

Spin Purification in Full-CI Quantum Monte Carlo via a First-Order Penalty Approach

Oskar Weser,* Niklas Liebermann, Daniel Kats, Ali Alavi, and Giovanni Li Manni*

Cite This: *J. Phys. Chem. A* 2022, 126, 2050–2060

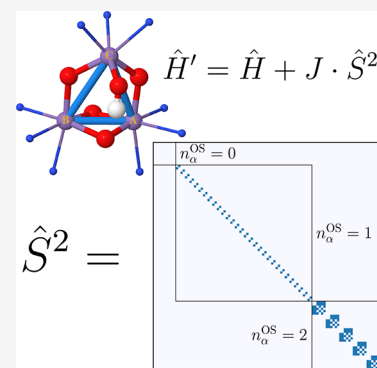
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: In this article, we demonstrate that a first-order spin penalty scheme can be efficiently applied to the Slater determinant based Full-CI Quantum Monte Carlo (FCIQMC) algorithm, as a practical route toward spin purification. Two crucial applications are presented to demonstrate the validity and robustness of this scheme: the ${}^1\Delta_g \leftarrow {}^3\Sigma_g$ vertical excitation in O_2 and key spin gaps in a $[Mn_3^{(IV)}O_4]$ cluster. In the absence of a robust spin adaptation/purification technique, both applications would be unattainable by Slater determinant based ground state methods, with any starting wave function collapsing into the higher-spin ground state during the optimization. This strategy can be coupled to other algorithms that use the Slater determinant based FCIQMC algorithm as configuration interaction eigensolver, including the Stochastic Generalized Active Space, the similarity-transformed FCIQMC, the tailored-CC, and second-order perturbation theory approaches. Moreover, in contrast to the GUGA-FCIQMC technique, this strategy features both spin projection and total spin adaptation, making it appealing when solving anisotropic Hamiltonians. It also provides spin-resolved reduced density matrices, important for the investigation of spin-dependent properties in polynuclear transition metal clusters, such as the hyperfine-coupling constants.



1. INTRODUCTION

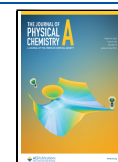
Strongly open-shell molecules present a number of challenges to quantum chemical methods, arising from the large number of nearly degenerate states with different total spin quantum number, S , which exist in such systems and are in general hard to resolve. In these systems, spin contamination is a major problem for an accurate description of their electronic spectrum. Such systems usually exhibit a strong *multireference* character, with numerous dominant electronic configurations featuring similar weights in the configuration interaction (CI) expansion. Furthermore, when a high-spin state is the ground-state, states of the same symmetry but with lower spin are impossible to obtain with ground state projective techniques. For these reasons, there has been much interest in recent years in developing spin-adapted approaches, which work in Hilbert spaces of *configuration state functions* (CSFs), rather than Slater determinants (SDs).^{1–16} In these approaches, \hat{S}^2 symmetry is explicitly enforced, ensuring zero spin contamination, and enabling the targeting of any desired spin state. The Graphical Unitary group approach (GUGA)^{17–26} is one such example of a fully spin-adapted approach, which was implemented within the stochastic full-CI quantum Monte Carlo (FCIQMC)^{9,27–30} and the Stochastic-CASSCF^{10,31,32} frameworks. Recently, we have discovered a strategy within GUGA that allows an unprecedented reduction of the multireference character (compression)^{32–35} of ground- and excited-state wave functions and the unique possibility to perform state-specific optimizations of ground- and excited-state wave functions.^{34,35}

These properties arise from a unique block-diagonal structure of the GUGA Hamiltonian matrix, even within the same spin-symmetry sector, that follows chemically/physically motivated molecular orbital transformations.³⁴ This strategy has been applied to exchange-coupled polynuclear transition metal clusters with a large number of localized open-shell orbitals^{32–35} and to one-dimensional Heisenberg and Hubbard model Hamiltonians.³⁶ In the latter cases, a connection with the concept of *alternancy symmetry* can be envisioned.^{37,38} Other sparse FCI solvers^{2,39–50} could also benefit from the enhanced sparsity of the Hamiltonian and wave functions that follow the above-mentioned strategy.

However, such sophisticated approaches to spin adaptation incur a number of complications related to their increased algorithmic complexity, including matrix element calculation and excitation generation process.^{9,32} In addition it is possible^{51–53} but complicated to describe spin projection properties in a spin adapted basis, which is, e. g., necessary for anisotropic Hamiltonians or the calculation of spin polarization. Furthermore, in systems with a more delocalized

Received: February 24, 2022

Published: March 17, 2022



character (i. e., covalency), the aforementioned compression advantages of the GUGA method are less evident. For these reasons, it is highly desirable to have a Slater determinant based approach to spin adaptation, which is possible via *spin-purification* concepts.

For cases with an even number of unpaired electrons, such as the oxygen dimer discussed later, it is also possible to place constraints on the spin by applying time-reversal symmetry and by working with pairs of spin-coupled functions.⁵⁴ This reduces the size of the Hilbert space by a factor of 2, while reducing any spin contamination, as in the reduced space either all even or odd spin states can be excluded. However, this strategy cannot separate singlet from quintet, nor can it operate in cases with an odd number of unpaired electrons.

The aim of the present article is to introduce one such method, based on a simple first-order spin-penalty approach, within the context of Slater determinant based FCIQMC. Spin purification techniques, including the first-order spin penalty approach, have recently been discussed in details by Levine and co-workers⁵⁵ and have already been utilized in the context of renormalization approaches.^{56,57} We build on the existing literature by explaining the origin for a range of optimal spin penalty parameters.

The remainder of the article is organized as follows. In the next section, we explain the theory behind the first-order spin-penalty method and explain the origin of the optimal spin penalty parameter. Section 3 contains two applications, the oxygen dimer and a trinuclear manganese cluster, to showcase the general applicability of this method in FCIQMC. Section 4 contains a summary and conclusion of our results. The Appendix (Section 5) contains detailed derivations of the convergence speed for different choices of the spin penalty parameter.

2. THEORY

2.1. First-Order Spin Penalty Method. We write the total spin operator, \hat{S}_z^2 , in terms of the spin projection, \hat{S}_z , and the ladder operators, \hat{S}_+ (raising) and \hat{S}_- (lowering) (see also ref 58), namely

$$\hat{S}_z^2 = \hat{S}_z(\hat{S}_z - 1) + \hat{S}_+\hat{S}_-. \quad (1)$$

Given two SDs, $|D_i\rangle$ and $|D_j\rangle$, the expression for $\langle D_i|\hat{S}_z^2|D_j\rangle$ is then given by

$$\langle D_i|\hat{S}_z^2|D_j\rangle = \begin{cases} \hat{S}_z(\hat{S}_z - 1) + n_\alpha^{\text{OS}} & |D_i\rangle = |D_j\rangle \text{ (diagonal elements),} \\ \text{sgn}(D_i, D_j) & |D_j\rangle = \hat{a}_{p\alpha}^\dagger \hat{a}_{q\beta} \hat{a}_{q\beta}^\dagger \hat{a}_{p\alpha} |D_i\rangle, p \neq q, \\ 0 & \text{else,} \end{cases} \quad (2)$$

where n_α^{OS} is the number of unpaired (open-shell, OS) α electrons. The off-diagonal elements are nonzero only for exchange excitations and are equal to $\text{sgn}(D_i, D_j) = \pm 1$, where the sign is given by the product of Fermionic phase factors.⁵⁸ Since exchange excitations require the same orbital configuration in both determinants, the \hat{S}_z^2 matrix features an interesting block-diagonal structure. Larger blocks are characterized by a common n_α^{OS} value, while the sub-blocks are characterized by a common occupation number vector. This block-diagonal and sparse structure (see Figure 1) is particularly well suited for FCIQMC.

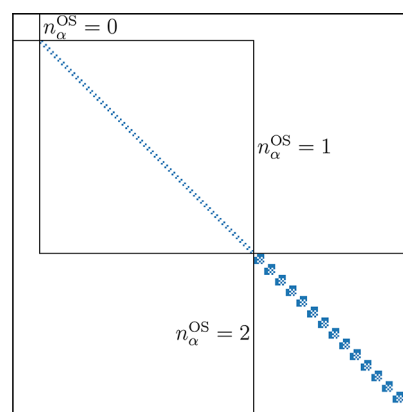


Figure 1. Block-diagonal structure of the \hat{S}_z^2 matrix in the SD basis of an (8,6) active space with $S_z = 0$, corresponding to a minimum active space for the singlet state of oxygen. White denotes zero, while blue denotes ± 1 entries. The diagonal is omitted.

In the first-order spin penalty approach, a modified Hamiltonian

$$\hat{H}' = \hat{H} + J \cdot \hat{S}^2, \quad 0 < J \in \mathbb{R} \quad (3)$$

is utilized. If J is chosen such that the low-spin state becomes the lowest state in the modified \hat{H}' Hamiltonian, ground state methods, including FCIQMC, will converge to that state. The on-the-fly evaluation of modified Hamiltonian matrix elements does not require additional memory, and has negligible runtime costs for the evaluation of the \hat{S}_z^2 correction. Since \hat{H} and \hat{S}^2 commute, the eigenstates of \hat{H}' are still eigenstates of \hat{H} and the eigenvalues of \hat{H} can be directly calculated from the corresponding eigenvalues of \hat{H}' by subtracting $J \cdot S(S + 1)$. Note that this subtraction can be performed in a well-defined manner only for converged eigensolutions. For unconverged, intermediate results, for example along FCIQMC dynamics and before stationary conditions are reached, it is necessary to evaluate directly the original Hamiltonian \hat{H} . In the present work we calculate the latter as an expectation value from the stochastically sampled one- and two-body reduced density matrices (RDMs).

2.2. Range of Optimal J Values. In the following, we discuss an optimal choice of the J value. For the unique J that makes all high-spin states energetically above or degenerate to the targeted spin state (*first flipping point*), a spin-symmetry-broken wave function is to be expected, which is an arbitrary admixture of the degenerate spin states in the modified Hamiltonian. This *flipping point* satisfies the following relation

$$J_{f,1} = \max_{S>S_{ls}} \frac{E_{ls,0} - E_{S,0}}{\Delta \langle \hat{S}_z^2 \rangle}, \quad (4)$$

where $E_{ls,0}$ and $E_{S,0}$ are the nonpenalized ground-state energies of the targeted low-spin (ls) and the higher spin (S) state and $\Delta \langle \hat{S}_z^2 \rangle = S(S + 1) - S_{ls}(S_{ls} + 1)$ corresponds to the difference in their spin expectation values. For J values larger than the first flipping point the desired low-spin state is obtained in the long-time limit of the FCIQMC dynamics. However, the speed of convergence and stability of the imaginary-time propagation in FCIQMC depends on how far J is from the first flipping point. At first glance, one could expect that, above the first flipping point, higher J values only improve the speed of convergence, because they increase the energy of the high-spin states, so an imaginary-time propagation with the modified Hamiltonian

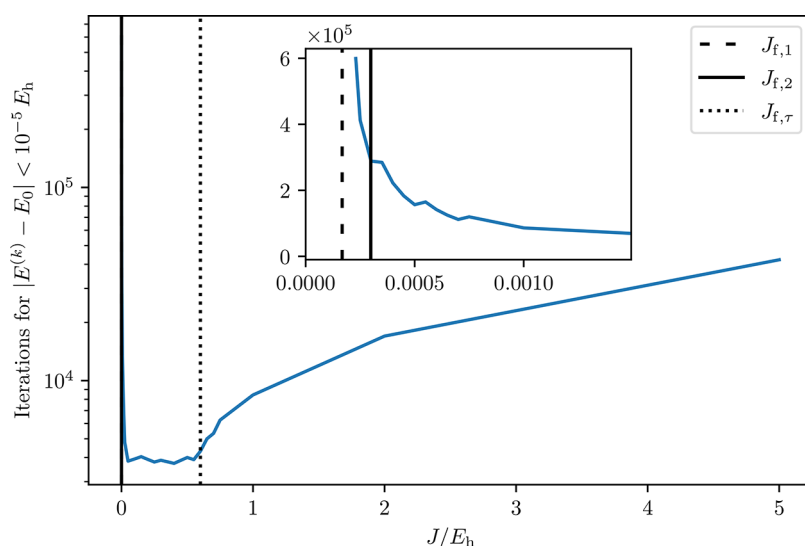


Figure 2. Number of iterations required to achieve convergence up to $1 \times 10^{-5} E_h$ using *deterministic* imaginary-time propagation for the $\Gamma^{(1/2)}$ state of the manganese cluster in an (9,9) active space. $J_{f,1}$ was calculated using eq 4, $J_{f,2}$ was calculated using eq 5, and $J_{f,\tau}$ was calculated using the spread $(E'_{\max} - E'_0)$, which was approximated from the spread of diagonal values of H' .

$\exp(-\tau(\hat{H}' - E_0))$ projects them out faster. But, as we (*vide infra*) and Levine et al.⁵⁵ observed, there exist J values beyond which the convergence deteriorates. Intuitively this can be understood, because for increasingly larger J values the modified Hamiltonian can be interpreted as mainly a \hat{S}^2 operator, corrected by a small perturbation represented by \hat{H} . Diagonalization of the pure \hat{S}^2 operator results in numerous degenerate eigenstates with equivalent spin eigenvalues (not energies), the lowest being $S(S+1) = 0$ or $S(S+1) = 3/4$, for even and odd numbers of electrons, respectively. This degeneracy is in part lifted by the Hamiltonian \hat{H} . However, for very large J values, the \hat{H} correction becomes relatively small compared to $J\hat{S}^2$, and projecting out the high-energy states of same spin becomes harder.

If we look at quantitative estimates for the speed of convergence (the detailed derivations are provided in the Appendix (Section 5)), an interval of J values exists inside which the speed of convergence is nearly constant and optimal (see Figure 2). The lower bound of this optimal range can be estimated by the *second flipping point*

$$J_{f,2} = \max_{S > S_{ls}} \frac{E_{ls,1} - E_{S,0}}{\Delta\langle \hat{S}^2 \rangle}, \quad (5)$$

where $E_{ls,1}$ is the energy of the first low-spin excited state of desired spin. The speed of convergence increases proportionally to the energy separation between the lowest and the second-lowest state. For $J_{f,1} < J < J_{f,2}$, we have $E'_{ls,0} < E'_{hs,0} < E'_{ls,1}$, and the spread between lowest and second to lowest energy state increases with J . For $J > J_{f,2}$, we have $E'_{ls,0} < E'_{ls,1} < E'_{hs,0}$, and the energy separation between lowest and second-lowest energy state is unaffected by J (under the assumption that they have the same spin multiplicity). Thus, for $J > J_{f,2}$, but still below the upper bound discussed in the following, the convergence is nearly independent of J (plateau in the speed of convergence, Figure 2).

For optimization techniques based on the imaginary-time Schrödinger equation, such as FCIQMC, the upper bound of the optimal range of J is given by the *τ -flipping point*, $J_{f,\tau}$, which denotes the point where the maximum time step starts to be

dominated by a $1/J$ proportionality. In the case of *deterministic* imaginary-time propagations we have $\Delta\tau = (E'_{\max} - E'_0)^{-1}$. If we assume that the highest energy state in \hat{H}' (for increasing $J > J_{f,1}$) and the energetically lowest spin-state have the same spin multiplicity, their energy difference will not be affected by J for a large range of J . For typical full-valence active space calculations this is a well-founded assumption. For $J > J_{f,\tau}$ the spread of \hat{H}' becomes increasingly dominated by the spread of spin expectation values, $E'_{\max} - E'_0 \approx J\Delta\langle \hat{S} \rangle_{\max}$, hence $\Delta\tau \approx (J\Delta\langle \hat{S} \rangle_{\max})^{-1}$. Thus, larger J values lead to smaller optimal $\Delta\tau$ values, with a consequent reduction of the speed of convergence.

In the case of *stochastic* imaginary-time propagation, as in FCIQMC, the $J_{f,\tau}$ is more complicated to find. The time step has to be chosen differently compared to the deterministic case to achieve stable dynamics. The conventional^{9,27} choice for $\Delta\tau$ in a SD basis is

$$\Delta\tau = \min_{i,j} \left(\frac{p_{\text{gen}}(i,j)}{H_{ij}} \right), \quad (6)$$

which not only depends on the spectrum of \hat{H} but also on the excitation generator in use and the determinant pairs i, j where spawns actually happened during a calculation. Note the implicit assumption, that the stochastic $\Delta\tau$ from eq 6 is smaller than the deterministically chosen $\Delta\tau$, that depends purely on the spread of the spectrum of \hat{H} . As in the deterministic case, there will be a $J_{f,\tau}$ from which the time step purely follows a $(1/J)$ dependency. If a weighted excitation generator is in use, the p_{gen} for exchange excitations will increase with J as do the H_{ij} and depending on the dynamics there will be different pairs i, j fulfilling the minimum of eq 6. Hence unlike the deterministic case there might be changes of $\Delta\tau$ already before reaching $J_{f,\tau}$. We also observe that for too large J the stochastic noise of the Monte Carlo simulation increases.

In summary, the convergence improves with increasing J for $J_{f,1} < J < J_{f,2}$. Also, there exists a $J_{f,\tau}$ after which a $1/J$ dependency of the time step follows, negatively affecting the speed of convergence. Between $J_{f,2}$ and $J_{f,\tau}$ there is a plateau of nearly optimal J values. We would like to point out (see Figure

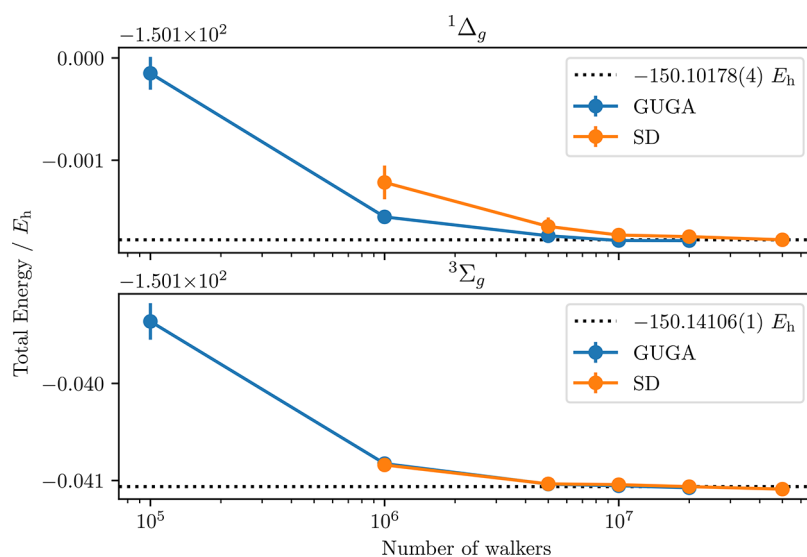


Figure 3. Total energy calculated from RDMs for the ${}^1\Delta_g$ (upper plot) and the ${}^3\Sigma_g$ (lower plot) spin states of the oxygen molecule.

2) that the convergence deteriorates slower for $J > J_{f,\tau}$ than for $J < J_{f,2}$; i. e., it is better to select a bit too high J values than too small ones. In addition a too large $J > J_{f,\tau}$ affects the convergence rate, while a too small $J < J_{f,1}$ leads to spin contamination. Therefore, it is advisable to choose J values inside the range but closer to $J_{f,\tau}$.

In practical applications, it is difficult to calculate the lower bound $J_{f,2}$ because it requires knowledge of the spin-state energetics, which are exactly the purpose of the calculation. However, the $J_{f,\tau}$ can be estimated from monitoring the time step and stochastic noise during FCIQMC *training runs* (using low walker populations) for different J values. The first flipping point $J_{f,1}$ can be found by monitoring the spin expectation value for different J ; if it converges to a high-spin state, the J is too small. We assume $J_{f,\tau} - J_{f,2} \approx J_{f,\tau} - J_{f,1}$ and select a trial J between $J_{f,1}$ and $J_{f,\tau}$. In the case of *stochastic* imaginary-time propagation, it is generally advised to also monitor the stochastic noise and reduce J accordingly.

3. APPLICATION

The robustness of the spin penalty method in SD-based FCIQMC has been explored in two crucial test-case applications. We investigated the vertical ${}^1\Delta_g \leftarrow {}^3\Sigma_g$ transition in the O_2 molecule using a full-CI expansion in a double- ζ quality basis set, and the vertical $\Gamma^{(1/2)} \leftarrow \Gamma^{(3/2)}$ (and $\Gamma^{(9/2)}$) transition in a $[\text{Mn}_3^{(\text{IV})}\text{O}_4]$ trinuclear cluster. In both cases, the ground state is the higher spin-state.

3.1. Oxygen Dimer. We used a distance of 1.203 Å, and correlated 16 electrons in the 28 orbitals of an ANO-RCC-VDZP basis.^{59,60} The Full-CI calculations were performed on the basis of the state-specific CASSCF(8,6) orbitals. A spin-pure calculation using GUGA-FCIQMC served as reference. The J parameter was set to $0.12 E_h$.

The $\langle \hat{H} \rangle$ expectation value calculated from RDMs is shown in Figure 3. The triplet converges faster than the singlet, and its total energy nearly matches the energy of the GUGA reference calculation for the same walker number. Generally, convergence with respect to the number of walkers in initiator-FCIQMC is mainly influenced by the compactness of the respective wave function. The triplet calculation started from a SD with $|M_S| = 1$, which is also the spin-pure configuration that

dominates the FCI wave function for this electronic state. On the contrary, in the case of the singlet spin state, a multideterminantal wave function is required to correctly describe the spin-pure reference space; therefore, the calculation converges slower with respect to the walker number than the GUGA-based one. However, the wall clock time to achieve the same quality of convergence is roughly comparable, since SD-based FCIQMC is generally faster for a given walker number, as discussed in the literature.^{9,33} For all walker populations, the spin expectation values have been used to confirm convergence to the correct spin state and monitor spin contamination. The deviation was larger for the singlet whose mean spin expectation value was 1.30×10^{-5} , compared to the theoretical 0.

3.2. Manganese Cluster. Two active spaces have been defined to test the spin penalty approach on the $[\text{Mn}_3^{(\text{IV})}\text{O}_4]$ trinuclear cluster (Figure 4). A small CAS(9,9) is utilized to

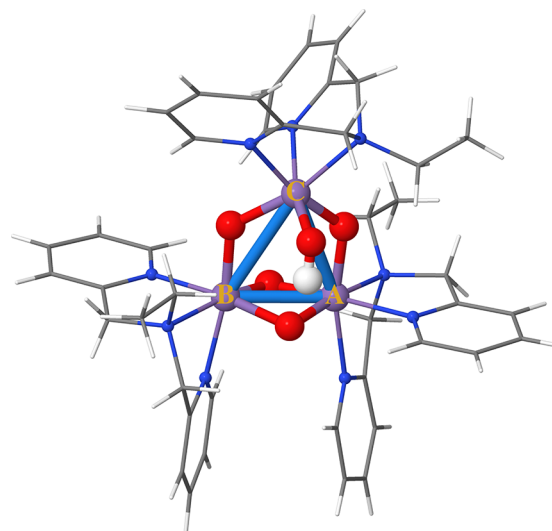
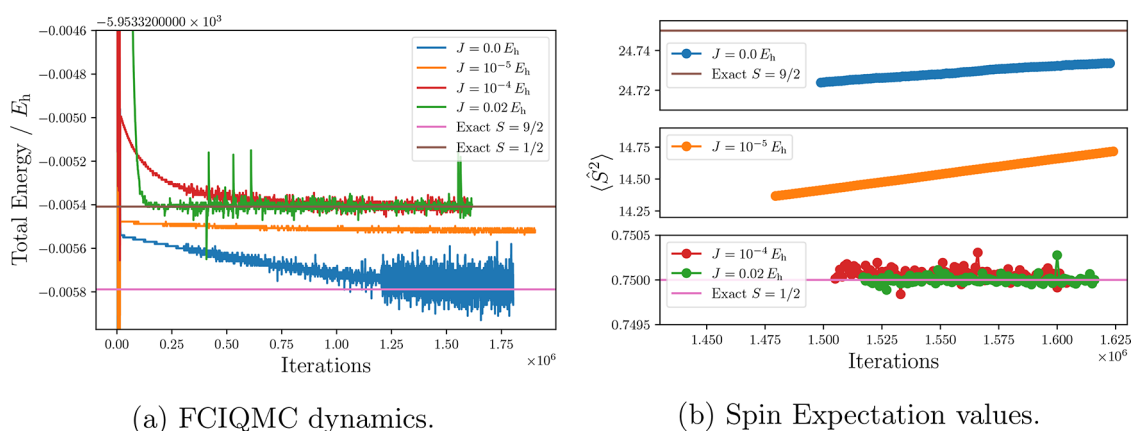


Figure 4. Structure of the $[\text{Mn}_3\text{O}_4]$ trinuclear model system extracted from ref 61. A, B, and C labels identify the Mn(IV) magnetic centers. Oxygen, nitrogen, carbon, and hydrogen atoms are labeled in red, blue, gray, and white, respectively.



(a) FCIQMC dynamics.

(b) Spin Expectation values.

Figure 5. (a) FCIQMC dynamics varying the amount of spin penalty. The projected energy shifted by the $J \cdot S(S + 1)$ value is reported, where S is the expected spin value. (b) Spin expectation values calculated from RDMs are shown. All simulations used 5×10^4 walkers.

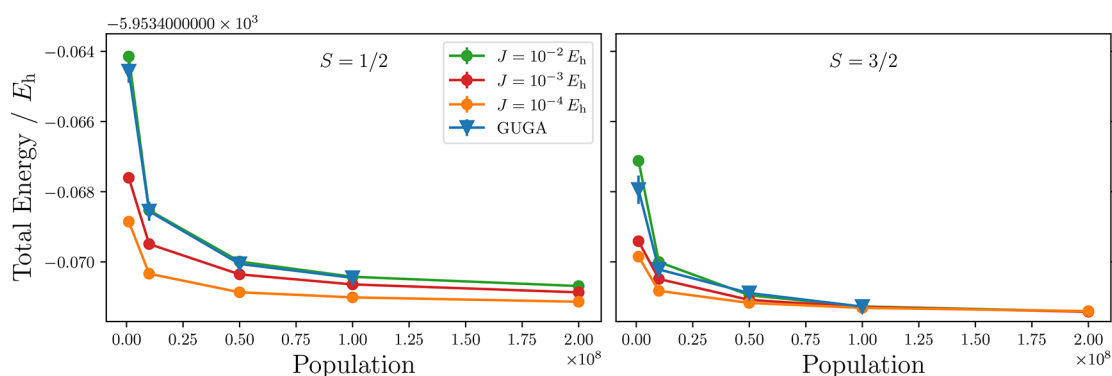


Figure 6. CAS(55,38) total energies, obtained as expectation values from one- and two-body RDMs, for the $S = 1/2$ (left) and $S = 3/2$ (right) spin states as a function of the walker population. The lower energies for smaller J values are not to be interpreted as a faster convergence with respect to walker number. They are a consequence of the admixing of the targeted spin state with higher spin states (details in the main text).

directly compare to the fully deterministic GUGA-based spin gap.^{62,63} A larger CAS(55,38) has been employed to demonstrate the numerical stability of the method in more realistic scenarios. The CAS(9,9) consists of the nine singly occupied t_{2g} orbitals on the three magnetic centers. The large active space consists of the 15 3d orbitals and their nine electrons, the 12 doubly occupied 2p orbitals of the bridging oxygen atoms, the 10 doubly occupied peripheral lone-pair orbitals pointing at the metal sites, and two doubly occupied orbitals of the $-OH$ group, of σ and π character. A similar active space has been previously chosen for a similar $[Mn_3O_4]$ cluster.^{35,64}

Through experimental investigation, Armstrong showed that the ground state of this system is a $\Gamma^{(3/2)}$ spin state.⁶¹ The too small CASSCF(9,9) erroneously predicts a $\Gamma^{(9/2)}$ ground state. Nonetheless, this small model calculation represents an interesting test case to explore the applicability of the spin penalty strategy. The larger CAS(55,38) describes qualitatively well the spin-state ordering, with a $S = 3/2$ ground state and a $S = 1/2$ state at slightly higher energy, in line with Armstrong's findings.

In Figure 5, we show the convergence behavior of the FCIQMC dynamics for different J applied to the CAS(9,9) wave function. For FCIQMC dynamics with $J = 0 E_h$ or too small spin penalties ($J = 1 \times 10^{-5} E_h$) the flipping point is not reached, and the FCIQMC dynamics converges to the high-spin state ($S = 9/2$), which is the ground state for the small CAS(9,9) model active space. For J values above the flipping

point, the low-spin state wave function is obtained. These results are confirmed by the total spin expectation value (Figure 5b). Speed of convergence increases for larger penalty values, and a large range of J values ($1 \times 10^{-4} E_h \leq J \leq 2 \times 10^{-2} E_h$) exists that provides stable and fast converging FCIQMC dynamics. Too large J values ($> 1 \times 10^{-1} E_h$) lead to convergence problems, which is in line with $J_{\text{crit}} = 0.1 E_h$.

For the CAS(55,38) model active space the competing doublet ($S = 1/2$) and quartet ($S = 3/2$) spin state wave functions have been optimized. GUGA-FCIQMC has been utilized as a reference. Three choices of J were used that permitted the characterization of the doublet spin state, namely $J = 1 \times 10^{-2} E_h$, $1 \times 10^{-3} E_h$, and $1 \times 10^{-4} E_h$. Figure 6 shows the energetics for the $S = 1/2$ and $S = 3/2$ states, as a function of the walker population. We notice that all dynamics are stable and fast converging. The choice of the large parameter, $J = 1 \times 10^{-2} E_h$, results in a nearly exact matching of the spin-purified total energy with the one obtained from the spin-adapted GUGA-FCIQMC approach, at the same walker population.

Lower J values result in lower total energies for low walker populations. The lower energies for smaller J values are not to be interpreted as a faster convergence of the spin penalty approach for lower J . Instead, considering that the spin expectation value for the smaller $J = 1 \times 10^{-3} E_h$ is higher than the expected value (Figure 7) we are brought to the conclusion that the unconverged wave function (low population) is in a broken-symmetry state, that results from the mixture of the

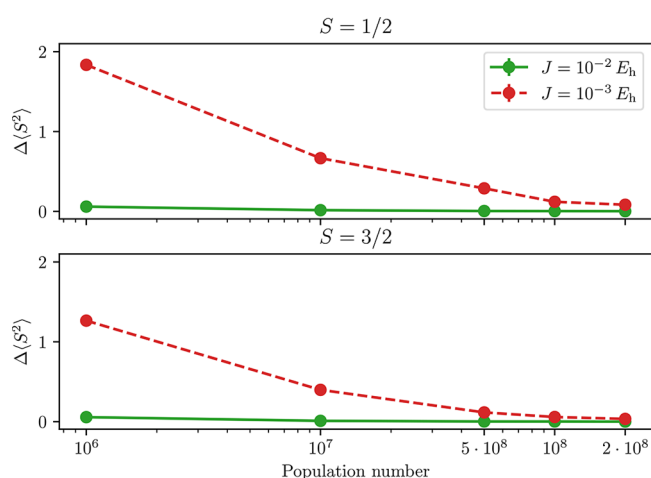


Figure 7. Spin contamination ($\Delta\langle\hat{S}^2\rangle = \langle\hat{S}^2\rangle - S(S+1)$) in the CAS(55,38) for different J and population numbers. The spin $\langle\hat{S}^2\rangle$ was calculated from RDMs. While the larger $J = 1 \times 10^{-2} E_h$ value provides dynamics with nearly exact spin expectation values for any chosen walker population, the smaller $J = 1 \times 10^{-3} E_h$ calculations converge to the correct spin expectation value more slowly and only for larger populations.

target spin state ($\Gamma^{(1/2)}$) and the higher spin states (for example $\Gamma^{(3/2)}$). Admixing the higher spin states artificially lowers the total energy. For larger walker populations and for larger J , the spin expectation value gets closer to the targeted value, eliminating any spin contamination from the optimized wave function.

It is worth noting that the 38 active orbitals have been localized and reordered to reach maximum compression of the GUGA wave function (see refs 33–35 for details). In Figure 11 of ref 34, we have shown that GUGA-FCIQMC converges faster than the Slater determinant FCIQMC counterpart, when using a tailored (in this case localized) MO basis. Thus, we expect that the chosen one-electron basis utilized for the CAS(55,38) calculation in general favors the GUGA-FCIQMC approach. However, we observe very similar convergence of the GUGA and the Slater determinant based spin-penalty approaches (for $J = 1 \times 10^{-2} E_h$). Moreover, it is interesting to notice that for an equivalent wall-clock time, the spin penalty approach can be run at higher walker population (2×10^8 walkers in the spin penalty method versus 1×10^8 walkers in GUGA) and reaches a lower total energy. These results suggest an overall better performance of the spin-purification approach. However, the GUGA strategy has two crucial advantages that we have documented in recent works:^{32–35} (a) within GUGA the space of the leading electronic configurations (CSFs) can be greatly reduced and directly connected to physical concepts, and (b) the GUGA CI Hamiltonian matrix has a unique *quasi-block-diagonal* structure, allowing for unprecedented *state-specific* optimizations of ground and/or excited states.

As a final remark, we observe that the spin-penalty strategy enables the combined S - and M_S -adaptation. This scheme is thus more flexible than the GUGA S -adaptation, and allows for the treatment of anisotropic Hamiltonians. One trades simplicity and universality for lower dimensionality when going from the M_S -adapted space to the S -adapted one. Moreover, while stochastically sampled higher-order density matrices are already available within the SD-based FCIQMC approach, allowing for multireference second order perturbation theory (PT2) methods,^{65,66} GUGA-FCIQMC three- and

four-body density matrices are not available, preventing for the moment GUGA-based PT2 strategies. Additionally, it is possible to envision spin-pure similarity-transformed FCIQMC calculations based on transcorrelated methods^{67–74} using the current spin penalty approach, while technical difficulties exist within the GUGA scheme, because of the presence of three-body interactions in the transcorrelated Hamiltonian.

4. CONCLUSION

In conclusion, we have demonstrated that spin purification based on a first-order spin penalty can be efficiently applied to the Slater-determinant based FCIQMC algorithm. We have also explained the origin of an optimal range of J values and that too large J parameters are to be avoided as they result in smaller time steps, deteriorating convergence. The method was successfully applied to calculate the ${}^1\Delta_g \leftarrow {}^3\Sigma_g^-$ transition in the O_2 dimer and the $\Gamma^{(1/2)} \leftarrow \Gamma^{(3/2)}$ (and $\Gamma^{(9/2)}$) transition in a $[Mn_3^{(IV)}O_4]$ trinuclear cluster model. The range of applicability of the spin penalty FCIQMC approach is very broad, including the coupling of large active space Stochastic-CASSCF and Stochastic-GASSCF wave functions to methods capable of recovering dynamic correlation outside the active space, such as MC-PDFT,^{75,76} and tailored-CC,⁷⁷ and crucially methods that require high-order interactions (for example in the form of three- and four-body RDMs) such as PT2 and similarity-transformed techniques. A large range of chemical systems and models for solid state materials can be investigated, including ferromagnetic superconductors of practical interest, such as UGe_2 ⁷⁸ and $URhGe$.⁷⁹ The method can also be extended to model Hamiltonians, such as the Hubbard model, often used to investigate spin interactions in strongly correlated materials. By this approach we are able to tackle anisotropic Hamiltonians and, as *spin-resolved* density matrices are available, spin-dependent properties, such as the hyperfine coupling tensors (pivotal in characterizing spin interactions in polynuclear transition metal clusters), are within reach. These aspects will be the subject of future work.

5. APPENDIX

In the following, we prove the existence of an optimal range of J values for the speed of convergence of the first-order spin penalty approach with the modified Hamiltonian

$$\hat{H}' = \hat{H} + J \cdot \hat{S}^2 \quad (7)$$

in FCIQMC and discuss strategies for selecting it. The derivation that follows applies to any other method that performs a stochastic or nonstochastic linearized imaginary-time propagation. We write the imaginary-time propagation of $\Psi(\tau = 0)$ as

$$\Psi(\tau) = \exp(-H\tau)\Psi(\tau = 0) \quad (8)$$

without renormalization. After linearization we arrive at the projector

$$\hat{G} = 1 + \Delta\tau(E_0 - \hat{H}), \quad (9)$$

where E_0 is the ground-state energy and the working equations arise from repeated application of \hat{G} which is nothing else than the *power method* for the operator \hat{G} . We will first discuss the speed of convergence of power methods and apply it to the linear projector \hat{G} .

5.1. Power Method

The speed of convergence for the power method is given by the following theorem.⁸⁰

Theorem 1 (Speed of convergence power method). Let $A \in \mathbb{R}^{n,n}$ be a real, symmetric matrix with positive eigenvalues λ_i sorted in decreasing order and

$$\lambda_1 > \lambda_2 \geq \dots \lambda_n \geq 0, \quad (10)$$

so the largest eigenvalue λ_1 is non-degenerate. The eigenvectors corresponding to λ_i are denoted with \mathbf{u}_i . We denote the starting guess with $\mathbf{x}^{(0)}$ and define the guess-vector and guess-eigenvalue of the k th iteration as

$$\mathbf{x}^{(k)} = A^k \mathbf{x}^{(0)} \quad (11)$$

$$\lambda^{(k)} = \langle \mathbf{x}^{(k)} | A \mathbf{x}^{(k)} \rangle. \quad (12)$$

If the starting guess has non-zero overlap with the eigenvector of highest eigenvalue ($\langle \mathbf{x}^{(0)} | \mathbf{u}_1 \rangle \neq 0$) then eqs 11 and 12 converge to \mathbf{u}_1 and λ_1 , respectively. The convergence rates are given by

$$|\lambda_1 - \lambda^{(k)}| \leq (\lambda_1 - \lambda_n) \left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \sin^2(\theta^{(0)}), \quad (13)$$

where $\theta^{(k)} = \angle(\mathbf{x}^{(k)}, \mathbf{u}_1)$.

For $J > J_{f,1}$ the lowest eigenstate \mathbf{u}_1 is the same in all cases. Since we also use the same starting guess in all cases the deviation angle of the starting guess $\theta^{(0)}$ is the same everywhere and we will omit the $\sin^2(\theta^{(0)})$ term from further equations.

For further analysis we would like to emphasize that eq 13 is an upper bound, and depending on the energetically higher eigenvalues, λ_i , $i > 2$, differences in convergence behavior may further occur.

We would like to showcase this with the three diagonal matrices from $\mathbb{R}^{4,4}$:

$$\begin{aligned} \mathbf{A} &= \text{diag}([1, 0.999995, 0.999990, 0]) \\ \mathbf{B} &= \text{diag}([1, 0.999990, 0.999982, 0]) \\ \mathbf{C} &= \text{diag}([1, 0.999990, 0.997309, 0]) \end{aligned} \quad (14)$$

and a starting vector

$$\mathbf{x}^{(0)} = \frac{1}{2}[1, 1, 1, 1]. \quad (15)$$

The λ_i , $1 \leq i \leq 3$ eigenvalues were not chosen randomly, but were calculated as the first three eigenvalues of the projector \hat{G} in the (9,9) active space on the manganese cluster for different values of J . \mathbf{A} corresponds to $J = (J_{f,1} + J_{f,2})/2$, \mathbf{B} corresponds to $J = J_{f,2} + 0.0001 E_{\text{hs}}$, and \mathbf{C} corresponds to $J = \frac{2}{3} J_{f,\tau}$. The matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} have the same eigenvectors and the same spread of $\lambda_1 - \lambda_4 = 1$. By the ratio of λ_2/λ_1 and theorem 1 it is expected that matrices \mathbf{B} and \mathbf{C} have the same convergence behavior and converge faster than \mathbf{A} .

In Figure 8, we see that the convergence in the long-time limit is indeed described by Theorem 1 and is only depending on the ratio of λ_2/λ_1 . Nevertheless in the beginning there is faster convergence for \mathbf{C} because the third component is projected out faster, due to a smaller λ_3 . So even for a constant ratio λ_2/λ_1 it is still advantageous to spread the distance between λ_1 and the other eigenvalues.

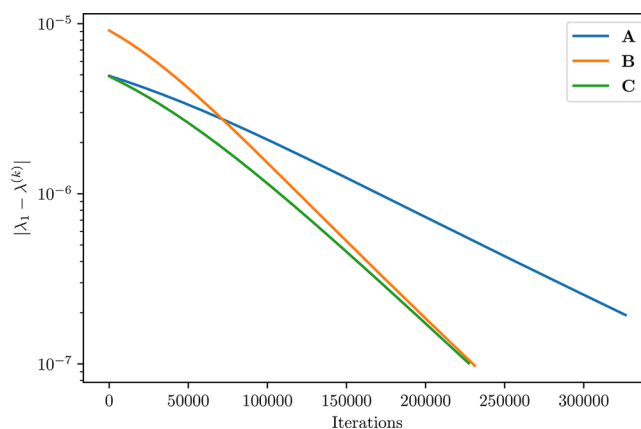


Figure 8. Error against the iteration number for the power method applied to matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} from eq 14.

5.2. Power Method Bounds in FCIQMC

In order to apply Theorem 1 to \hat{G} we need to assume

$$\Delta\tau \leq \frac{1}{E_{\text{max}} - E_0} \quad (16)$$

and that the ground-state is non-degenerate. The assumption from eq 16 is usually fulfilled in FCIQMC, since the reduction of stochastic noise often requires even smaller time steps. We denote the eigenvalues of \hat{G} with λ_i . By construction we have $1 = \lambda_1$, and by assumption (eq 16), we have $\lambda_n \geq 0$ which gives in total $1 = \lambda_1 > \lambda_2 \geq \dots \lambda_n \geq 0$. Note that the highest eigenvalue λ_1 of \hat{G} belongs to the state with the lowest eigenvalue of H .

We will write the eigenvalues of the modified Hamiltonian and resulting variables like the time step with a prime superscript, $'$. To simplify the discussion herein, we assume deterministic imaginary-time propagation and that $\Delta\tau$ is chosen maximal

$$\Delta\tau = \frac{1}{E_{\text{max}} - E_0}, \quad (17)$$

for its respective Hamiltonian and value of J . We also assume

$$E_0 = E_{\text{hs},0} < E_{\text{ls},0} < E_{\text{ls},1} \quad (18)$$

for the original Hamiltonian \hat{H} .

5.2.1. Case: $J_{f,1} < J < J_{f,2}$. The ordering of eigenvalues of \hat{H}' is given by

$$E'_0 = E'_{\text{ls},0} < E'_{\text{hs},0} < E'_{\text{ls},1}. \quad (19)$$

We want to determine the relevant eigenvalues of

$$\hat{G} = 1 + \Delta\tau'(E'_0 - \hat{H}'). \quad (20)$$

By construction we have

$$\begin{aligned} \lambda_1 &= 1, \\ \lambda_2 &= 1 + \Delta\tau'(E'_0 - E'_{\text{hs},0}) \\ &= 1 - \Delta\tau'(\Delta E_{\text{hs}} + J\Delta\langle \hat{S} \rangle), \end{aligned} \quad (21)$$

with $\Delta E_{\text{hs}} = E_{\text{hs},0} - E_{\text{ls},0}$ and $\Delta\langle \hat{S} \rangle$ being the difference of \hat{S} expectation value between the high- and low-spin state. Because of the assumption for $\Delta\tau$ (eq 17) we get

$$\lambda_n = 1 + \frac{1}{E'_{\text{max}} - E'_0}(E'_0 - E'_{\text{max}}) = 0. \quad (22)$$

In total, we obtain $(\lambda_1 - \lambda_n) = 1$ and $\frac{\lambda_2}{\lambda_1} = \lambda_2$. As discussed in the main text we assume that the time step is not yet affected by J , so $\Delta\tau' \approx \Delta\tau$. We conclude that the convergence is bounded as

$$\begin{aligned} |\lambda_1 - \lambda^{(k)}| &\leq |\lambda_2|^{2k} \\ &= |1 - \Delta\tau(\Delta E_{\text{hs}} + J\Delta\langle\hat{S}\rangle)|^{2k}, \end{aligned} \quad (23)$$

which gets tighter with increasing J .

5.2.2. Case: $J_{f,2} < J < J_{f,\tau}$ The ordering of eigenvalues of \hat{H}' is given by

$$E'_0 = E'_{\text{ls},0} < E'_{\text{ls},1} < E'_{\text{hs},0}. \quad (24)$$

Since the energetic spacing between eigenstates of same spin is not affected by $J\hat{S}^2$ we get

$$\begin{aligned} \lambda_1 &= 1, \\ \lambda_2 &= 1 + \Delta\tau'(E'_{\text{ls},0} - E'_{\text{ls},1}) \\ &= 1 + \Delta\tau'(E_{\text{ls},0} - E_{\text{ls},1}) \\ &= 1 - \Delta\tau'\Delta E_{\text{ls}}, \end{aligned} \quad (25)$$

with $\Delta E_{\text{ls}} = (E_{\text{ls},1} - E_{\text{ls},0})$.

As in the previous case and by definition of $J_{f,\tau}$ we assume that the time step is not yet affected by J , so $\Delta\tau' \approx \Delta\tau$. As before we conclude that the convergence is bounded as

$$\begin{aligned} |\lambda_1 - \lambda^{(k)}| &\leq |\lambda_2|^{2k} \\ &= |1 - \Delta\tau\Delta E_{\text{ls}}|^{2k}, \end{aligned} \quad (26)$$

which is independent of J . Note that the other eigenvalues of \hat{G} like λ_3 are still affected by J , and the convergence is still slightly improved by increasing J .

5.2.3. Case: $J > J_{f,\tau}$ A tilde will be used for this case. The ordering of the eigenvalues is given by

$$\tilde{E}'_0 = \tilde{E}'_{\text{ls},0} < \tilde{E}'_{\text{ls},1} \ll \tilde{E}'_{\text{hs},0} \quad (27)$$

The energetic spacing between the two lowest states is independent of J and we get:

$$\begin{aligned} \tilde{\lambda}_1 &= 1, \\ \tilde{\lambda}_2 &= 1 + \Delta\tilde{\tau}'(\tilde{E}'_{\text{ls},0} - \tilde{E}'_{\text{ls},1}) \\ &= 1 + \Delta\tilde{\tau}'(E_{\text{ls},0} - E_{\text{ls},1}) \\ &= 1 - \Delta\tilde{\tau}'\Delta E_{\text{ls}}. \end{aligned} \quad (28)$$

Again, we assume a maximum choice for the time step, so

$$\Delta\tilde{\tau}' = (\tilde{E}'_{\text{max}} - \tilde{E}'_0)^{-1}. \quad (29)$$

As before, we conclude

$$\tilde{\lambda}_n = 0. \quad (30)$$

For very large J , the spread of eigenvalues is dominated by the spacing between the spin eigenvalues, so we get

$$(\tilde{E}'_{\text{max}} - \tilde{E}'_0) \approx J\Delta\langle\hat{S}\rangle_{\text{max}}, \quad (31)$$

where $\Delta\langle\hat{S}\rangle_{\text{max}}$ is the difference between the lowest and highest possible spin state for that system. We get

$$\Delta\tilde{\tau}' = (\tilde{E}'_{\text{max}} - \tilde{E}'_0)^{-1} \approx (J\Delta\langle\hat{S}\rangle_{\text{max}})^{-1}. \quad (32)$$

This means that the convergence is bounded as

$$\begin{aligned} |\tilde{\lambda}_1 - \tilde{\lambda}^{(k)}| &\leq |\tilde{\lambda}_2|^{2k} \\ &= \left| \left(1 - \frac{\Delta E_{\text{ls}}}{J\Delta\langle\hat{S}\rangle_{\text{max}}} \right) \right|^{2k}, \end{aligned} \quad (33)$$

which gets tighter with decreasing J .

5.3. Analyzing All Cases Together

We summarize that the convergence improves with increasing J for $J_{f,1} < J < J_{f,2}$. For $J > J_{f,\tau}$ the convergence deteriorates with increasing J . Between $J_{f,2}$ and $J_{f,\tau}$ there is a plateau of near-optimal J values.

If we take the derivative of the two bounds $(1 - \Delta\tau(\Delta E_{\text{hs}} + J\Delta\langle\hat{S}\rangle))$ and $\left(1 - \frac{\Delta E_{\text{ls}}}{J\Delta\langle\hat{S}\rangle_{\text{max}}}\right)$ after J we can see that in the case of $J > J_{f,\tau}$ the rate of change is smaller. This means that when in doubt it is better to choose $J > J_{f,\tau}$ than to choose $J < J_{f,2}$; thus, it is better to “overshoot”.

It might be tempting to equate the two upper bounds from the separate cases to determine a unique J_{opt} . This is not a defined operation because both expressions were separately derived under the assumption of either high or low J . In the scope of one of the bounds, the other one is not valid.

5.4. Numerical Tests for J Choice

In order to numerically confirm our derivations for the choice of optimal J values and to stay as close as possible to the assumptions in our previous derivation, we also performed a fully deterministic imaginary-time evolution for the systems that were discussed in our main text. The fully deterministic propagation was achieved by defining a semi-stochastic space for FCIQMC that ranged over the whole Hilbert space. The spacing of eigenvalues for the calculation of $\Delta\tau$ (eq 17) was estimated from the spread of diagonal values of the Hamiltonian matrix.

In Figure 9, we see the number of iterations required to achieve convergence up to $1 \times 10^{-5} E_h$ in a deterministic imaginary-time propagation for the $^1\Delta_g$ state of oxygen in the minimal (8,6) active space. If we keep the doubly logarithmic scale in mind, we see the previous derivations numerically

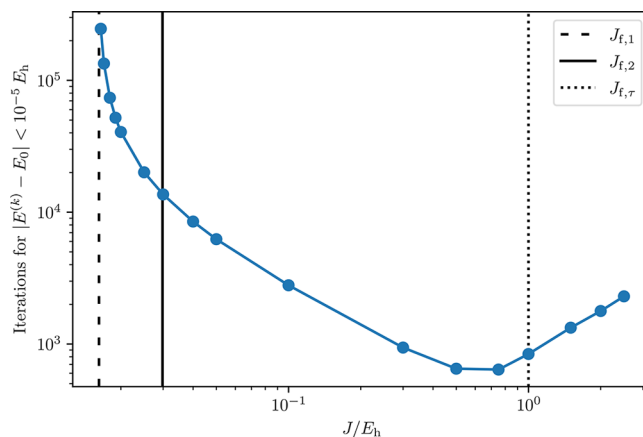


Figure 9. Number of iterations required to achieve convergence up to $1 \times 10^{-5} E_h$ in a fully deterministic imaginary-time propagation for the $^1\Delta_g$ state of oxygen in an (8,6) active space.

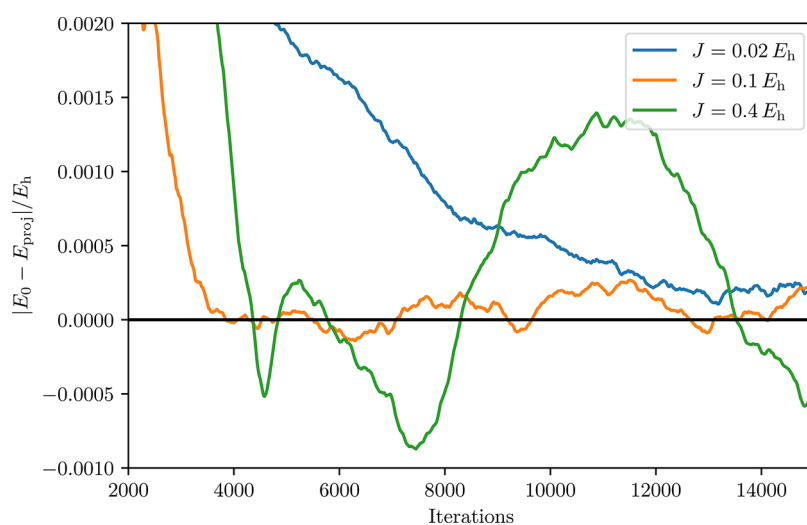


Figure 10. Moving average (window length is 200) of the projected energy of FCIQMC for different J values for the $^1\Delta_g$ state of oxygen in an (8,6) active space.

confirmed. For $J_{f,1} < J < J_{f,2}$, the convergence improves drastically with increasing J . For $J_{f,2} < J < J_{f,\tau}$ a plateau forms, but it is still better to choose J closer to $J_{f,\tau}$. For $J > J_{f,\tau}$ the convergence deteriorates with increasing J .

For a stochastic imaginary-time propagation the criterion of convergence up to $1 \times 10^{-5} E_h$ cannot be easily evaluated. Instead the moving average of the projected energy for three different J values is drawn in Figure 10. For this system, we had $J_{f,2} = 0.03 E_h$ and $J_{f,\tau} \approx 0.07 E_h$. The blue line describes a dynamic with $J < J_{f,2}$, and the convergence improves considerably when going to higher J values as in the orange line. For values considerably above $J_{f,\tau}$ (green line), the quality of the calculation decreases again.

AUTHOR INFORMATION

Corresponding Authors

Oskar Weser – Max-Planck-Institute for Solid State Research, 70569 Stuttgart, Germany; orcid.org/0000-0001-5503-1195; Email: oskar.weser@gmail.com

Giovanni Li Manni – Max-Planck-Institute for Solid State Research, 70569 Stuttgart, Germany; orcid.org/0000-0002-3666-3880; Email: g.limanni@fkf.mpg.de

Authors

Niklas Liebermann – Max-Planck-Institute for Solid State Research, 70569 Stuttgart, Germany

Daniel Kats – Max-Planck-Institute for Solid State Research, 70569 Stuttgart, Germany; orcid.org/0000-0002-7274-0601

Ali Alavi – Max-Planck-Institute for Solid State Research, 70569 Stuttgart, Germany; Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpca.2c01338>

Funding

Open access funded by Max Planck Society.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Pradipta Samanta, Kai Guther, and Werner Dobrautz for valuable scientific discussions. The authors acknowledge the Max Planck Society for financial support.

REFERENCES

- (1) Sharma, S.; Chan, G. K.-L. Spin-adapted density matrix renormalization group algorithms for quantum chemistry. *J. Chem. Phys.* **2012**, *136*, 124121–124137.
- (2) Ma, D.; Li Manni, G.; Gagliardi, L. The generalized active space concept in multiconfigurational self-consistent field methods. *J. Chem. Phys.* **2011**, *135*, 044128–044138.
- (3) Li Manni, G.; Ma, D.; Aquilante, F.; Olsen, J.; Gagliardi, L. SplitGAS Method for Strong Correlation and the Challenging Case of Cr_2 . *J. Chem. Theory Comput.* **2013**, *9*, 3375–3384.
- (4) Li, Z.; Chan, G. K.-L. Spin-Projected Matrix Product States: Versatile Tool for Strongly Correlated Systems. *J. Chem. Theory Comput.* **2017**, *13*, 2681–2695.
- (5) Qiu, Y.; Henderson, T. M.; Zhao, J.; Scuseria, G. E. Projected coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 064111–064126.
- (6) Tsuchimochi, T.; Ten-no, S. L. Extending spin-symmetry projected coupled-cluster to large model spaces using an iterative null-space projection technique. *J. Comput. Chem.* **2019**, *40*, 265–278.
- (7) Shepard, R.; Gidofalvi, G.; Brozell, S. R. The multifacet graphically contracted function method. I. Formulation and implementation. *J. Chem. Phys.* **2014**, *141*, 064105–064130.
- (8) Gidofalvi, G.; Brozell, S. R.; Shepard, R. Wave function analysis with Shavitt graph density in the graphically contracted function method. *Theor. Chem. Acc.* **2014**, *133*, 1512–1522.
- (9) Dobrautz, W.; Smart, S. D.; Alavi, A. Efficient formulation of full configuration interaction quantum Monte Carlo in a spin eigenbasis via the graphical unitary group approach. *J. Chem. Phys.* **2019**, *151*, 094104–094136.
- (10) Guther, K.; Anderson, R. J.; Blunt, N. S.; Bogdanov, N. A.; Cleland, D.; Dattani, N.; Dobrautz, W.; Ghanem, K.; Jeszenszki, P.; Liebermann, N.; et al. NECI: N-Electron Configuration Interaction with an emphasis on state-of-the-art stochastic methods. *J. Chem. Phys.* **2020**, *153*, 034107–034131.
- (11) Zhang, N.; Liu, W.; Hoffmann, M. R. Iterative Configuration Interaction with Selection. *J. Chem. Theory Comput.* **2020**, *16*, 2296–2316.
- (12) Chilkuri, V. G.; Neese, F. Comparison of many-particle representations for selected-CI I: A tree based approach. *J. Comput. Chem.* **2021**, *42*, 982–1005.

- (13) Chilkuri, V. G.; Neese, F. Comparison of Many-Particle Representations for Selected Configuration Interaction: II. Numerical Benchmark Calculations. *J. Chem. Theory Comput.* **2021**, *17*, 2868–2885.
- (14) Mato, J.; Gordon, M. S. A general spin-complete spin-flip configuration interaction method. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2615–2626.
- (15) Paldus, J. Matrix elements of unitary group generators in many-fermion correlation problem. I. tensorial approaches. *J. Math. Chem.* **2021**, *59*, 1–36.
- (16) Paldus, J. Matrix elements of unitary group generators in many-fermion correlation problem. II. Graphical methods of spin algebras. *J. Math. Chem.* **2021**, *59*, 37–71.
- (17) Paldus, J. Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems. *J. Chem. Phys.* **1974**, *61*, 5321–5330.
- (18) Paldus, J. A pattern calculus for the unitary group approach to the electronic correlation problem. *Int. J. Quantum Chem.* **1975**, *9*, 165–174.
- (19) Paldus, J. Unitary-group approach to the many-electron correlation problem: Relation of Gelfand and Weyl tableau formulations. *Phys. Rev. A* **1976**, *14*, 1620–1625.
- (20) Paldus, J. *Electrons in Finite and Infinite Structures*; Springer US: 1977; pp 411–429.
- (21) Shavitt, I. Graph theoretical concepts for the unitary group approach to the many-electron correlation problem. *Int. J. Quantum Chem.* **1977**, *12*, 131–148.
- (22) Shavitt, I. Matrix Element Evaluation in the Unitary Group Approach to the Electron Correlation Problem. *Int. J. Quantum Chem.* **1978**, *14*, 5–32.
- (23) Paldus, J.; Boyle, M. J. Unitary Group Approach to the Many-Electron Correlation Problem via Graphical Methods of Spin Algebras. *Phys. Scr.* **1980**, *21*, 295–311.
- (24) Shavitt, I. In *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*; Hinze, J., Ed.; Springer: Berlin and Heidelberg, Germany, 1981; pp 51–99.
- (25) Paldus, J. In *Unitary Group Approach to Many-Electron Correlation Problem*; Hinze, J., Ed.; Springer: Berlin and Heidelberg, Germany, 1981; pp 1–50.
- (26) Paldus, J. In *Theoretical Chemistry Advances and Perspectives*; Eyring, H., Ed.; Academic Press: New York, 1976; Vol. 2.
- (27) Booth, G. H.; Thom, A. J. W.; Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.* **2009**, *131*, 054106–054115.
- (28) Overy, C.; Booth, G. H.; Blunt, N. S.; Shepherd, J. J.; Cleland, D.; Alavi, A. Unbiased Reduced Density Matrices and Electronic Properties from Full Configuration Interaction Quantum Monte Carlo. *J. Chem. Phys.* **2014**, *141*, 244117–244127.
- (29) Blunt, N. S.; Smart, S. D.; Booth, G. H.; Alavi, A. An Excited-State Approach Within Full Configuration Interaction Quantum Monte Carlo. *J. Chem. Phys.* **2015**, *143*, 134117–134125.
- (30) Blunt, N. S.; Booth, G. H.; Alavi, A. Density matrices in full configuration interaction quantum Monte Carlo: Excited states, transition dipole moments, and parallel distribution. *J. Chem. Phys.* **2017**, *146*, 244105–244116.
- (31) Li Manni, G.; Smart, S. D.; Alavi, A. Combining the Complete Active Space Self-Consistent Field Method and the Full Configuration Interaction Quantum Monte Carlo within a Super-CI Framework, with Application to Challenging Metal-Porphyrins. *J. Chem. Theory Comput.* **2016**, *12*, 1245–1258.
- (32) Dobrutz, W.; Weser, O.; Bogdanov, N. A.; Alavi, A.; Li Manni, G. Spin-Pure Stochastic-CASSCF via GUGA-FCIQMC Applied to Iron–Sulfur Clusters. *J. Chem. Theory Comput.* **2021**, *17*, 5684–5703.
- (33) Li Manni, G.; Dobrutz, W.; Alavi, A. Compression of Spin-Adapted Multiconfigurational Wave Functions in Exchange-Coupled Polynuclear Spin Systems. *J. Chem. Theory Comput.* **2020**, *16*, 2202–2215.
- (34) Li Manni, G.; Dobrutz, W.; Bogdanov, N. A.; Guther, K.; Alavi, A. Resolution of Low-Energy States in Spin-Exchange Transition-Metal Clusters: Case Study of Singlet States in [Fe(III)₄S₄] Cubanes. *J. Phys. Chem. A* **2021**, *125*, 4727–4740.
- (35) Li Manni, G. Modeling magnetic interactions in high-valent trinuclear [Mn₃(IV)O₄]₄⁺ complexes through highly compressed multi-configurational wave functions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 19766–19780.
- (36) Dobrutz, W.; Katukuri, V. M.; Bogdanov, N. A.; Kats, D.; Li Manni, G.; Alavi, A. Combined unitary and symmetric group approach applied to low-dimensional spin systems. *ArXiv* **2021**; arXiv:2112.09594.
- (37) Koutecký, J.; Paldus, J.; Čížek, J. Alternancy symmetry: A unified viewpoint. *J. Chem. Phys.* **1985**, *83*, 1722–1735.
- (38) Čížek, J.; Hashimoto, K.; Paldus, J.; Takahashi, M. Relationship between Lieb and Wu Approach and Standard Configuration Interaction Method for the B_{2u} State of the Hubbard Model of Benzene. *Isr. J. Chem.* **1991**, *31*, 423–426.
- (39) Holmes, A. A.; Tubman, N. M.; Umrigar, C. J. Heat-Bath Configuration Interaction: An Efficient Selected Configuration Interaction Algorithm Inspired by Heat-Bath Sampling. *J. Chem. Theory Comput.* **2016**, *12*, 3674–3680.
- (40) Tubman, N. M.; Lee, J.; Takeshita, T. Y.; Head-Gordon, M.; Whaley, K. B. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.* **2016**, *145*, 044112–044118.
- (41) Garniron, Y.; Scemama, A.; Loos, P.-F.; Caffarel, M. Hybrid stochastic-deterministic calculation of the second-order perturbative contribution of multireference perturbation theory. *J. Chem. Phys.* **2017**, *147*, 034101–034109.
- (42) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks. *J. Chem. Theory Comput.* **2018**, *14*, 4360–4379.
- (43) Garniron, Y.; Scemama, A.; Giner, E.; Caffarel, M.; Loos, P.-F. Selected configuration interaction dressed by perturbation. *J. Chem. Phys.* **2018**, *149*, 064103–064108.
- (44) Scemama, A.; Benali, A.; Jacquemin, D.; Caffarel, M.; Loos, P.-F. Excitation energies from diffusion Monte Carlo using selected configuration interaction nodes. *J. Chem. Phys.* **2018**, *149*, 034108–034115.
- (45) Liu, W.; Hoffmann, M. R. iCI: Iterative CI toward full CI. *J. Chem. Theory Comput.* **2016**, *12*, 1169–1178.
- (46) Liu, W.; Hoffmann, M. R. SDS: the static–dynamic–static framework for strongly correlated electrons. *Theor. Chem. Acc.* **2014**, *133*, 1481–1492.
- (47) Ma, D.; Li Manni, G.; Olsen, J.; Gagliardi, L. Second-Order Perturbation Theory for Generalized Active Space Self-Consistent-Field Wave Functions. *J. Chem. Theory Comput.* **2016**, *12*, 3208–3213.
- (48) Weser, O.; Freitag, L.; Guther, K.; Alavi, A.; Li Manni, G. Chemical insights into the electronic structure of Fe(II) porphyrin using FCIQMC, DMRG, and generalized active spaces. *Int. J. Quantum Chem.* **2021**, *121*, 26454–26467.
- (49) Weser, O.; Guther, K.; Ghanem, K.; Li Manni, G. Stochastic Generalized Active Space Self-Consistent Field: Theory and Application. *J. Chem. Theory Comput.* **2022**, *18*, 251–272.
- (50) Odoh, S. O.; Li Manni, G.; Carlson, R. K.; Truhlar, D. G.; Gagliardi, L. Separated-Pair Approximation and Separated-Pair Pair-Density Functional Theory. *Chem. Sci.* **2016**, *7*, 2399–2413.
- (51) Gould, M. D.; Paldus, J. Spin-dependent Unitary Group Approach. I. General Formalism. *J. Chem. Phys.* **1990**, *92*, 7394–7401.
- (52) Gould, M. D.; Paldus, J.; Chandler, G. S. Unitary Group Approach to Reduced Density Matrices. *J. Chem. Phys.* **1990**, *93*, 4142–4153.
- (53) Gould, M. D.; Battle, J. S. Spin-dependent Unitary Group Approach. II. Derivation of Matrix Elements for Spin-dependent Operators. *J. Chem. Phys.* **1993**, *99*, 5961–5975.
- (54) Booth, G. H.; Cleland, D. M.; Thom, A. J. W.; Alavi, A. Breaking the Carbon Dimer. *J. Chem. Phys.* **2011**, *135*, 084104–084117.

- (55) Fales, B. S.; Hohenstein, E. G.; Levine, B. G. Robust and Efficient Spin Purification for Determinantal Configuration Interaction. *J. Chem. Theory Comput.* **2017**, *13*, 4162–4172.
- (56) Marti, K. H.; Ondík, I. M.; Moritz, G.; Reiher, M. Density matrix renormalization group calculations on relative energies of transition metal complexes and clusters. *J. Chem. Phys.* **2008**, *128*, 014104–014116.
- (57) Li, Z.; Chan, G. K.-L. Hilbert space renormalization for the many-electron problem. *J. Chem. Phys.* **2016**, *144*, 084103–084125.
- (58) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd.: 2000.
- (59) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. Density Matrix Averaged Atomic Natural Orbital (ANO) Basis Sets for Correlated Molecular Wave Functions. *Theor. Chem. Acc.* **1990**, *77*, 291–306.
- (60) Roos, B. O.; Malmqvist, P.-Å. Relativistic Quantum Chemistry: The Multiconfigurational Approach. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2919–2927.
- (61) Pal, S.; Chan, M. K.; Armstrong, W. H. Ground spin state variability in manganese oxo aggregates. Demonstration of an $S = 3/2$ ground state for $[\text{Mn}_3\text{O}_4(\text{OH})(\text{bpea})_3](\text{ClO}_4)_3$. *J. Am. Chem. Soc.* **1992**, *114*, 6398–6406.
- (62) Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; De Vico, L.; Fdez. Galván, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; et al. Molcas 8: New Capabilities for Multiconfigurational Quantum Chemical Calculations Across the Periodic Table. *J. Comput. Chem.* **2016**, *37*, 506–541.
- (63) Fdez. Galván, I.; Vacher, M.; Alavi, A.; Angeli, C.; Aquilante, F.; Autschbach, J.; Bao, J. J.; Bokarev, S. I.; Bogdanov, N. A.; Carlson, R. K.; et al. OpenMolcas: From Source Code to Insight. *J. Chem. Theory Comput.* **2019**, *15*, 5925–5964.
- (64) Sarneski, J. E.; Thorp, H. H.; Brudvig, G. W.; Crabtree, R. H.; Schulte, G. K. Assembly of high-valent oxomanganese clusters in aqueous solution. Redox equilibrium of water-stable $\text{Mn}_3\text{O}_4^{4+}$ and $\text{Mn}_2\text{O}_2^{3+}$ complexes. *J. Am. Chem. Soc.* **1990**, *112*, 7255–7260.
- (65) Anderson, R. J.; Shiozaki, T.; Booth, G. H. Efficient and stochastic multireference perturbation theory for large active spaces within a full configuration interaction quantum Monte Carlo framework. *J. Chem. Phys.* **2020**, *152*, 054101–054116.
- (66) Halson, J. J.; Anderson, R. J.; Booth, G. H. Improved stochastic multireference perturbation theory for correlated systems with large active spaces. *Mol. Phys.* **2020**, *118*, e1802072–1802084.
- (67) Boys, S. F.; Handy, N. C.; Linnett, J. W. The determination of energies and wavefunctions with full electronic correlation. *Proc. R. Soc. A* **1969**, *310*, 43–61.
- (68) Boys, S. F.; Handy, N. C.; Linnett, J. W. A calculation for the energies and wavefunctions for states of neon with full electronic correlation accuracy. *Proc. R. Soc. A* **1969**, *310*, 63–78.
- (69) Boys, S. F.; Handy, N. C.; Linnett, J. W. A first solution, for LiH, of a molecular transcorrelated wave equation by means of restricted numerical integration. *Proc. R. Soc. A* **1969**, *311*, 309–329.
- (70) Handy, N. C. Energies and Expectation Values for Be by the Transcorrelated Method. *J. Chem. Phys.* **1969**, *51*, 3205–3212.
- (71) Handy, N. The transcorrelated method for accurate correlation energies using gaussian-type functions: examples on He, H_2 , LiH and H_2O . *Mol. Phys.* **1972**, *23*, 1–27.
- (72) Luo, H.; Alavi, A. Combining the Transcorrelated Method with Full Configuration Interaction Quantum Monte Carlo: Application to the Homogeneous Electron Gas. *J. Chem. Theory Comput.* **2018**, *14*, 1403–1411.
- (73) Cohen, A. J.; Luo, H.; Guther, K.; Dobrautz, W.; Tew, D. P.; Alavi, A. Similarity transformation of the electronic Schrödinger equation via Jastrow factorization. *J. Chem. Phys.* **2019**, *151*, 061101–061105.
- (74) Guther, K.; Cohen, A. J.; Luo, H.; Alavi, A. Binding curve of the beryllium dimer using similarity-transformed FCIQMC: Spectroscopic accuracy with triple-zeta basis sets. *J. Chem. Phys.* **2021**, *155*, 011102–011109.
- (75) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669–3680.
- (76) Gagliardi, L.; Truhlar, D. G.; Li Manni, G.; Carlson, R. K.; Hoyer, C. E.; Bao, J. L. Multiconfiguration Pair-Density Functional Theory: A New Way to Treat Strongly Correlated Systems. *Acc. Chem. Res.* **2017**, *50*, 66–73.
- (77) Vitale, E.; Alavi, A.; Kats, D. FCIQMC-Tailored Distinguishable Cluster Approach **2020**, *16*, 5621.
- (78) Saxena, S.; Agarwal, P.; Ahilan, K.; Grosche, F. M.; Haselwimmer, R. K. W.; Steiner, M. J.; Pugh, E.; et al. Superconductivity on the border of itinerant-electron ferromagnetism in UGe_2 . *Nature* **2000**, *406*, 587–592.
- (79) Aoki, D.; Huxley, A.; Ressouche, E.; Braithwaite, D.; Flouquet, J.; Brison, J.-P.; Lhotel, E.; Paulsen, C. Coexistence of superconductivity and ferromagnetism in URhGe . *Nature* **2001**, *413*, 613–616.
- (80) Arbenz, P. *Lecture Notes on Solving Large Scale Eigenvalue Problems*; <https://people.inf.ethz.ch/arbenz/ewp/Lnotes/lsevp.pdf>; Computer Science Department ETH Zürich: 2016.