and TC-FCIQMC calculations using HF orbitals (restricted open-shell HF orbitals in the case of open-shell systems) expanded in the standard cc-pVxZ family of basis sets.[48] The quality of energy differences is then assessed using the atomization energy of the molecules. In order to be able to determine whether the methodology is capable of delivering chemically accurate relative energies, i.e., incurring an error of less than 1 kcal/mol = 1.6 mHa, we aim to keep each of the different errors well below this threshold; we comment on the magnitude of the expected error from each different source in the subsections that follow. We expect a total bias in our resulting relative energies of less than 0.5 mHa.

For all of our calculations, we generate our orbitals and integration grids using PYSCF,[52] optimize Jastrow factors using the CASINO continuum QMC package,[5] compute TC matrix elements using the TCHINT library developed by the authors,[53] and perform FCIQMC and TC-FCIQMC calculations with the NECI FCIQMC package.[31] We report FCIQMC energies obtained by projection onto the HF determinant, as per Eq. (6).

In the result of this section, we will discuss the effect of modifying each part of this calculation pipeline on the final TC-FCIQMC energies; the aspects not being discussed are assumed to operate as per our final recommendations unless otherwise stated, e.g., when we discuss the use of variance or energy minimization, we use integration grids of $l_{\mathrm{grid}} = 2$, while when we discuss grid density, we use variance-minimized Jastrow factors. In particular, note that the reported FCIQMC energies have been extrapolated to the full CI (FCI) limit as described in Sec. III D 1, unless explicitly stated otherwise.

A. Jastrow factor

In continuum quantum Monte Carlo calculations, the Jastrow factor for a molecule consisting of N_{n} nuclei and N_{e} electrons is usually constructed as the sum of isotropic electron–electron, electron–nucleus, and electron–electron–nucleus terms,[54]

$$J = \sum_{i<j}^{N_{\mathrm{e}}} u(r_{ij}) + \sum_{i}^{N_{\mathrm{e}}}\sum_{I}^{N_{\mathrm{n}}} \chi(r_{iI}) + \sum_{i<j}^{N_{\mathrm{e}}}\sum_{I}^{N_{\mathrm{n}}} f(r_{ij}, r_{iI}, r_{jI}). \tag{12}$$

TABLE I. Electronic ground states and equilibrium bond lengths used for the molecules considered in this work, following Ref. 49.

System	State	r_{eq} (Å)
C_2	$^1\Sigma_g^+$	1.2425
CN	$^2\Sigma^+$	1.1718
N_2	$^1\Sigma_g^+$	1.0977
O_2	$^3\Sigma_g^-$	1.2075

04 July 2023 12:30:16

© Author(s) 2023

As in Ref. 55, we express each of these terms as a natural power expansion in the relevant inter-particle distances,

$$u(r_{ij}) = t(r_{ij}, L_u)\sum_k a_k r_{ij}^k,$$
$$\chi(r_{iI}) = t(r_{iI}, L_\chi)\sum_k b_k r_{iI}^k, \tag{13}$$
$$f(r_{ij}, r_i, r_j) = t(r_{iI}, L_f)t(r_{jI}, L_f)\sum_{k,l,m} c_{klm} r_{ij}^k r_{iI}^l r_{jI}^m,$$

where $\{a_k\}$, $\{b_k\}$, and $\{c_{klm}\}$ are linear parameters, L_u, L_χ, and L_f are cutoff lengths, $t(r, L) = (1 - r/L)^3\Theta(r - L)$ is a cutoff function, and $\Theta(r - L)$ is the Heaviside step function.

In essence, the VMC and DMC methods sample real-space electronic configurations $\{\mathbf{R}\}$ following an appropriate distribution based on an analytic trial wave function $\Psi_\mathrm{T}(\mathbf{R})$ and produce a variational estimate of the total energy, which is an average of the local energy $E_\mathrm{L}(\mathbf{R}) = \Psi_\mathrm{T}^{-1}(\mathbf{R})\hat{H}(\mathbf{R})\Psi_\mathrm{T}(\mathbf{R})$, at the sampled configurations. Both the electron–electron and electron–nucleus Kato cusp conditions[1] are of crucial importance in suppressing extreme outliers in the local energy samples, making it possible to obtain meaningful wave-function parameter sets from VMC-based optimization. It is standard practice to apply the electron–electron cusp condition on the u term of the Jastrow factor, while the electron–nucleus cusp condition is enforced by modifying the $l = 0$ component, $\phi(r)$, of cuspless molecular orbitals near nuclei so that they exhibit a cusp.[56] This has been found to be a better approach than applying the electron–nucleus cusp condition via the parameters in the χ term of the Jastrow factor.[5,55,56]

However, in TC-FCIQMC, it is preferable to use unmodified molecular orbitals obtained from standard basis sets. It would be possible to optimize the Jastrow factor parameters in VMC in the presence of cusp-corrected orbitals and use them in TC-FCIQMC with a cusp-uncorrected orbital basis, but the Jastrow factor would then be sub-optimal by construction in the latter calculation. Instead, we recast the cusp-correction scheme of Ref. 56 as an electron–nucleus Jastrow factor term Λ, to be added to (rather than replacing) the χ term in Eq. (12).

We construct our cusp-correcting Jastrow factor term as

$$\Lambda(r) = \left[\ln \tilde{\phi}(r) - \ln \phi(r)\right]\Theta(r - r_\mathrm{c}), \tag{14}$$

where, using the notation of Ref. 56, r_c is the cutoff radius, $\phi(r)$ is the $l = 0$ component of the target orbital at the desired nucleus, and $\tilde{\phi}(r)$ is its cusp-corrected counterpart,

$$\tilde{\phi}(r) = e^{\sum_{l=0}^{4}\lambda_l r^l} + C, \quad r < r_\mathrm{c}. \tag{15}$$

Here, $\{\lambda_l\}$ are parameters determining the shape of the corrected orbital and parameter C is only set to a non-zero value in the presence of nodes of $\phi(r)$ near the nucleus.

Applying continuity and differentiability conditions (see Eq. 14 of Ref. 56) leaves λ_0 and r_c as the only free parameters in Eqs. (14) and (15). Reference 56 describes an approach to obtaining reasonable values for these parameters, which we use as initial values to be refined by the subsequent VMC optimization procedure. In practice, we evaluate $\phi(r)$ in $\Lambda(r)$ by spline interpolation of tabulated data. Figure 1 illustrates the effect of using a Λ term in practice.

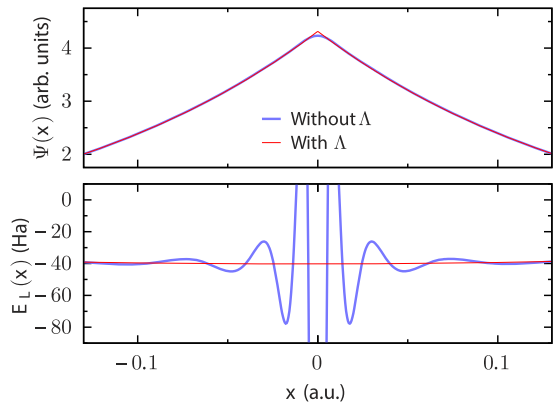

FIG. 1. Wave function value and local energy as a function of the x coordinate of an electron in a carbon atom as it crosses the nucleus at $x = 0$, both without and with a cusp-correcting Λ Jastrow factor term applied to a HF wave function using the cc-pVDZ basis.

For reference, in our calculations, we optimize a total of 44 Jastrow factor parameters for the atoms and homonuclear dimers and 80 parameters for the CN dimer; we keep the L_u, L_χ, and L_f cutoff lengths fixed at sensible values for simplicity.

B. Jastrow factor optimization

The VMC method not only is capable of evaluating the variational energy associated with a trial wave function Ψ_T,

$$E_\mathrm{VMC} = \frac{\langle\Psi_\mathrm{T}|\hat{H}|\Psi_\mathrm{T}\rangle}{\langle\Psi_\mathrm{T}|\Psi_\mathrm{T}\rangle}, \tag{16}$$

but also provides a variational framework in which any parameters $\boldsymbol{\alpha}$ present in this trial wave function can be optimized; note that, in practice, $|\Psi_\mathrm{T}\rangle = e^J|D_\mathrm{HF}\rangle$, and $\boldsymbol{\alpha}$ are the optimizable parameters in J in this work. Wave function optimization is usually carried out using a correlated-sampling approach in which a set of n_opt electronic real-space configurations $\{\mathbf{R}_i\}_{i=1}^{n_\mathrm{opt}}$ distributed according to the initial wave function squared $|\Psi_\mathrm{T}(\mathbf{R};\ \boldsymbol{\alpha}_0)|^2$ is generated, and then, a target function is minimized by varying $\boldsymbol{\alpha}$ at fixed $\{\mathbf{R}_i\}$. One such target function is the "variance of the VMC energy,"[57,58]

$$\sigma_\mathrm{VMC}^2 = \frac{\langle\Psi_\mathrm{T}|(\hat{H} - E_\mathrm{VMC})^2|\Psi_\mathrm{T}\rangle}{\langle\Psi_\mathrm{T}|\Psi_\mathrm{T}\rangle}, \tag{17}$$

which reaches its lower bound of zero when the trial wave function is an eigenstate of the Hamiltonian. In practice, minimizing σ_VMC^2 yields variational energies affected by large random fluctuations, as demonstrated below. In continuum QMC methods, modifications have been devised to circumvent this problem, such as weight limiting, unweighted variance minimization, or the minimization of other measures of spread, such as the median absolute deviation from the median energy.[5] The computational cost of optimizing Jastrow factors within VMC scales as a small power of system size, typically estimated to be $\mathcal{O}(N_\mathrm{e}^3)$.

04 July 2023 12:30:16

© Author(s) 2023

1. Minimizing the variance of the reference energy

In the context of the TC method, the reference energy,

$$E_{\mathrm{ref}} = \langle D_{\mathrm{HF}}|e^{-J}\hat{H}e^{J}|D_{\mathrm{HF}}\rangle, \tag{18}$$

is of particular significance since it represents the starting point of the calculation, e.g., the energy at $\tau = 0$ of a TC-FCIQMC calculation or the zeroth-order contribution to the TC coupled cluster energy. We refer to its associated variance,

$$\sigma^2_{\mathrm{ref}} = \langle D_{\mathrm{HF}}|e^{-J}|\hat{H} - E_{\mathrm{ref}}|^2 e^{J}|D_{\mathrm{HF}}\rangle, \tag{19}$$

as the "variance of the reference energy," which can be easily evaluated for a finite VMC sample of size n_{opt} as the sample variance of the Slater–Jastrow energy over the Hartree–Fock distribution,

$$S^2_{\mathrm{ref}} = \frac{1}{n_{\mathrm{opt}} - 1}\sum_{n=1}^{n_{\mathrm{opt}}}\left|\frac{\hat{H}(\mathbf{R}_n)\Psi_{\mathrm{SJ}}(\mathbf{R}_n)}{\Psi_{\mathrm{SJ}}(\mathbf{R}_n)} - \bar{E}_{\mathrm{ref}}\right|^2, \tag{20}$$

which tends to σ^2_{ref} as $n_{\mathrm{opt}} \to \infty$, where $\Psi_{\mathrm{SJ}} = e^{J}D_{\mathrm{HF}}$ is the Slater–Jastrow wave function, $\{\mathbf{R}_n\}_{n=1}^{n_{\mathrm{opt}}}$ are electronic configurations distributed according to D^2_{HF}, and the VMC estimate of the reference energy is

$$\bar{E}_{\mathrm{ref}} = \frac{1}{n_{\mathrm{opt}}}\sum_{n=1}^{n_{\mathrm{opt}}}\frac{\hat{H}(\mathbf{R}_n)\Psi_{\mathrm{SJ}}(\mathbf{R}_n)}{\Psi_{\mathrm{SJ}}(\mathbf{R}_n)}. \tag{21}$$

It should be noted that the variance of the reference energy has been used as a target function for the optimization of Jastrow factors in earlier work on the TC method, e.g., in Refs. 9 and 14, albeit in somewhat different theoretical frameworks.

To understand the physical significance of the variance of the reference energy, note that Eq. (19) can be written in second quantized form as

$$\sigma^2_{\mathrm{ref}} = \sum_{I\neq \mathrm{HF}}\left|\langle D_I|\hat{H}_{\mathrm{TC}}|D_{\mathrm{HF}}\rangle\right|^2, \tag{22}$$

where I runs over a complete basis set. Minimizing σ^2_{ref} amounts to minimizing the coupling of the HF determinant with the remainder of the space, which in the TC-FCIQMC method reduces the spawning rate from the HF determinant to its connected excited-state determinants, increasing the amplitude of the HF determinant in the resulting Φ. In other words, if the Slater–Jastrow wave function were an exact eigenstate of $\hat{H}$, a TC-FCIQMC simulation starting from the HF determinant would immediately converge to the strictly single-determinant solution. Although this ideal scenario cannot be achieved in practice, it nevertheless illustrates the benefits of obtaining a relatively single-reference CI solution by minimizing the variance of the reference energy. We expect that this increased single-reference character will benefit other approaches as well, such as transcorrelated coupled-cluster methods.

We, therefore, investigate the performance of minimizing σ^2_{ref} as an alternative to minimizing σ^2_{VMC} tailored to the TC method. In Fig. 2, we compare the VMC energy and variance obtained by both variance minimization methods, along with energy-minimized[59–61] results for reference, for the systems considered in this work using $n_{\mathrm{opt}} = 10^5$ VMC configurations. Minimizing the variance of the VMC energy produces lower average values of σ^2_{VMC}, as one would expect, but also erratic VMC energies with very large standard deviations (up to ~50 mHa in our tests). Minimizing the variance of the reference energy, on the other hand, produces values of σ^2_{VMC}, which are only slightly higher on average than those obtained from minimizing the variance of the VMC energy (1%–5% in our tests) while producing more stable VMC energies with much smaller standard deviations (up to ~3 mHa in our tests). We, therefore, do not use "regular" variance minimization since it introduces large stochastic noise, making it unsuitable for optimizing Jastrow factors, and from this point on we use the term "variance minimization" to refer to the minimization of the variance of the reference energy.

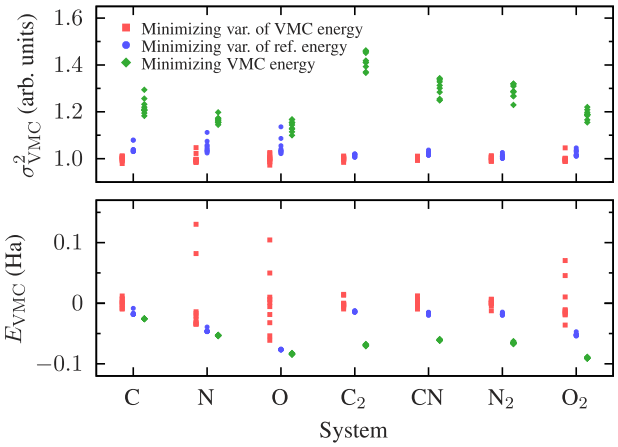

FIG. 2. Variance of the VMC energy (top) and VMC energy (bottom) of the systems considered in this work using the cc-pVTZ basis and Jastrow factors obtained by minimizing the variance of the VMC energy (red squares), the variance of the reference energy (blue circles), or the VMC energy (green diamonds) in each of ten independent optimization runs with $n_{\mathrm{opt}} = 10^5$ VMC configurations. To ease comparisons, variances have been rescaled and energies shifted by their average values from minimizing the variance of the VMC energy (i.e., the red squares average to a variance of 1 and an energy of 0 in the plot). The subpar ability of standard variance minimization to yield consistent VMC energies is evident in the bottom panel.

2. Using an adequate sample size

In continuum QMC calculations, Jastrow factors are usually optimized using relatively few VMC configurations, with n_{opt} typically in the tens of thousands. However, it has been noted that substantially larger values of n_{opt} in the hundreds of thousands are needed to converge expectation values other than the VMC energy with respect to the wave function parameters.[62]

The level of convergence of any specific expectation value at a specific value of n_{opt} can be easily assessed by performing multiple optimization runs with different random seeds and evaluating the standard deviation of the results. In practice, we find that we can use the uncertainty on the VMC estimate of the reference energy $\bar{E}_{\mathrm{ref}}$ as a proxy for the standard deviation on the TC-FCIQMC energy. This is a reasonable replacement because (i) the standard deviation of the TC-FCIQMC energy is not larger than the standard deviation of the reference energy, as shown in Fig. 3 (it is usually significantly smaller thanks to the ability of TC-FCIQMC to compensate for the presence of a bias in E_{ref} via the correlation energy), and (ii) the

04 July 2023 12:30:16

© Author(s) 2023

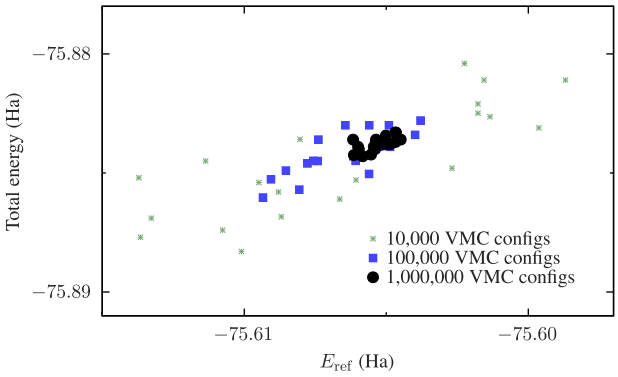

FIG. 3. TC-FCIQMC energy of the C_2 molecule obtained using $n_w = 10^6$ walkers with the cc-pVDZ basis as a function of the reference energy E_{ref} for multiple independent Jastrow factor parameter sets obtained by variance minimization using three different VMC sample sizes. The horizontal spread is about 1.8 times larger than the vertical spread, in line with the expectation that the standard deviation of the TC-FCIQMC energy is smaller than that of the reference energy.

standard deviation of the reference energy is not larger than the statistical uncertainty on the VMC estimate of the reference energy $\bar{E}_{\mathrm{ref}}$ obtained with n_{opt} configurations (it is usually significantly smaller due to the use of correlated sampling in the optimization procedure).

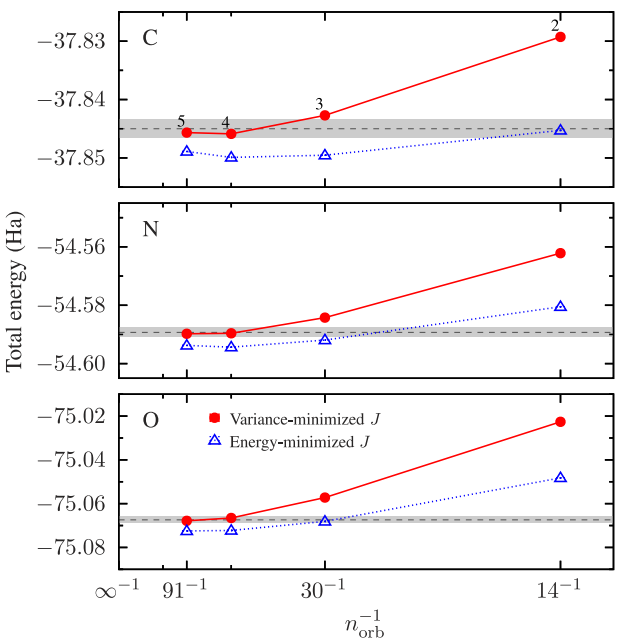

FIG. 4. Total energy of the C, N, and O atoms as a function of the reciprocal of the number of molecular orbitals in the cc-pVxZ basis set. The non-variational behavior of up to about 5 mHa is evident for the energy-minimized Jastrow factors, for which convergence to the exact energy as a function of basis-set size is rather slow. The shaded areas represent ±1 kcal/mol around the exact non-relativistic total energy from Ref. 50. The points in the top panel are annotated with the basis set cardinal number x.

We have verified that these inequalities hold for all the systems studied in this work.

For the atoms and molecules considered in the present work, we find that $n_{\mathrm{opt}} = 2 \times 10^7$ yields TC-FCIQMC energies with standard deviations of less than 0.1 mHa. While this value of n_{opt} is three orders of magnitude greater than typical values used in "regular" VMC calculations, the optimization stage typically takes tens of core-hours, representing an insignificant part of the total computational expense of TC-FCIQMC runs.

3. Energy minimization

The obvious alternative to variance minimization is minimizing the VMC energy itself,[59–61] which, as demonstrated in Fig. 2, results in lower VMC energies but higher VMC variances. Energy minimization yields wave functions, that minimize the statistical fluctuations of the local energy in DMC calculations,[63] and is typically the optimization method of choice in the context of continuum QMC methods, but for other purposes, it is unclear whether the

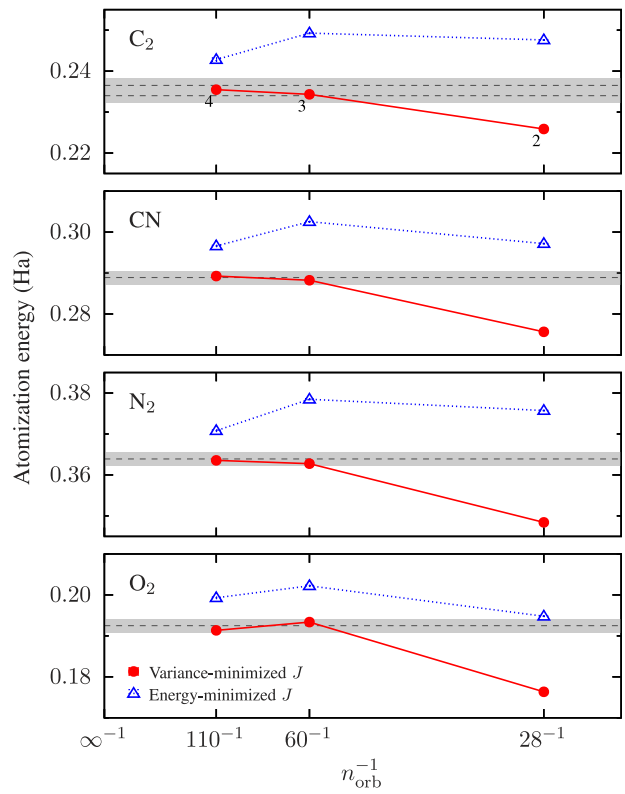

FIG. 5. Atomization energy of the C_2, CN, N_2, and O_2 molecules as a function of the reciprocal of the number of molecular orbitals using the cc-pVxZ family of basis sets and Jastrow factors obtained by variance minimization and energy minimization. The FCI limit of the atomization energies for energy-minimized Jastrow factors is estimated from 10^6-walker results assuming that the initiator error is identical to that using variance-minimized Jastrow factors at the same population. The shaded areas represent ±1 kcal/mol around the theoretical estimate of the non-relativistic atomization energy of Ref. 49; the distinct estimate of Ref. 50 is also shown for C_2. The points in the top panel are annotated with the basis set cardinal number x.

04 July 2023 12:30:16

© Author(s) 2023

resulting wave functions provide a better description of the system than those produced by variance minimization.

In Fig. 4, we compare the convergence with basis-set size of the TC-FCIQMC total energies of the C, N, and O atoms using energy- and variance-minimized Jastrow factors. Variance minimization appears to produce wave functions that converge quickly and largely variationally to the basis set limit, while energy-minimized wave functions tend to yield non-variational TC-FCIQMC energies that converge very slowly to the basis set limit.

In Fig. 5, we plot the atomization energies of the C_2, CN, N_2, and O_2 molecules as a function of reciprocal basis-set size, again demonstrating that variance-minimized Jastrow factors exhibit favorable convergence properties. Taking into account the evidence depicted in Figs. 4 and 5, we use variance minimization to optimize our Jastrow factors for subsequent use with the TC method.

C. Matrix element evaluation

Once the parameters in the Jastrow factor have been obtained, we proceed to compute the TC contributions to the two- and three-body terms of the Hamiltonian, K_{ij}^{ab} and L_{ijk}^{abc}, respectively, by numerical integration on a grid. For this task, we use Treutler–Ahlrichs integration grids,[64,65] which are atom-centered grids constructed as the combination of a radial grid running up to the Bragg radius and a Lebedev angular grid. This type of integration grid is commonly used in density functional theory calculations for which grids are usually "pruned" by reducing the number of angular grid points near the atoms since typical integrands are spherically symmetric around them, but this is not the case for our integrands, so we use "unpruned" grids.

We obtain the integration grids using PYSCF,[52] which provides an integer parameter, l_{grid}, controlling overall grid density. We test for grid errors by evaluating TC-FCIQMC energies at $l_{\text{grid}} = 0$–5 and defining the integration error as the difference of each of these results with the value obtained by linear extrapolation of the energies at $l_{\text{grid}} = 2$–5 to the $1/n_{\text{grid}} \to 0$ limit.

First, we investigate whether using basis sets of higher cardinal numbers requires finer grid densities to handle the sharper features in the corresponding orbitals. In Fig. 6, we plot the absolute integration error in the total energy of the C, N, and O atoms as a function of $1/n_{\text{grid}}$ using the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets. If larger basis sets incurred a larger integration error, one would expect the cc-pVQZ results to consistently exhibit larger integration errors in Fig. 6, but, instead, we find that the choice of basis set has little to no effect on the integration error, indicating that the grid density need not be adjusted when the basis set changes.

We now focus on the convergence of the integration error with grid point density in total energies and energy differences. In Fig. 7, we plot the absolute integration error in the total energies of the molecules, in those of the atoms they contain, and in the corresponding atomization energies using the cc-pVDZ basis. We find

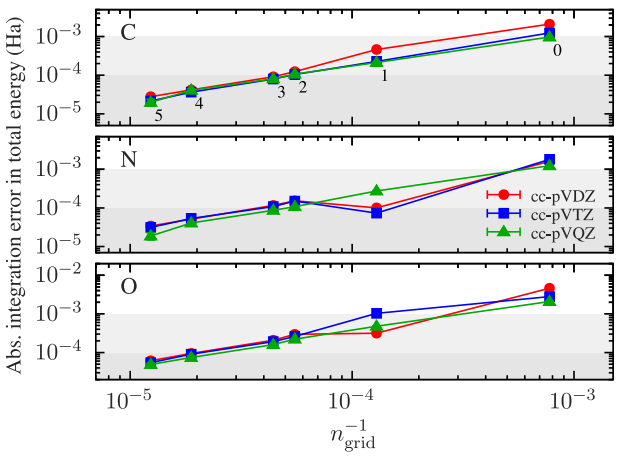

FIG. 6. Absolute integration error in the total energy of the C, N, and O atoms as a function of $1/n_{\text{grid}}$ using different basis sets in the cc-pVxZ family. The gray areas correspond to integration errors of less than 1 and 0.1 mHa. The points in the top panel are annotated with the value of PYSCF's grid density parameter l_{grid}. Notice the logarithmic scale on both axes.

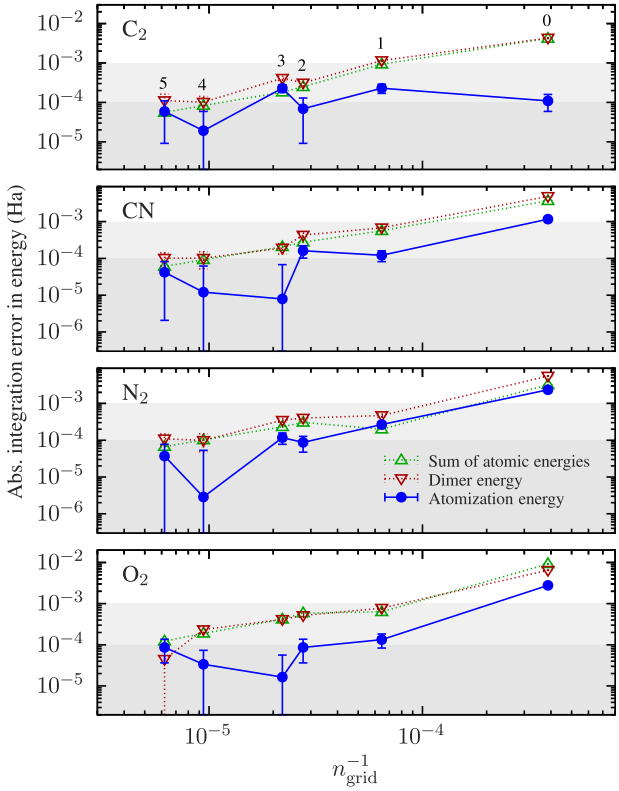

FIG. 7. Absolute integration error in the energy as a function of the reciprocal of the number of grid points used for the evaluation of the TC integrals. Results are shown for the C_2, CN, N_2, and O_2 molecules at 10^5 walkers using the cc-pVDZ basis set, along with the absolute integration error in the sum of the total energies of the atoms conforming each molecule at walker number convergence and the absolute integration error in the atomization energy of the molecule obtained as the difference between both total energies. The gray areas correspond to integration errors of less than 1 and 0.1 mHa. The points in the top panel are annotated with the value of PYSCF's grid density parameter l_{grid}. Notice the logarithmic scale on both axes. These results demonstrate that $l_{\text{grid}} = 2$ is sufficient to achieve sub-mHa accuracy in total energies and sub-0.1-mHa accuracy in relative energies.

04 July 2023 12:30:16

© Author(s) 2023

substantial integration-error cancellation in energy differences, with the atomization energy for all four molecules reaching 0.1 mHa for the $l_{\text{grid}} = 2$ grid, which is a 60-point radial grid combined with a 302-point angular grid totaling 18 120 points per atom, which we, therefore, use throughout this work. The total energies incur integration errors of less than 1 mHa for all systems considered using $l_{\text{grid}} = 2$.

D. TC-FCIQMC calculations

1. Convergence with the number of walkers

The initiator approximation[30] is a tool to prevent uncontrolled walker growth in FCIQMC calculations, allowing to stabilize the walker population around an arbitrary target value n_{w}. This approximation incurs a bias with respect to the FCI limit, which goes to zero as $n_{\text{w}} \to \infty$. Typically, the target population is increased until the FCIQMC energy appears to converge to the target precision. We find this approach to work remarkably well for the atoms considered in this work, for which we use $n_{\text{w}} = 5 \times 10^6$ and 5×10^5 in our non-TC and TC-FCIQMC calculations, respectively, but for the molecules, we find the level of convergence up to $n_{\text{w}} = 10^8$ to be unsatisfactory for our purposes.

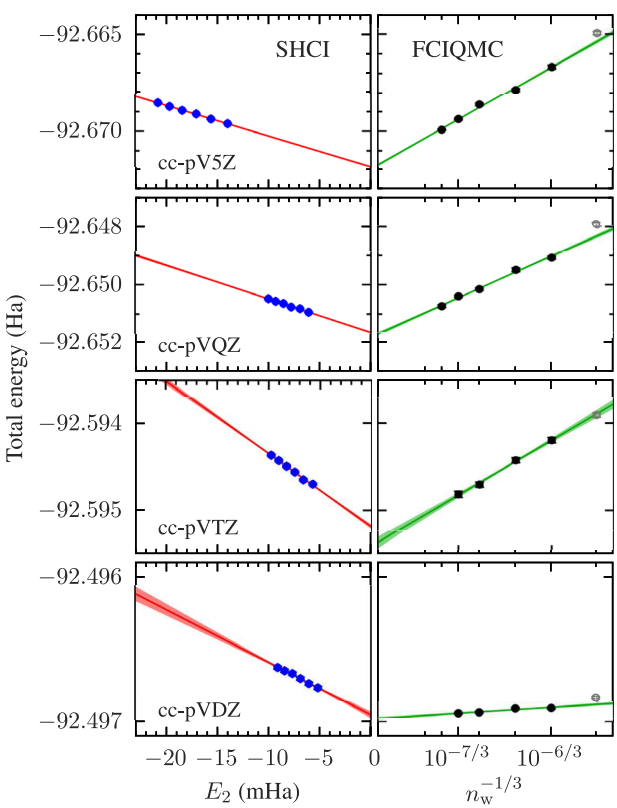

FIG. 8. Extrapolation to the FCI limit of (non-TC) SHCI energies with respect to the second-order PT correction E_2 (left) and of FCIQMC energies with respect to $n_{\text{w}}^{-1/3}$ (right) for the CN molecule. The grayed-out points are not used in the FCIQMC extrapolations. The translucent areas represent the uncertainty in the values of the fit at each point caused by the statistical uncertainty in the individual energies.

For the semi-stochastic heat-bath CI (SHCI) method, it has been shown that one can obtain accurate estimates of total energies in the FCI limit by linear extrapolation of SHCI energies to the limit in which the second-order perturbation-theory (PT) correction E_2 is zero.[66,67] In order to obtain benchmark energies to compare FCIQMC results against, we perform SHCI calculations using the DICE code[68,69] for the molecules considered in this work with variational spaces of sizes such that $|E_2|$ is not much larger than about 10 mHa;[66,67] we only exceed this threshold for the cc-pV5Z calculations due to memory constraints.

We then perform non-TC FCIQMC calculations using between 10^6 and 2×10^8 walkers, and we empirically find that extrapolating the FCIQMC energy linearly in $n_{\text{w}}^{-1/3}$ to $n_{\text{w}}^{-1/3} \to 0$ yields results in excellent agreement with the FCI limits from SHCI. We demonstrate this for the CN molecule in Fig. 8; the results for the other molecules are similarly accurate, with FCI limits obtained from FCIQMC differing by 0.4 mHa or less from their SHCI counterparts throughout. Note that in these extrapolations we do not include results at $n_{\text{w}} < 10^7$ walkers, which we deem to lie outside of the asymptotic regime.

The theoretical foundation and validity of this extrapolation technique will be studied in depth in future work. Here, we limit ourselves to empirically verifying that it yields accurate FCI limits for the systems considered, and we make the sole assumption that the technique continues to produce equally accurate results for these molecules when the TC Hamiltonian is used. We expect the extrapolation of TC-FCIQMC energies to require smaller values of n_{w} than in the non-TC case due to the increased compactness of the CI wave function with the TC Hamiltonian.

The more compact the CI wave function is the easier it is, for FCIQMC to sample the wave function accurately, and the smaller the initiator error becomes at a fixed population. The TC method has been shown to make CI wave functions more compact for the two-dimensional Hubbard model.[33] Let $\{c_I\}$ be the L^2-normalized coefficients of the CI wave function such that $\sum_I c_I^2 = 1$. The quantity

$$\xi = \frac{c_{\text{HF}}^{(\text{TC})} - c_{\text{HF}}^{(\text{non-TC})}}{1 - c_{\text{HF}}^{(\text{non-TC})}} \tag{23}$$

is then a measure of the enhancement in the compactness of the wave function, going from 0 for no enhancement to 1 if the TC wave function becomes exactly single-determinantal. From the data in Ref. 33, the TC method yields a maximum $\xi = 0.64$ for the 18-site two-dimensional Hubbard model. We find that the values of ξ for atomic and molecular systems are not dissimilar from this; see Table II.

Having established the more compact nature of the CI wave function with the TC Hamiltonian, we deem it appropriate to run the

TABLE II. Enhancement of the compactness of the CI wave function, ξ in Eq. (23), between our non-TC and TC FCIQMC calculations.

	C	N	O	C_2	CN	N_2	O_2
cc-pVDZ	0.46	0.63	0.71	0.14	0.23	0.38	0.53
cc-pVTZ	0.45	0.61	0.69	0.15	0.24	0.40	0.55
cc-pVQZ	0.44	0.60	0.69	0.15	0.24	0.41	0.57

04 July 2023 12:30:16

© Author(s) 2023

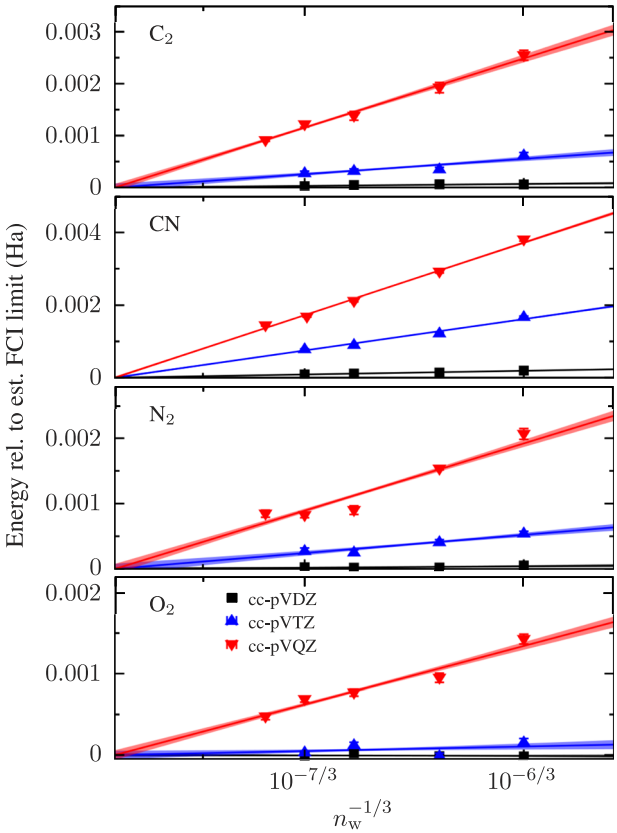

FIG. 9. Estimated initiator error in the TC-FCIQMC energies as a function of $n_w^{-1/3}$ for the C_2, CN, N_2, and O_2 molecules with the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets. The translucent areas represent the uncertainty in the values of the fit at each point propagated from the statistical uncertainty in the individual energies.

TC-FCIQMC calculations using between 10^6 and 10^7 walkers for the cc-pVDZ and cc-pVTZ bases and between 10^6 and 2×10^7 walkers with the cc-pVQZ basis. In Fig. 9, we plot the resulting TC-FCIQMC energy relative to its corresponding FCI limit, i.e., the initiator error, as a function of $n_w^{-1/3}$ for each of the molecules and basis sets considered. As anticipated, the initiator error is, to a good approximation, proportional to $n_w^{-1/3}$, and we expect the extrapolated FCI limits to incur errors no larger than 0.4 mHa as in the non-TC case.

2. Neglecting three-body excitations

Sampling in FCIQMC calculations involves spawning walkers from occupied determinants onto determinants connected to them by the Hamiltonian. The $\hat{L}$ matrix not only contributes to two-body excitations via elements of the form L_{ijk}^{ibc} but also connects determinants, which differ by a pure three-body excitation via L_{ijk}^{abc}, which represents a huge increase in the connectivity of the Hilbert space with respect to the non-TC problem. However, the matrix elements L_{ijk}^{abc} are usually very small in magnitude, so neglecting pure three-body excitations is a valid approximation to consider.[70]

This approximation effectively implies that L_{ijk}^{abc} can be regarded as zero when all indices are distinct, except when three or more indices correspond to spatial orbitals, which are occupied in the HF determinant, since these matrix elements are required during the evaluation of the projected energy. Therefore, neglecting pure three-body excitations reduces the amount of storage needed to hold $\hat{L}$ from $\mathcal{O}(n_{\text{orb}}^6)$ to $\mathcal{O}(n_{\text{orb}}^5) + \mathcal{O}(N_e^3 n_{\text{orb}}^3)$; we report the specific reduction factors obtained for the molecules considered in this work in Table III.

TABLE III. $\hat{L}$ matrix storage reduction factor from neglecting pure three-body excitations, computed as the number of non-zero matrix elements in the full $\hat{L}$ matrix divided by the number of non-zero matrix elements with repeated indices or three or more indices corresponding to orbitals occupied in the HF determinant.

	C	N	O	C_2	CN	N_2	O_2
cc-pVDZ	1.23	1.17	1.17	1.87	1.78	1.78	1.58
cc-pVTZ	2.04	2.02	1.93	3.72	3.66	3.66	3.46
cc-pVQZ	3.31	3.44	3.13	6.60	6.54	6.57	6.41

Two-body excitations are, in practice, more expensive to attempt than triple excitations, so neglecting the latter actually *increases* the cost per step of the calculation. However, neglecting pure three-body excitations allows the TC-FCIQMC time step to be larger, resulting in reduced serial correlation in the statistics, which enables reaching the target accuracy in fewer steps, and one can generally expect a net cost reduction from this approximation. We report the specific reduction factors found for the molecules considered in this work in Table IV.

In Table V, we show the error in the atomization energy of the molecules incurred by neglecting pure three-body excitations. We find that this approximation results in errors of the order of ~0.3 mHa at the cc-pVTZ level, which is a relatively small bias considering the substantial storage and cost benefits of the approximation. Note that we do not use this approximation in the main results presented in Sec. IV.

TABLE IV. Reduction factor in the walltime required to advance one unit of imaginary time at fixed population from neglecting pure three-body excitations in the TC-FCIQMC calculation.

	C	N	O	C_2	CN	N_2	O_2
cc-pVDZ	0.9	1.0	1.1	1.7	1.2	1.5	1.6
cc-pVTZ	1.0	1.0	0.8	2.4	1.0	1.8	2.0
cc-pVQZ	1.5	1.5	1.0	3.1	0.9	1.9	2.3

TABLE V. Error in the atomization energy of the molecules considered in this work incurred by neglecting pure three-body excitations from the FCIQMC dynamics, in mHa.

	C_2	CN	N_2	O_2
cc-pVDZ	−0.62(2)	−0.46(0)	−0.56(2)	−0.55(2)
cc-pVTZ	−0.36(5)	−0.30(2)	−0.32(5)	−0.20(3)
cc-pVQZ	−0.45(6)	−0.21(2)	−0.32(7)	−0.27(5)

04 July 2023 12:30:16

© Author(s) 2023

TABLE VI. Total energies in Ha obtained for the atoms and molecules considered in this work, along with benchmark non-relativistic results. The statistical uncertainties arising from Monte Carlo sampling are smaller than 0.0001 Ha in all cases.

		C	N	O	C_2	CN	N_2	O_2
Non-TC	cc-pVDZ	−37.7619	−54.4801	−74.9117	−75.7320	−92.4970	−109.2809	−149.9915
	cc-pVTZ	−37.7900	−54.5252	−74.9853	−75.8094	−92.5954	−109.4014	−150.1554
	cc-pVQZ	−37.8126	−54.5535	−75.0236	−75.8578	−92.6517	−109.4653	−150.2362
	cc-pV5Z	−37.8199	−54.5627	−75.0369	−75.8752	−92.6717	−109.4881	−150.2655
	cc-pV6Z	−37.8263	−54.5697	−75.0447				
TC	cc-pVDZ	−37.8293	−54.5622	−75.0226	−75.8844	−92.6671	−109.4727	−150.2216
	cc-pVTZ	−37.8427	−54.5842	−75.0572	−75.9197	−92.7152	−109.5312	−150.3078
	cc-pVQZ	−37.8459	−54.5896	−75.0665	−75.9272	−92.7247	−109.5428	−150.3244
	cc-pV5Z	−37.8457	−54.5898	−75.0678				
Reference 49					−75.9240	−92.7232	−109.5425	−150.3273
Reference 50		−37.8450	−54.5893	−75.0674	−75.9265		−109.5427	−150.3274
Reference 51		−37.8450	−54.5893	−75.0674		−92.7229	−109.5425	−150.3275

IV. RESULTS AND DISCUSSION

We now analyze the accuracy of the TC method with tailored Jastrow factors by comparing the convergence of the results as a function of basis set size with non-TC results and with benchmark complete basis-set limit (CBS) values from Refs. 49–51.

In Table VI, we list the total energies that we obtain for each system and basis set. We find the TC total energies to be remarkably accurate already at the cc-pVQZ the basis-set level, differing by less than 2 mHa per atom from benchmark CBS values, while the non-TC total energies still miss the benchmarks by 25–30 mHa per atom with the cc-pV5Z basis set and 20 mHa per atom at the cc-pV6Z level. The TC total energies exhibit slight non-variational convergence, with the atomic energies reaching values 0.5 mHa below the benchmark before increasing again toward it for larger basis-set sizes. While non-variationality is not a desirable feature in a method, the amount by which the TC results dip below the basis-set limit is sufficiently tiny for this issue to be entirely ignored in practice.

From a chemical perspective, relative energies are more important than total energies. The atomization energies of the C_2, CN, N_2, and O_2 molecules obtained from the total energies in Table VI are given in Table VII and plotted in Fig. 10. Again, we find that the TC results exhibit much faster convergence with basis-set size than their non-TC counterparts, with TC atomization energies being chemically accurate with respect to the CBS benchmarks already at the

TABLE VII. Atomization energies in mHa obtained for the molecules considered in this work, along with benchmark non-relativistic results. The statistical uncertainties arising from Monte Carlo sampling are smaller than 0.1 mHa in all cases.

		C_2	CN	N_2	O_2
Non-TC	cc-pVDZ	208.2	255.0	320.7	168.0
	cc-pVTZ	229.4	280.1	351.0	184.9
	cc-pVQZ	232.6	285.6	358.4	189.0
	cc-pV5Z	235.5	289.2	362.6	191.7
TC	cc-pVDZ	225.9	275.6	348.4	176.4
	cc-pVTZ	234.3	288.2	362.7	193.4
	cc-pVQZ	235.5	289.3	363.6	191.4
Reference 49		234.0	288.9	363.9	192.5
Reference 50		236.5		364.1	192.6
Reference 51			288.6	363.9	192.7

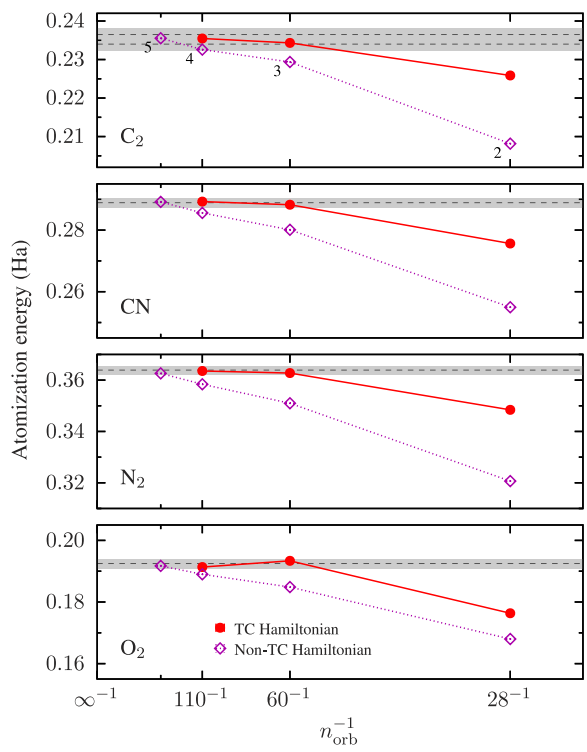

FIG. 10. Atomization energy of the C_2, CN, N_2, and O_2 molecules obtained with FCIQMC and TC-FCIQMC as a function of the reciprocal of the number of molecular orbitals using the cc-pV*x*Z family of basis sets. The points in the top panel are annotated with the basis set cardinal number *x*. The shaded areas represent ±1 kcal/mol around the theoretical estimate of the non-relativistic atomization energy of Ref. 49; the distinct estimate of Ref. 50 is also shown for C_2.

04 July 2023 12:30:16

© Author(s) 2023

cc-pVTZ basis-set level, matching the accuracy of non-TC atomization energies obtained using the cc-pV5Z basis set. The TC method, therefore, provides an advantage of about two cardinal numbers for the computation of relative energies.

The application of the TC method to quantum chemical methods, in general, could be presumed to be problematic because any theoretical guarantee of cancellation of errors in energy differences disappears with the introduction of separately optimized Jastrow factors for each system. However, the fact that in our results the relative energy converges at smaller basis-set sizes than the total energy implies that substantial error cancellation is at play in practice. We take this as evidence that the TC method with tailored Jastrow factors does not fully suppress advantageous error cancellation from the underlying methodology.

It is important to note that neural-network based trial wave functions proposed in recent years for use with VMC and DMC,[71,72] while promising, do not achieve chemical accuracy reliably. For example, an atomization energy of N_2 of 361.2(2) mHa is obtained from the results provided by Ref. 71, which is 2.7(2) mHa away from the benchmark; the slight difference in the bond length used for that calculation only accounts for 0.1 mHa of the energy difference. By contrast, our TC-FCIQMC atomization energy obtained with the cc-pVTZ basis is only 1.2 mHa away from the benchmark.

V. CONCLUSIONS

We present a method to optimize flexible Jastrow factors of a form commonly used in continuum quantum Monte Carlo methods for application in the TC method, which we have tested within TC-FCIQMC. Minimizing the variance of the reference energy is shown to be an especially good fit for the TC method since it maximizes the single-reference character of the CI wave function, and we have demonstrated how this method outperforms standard energy minimization in this context. Various approximate aspects of the calculations have been considered, and care has been taken to ensure that we can produce relative energies within significantly less than 1 mHa of their FCI limit.

Our results show that the TC method with tailored Jastrow factors delivers remarkably accurate total energies and gives relative energies with a cc-pVTZ basis set, which rivals the accuracy of non-TC relative energies with the much larger cc-pV5Z basis set. We expect that future work in this topic, which will likely include technical enhancements, efficient approximations for dealing with the three-body integrals, deterministic optimization methods, the use of polyatomic and spin-dependent Jastrow factor terms, and methodological adaptations for strongly correlated systems, will enable further reductions in the basis-set sizes required to perform accurate quantum chemistry calculations with the TC method.

ACKNOWLEDGMENTS

The authors would like to thank H. Luo, D. Kats, K. Liao, and E. Christlmaier for useful discussions and K. Guther and K. Ghanem for their early work on the TCHINT library. P.L.R. and A.A. acknowledge the support from the European Centre of Excellence in Exascale Computing TREX, funded by the Horizon 2020 program of the European Union under Grant No. 952165. W.D. acknowledges the funding from the Horizon Europe research and innovation program of the European Union under the Marie Skłodowska–Curie Grant Agreement No. 101062864. Any views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

J. Philip Haupt: Investigation (equal); Software (equal); Writing – review & editing (equal). **Seyed Mohammadreza Hosseini**: Investigation (equal). **Pablo López Ríos**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Werner Dobrautz**: Software (equal); Writing – review & editing (equal). **Aron Cohen**: Software (equal). **Ali Alavi**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); 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] T. Kato, "On the eigenfunctions of many-particle systems in quantum mechanics," Commun. Pure Appl. Math. **10**, 151 (1957).
[2] R. T. Pack and W. B. Brown, "Cusp conditions for molecular wavefunctions," J. Chem. Phys. **45**, 556 (1966).
[3] E. A. Hylleraas, "Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium," Z. Phys. **54**, 347 (1929).
[4] R. Jastrow, "Many-body problem with strong forces," Phys. Rev. **98**, 1479 (1955).
[5] 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).
[6] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Rev. Mod. Phys. **73**, 33 (2001).
[7] S. F. Boys and N. C. Handy, "The determination of energies and wavefunctions with full electronic correlation," Proc. R. Soc. A **310**, 43 (1969).
[8] N. C. Handy, "Energies and expectation values for Be by the transcorrelated method," J. Chem. Phys. **51**, 3205 (1969).
[9] N. C. Handy, "On the minimization of the variance of the transcorrelated Hamiltonian," Mol. Phys. **21**, 817 (1971).
[10] M. Nooijen and R. J. Bartlett, "Elimination of Coulombic infinities through transformation of the Hamiltonian," J. Chem. Phys. **109**, 8232 (1998).
[11] S. Ten-no, "A feasible transcorrelated method for treating electronic cusps using a frozen Gaussian geminal," Chem. Phys. Lett. **330**, 169 (2000).
[12] S. Ten-no, "Three-electron integral evaluation in the transcorrelated method using a frozen Gaussian geminal," Chem. Phys. Lett. **330**, 175 (2000).

04 July 2023 12:30:16

© Author(s) 2023

[13]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 (2002).
[14]N. Umezawa and S. Tsuneyuki, "Transcorrelated method for electronic systems coupled with variational Monte Carlo calculation," J. Chem. Phys. **119**, 10015 (2003).
[15]N. Umezawa and S. Tsuneyuki, "Ground-state correlation energy for the homogeneous electron gas calculated by the transcorrelated method," Phys. Rev. B **69**, 165102 (2004).
[16]R. Sakuma and S. Tsuneyuki, "Electronic structure calculations of solids with a similarity-transformed Hamiltonian," J. Phys. Soc. Jpn. **75**, 103705 (2006).
[17]S. Tsuneyuki, "Transcorrelated method: Another possible way towards electronic structure calculation of solids," Prog. Theor. Phys. Suppl. **176**, 134 (2008).
[18]M. Ochi and S. Tsuneyuki, "Optimization of the Jastrow factor in the correlated wave function of electrons using the first-principles transcorrelated method for solid-state calculations," J. Phys. Conf. Ser. **454**, 012020 (2013).
[19]W. Klopper and W. Kutzelnigg, "Møller-Plesset calculations taking care of the correlation CUSP," Chem. Phys. Lett. **134**, 17 (1987).
[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 (1991).
[21]W. Klopper, "Orbital-invariant formulation of the MP2-R12 method," Chem. Phys. Lett. **186**, 583 (1991).
[22]J. Noga, W. Kutzelnigg, and W. Klopper, "CC-R12, a correlation cusp corrected coupled-cluster method with a pilot application to the Be_2 potential curve," Chem. Phys. Lett. **199**, 497 (1992).
[23]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 (1994).
[24]W. Klopper and C. C. M. Samson, "Explicitly correlated second-order Møller-Plesset methods with auxiliary basis sets," J. Chem. Phys. **116**, 6397 (2002).
[25]E. F. Valeev, "Improving on the resolution of the identity in linear R12 ab initio theories," Chem. Phys. Lett. **395**, 190 (2004).
[26]S. Ten-no, "Initiation of explicitly correlated Slater-type geminal theory," Chem. Phys. Lett. **398**, 56 (2004).
[27]L. Kong, F. A. Bischoff, and E. F. Valeev, "Explicitly correlated R12/F12 methods for electronic structure," Chem. Rev. **112**, 75 (2012).
[28]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 (2018).
[29]G. H. Booth, A. J. W. Thom, and A. Alavi, "Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in slater determinant space," J. Chem. Phys. **131**, 054106 (2009).
[30]D. Cleland, G. H. Booth, and A. Alavi, "Survival of the fittest: Accelerating convergence in full configuration interaction quantum Monte Carlo," J. Chem. Phys. **132**, 041103 (2010).
[31]K. Guther, R. J. Anderson, N. S. Blunt, N. A. Bogdanov, D. Cleland, N. Dattani, W. Dobrautz, K. Ghanem, P. Jeszenszki, N. Liebermann, G. L. Manni, A. Y. Lozovoi, H. Luo, D. Ma, F. Merz, C. Overy, M. Rampp, P. K. Samanta, L. R. Schwarz, J. J. Shepherd, S. D. Smart, E. Vitale, O. Weser, G. H. Booth, and A. Alavi, "NECI: *N*-electron configuration interaction with an emphasis on state-of-the-art stochastic methods," J. Chem. Phys. **153**, 034107 (2020).
[32]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).
[33]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).
[34]T. Schraivogel, A. J. Cohen, A. Alavi, and D. Kats, "Transcorrelated coupled cluster methods," J. Chem. Phys. **155**, 191101 (2021).
[35]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).
[36]T. Schraivogel, E. M. Christlmaier, P. López Ríos, A. Alavi, and D. Kats, "Transcorrelated coupled cluster methods. II. Molecular systems," J. Chem. Phys. **158**, 214106 (2023).
[37]E. M. 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. (in press).
[38]S. F. Boys and N. C. Handy, "A calculation for the energies and wavefunctions for states of neon with full electronic correlation accuracy," Proc. R. Soc. London A **310**, 63 (1969).
[39]K. E. Schmidt and J. W. Moskowitz, "Correlated Monte Carlo wave functions for the atoms He through Ne," J. Chem. Phys. **93**, 4172 (1990).
[40]K. Guther, A. J. Cohen, H. Luo, and A. Alavi, "Binding curve of the beryllium dimer using similarity-transformed FCIQMC: Spectroscopic accuracy with triple-zeta basis sets," J. Chem. Phys. **155**, 011102 (2021).
[41]P. Jeszenszki, H. Luo, A. Alavi, and J. Brand, "Accelerating the convergence of exact diagonalization with the transcorrelated method: Quantum gas in one dimension with contact interactions," Phys. Rev. A **98**, 053627 (2018).
[42]P. Jeszenszki, U. Ebling, H. Luo, A. Alavi, and J. Brand, "Eliminating the wavefunction singularity for ultracold atoms by a similarity transformation," Phys. Rev. Res. **2**, 043270 (2020).
[43]K. Liao, H. Zhai, E. M. 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 (2023).
[44]A. Baiardi and M. Reiher, "Transcorrelated density matrix renormalization group," J. Chem. Phys. **153**, 164115 (2020).
[45]A. Ammar, A. Scemama, and E. Giner, "Extension of selected configuration interaction for transcorrelated methods," J. Chem. Phys. **157**, 134107 (2022).
[46]I. O. Sokolov, W. Dobrautz, H. Luo, A. Alavi, I. Tavernelli, "Orders of magnitude reduction in the computational overhead for quantum many-body problems on quantum computers via an exact transcorrelated method," arXiv:2201.03049 [quant-ph] (2022).
[47]W. Dobrautz, I. O. Sokolov, K. Liao, P. López Ríos, M. Rahm, A. Alavi, and I. Tavernelli, "Ab initio transcorrelated method enabling accurate quantum chemistry on noisy near-term quantum hardware," arXiv:2303.02007 [quant-ph] (2023).
[48]T. H. Dunning, Jr., "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007 (1989).
[49]D. Feller, K. A. Peterson, and D. A. Dixon, "A survey of factors contributing to accurate theoretical predictions of atomization energies and molecular structures," J. Chem. Phys. **129**, 204105 (2008).
[50]L. Bytautas and K. Ruedenberg, "Correlation energy extrapolation by intrinsic scaling. IV. Accurate binding energies of the homonuclear diatomic molecules carbon nitrogen, oxygen, and fluorine," J. Chem. Phys. **122**, 154110 (2005).
[51]M. E. Harding, J. Vázquez, B. Ruscic, A. K. Wilson, J. Gauss, and J. F. Stanton, "High-accuracy extrapolated *ab initio* thermochemistry. III. Additional improvements and overview," J. Chem. Phys. **128**, 114111 (2008).
[52]Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, "Recent developments in the P_{Y}SCF program package," J. Chem. Phys. **153**, 024109 (2020).
[53]Transcorrelated Hamiltonian integral library tchint to be released; available from the authors upon reasonable request.
[54]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).
[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]A. Ma, M. D. Towler, N. D. Drummond, and R. J. Needs, "Scheme for adding electron-nucleus cusps to Gaussian orbitals," J. Chem. Phys. **122**, 224322 (2005).

04 July 2023 12:30:16

© Author(s) 2023

[57] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, "Optimized trial wave functions for quantum Monte Carlo calculations," Phys. Rev. Lett. **60**, 1719 (1988).
[58] 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 (1999).
[59] 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).
[60] J. Toulouse and C. J. Umrigar, "Optimization of quantum Monte Carlo wave functions by energy minimization," J. Chem. Phys. **126**, 084102 (2007).
[61] 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).
[62] G. G. Spink, P. López Ríos, N. D. Drummond, and R. J. Needs, "Trion formation in a two-dimensional hole-doped electron gas," Phys. Rev. B **94**, 041410(R) (2016).
[63] D. M. Ceperley, "The statistical error of Green's function Monte Carlo," J. Stat. Phys. **43**, 815 (1986).
[64] A. D. Becke, "A multicenter numerical integration scheme for polyatomic molecules," J. Chem. Phys. **88**, 2547 (1988).
[65] O. Treutler and R. Ahlrichs, "Efficient molecular numerical integration schemes," J. Chem. Phys. **102**, 346 (1995).
[66] J. E. T. Smith, B. Mussard, A. A. Holmes, and S. Sharma, "Cheap and near exact CASSCF with large active spaces," J. Chem. Theory Comput. **13**, 5468 (2017).
[67] A. A. Holmes, C. J. Umrigar, and S. Sharma, "Excited states using semistochastic heat-bath configuration interaction," J. Chem. Phys. **147**, 164111 (2017).
[68] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, "Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory," J. Chem. Theory Comput. **13**, 1595 (2017).
[69] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, "Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling," J. Chem. Theory Comput. **12**, 3674 (2016).
[70] W. Dobrautz, A. J. Cohen, A. Alavi, and E. Giner, "Performance of a one-parameter correlation factor for transcorrelation: Study on a series of second row atomic and molecular systems," J. Chem. Phys. **156**, 234108 (2022).
[71] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, "Ab initio solution of the many-electron Schrödinger equation with deep neural networks," Phys. Rev. Res. **2**, 033429 (2020).
[72] J. Hermann, Z. Schätzle, and F. Noé, "Deep-neural-network solution of the electronic Schrödinger equation," Nat. Chem. **12**, 891 (2020).

04 July 2023 12:30:16

© Author(s) 2023