Excited electronic states of molecules and solid-state systems play a crucial role in determining their chemical properties, including their response to irradiation by light and their behavior at finite temperature. But calculating excited-state properties from first principles often proves challenging, particularly for systems of strongly correlated electrons. Rigorous full configuration interaction (FCI) calculations for such systems require the treatment of a number of electronic configurations that scales combinatorially with the number of electrons and the number of orbitals they may occupy, which has prevented their application to all but the smallest chemical systems. This has led to the development of active space techniques, such as complete active space configuration interaction (CASSCF) or its orbital-optimized extension via a self-consistent field (CASSCF), which are limited to systems containing about 20 electrons occupying 20 spatial orbitals. Larger active spaces can be accurately treated with the density matrix renormalization group (DMRG) and selected CI methods, which can be used to calculate excited states of systems containing about 50 active orbitals.

Quantum Monte Carlo (QMC) methods enable an alternative, efficient treatment of strong electronic correlation in large systems. QMC methods leverage stochastic sampling to iteratively evolve a quantum state toward the minimum of an energy functional in a specified manifold. Because QMC is fundamentally a stochastic minimization procedure, it is used routinely to calculate ground-state properties. In some cases, this procedure can be straightforwardly extended to calculate excited-state properties if symmetry considerations are used to exclude the ground state from the manifold of possible solutions.

motivated by these challenges, we recently introduced an alternative, general stochastic approach to calculating matrix eigenvalues, referred to as subspace iteration with repeated random sparsification. Introducing randomness into standard subspace iteration presents issues related to maintaining orthogonality among multiple vectors and estimating energy eigenvalues as the iteration proceeds. These necessitate careful algorithmic choices that are described in detail in ref. In particular, we choose to randomize a nonstandard, asymmet-

**Full Configuration Interaction Excited-State Energies in Large Active Spaces from Randomized Subspace Iteration**

Samuel M. Greene,\(^1\) Robert J. Webber,\(^2\) a) James E. T. Smith,\(^3\) Jonathan Weare,\(^2,\) b) and Timothy C. Berkelbach\(^1,\) c)

\(^1\)Department of Chemistry, Columbia University, New York, New York 10027, United States
\(^2\)Courant Institute of Mathematical Sciences, New York University, New York, New York 10012, United States
\(^3\)Center for Computational Quantum Physics, Flatiron Institute, New York, New York 10010, United States

We present a stable and systematically improvable quantum Monte Carlo (QMC) approach to calculating excited-state energies, which we implement using our fast randomized iteration method for the full configuration interaction problem (FCI-FRI). Unlike previous excited-state quantum Monte Carlo methods, our approach, which is an asymmetric variant of subspace iteration, avoids the use of dot products of random vectors and instead relies upon trial vectors to maintain orthogonality and estimate eigenvalues. By leveraging recent advances, we apply our method to calculate ground- and excited-state energies of strongly correlated molecular systems in large active spaces, including the carbon dimer with 8 electrons in 108 orbitals \((8e,108o)\), an oxo-Mn(salen) transition metal complex \((28e,28o)\), ozone \((18e,87o)\), and butadiene \((22e,82o)\). In the majority of these test cases, our approach yields total excited-state energies that agree with those from state-of-the-art methods—including heat-bath CI, the density matrix renormalization group approach, and FCIQMC—to within sub-milliHartree accuracy. In all cases, estimated excitation energies agree to within about 0.1 eV.
ric subspace iteration. Unlike previous approaches, our randomized subspace iteration is stable and avoids the use of dot products of random vectors. We use approximate eigenvectors (i.e. trial vectors) to estimate energies and maintain orthogonality. In ref 33, we applied this approach to the full configuration interaction problem using a simple implementation of our fast randomized iteration approach (FCI-FRI)\(^{35-37}\) to accelerate matrix-vector multiplications. Like FCIQMC and other discrete-space QMC methods, FCI-FRI is a stochastic implementation of an iterative linear algebra scheme involving sequential matrix-vector multiplication operations. Stochastic sampling is used to impose sparsity in vectors and matrices, thereby reducing the computational cost of these operations. In ref 33, we successfully applied our randomized subspace iteration scheme to calculate ground- and excited-state energies of three small molecular systems.\(^{33}\) However, the simple flavor of FCI-FRI that we used prevented us from studying larger systems. In this work, we apply several strategies, described below, that enable application to larger, more challenging molecular systems, including some strategies developed for ground-state FCI-FRI calculations.\(^{37}\) We will refer to the resulting excited-state scheme as FCI-FRI, but we emphasize that the general, randomized subspace iteration described in ref 33 can be applied with other QMC techniques, including FCIQMC. To our knowledge, the FCI-FRI calculations presented here are the largest to date performed with discrete-space QMC for excited states in the same symmetry class.

We focus on the calculation of eigenvalues of a matrix representation \(\mathbf{H}\) of the Hamiltonian operator for \(N\) interacting electrons in a discrete many-particle basis constructed from \(M\) single-particle orbitals. In this work, we use orbitals obtained from an approximate CASSCF calculation using the heat-bath configuration interaction (HCI) algorithm,\(^{38}\) as described in Appendix A. In contrast to our previous FCI-FRI papers, in which we used a discrete basis of Slater determinants, here we instead use spin-coupled functions.\(^{39}\) This imposes a block-diagonal structure on \(\mathbf{H}\), effectively reducing the dimension of the eigenproblem and thereby reducing the computational cost of our method. Each block contains only eigenstates with a particular spin parity, i.e. for which the spin \(S\) is either even or odd. The number of spin-coupled functions for a given spin parity, denoted generically as \(N_{\text{CI}}\), scales as \(O(M \text{ choose } N)\). The number of nonzero elements in each column of \(\mathbf{H}\), which determines the cost of performing sparse matrix-vector multiplication operations, scales as \(O(N^2M^2)\). Definitions of spin-coupled functions and formulas for elements of \(\mathbf{H}\) in this basis are provided in Appendix B. For some systems, we leveraged point-group symmetry to impose additional block-diagonal structure and further reduce the effective dimension. Symmetry labels, Hamiltonian matrix elements, and Hartree-Fock orbitals were calculated using the PySCF software.\(^{40}\)

We first describe subspace iteration applied to the calculation of the \(N_{\text{eig}}\) lowest-energy eigenvalues in a single block of \(\mathbf{H}\). In each iteration \(\tau\), one calculates a \(N_{\text{CI}} \times N_{\text{eig}}\) iterate matrix \(\mathbf{X}^{(\tau)}\), whose columns each approximate an eigenvector of \(\mathbf{H}\). Starting from an initial matrix \(\mathbf{X}^{(0)} = \mathbf{U}\), subsequent iterates are constructed as\(^{41,42}\)

\[
\mathbf{X}^{(\tau+1)} = (\mathbf{I} - \varepsilon \mathbf{H}) \mathbf{X}^{(\tau)} \left[ \mathbf{G}^{(\tau)} \right]^{-1}
\]

where \(\varepsilon\) is a small, positive number, \(\mathbf{I}\) is the identity matrix, and \(\mathbf{G}^{(\tau)}\) is chosen to approximately enforce orthonormality among columns of iterates, as described in Appendix C. For \(N_{\text{eig}} = 1\), this is equivalent to power iteration, which underlies many discrete-space QMC methods. Our approach to orthonormalization (via multiplication by \(\mathbf{G}^{(\tau)}\)) is similar to the use of a dynamically adjusted energy shift in other QMC methods\(^{20,43}\) but was found to offer better stability for our excited-state calculations. If \(\varepsilon\) is sufficiently small, the iterates will converge to the space of the \(N_{\text{eig}}\) lowest-energy eigenvectors as \(\tau \to \infty\)\(^{44}\) and energy eigenvalues can be calculated.

Applying this subspace iteration in its deterministic form to the chemical systems of interest in this work is intractable because of the size of \(N_{\text{CI}}\) and the associated memory and CPU costs. The FCI-FRI framework addresses this challenge by using stochastic compression to impose sparsity and reduce the cost of matrix multiplication. Defining the compression operator \(\Phi\), a stochastically compressed vector \(\Phi(x)\) has elements that equal those of the input vector \(x\) in expectation (i.e. \(\mathbb{E}[\Phi(x)] = x\)) and has at most \(m\) nonzero elements, where \(m\) is a tunable parameter. We use a specific stochastic compression scheme known as pivotal compression. In many applications, this scheme achieves low statistical error, as confirmed by both theoretical analysis and numerical experiments.\(^{33}\) In this scheme, a number \(d\) of the largest-magnitude elements in the input vector \(x\) are left unchanged in compression, where \(d\) is determined by an algorithm that depends both on \(m\) and on the relative magnitudes of elements in \(x\). A number \((m - d)\) of the remaining elements are randomly selected to be nonzero in the compressed vector, according to a pivotal resampling scheme that enforces statistical correlations among the elements.\(^{45-47}\) The probability of selecting each element is proportional to its magnitude. Elements not selected are zero in \(\Phi(x)\). In general, the statistical error incurred in compression decreases as \(m\) is increased, and it is zero if \(m\) equals or exceeds the number of nonzero elements in \(x\). This compression scheme is used to calculate the matrix product \((1 - \varepsilon \mathbf{H}) \mathbf{X}^{(\tau)}\) in tandem with our “unnormalized” heat-bath Power-Pitzer Hamiltonian factorization scheme\(^{37,48}\) and the initiator approximation.\(^{25,49}\) With compression performed in this way, the computational scaling of a single iteration of eq 1 is reduced from \(O(N_{\text{eig}}N^2M^2N_{\text{CI}})\) to \(O(N_{\text{eig}}Mm)\). Details of the pivotal compression scheme are provided in ref 33 and details of the current FCI-FRI implementation are provided in Sections 4 and 5 of the Supporting Information.

In our randomized subspace iteration, we use averaging to obtain eigenvalue estimates, as the random iterates only represent the lowest-energy eigenvectors on average. At regular intervals (every 100 iterations in this work), we evaluate and store the small \(N_{\text{eig}} \times N_{\text{eig}}\) matrices \(\mathbf{U}^T \mathbf{H} \mathbf{X}^{(\tau)}\) and \(\mathbf{U}^T \mathbf{X}^{(\tau)}\). Denoting averages of these matrices as \(\langle \mathbf{U}^T \mathbf{H} \mathbf{X}^{(\tau)} \rangle\) and \(\langle \mathbf{U}^T \mathbf{X}^{(\tau)} \rangle\), we solve the generalized eigenvalue problem

\[
\langle \mathbf{U}^T \mathbf{H} \mathbf{X}^{(\tau)} \rangle \mathbf{w} = \langle \mathbf{U}^T \mathbf{X}^{(\tau)} \rangle \mathbf{w} \mathbf{A}
\]
to obtain a diagonal matrix $\Lambda$ of eigenvalue estimates. The standard error of each estimate can be approximated as described in Section 6 of the Supporting Information. Eq (2) can be understood as a generalization of the projected energy estimator commonly used in other QMC methods \(^{18,20}\) to multiple eigenvalues. A related eigenvalue estimator is also used in the QMC method first proposed in ref 50, although in that method elements of the analogous matrices are evaluated by analyzing correlations within a single trajectory; here, we instead use multiple orthogonal trajectories. Initial iterations are excluded from the averages in eq (2) in order to ensure sufficient equilibration. This particular approach to estimating eigenvalues was chosen to mitigate statistical biases arising from nonlinearities in the eigenvalue equation.\(^{33}\)

Eigenvalue estimates obtained by eq (2) are exact in two limiting cases: with infinitely many samples (in which case $(\mathbf{X}^T\mathbf{X})^{-1}$ exactly spans the eigenvectors) or for any eigenvector exactly contained in the column span of the matrix $U$ of trial vectors. Due to the latter property, we expect better eigenvalue estimates when more accurate trial vectors are used. In this work, we use accurate, multideterminantal trial vectors obtained by HCI.\(^{51}\) Specifically, we work in a basis of approximate HCISCF orbitals\(^{38}\) and use up to 6 million determinants per trial vector. This number of determinants $N_d$ is limited by the cost of evaluating the matrices in the eigenvalue eq 2, which has a computational scaling of $O(N_dN^2M^2)$. Further details about our orbital basis and trial vectors are described in Section 1 of the Supporting Information.

We first applied our FCI-FRI subspace iteration to calculate eigenenergies of the carbon dimer ($C_2$). Others have previously used $C_2$ as a rigorous test case for new quantum chemistry methods due to the significant multireference character of its lowest-energy eigenstates.\(^{51,52}\) Here we use a large cc-pVQZ basis\(^{53}\) and correlate all valence electrons in all orbitals (freezing core electrons), resulting in a CI problem of 8 electrons in 108 spatial orbitals (8e,108o). We focus here on the Hamiltonian block containing states with even-spin (singlet, quintet, etc.) and $A_g$ symmetry in the $D_2h$ point group. Although it is possible to impose additional block-diagonal structure on the Hamiltonian by leveraging the full $D_{oh}$ symmetry of $C_2$,\(^{51}\) we do not employ such an approach here.

Table I shows eigenenergy estimates for the six lowest-energy states within this Hamiltonian block for $C_2$ at its equilibrium geometry, i.e. with an internuclear separation $r_{C-C}$ of 1.24253 Å. These were obtained by applying FCI-FRI with an initiator threshold of $n_a = 1$ and four different values of $m$, the number of nonzero elements used in stochastic compression operations for each iterate column. Standard error estimates for all calculations with $m \geq 1 \times 10^6$ are less than 0.1 m$E_h$, while error estimates for $m = 1 \times 10^5$ are 0.66 m$E_h$ or less. Therefore, the discrepancies between estimates for different values of $m$ primarily result from the statistical biases associated with our normalization and orthogonalization procedures, as well as the initiator approximation. The magnitudes of all of these biases are expected to decrease with increasing $m$, as evidenced here by the convergence in the energy for each state as $m$ is increased.

Estimates obtained using $m = 1 \times 10^6$ differ from those obtained using $m = 8 \times 10^6$ by less than 4 m$E_h$, suggesting that our approach can yield results of acceptable accuracy even for calculations not yet converged with respect to $m$. Table I also presents results from three previous state-of-the-art calculations on $C_2$ in the same cc-pVQZ basis, using FCIQMC,\(^{31}\) FCI/QMC,\(^{31}\) and DMRG.\(^{52}\) Because these previous calculations leveraged the full $D_{oh}$ symmetry of $C_2$, they excluded some of the eigenstates considered in this work. At the greatest value of $m$ we tested, $m = 8 \times 10^6$, our total energy estimates differ from these previous calculations by at most 0.3 m$E_h$, leading to excitation energies that agree to 0.01 eV or better.

We additionally considered $C_2$ in the same cc-pVQZ basis at a nonequilibrium “stretched” geometry, with $r_{C-C} = 2.0$ Å. Differences between subsequent eigenenergies are smaller at this geometry than at the equilibrium geometry, which makes it more difficult to obtain accurate energy estimates using randomized subspace iteration.\(^{33}\) This system therefore constitutes an even more rigorous test for our approach than equilibrium $C_2$. Accordingly, we could not obtain energy estimates with $m = 1 \times 10^6$ due to numerical instabilities caused by the statistical error, a phenomenon that we have observed previously.\(^{33}\) Standard errors for all other estimates are less than 0.1 m$E_h$. Energies obtained with $m = 8 \times 10^6$ differed from those from previous calculations by at most 0.3 m$E_h$.

Estimates for the $4^1A_g$ state exhibit the greatest sensitivity to the value of $m$. The energy calculated with $m = 1 \times 10^6$ differs from that from $m = 8 \times 10^6$ calculation by 23.92 m$E_h$. We suspect that the increased sensitivity of estimates for this state is caused by the presence of a higher-lying state close in energy. In the deterministic implementation of our algorithm, each energy estimate $\mathbf{\lambda}_k$ converges at a rate proportional to $[(1-\varepsilon E_k)/(1-\varepsilon E_{(N_s+1)})]^T$, where $E_k$ denotes the exact eigenenergy for the $k^{\text{th}}$ state.\(^{33}\) Consequently, the rate of convergence is determined by the energy gap between the considered low-energy subspace and the higher-lying eigenenergies. Although a similar analysis for the randomized algorithm is more complicated, one can reasonably expect that eigenvalues that converge more slowly in the deterministic algorithm are more susceptible to statistical fluctuations in the randomized algorithm, which are larger at lesser values of $m$. These statistical fluctuations can give rise to a greater statistical bias, as is observed here.

As another application of our approach, we considered an oxo-Mn(salen) complex. Manganese(salen) complexes in general are used as catalysts for enantioselective epoxidation of alkenes.\(^{54–58}\) The mechanisms of such reactions are not yet fully understood, as various mechanisms have been observed under different reaction conditions and for different alkene reactants.\(^{59–64}\) A previous theoretical study has suggested that the spin state of the catalyst can play a crucial role in determining the mechanism.\(^{65}\) Because the catalyst’s singlet, triplet, and quintet states are often close in energy, accurately predicting their relative energies at a given geometry is a crucial prerequisite for mechanistic studies.\(^{66}\) Because these states also exhibit strong multireference character, accurately calculating their energies by electronic structure theory is difficult. This challenge has prompted several theoretical in-
vestigations of the electronic structure of these catalysts.  

Here, following previous theoretical studies, we calculate the eigenenergies of a model complex (Figure 1) with a similar geometry reported in ref 66. Following previous studies, we used the 6-31G* basis and treated only an active subset of the singlet. Previous calculations on this system used active spaces of 5.0 kcal/mol, in good agreement with our FCI-FRI predictions. Although we cannot rule out that our HCl or FCI-FRI results are unconverged, their mutual agreement is encouraging.

We next applied our method to ozone (O₃), which plays important roles in Earth’s atmosphere due to its presence in smog and its role in scattering ultraviolet light in the stratosphere. Ozone has proven particularly challenging for conventional electronic structure methods, in part due to its multireference nature and the large magnitude of its correlation energy relative to that of other energetic properties. Nevertheless, previous theoretical studies have suggested the presence of a metastable “ring minimum” structure of D₃h symmetry as an intermediate in the photochemical decomposition of ozone to O₂ + O. This structure has yet to be observed experimentally, and whether or not it is a stable intermediate depends crucially on its energy relative to the equilibrium structure. In order to investigate possible formation pathways for this metastable structure, Chien et al. used HCl to calculate the energies of the two lowest-energy electronic states at the equilibrium geometry (denoted OM), the metastable geometry (RM), and the transition state separating the two (TS). Here we calculate energies at these same geometries (as reported in ref 79) and compare our results. We considered the singlet ground state and first singlet excited state of A₁ symmetry. Following ref 79, we used a cc-pVTZ basis and correlated all valence electrons in all possible orbitals, yielding a (18s,28s) active space. Estimates of the energies of the two lowest-energy even-spin eigenstates at each geometry, obtained using FCI-FRI, are shown in Figure 2.

We report our results in terms of energy differences, rather than absolute energies, to enable direct comparisons with ref 79. Because energy differences were reported to an accuracy of 0.01 eV in ref 79, we represent these results as shaded regions with widths of 0.01 eV in Figure 2. Standard errors for all energy difference estimates from FCI-FRI are less than 6 × 10⁻⁵ eV. At the OM geometry, our estimated energy difference calculated with m = 80 million agreed to within 0.01 eV (0.4 mEₜₕ) of the HCl result from ref 79.
between FCI-FRI and HCI estimates were greater at the RM geometry: our best estimate (at \( m = 80 \) million) differed from the HCI energy difference by 0.18 eV (6.6 \( mE_b \)), but FCI-FRI results can be seen to be unconverted with respect to \( m \). Estimates of the difference between the ground-state energies at the OM and RM geometries are less sensitive to the value of \( m \). All estimates, except the one at \( m = 20 \) million, agree with each other to within 0.002 eV and with the HCI result to within 0.01 eV. These results reflect the fact that the energy of the RM 2 \(^1\)A\(_1\) state is most sensitive to the value of \( m \), perhaps due to its sign structure or sparsity structure.

Applying our method to estimate the two lowest-energy eigenvalues at the TS geometry yielded a conjugate pair of complex eigenvalue estimates, as is possible since the matrices \((U^\dagger HX^{(\tau)})_s\) and \((U^\dagger X^{(\tau)})_s\) in the eigenvalue equation (2) are real and nonsymmetric. The appearance of complex eigenvalues is likely due to the fact that the difference of the two lowest-energy eigenvalues for this system (0.01 eV as estimated by HCI\(^79\)) is substantially less than that for the OM or RM systems. In order to better understand how small energy differences can lead to complex eigenvalue estimates, we first recognize that the estimates obtained by solving eq (2) are equivalent to the eigenvalues of the matrix \((U^\dagger HX^{(\tau)})_s(U^\dagger X^{(\tau)})_s^{\dagger}\). Denoting the entries of this \( 2 \times 2 \) matrix as \( \alpha_{ij} \), we recall that its eigenvalues are real as long as \((\alpha_{11} - \alpha_{22})^2 + 4\alpha_{21}\alpha_{12} > 0 \). We numerically find that this inequality is violated for the matrices obtained from our calculations with the TS geometry but not with the OM and RM geometries, mainly because the two diagonal elements of \((U^\dagger HX^{(\tau)})_s(U^\dagger X^{(\tau)})_s^{\dagger}\) are closer in value in our TS calculations than in our OM and RM calculations. This likely reflects the near-degeneracy of the exact eigenvalues for the TS system. The FCI-FRI eigenvalue estimates for the TS system will be real for \( m \) sufficiently large, since our method becomes exact as \( m \) approaches the dimension of the Hamiltonian matrix. Indeed, we find that as \( m \) is increased from 10 million to 40 million, the magnitude of the imaginary part of each eigenvalue decreases. Considering only the real part of our TS eigenvalue estimates at \( m = 40 \) million and subtracting our corresponding OM 1 \(^1\)A\(_1\) energy estimate, we estimate the difference of the TS and OM ground-state energies to be 2.44 eV. This compares favorably to the 2.41 eV difference.
The excited-state properties of such materials critically determine their performance in optoelectronic applications. However, accurately characterizing the excited-state electronic structure of even simple conjugated molecules by theoretical means has proven challenging. For these reasons, conjugated organic systems are rigorous tests of electronic structure methods, and their accuracy has important implications for next-generation materials. The excited-state electronic structure of trans-butadiene has been studied by HCl, DMRG, and FCIQMC. For consistency between these literature results and our calculations, we use the geometry reported in ref 80 and the ANO-L-pVDZ basis and correlated all valence electrons in all orbitals, yielding a (22e,82o) active space.

Trial vectors were calculated as described above and used in subsequent FCI-FRI calculations of the two lowest-energy eigenvalues. Results from these calculations are presented in Figure 3. Analysis of our results in comparison to previous calculations suggests that this is a more difficult system than those discussed above. Our estimates of the 1 \(^1\)Ag ground-state energy are relatively insensitive to the value of m, differing by only 0.1 mEh. Although our estimate at m = 40 million differs from the FCIQMC estimate by only 0.4 mEh, it exhibits greater discrepancies (8 mEh) with HCl and DMRG estimates. (We compare to the DMRG estimate reported in ref 79, which was obtained by extrapolating the results in ref 15.) The uncertainties in these estimates are all reported to be less than 0.1 mEh, so these discrepancies suggest the presence of systematic errors. The authors of ref 79 speculated that the FCIQMC energy is an overestimate due to errors from the initiator approximation and to potential inaccuracies in the reported uncertainty. Given the similarities between our approach and FCIQMC, as well as the similarities in our estimates, it is likely that our energy is an overestimate for the same reasons.

Estimates of the 2 \(^1\)Ag eigenvalue exhibit greater sensitivity to the value of m. Our estimate at m = 40 million is 73 mEh greater than at m = 10 million. This estimate also differs from the HCl estimate by 14 mEh, likely due to the reasons discussed above. Nonetheless, at our largest value of m, the FCI-FRI excitation energy is calculated to be 6.72 eV, which is in reasonable agreement with the HCl value of 6.58 eV.

In summary, we presented a general and systematically improvable strategy for calculating excited-state energies of strongly correlated electronic systems in large active spaces within the FCI-FRI framework. Unlike previous “replica” methods for excited-state calculations, our asymmetric, randomized subspace iteration avoids the calculation of dot products of random vectors and instead uses approximate trial vectors to enforce orthogonality and estimate energies. We expect that this feature will enable the reliable estimation of excited-state energies for large systems. Applying our method to the carbon dimer (C\(_2\)) in a cc-pVQZ basis at two different geometries yielded energy estimates within 0.3 mEh of those from previous calculations. Estimates for a oxo-Mn(salen) complex differed from our independent HCl calculations by up to 2 mEh. Discrepancies for the ozone and butadiene molecules were greater (7 mEh and 14 mEh, respectively), possibly due to strong correlation effects in these systems. These total energy discrepancies translate to excitation energy discrepancies of about 0.1 eV or less.

Additional developments could enable further reductions in the errors and computational cost of FCI-FRI, thereby enabling its application to systems even larger than those considered here. The largest calculations in this work required 7-14 days of execution time on 448 cores to achieve reliable convergence. The time required to calculate the matrix products \(U^T X^{(T)}\) constituted a significant portion of overall execution time, due to the large number of nonzero elements in both U and X\(^{(T)}\). Future work could involve developing strategies for reducing this cost, for example by factorizing the matrix H. It may be possible to obtain more compact forms of the trial vectors without sacrificing accuracy, such as through transformations of the single-particle basis.

More systematic research is needed to better understand how the various parameters used in our calculations affect the accuracy of our estimates. Such investigation could lead to techniques for automating the selection of certain parameters, following previous preliminary studies in the context of FCIQMC. We found that insufficient sampling in our approach leads to an uncontrolled increase in the condition number of the matrix \(U^T X^{(T)}\), so monitoring this condition number could serve as a diagnostic tool. This could also lead to strategies for choosing different initiator thresholds for each state of interest in a calculation, thus enabling further reductions in the bias introduced by the initiator approximation. Notwithstanding these remaining challenges, the results of our calculations in this work suggest the effectiveness of the features of both randomized subspace iteration and FCI-FRI for enabling the treatment of large chemical systems.
the generality of our approach to excited-state calculations, some of these features could be incorporated into other QMC schemes, such as FCIQMC, auxiliary-field QMC, or diffusion Monte Carlo.

**ACKNOWLEDGMENTS**

The authors thank Aaron Dinner, Michael Lindsey, and Verena Neufeld for useful conversations and Benjamin Pritchard for his suggestions for improving the readability and performance of our code. S.M.G. was supported by a software fellowship from the Molecular Sciences Software Institute, which is funded by the U.S. National Science Foundation grant OAC-1547580. R.W. is supported by New York University’s Dean’s Dissertation Fellowship and by the National Science Foundation through award DMS-1646339. J.W. acknowledges support from the Advanced Scientific Computing Research Program within the DOE Office of Science through award DE-SC0020427. The Flatiron Institute is a division of the Simons Foundation.

**Appendix A: Orbital and Trial Vector Construction**

Because our goal in this work is to estimate eigenvalues of challenging strongly correlated systems to high accuracy, we seek to obtain more accurate trial vectors than in our previous studies. To this end, we employ a selected configuration interaction method, as is increasingly done in other QMC methods. Generically, selected configuration interaction involves constructing a variational subspace of Slater determinants determined to contribute significantly to the eigenvectors for our purposes as well. After calculating the eigenvectors for our purposes as well. After calculating the second-order perturbation theory contribution to the energy, HCl has already been demonstrated to yield eigenenergies of sub-millihartree accuracy for a variety of challenging chemical systems, so we expect it to yield accurate eigenvectors for our purposes as well. After calculating the eigenvectors of interest via a variational HCl calculation, we project them into a basis of spin-coupled functions according to eq (B1) or (B2).

Here we provide more specific details on how we chose the orbital basis and calculated the matrix $U$ of trial vectors for each of the chemical systems considered in this work. We begin by performing an inexpensive state-averaged HCISCF calculation, i.e., with a small variational subspace. The value of $\epsilon_1$, which determines the size of this subspace in HCl, was chosen to be 5 mEh (millihartrees) for the oxo-Mn(Salen) system and 0.2 mEh for ozone and butadiene; for C$_2$, we skipped the HCISCF step. The initial active space orbitals were obtained from unrestricted Hartree-Fock natural orbitals (for oxo-Mn(Salen)) or second-order Möller-Plesset natural orbitals (for ozone and butadiene). We then transformed all active orbitals into the basis of natural orbitals calculated at HCISCF convergence (or from a one-shot, inexpensive HCl calculation with $\epsilon_1 = 5$ mEh in the case of C$_2$). This orbital basis was used for more accurate HCl calculations (to obtain trial vectors) and in subsequent FCI-FRI calculations because both methods benefit from the sparsity it provides. The value of $\epsilon_1$ and the size of the resulting variational subspace for these more accurate HCl calculations are given in Table II for each system studied. These subspaces are smaller than those typically used in state-of-the-art HCl calculations because otherwise the cost of evaluating the matrix product $U^T H X$ is prohibitive. (Note that the matrix $U^T H$ is too large to store and so we reevaluate its entries on-the-fly.)

**Appendix B: Spin-Coupled Functions and the Hamiltonian Matrix**

This section provides formulas for elements of the Hamiltonian matrix in a basis of spin-coupled functions, which are defined as linear combinations of Slater determinants. Here we use the notation $|J>$ to denote a generic spin-coupled function, constructed from a Slater determinant $|J>$ and possibly $\hat{T}|J>$, where $\hat{T}$ is the time-reversal operator that exchanges spin-up and spin-down electrons. Spin-coupled functions in the even spin parity block of the Hamiltonian are denoted $|J>_e$ and have the form

$$|J>_e = \begin{cases} |J>, & \hat{T}|J> = |J> \\ 2^{-1/2} (|J> + \hat{T}|J>), & \hat{T}|J> \neq |J> \end{cases} \tag{B1}$$

Spin-coupled functions in the odd spin parity block are denoted $|J>_o$ and have the form

$$|J>_o = 2^{-1/2} (|J> - \hat{T}|J>) \tag{B2}$$

These can only be comprised of Slater determinants $|J>$ for which $\hat{T}|J> \neq |J>$. Combining eqs (B1) and (B2) with the Slater-Condon rules yields formulas for the elements of the Hamiltonian matrix $H$. 

**TABLE II. Parameters defining the sizes of the variational subspaces for HCl calculations used to construct trial vectors. $\epsilon_1$ indicates the parameter in HCl that determines the size of the subspace, as defined in ref 16, and $N_{HCl}$ denotes the number of determinants in the final subspace.**

| System                  | $\epsilon_1$ (mEh) | $N_{HCl}$/10$^6$ |
|-------------------------|---------------------|------------------|
| equilibrium C$_2$        | 0.10                | 3.50             |
| stretched C$_2$         | 0.20                | 1.49             |
| oxo-Mn(Salen)           | 0.30                | 1.59             |
| ozone (OM)              | 0.20                | 2.67             |
| ozone (RM)              | 0.15                | 5.62             |
| ozone (TS)              | 0.15                | 5.30             |
| butadiene               | 0.10                | 5.29             |
Diagonal elements are given as
\[
\langle J| \hat{T}| J \rangle = \sum_{j \in J} \langle i| J \rangle + \frac{1}{2} \sum_{i,j \in J} \langle ij| ii \rangle + z \langle J| \hat{T}| J \rangle
\]
where \( h_{ij} \) denotes an element of the one-electron component of the Hamiltonian, \( \langle ij| ii \rangle \) denotes an antisymmetrized two-electron integral, and \( j \in | J \rangle \) denotes the constraint that the orbital \( j \) is occupied in \( | J \rangle \). The variable \( z \) is \( +1 \) in the even spin parity block and \( -1 \) in the odd spin parity block. The last term, \( z \langle J| \hat{T}| J \rangle \), is nonzero only if \( | J \rangle \) and \( \hat{T}| J \rangle \) differ by a double excitation. Denoting the occupied orbitals defining this excitation as \( i \) and \( j \) and the virtual orbitals \( a \) and \( b \), and defining \( | L \rangle = \hat{T}| J \rangle \), the last term can be evaluated as
\[
\langle J| \hat{T}| L \rangle = \gamma^a_{ia} \gamma^b_{j b} \langle ab| ij \rangle
\]
where \( \gamma_{ia} \) is the number of occupied orbitals in between orbitals \( i \) and \( a \) in \( | J \rangle \), as determined by a consistent ordering of orbitals among all Slater determinants.\(^\text{39} \) Off-diagonal elements of \( \hat{H} \) are given as
\[
\langle J| \hat{H}| K \rangle = N^J_J N^K_K \left( \langle J| \hat{H}| K \rangle + z \langle J| \hat{T}| K \rangle \right)
\]
Eq (B5) follows from the observations that \( \hat{T} \) and \( \hat{H} \) commute and that \( \hat{T}^2 \langle J \rangle = \langle J \rangle \). The matrix elements \( \langle J| \hat{H}| K \rangle \) and \( \langle J| \hat{T}| K \rangle \) can be evaluated according to standard Slater-Condon rules. The generic matrix element \( \langle J| \hat{H}| L \rangle \) is given in eq (B4) if \( | J \rangle \) and \( | L \rangle \) differ by a double excitation, and as
\[
\langle J| \hat{H}| L \rangle = \gamma^J_{ia} \left( h_{ia} + \sum_{j \in J} \langle ij| a j \rangle \right)
\]
if they differ by a single excitation involving an occupied orbital \( i \) and virtual orbital \( a \).

Appendix C: Orthonormalization of Iterate Columns

This section describes how we choose the matrices \( \mathbf{G}^{(\tau)} \) in eq (1) to alleviate numerical issues associated with linear dependence among the columns of the iterates \( \mathbf{X}^{(\tau)} \). If \( \mathbf{G}^{(\tau)} = \mathbf{I} \), then our randomized iteration is statistically unbiased, but the norms of the iterate columns converge to either 0 or \( \infty \) as \( \tau \to \infty \), and the columns become increasingly linearly dependent as they all approach the ground-state eigenvector. These numerical issues would render it impossible to obtain accurate eigenvalue estimates, so we instead construct \( \mathbf{G}^{(\tau)} \) as follows. In most iterations, \( \mathbf{G}^{(\tau)} \) is chosen to be \( \mathbf{N}^{(\tau)} \), a diagonal matrix with elements
\[
\mathbf{N}^{(\tau)}_{kk} = \left( \frac{\| \mathbf{X}^{(\tau)}_k \|_1}{\| \mathbf{X}^{(\tau-1)}_k \|_1} \right)^\alpha \left( \mathbf{N}^{(\tau-1)}_{kk} \right)^{1-\alpha}
\]
where \( \mathbf{X}^{(\tau)}_k \) denotes the \( k \)-th column of \( \mathbf{X}^{(\tau)} \), \( \| \cdot \|_1 \) denotes the \( \ell_1 \)-norm of a vector (the sum of the magnitudes of its elements), and \( \alpha \) is a tunable parameter. \( \mathbf{N}^{(0)} \) is initialized as the identity matrix. With this choice of \( \mathbf{G}^{(\tau)} \), setting \( \alpha = 1 \) would ensure that the column norms of iterates remain constant as the iteration proceeds. However, in the randomized implementation of this method, this introduces a statistical bias arising from the nonlinear dependence of \( \mathbf{N}^{(\tau)} \) on random variables, i.e. the iterate column norms. We therefore choose \( \alpha < 1 \) so that \( \mathbf{N}^{(\tau)} \) depends less strongly on these random variables. This causes the norms to fluctuate, but still prevents them from tending to 0 or \( \infty \) while reducing the magnitude of this bias. Numerical tests indicated that \( \alpha = 0.5 \) is a suitable choice, so it was used for all calculations presented here.

At intervals of \( \Delta \) iterations, we construct \( \mathbf{G}^{(\tau)} \) differently in order to also maintain linear independence of the iterate columns. In these iterations, \( \mathbf{G}^{(\tau)} \) is instead chosen to be \( \mathbf{N}^{(\tau)} \mathbf{D}^{(\tau)} \mathbf{R}^{(\tau)} \), where \( \mathbf{N}^{(\tau)} \) is defined as above and \( \mathbf{R}^{(\tau)} \) is the upper triangular factor of a QR factorization of \( \mathbf{U}^\tau \mathbf{X}^{(\tau)} \). This choice of \( \mathbf{G}^{(\tau)} \) enforces orthogonality of the iterate columns within the span of the trial vectors \( \mathbf{U} \). Since inclusion of the factor \( \mathbf{R}^{(\tau)} \) in \( \mathbf{G}^{(\tau)} \) also introduces a normalization constraint, the diagonal matrix \( \mathbf{D}^{(\tau)} \) is chosen to remove that constraint and ensure that normalization is controlled only via the matrix \( \mathbf{N}^{(\tau)} \). This reduces the bias associated with orthogonalization. Elements of \( \mathbf{D}^{(\tau)} \) are
\[
\mathbf{D}^{(\tau)}_{kk} = \left( \frac{\| \mathbf{X}^{(\tau)} \mathbf{R}^{(\tau)} \mathbf{U}^{-1} \mathbf{X}^{(\tau)}_k \|_1}{\| \mathbf{X}^{(\tau)}_k \|_1} \right)^{\frac{1}{2}}
\]
Since elements of \( \mathbf{R}^{(\tau)} \) depend nonlinearly on the random iterates, this orthogonalization procedure also introduces a statistical bias. This strategy differs slightly from that employed in ref 33, where we instead applied QR factorization to \( \mathbf{U}^\tau (\mathbf{I} - \varepsilon \mathbf{H}) \mathbf{X}^{(\tau)} \). We found that the alternative strategy employed here made little difference to our final results and enabled reductions in the computational cost of our implementation. By monitoring the condition number of \( \mathbf{U}^\tau \mathbf{X}^{(\tau)} \), we found \( \Delta = 1000 \) to be a reasonable choice for the systems discussed here, but our results are relatively unchanged by more frequent orthogonalization.

Appendix D: Multiplication and Compression Involving the Hamiltonian Matrix

This section describes how we apply stochastic compression to calculate the matrix product \( (\mathbf{I} - \varepsilon \mathbf{H}) \mathbf{X}^{(\tau)} \) in our asymmetric subspace iteration. The simplest approach involves stochastically compressing each column of the iterate \( \mathbf{X}^{(\tau)} \) and replacing \( \mathbf{X}^{(\tau)} \) with the resulting matrix. This corresponds to the procedure described in ref 33. If \( \mathbf{X}^{(\tau)} \) is compressed to \( m \) nonzero elements per column and the resulting sparsity structure is leveraged, the memory and CPU cost of the subsequent matrix multiplication operation scales as \( \mathcal{O}(N^2 M^2 m N_{\text{eigen}}) \). Thus, one can control the cost of multiplication by tuning \( m \). In practice, however, \( m \) cannot be chosen to be arbitrarily small. As has been demonstrated previously in the context...
of related methods, statistical error can increase very rapidly as \( m \) is decreased,\textsuperscript{25,36,94–97} rendering it impossible to achieve accurate energy estimates, even after averaging over many iterations. For the large quantum chemistry problems of interest in this work, the values of \( m \) required for acceptable accuracy render this approach too computationally expensive.

In order to further reduce the cost of performing these multiplications operations, we employ a factorization strategy\textsuperscript{37} related to those developed previously in the context of full configuration interaction quantum Monte Carlo (FCIQMC) methods.\textsuperscript{39,48} In describing this strategy, it will be useful to introduce notation denoting \textit{compositions} of compression operations and matrix multiplications: for example, \((\mathbf{H} \circ \Phi)\mathbf{x}\) indicates the vector obtained by first stochastically compressing \(\mathbf{x}\) and then multiplying the resulting compressed vector by \(\mathbf{H}\). Within this factorization strategy, the column of each iterate is calculated by applying a sequence of matrix multiplication and compression operations to the corresponding column of the previous iterate:

\[
\mathbf{X}^{(\tau+1)}_k = (\mathbf{P}_{\text{diag}} + \mathbf{B}^{(\tau,k)} \circ \Phi \circ \mathbf{Q}^{(5)} \circ \Phi \circ \mathbf{Q}^{(4)} \circ \Phi \circ \mathbf{Q}^{(3)} \circ \Phi \circ \mathbf{Q}^{(2)} \circ \Phi \circ \mathbf{Q}^{(1)}) \Phi \left( \mathbf{X}^{(\tau)}_k \right)
\]  

(D1)

Each of the matrices \(\mathbf{Q}^{(1)}, \mathbf{Q}^{(2)}, \ldots, \mathbf{Q}^{(5)}\) is constructed to have less than \(O(M)\) nonzero elements per column. Because the vector resulting after each compression operation has at most \(m\) nonzero elements, the cost of the multiplications operations involving these matrices is limited to \(O(Mm)\). The matrix \(\mathbf{B}^{(\tau,k)}\), described in more detail below, depends on \(\mathbf{X}^{(\tau)}_k\) and has \(O(1)\) nonzero elements per column. \(\mathbf{P}_{\text{diag}}\) is a diagonal matrix containing the diagonal elements of \((1 - \epsilon \mathbf{H})\). The matrix-vector product \(\mathbf{P}_{\text{diag}} \circ \Phi\), can be formed at \(O(m)\) cost. We emphasize that none of the matrices used in this factorization are stored explicitly, and that elements are instead evaluated in the course of each multiplication operation, in order to ensure memory efficiency.

Although it is possible to choose the matrices \(\mathbf{B}^{(\tau,k)}\) and \(\mathbf{Q}^{(1)}, \mathbf{Q}^{(2)}, \ldots, \mathbf{Q}^{(5)}\) in eq (D1) such that \(\mathbf{X}^{(\tau+1)}_k\) equals \((1 - \epsilon \mathbf{H})\mathbf{X}^{(\tau)}\) in expectation, it is often advantageous to relax this requirement and construct \(\mathbf{B}^{(\tau,k)}\) differently, according to an approach known as the initiator approximation, originally developed for FCIQMC.\textsuperscript{25,40} This approximation introduces a bias in the resulting eigenvalue estimates. We use this approach for the calculations presented here, as it was previously found to greatly reduce statistical error in ground-state FCIQMC and FCI-FRI calculations, thus enabling the application of these methods to larger chemical systems. The initiator approximation involves constructing \(\mathbf{B}^{(\tau,k)}\) such that, in the course of the matrix multiplications in eq (D1), only elements in \(\mathbf{X}^{(\tau)}_k\) with magnitudes greater than an initiator threshold are allowed to contribute to elements of \(\mathbf{X}^{(\tau+1)}_k\) that are zero in \(\mathbf{X}^{(\tau)}_k\). Specific formulas for the elements of \(\mathbf{B}^{(\tau,k)}\), as well as those of \(\mathbf{Q}^{(1)}, \mathbf{Q}^{(2)}, \ldots, \mathbf{Q}^{(5)}\), are provided in Section E.

Applications of the initiator approximation in an FCIQMC context used a fixed, user-specified value for the initiator threshold, \(n_e\). Such an approach presents an issue for our particular FCI-FRI method for multiple eigenvalue calculations. The column norms of iterates can become very different as the iteration proceeds, in which case there can be very different numbers of elements with magnitudes greater than \(n_e\) in each column. We therefore used a different threshold \(t_k\) for each column \(k\), scaled by the column norm, to ensure that the

\[
t_k = n_e \frac{\|\mathbf{X}^{(\tau)}_k\|}{m^{-1}}
\]  

(D2)

In many implementations of FCIQMC, the number of samples used for stochastic matrix-vector multiplication (analogous to \(m\)) is approximately equal to the \(\ell_1\)-norm of the vector being multiplied. In this case, our implementation of the initiator approximation is equivalent to previous implementations. In practice, the \(\ell_1\)-norms of iterate columns in FCI-FRI are less than those of iterates in FCIQMC (and correspondingly less than \(m\)), since our algorithm does not require an initial “population growth phase.”\textsuperscript{94} Note that \(t_k\) approaches 0 as \(m\) is increased, in which case \(\mathbf{X}^{(\tau+1)}\) approaches \((1 - \epsilon \mathbf{H})\mathbf{X}^{(\tau)}\) in expectation, and the bias introduced by the initiator approximation approaches 0. This behavior also parallels that of the initiator approximation as commonly applied to FCIQMC.

Appendix E: Hamiltonian Matrix Factorization

This section provides formulas for the elements of the matrices \(\mathbf{B}^{(\tau,k)}\) and \(\mathbf{Q}^{(1)}, \mathbf{Q}^{(2)}, \ldots, \mathbf{Q}^{(5)}\) used to generate the iterate \(\mathbf{X}^{(\tau+1)}\) from \(\mathbf{X}^{(\tau)}\). We refer to these matrices as factors of \(\mathbf{H}\) because off-diagonal elements of the matrix \(\mathbf{B}^{(\tau,k)}\mathbf{Q}_i\mathbf{Q}^{(4)}\mathbf{Q}^{(3)}\mathbf{Q}^{(2)}\mathbf{Q}^{(1)}\) approximately equal those of \(-\epsilon \mathbf{H}\). This scheme is based on the modified heat-bath Power-Pitzer factorization scheme introduced in ref 37 and includes modifications to enable its application in a basis of spin-coupled functions.

We begin by defining the notation used to index elements of these matrices. One of the Slater determinants defining each spin-coupled function \(|\bar{J}\rangle\) is arbitrarily chosen as its \textit{representative} Slater determinant and is denoted \(|\bar{J}\rangle_{\text{rep}}\). Matrices are indexed by excitations from these representative Slater determinants. For example, \((\bar{J}, 1, i, a)\) denotes a single excitation involving occupied orbital \(i\) and virtual orbital \(a\) from \(|\bar{J}\rangle_{\text{rep}}\), and \((\bar{J}, 1, i, j, a, b)\) denotes a double excitation involving occupied orbitals \(i\) and \(j\) and virtual orbitals \(a\) and \(b\) from \(|\bar{J}\rangle_{\text{rep}}\).

Elements of the matrices in the factorization are defined in terms of a matrix \(\mathbf{D}\) and vectors \(\mathbf{S}\) and \(\mathbf{Y}\), precomputed at the
beginning of each calculation. The $2M \times 2M$ matrix $D$ has elements

$$D_{pq} = (1 - \delta_{pq}) \sum_{r,s} |\langle pq | rs \rangle|$$  \hspace{1cm} \text{(E1)}

where $p$ and $q$ represent indices of spin orbitals, $\delta_{pq}$ is a Kronecker delta, and $|\langle pq | rs \rangle|$ is an antisymmetrized two-electron integral. Due to spin symmetries present in the two-electron integrals, $D$ has only $M^2 + \binom{M}{2}$ unique elements that need to be stored. The vector $S$ has elements

$$S_r = \frac{\sum_q D_{rq}}{\sum_p D_{pq}}$$  \hspace{1cm} \text{(E2)}

and $Y$ has elements

$$Y_i = \sum_a |\langle i a | a l \rangle|^{1/2}$$  \hspace{1cm} \text{(E3)}

For all matrices defined in the remainder of this section, the values of matrices not specified below are 0. The row space of $Q^{(3)}$ has dimension $2N_{\text{CI}}$ and contains elements corresponding to either a generic single excitation, $(J, 1)$ or double excitation $(J, 2)$ from the reference determinant for each spin-coupled function $|\tilde{J}\rangle$. Elements for single excitations are specified as

$$Q^{(1)}_{(K,1),J} = \frac{n_s}{n_s + n_d} \delta_{KJ}$$  \hspace{1cm} \text{(E4)}

and those for double excitations as

$$Q^{(1)}_{(K,2),J} = \frac{n_d}{n_s + n_d} \delta_{KJ}$$  \hspace{1cm} \text{(E5)}

where $n_s$ and $n_d$ are the number of symmetry-allowed single and double excitations from the Hartree-Fock determinant, respectively, and $\delta_{KJ}$ is a Kronecker delta. Here, and in the remainder of this section, we consider only spin and spatial (point-group) symmetries when determining which excitations are allowed.

Indices for the row space of $Q^{(2)}$ include an occupied orbital index $i$. Single-excitation elements are specified as

$$Q^{(2)}_{(J,1),a,(J,1)} = \left( n_{J\text{oc}}^{(1)} \right)^{-1}$$  \hspace{1cm} \text{(E6)}

where $n_{J\text{oc}}^{(1)}$ is the number of occupied orbitals in $|\tilde{J}\rangle_{\text{rep}}$ for which there is at least one virtual orbital of the same symmetry. Double-excitation elements are specified as

$$Q^{(2)}_{(J,2),a,(J,2)} = S_i$$  \hspace{1cm} \text{(E7)}

where the index $i$ is constrained to be any of the occupied orbitals in $|\tilde{J}\rangle_{\text{rep}}$ except the first.

Elements of $Q^{(3)}$ corresponding to single excitations include a virtual orbital index $a$:

$$Q^{(3)}_{(J,1,a),(J,1,a)} = \left( n_{J\text{virt}}^{(1)} \right)^{-1}$$  \hspace{1cm} \text{(E8)}

Here, $n_{J\text{virt}}^{(1)}(i)$ denotes the number of virtual orbitals in $|\tilde{J}\rangle_{\text{rep}}$ with the same symmetry as the occupied orbital $i$. Elements corresponding to double excitations are indexed differently:

$$Q^{(3)}_{(J,2,i),(J,2,i)} = D_{ij} S_i^{-1}$$  \hspace{1cm} \text{(E9)}

Here, the index of the second occupied orbital $j$ in the double excitation is constrained to be less than that of the first ($i$).

All elements in $Q^{(4)}$ corresponding to single excitations are $1$:

$$Q^{(4)}_{(J,1,a),(J,1,a)} = 1$$  \hspace{1cm} \text{(E10)}

Double-excitation elements are given as

$$Q^{(5)}_{(J,2,i,a),(J,2,i,a)} = |\langle i a | a l \rangle|^{1/2} X_{ij}^{-1}$$  \hspace{1cm} \text{(E11)}

where $a$ is constrained to be any virtual orbital in $|\tilde{J}\rangle_{\text{rep}}$ except the first.

In analogy to eq (E10), the values of single-excitation elements in $Q^{(5)}$ are $1$. Double-excitation elements in $Q^{(5)}$ are

$$Q^{(5)}_{(J,2,i,a),(J,2,j,a)} = |\langle i a | b a \rangle|^{1/2} X_{ij}^{-1}$$  \hspace{1cm} \text{(E12)}

where $b$ is constrained to be any virtual orbital in $|\tilde{J}\rangle_{\text{rep}}$ with an index less than $a$ for which the direct symmetry product $\Gamma_i \otimes \Gamma_j$ is equal to $\Gamma_a \otimes \Gamma_k$. Further details on computing direct symmetry products can be found in refs 36 and 98.

Multiplication by the final matrix in the factorization, $B^{(r\tau)}$, serves to sum elements corresponding to excitations that map to the same spin-coupled function while enforcing the initiator approximation. Elements of $B^{(r\tau)}$ corresponding to single excitations are specified as

$$B^{(r\tau)}_{K,(J,1,a)} = \begin{cases} 0 & X_{K,K}^{(r\tau)} = 0 \text{ and } |X_{J,K}^{(r\tau)}| < t_k \\ -\frac{\varepsilon H_{K,j}}{Q_{(J,1,a),J}} & \text{otherwise} \end{cases}$$  \hspace{1cm} \text{(E13)}

for the spin-coupled function $|\tilde{K}\rangle$ connected to $|\tilde{J}\rangle$ by a single excitation involving occupied orbital $i$ and virtual orbital $a$ (i.e. for which $\langle \tilde{K}| \epsilon_i^{(r\tau)} |\tilde{J}\rangle \neq 0$). The variable $t_k$ is defined in eq (D2). The matrix $Q$ is defined as the product $Q^{(5)} Q^{(4)} Q^{(3)} Q^{(1)}$. Its elements can be calculated inexpensively on the fly due to the sparse structure of $Q^{(3)}$, $Q^{(2)}$, ..., $Q^{(5)}$. Elements of $B^{(r\tau)}$ corresponding to double excitations are

$$B^{(r\tau)}_{K,(J,2,i,a,b)} = \begin{cases} 0 & X_{K,k}^{(r\tau)} = 0 \text{ and } |X_{J,k}^{(r\tau)}| < t_k \\ -\frac{\varepsilon H_{K,j}}{Q_{(J,2,i,a,b),J}} & \text{otherwise} \end{cases}$$  \hspace{1cm} \text{(E14)}

for $\langle \tilde{K}| \epsilon_i^{(r\tau)} |\tilde{J}\rangle \neq 0$.

**Appendix F: Estimating the standard error in eigenvalue estimates**

This section describes our approach to approximating the standard error associated with each eigenvalue estimate $A_{kk}$
obtained by solving eq (2) in the main text. Because subsequent iterates are correlated, these errors can be estimated by applying suitable Markov chain Monte Carlo error estimation techniques99 to the scalar-valued trajectory

\[ z_k' = (U^\dagger H X^\ominus)^\dagger - \Lambda_k (U^\dagger X^\ominus)^\dagger w_k \]  

where \( z_k \) and \( w_k \) represent the left and right generalized eigenvectors, respectively, corresponding to \( \Lambda_k \). We used the emcee software package29 to estimate standard errors and associated autocorrelation times. For most of the systems considered in this work, these autocorrelation times exceeded 100 iterations, suggesting that evaluating \( U^\dagger H X^\ominus \) and \( U^\dagger X^\ominus \) more frequently would not significantly change eigenvalue estimates or their associated standard errors. In our implementation, the computational cost of each evaluation scales with the number of nonzero elements in \( U \) and can be quite significant in practice. Evaluating these matrices less frequently enables us to afford the increased cost associated with using more accurate trial vectors. These long autocorrelation times also render it difficult to converge our standard error estimates using the default parameters in the emcee software. However, given that the range of values in any single trajectory (i.e. the difference between the maximum and minimum) is usually less than 0.5 m\( \text{e}_\text{h} \), we believe our error estimates to be sufficiently accurate for the comparisons reported in this work.

REFERENCES

1. B. Mennucci, “The simulation of UV-Vis spectroscopy with computational methods,” in Computational Spectroscopy, edited by J. Grunenberg (John Wiley & Sons, Ltd, 2010) Chap. 5, pp. 151–171, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527633272.ch5.
2. E. J. Beard, G. Sivaraman, A. Vázquez-Mayagotía, V. Vishwanath, and J. M. Cole, “Comparative dataset of experimental and computational attributes of UV/vis absorption spectra,” Sci. Data 6, 307 (2019).
3. L. Gonzalez and R. Lindh, eds., Quantum Chemistry and Dynamics of Excited States: Methods and Applications (John Wiley & Sons, Ltd, 2010) Chap. 5, pp. 09-108 (2005), proceedings of the 30th International Congress of Theoretical Chemists of Latin Expression.
4. M. Navarrete-Miguel, J. Segarra-Martí, A. Francés-Monerris, A. Giusani, P. Farahani, B.-W. Ding, A. Monari, Y.-J. Liu, and D. Roca-Sanjuan, “Quantum chemistry of the excited state: recent trends in methods developments and applications,” in Photochemistry: Volume 46, Vol. 46 (The Royal Society of Chemistry, 2019) pp. 28–77.
5. J. W. Park and T. Shiozaki, “On-the-fly CASPT2 surface hopping dynamics,” Journal of Chemical Theory and Computation, J. Chem. Theory Comput. 13, 3676–3683 (2017).
6. J. E. Botnick, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi, “Understanding the surface hopping view of electronic transitions and decoherence,” Annu. Rev. Phys. Chem. 67, 387–417 (2016), https://doi.org/10.1146/annurev-physchem-040215-112245.
7. G. Barbutti, “Nonadiabatic dynamics with trajectory surface hopping method,” WIREs Comput. Mol. Sci. 1, 620–633 (2011), https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.64.
8. G. Harsha, T. M. Henderson, and G. E. Scuseria, “Thermofield theory for finite-temperature quantum chemistry,” J. Chem. Phys. 150, 154109 (2019), https://doi.org/10.1063/1.5089560.
algorithm with application to the C calculations. I. the atoms boron through neon and hydrogen,” J. Chem. Phys.

D. M. Ceperley and B. Bernu, “The calculation of excited state properties of matrices,” in Numerical Linear Algebra, edited by J. J. Miller (Academic Press, 1975) pp. 185 – 196.

C. J. Umrigar, M. P. Nightingale, and K. J. Runge, “A diatomic epoxidation of unfunctionalized olefins,” Tetrahedron Lett.

L.-H. Lim and J. Weare, “Fast randomized iteration: Diagonalization of matrices,” in Topics in Numerical Analysis II, edited by G. W. Stewart, “Methods of simultaneous iteration for calculating eigenvalues and eigenvectors of matrices,” in Numerical Linear Algebra, edited by J. J. Miller (Academic Press, 1975) pp. 185 – 196.

C. J. Umrigar, M. P. Nightingale, and K. J. Runge, “A diatomic epoxidation of unfunctionalized olefins,” Tetrahedron Lett.

G. Stewart, “Methods of simultaneous iteration for calculating eigenvalues and eigenvectors of matrices,” in Numerical Linear Algebra, edited by J. J. Miller (Academic Press, 1975) pp. 185 – 196.

C. J. Umrigar, M. P. Nightingale, and K. J. Runge, “A diatomic epoxidation of unfunctionalized olefins,” Tetrahedron Lett.

G. Stewart, “Methods of simultaneous iteration for calculating eigenvalues and eigenvectors of matrices,” in Numerical Linear Algebra, edited by J. J. Miller (Academic Press, 1975) pp. 185 – 196.

C. J. Umrigar, M. P. Nightingale, and K. J. Runge, “A diatomic epoxidation of unfunctionalized olefins,” Tetrahedron Lett.

G. Stewart, “Methods of simultaneous iteration for calculating eigenvalues and eigenvectors of matrices,” in Numerical Linear Algebra, edited by J. J. Miller (Academic Press, 1975) pp. 185 – 196.

C. J. Umrigar, M. P. Nightingale, and K. J. Runge, “A diatomic epoxidation of unfunctionalized olefins,” Tetrahedron Lett.

G. Stewart, “Methods of simultaneous iteration for calculating eigenvalues and eigenvectors of matrices,” in Numerical Linear Algebra, edited by J. J. Miller (Academic Press, 1975) pp. 185 – 196.

C. J. Umrigar, M. P. Nightingale, and K. J. Runge, “A diatomic epoxidation of unfunctionalized olefins,” Tetrahedron Lett.
C. Daday, S. Smart, G. H. Booth, A. Alavi, and C. Filippi, “Full configuration interaction excitations of ethene and butadiene: Resolution of an ancient question,” J. Chem. Theory Comput. 8, 4441–4451.

S. P. Ansari and F. Ali, “Conjugated organic polymers for optoelectronic devices,” in Functional Polymers, edited by M. A. Jafar Mazumder, H. Sheardown, and A. Al-Ahmed (Springer International Publishing) pp. 1–40.

K. Oka, B. Winther-Jensen, and H. Nishide, “Organic π-conjugated polymers as photocathode materials for visible-light-enhanced hydrogen and hydrogen peroxide production from water,” Adv. Energy Mater. 11, 2003724.

P. Tavan and K. Schulten, “Electronic excitations in finite and infinite polyenes,” Phys. Rev. B 36, 4337–4358 (1987).

J. D. Watts, S. R. Gwaltney, and R. J. Bartlett, “Coupled-cluster calculations of the excitation energies of ethylene, butadiene, and cyclopentadiene,” J. Chem. Phys. 105, 6979–6988 (1996).

J. H. Starcke, M. Wormit, J. Schirmer, and A. Dreuw, “How much double excitation character do the lowest excited states of linear polyenes have?” Chem. Phys. 329, 39–49, electron Correlation and Multimode Dynamics in Molecules.

P.-O. Widmark, P.-Å. Malmqvist, and B. O. Roos, “Density matrix averaged atomic natural orbital (ano) basis sets for correlated molecular wave functions,” Theor. Chim. Acta 77, 291–306.

J. S. Spencer and A. J. W. Thom, “Developments in stochastic coupled cluster theory: The initiator approximation and application to the uniform electron gas,” J. Chem. Phys. 144, 084108 (2016).

M. Motta and S. Zhang, “Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method,” Wiley Interdiscip. Rev.: Comput. Mol. Sci., 1364 (2018).

A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.-F. Loos, “Excitation energies from diffusion Monte Carlo using selected configuration interaction nodes,” J. Chem. Phys. 149, 034108 (2018).

M. Caffarel, T. Appelcourt, E. Giner, and A. Scemama, “Using CIPSI nodes in diffusion Monte Carlo,” in Recent Progress in Quantum Monte Carlo, ACS Symposium Series, Vol. 1234 (American Chemical Society, 2016) pp. 15–46.

Dice,” (2021), https://github.com/sanshar/Dice (accessed May 4, 2021).

J. J. Eriksen, T. A. Anderson, J. D. Deustua, K. Ghanem, D. Hait, M. R. Hoffmann, S. Lee, D. S. Levine, I. Magoulas, J. Shen, N. M. Tubman, K. B. Whaley, E. Xu, Y. Yao, N. Zhang, A. Alavi, G. K.-L. Chan, M. Head-Gordon, W. Liu, P. Piecuch, S. Sharma, S. L. Ten-no, C. J. Umrigar, and J. Gauss, “The ground state electronic energy of benzene,” The Journal of Physical Chemistry Letters, J. Phys. Chem. Lett. 11, 8922–8929 (2020).

Y. Yao, E. Giner, J. Li, J. Toulouse, and C. J. Umrigar, “Almost exact energies for the Gaussian-2 set with the semistochastic heat-bath configuration interaction method,” J. Chem. Phys. 153, 124117 (2020).

J. S. Spencer, N. S. Blunt, and W. M. Foulkes, “The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method,” J. Chem. Phys. 136, 054110 (2012).

M. H. Kolodrubetz, J. S. Spencer, B. K. Clark, and W. M. C. Foulkes, “The effect of quantization on the full configuration interaction quantum Monte Carlo sign problem,” J. Chem. Phys. 138, 024110 (2013).

J. J. Shepherd, G. E. Scuseria, and J. S. Spencer, “Sign problem in full configuration interaction quantum Monte Carlo: Linear and sublinear representation regimes for the exact wave function,” Phys. Rev. B 90, 155130 (2014).

W. A. Vigor, J. S. Spencer, M. J. Bearpark, and A. J. W. Thom, “Understanding and improving the efficiency of full configuration interaction quantum Monte Carlo,” J. Chem. Phys. 144, 094110 (2016).

G. H. Booth, S. D. Smart, and A. Alavi, “Linear-scaling and parallelisable algorithms for stochastic quantum chemistry,” Mol. Phys. 112, 1855–1869 (2014).

D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC Hammer," Preprint at https://arxiv.org/abs/1202.3665 (2013), (accessed January 26, 2022).