RESEARCH ARTICLE SUMMARY

QUANTUM CHEMISTRY

Accurate computation of quantum excited states with neural networks

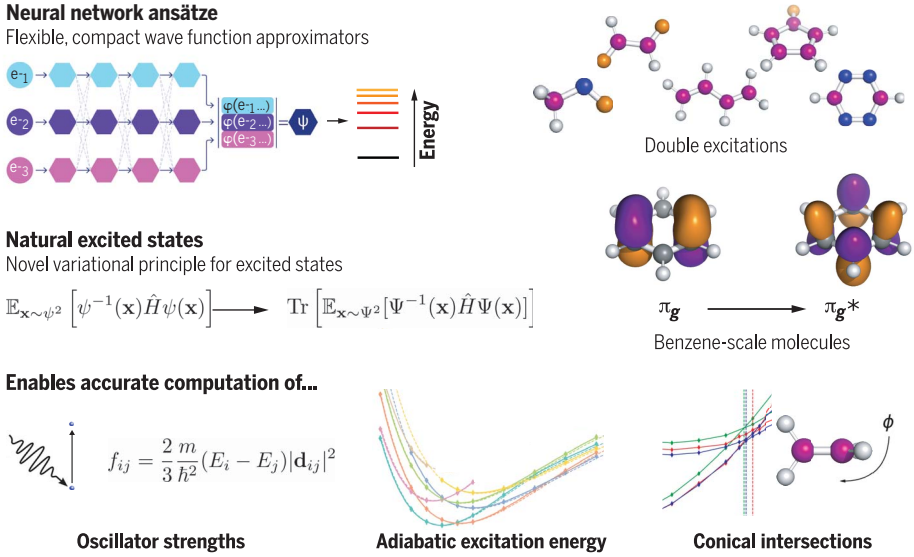
David Pfau*, Simon Axelrod, Halvard Sutterud, Ingrid von Glehn, James S. Spencer

INTRODUCTION: Understanding the physics of how matter interacts with light requires accurate modeling of electronic excited states of quantum systems. This underpins the behavior of photocatalysts, fluorescent dyes, quantum dots, light-emitting diodes (LEDs), lasers, solar cells, and more. Existing quantum chemistry methods for excited states can be much more inaccurate than those for ground states, sometimes qualitatively so, or can require prior knowledge targeted to specific states. Neural networks combined with variational Monte Carlo (VMC) have achieved remarkable accuracy for ground state wave functions for a range of systems, including spin models, molecules, and condensed matter systems. Although VMC has been used to study excited states, prior approaches have limitations that make it difficult or impossible to use them with neural networks and often have many free parameters that require tuning to achieve good results.

RATIONALE: We combine the flexibility of neural network ansätze with a mathematical insight that allows us to convert the problem of finding excited states of a system to one of finding the ground state of an expanded system, which can then be tackled with standard VMC. We call this approach natural excited states VMC (NES-VMC). Linear independence of the excited states is automatically imposed through the functional form of the ansatz. The energy and other observables of each excited state are obtained from diagonalizing the matrix of Hamiltonian expectation values taken over the single-state ansätze, which can be accumulated with no additional cost. Crucially, this approach has no free parameters to tune and needs no penalty terms to enforce orthogonalization. We examined the accuracy of this approach with two different neural network architectures—the FermiNet and Psiformer.

RESULTS: We demonstrated our approach on benchmark systems ranging from individual atoms up to molecules the size of benzene. We validated the accuracy of NES-VMC on first-row atoms, closely matching experimental results, and on a range of small molecules, obtaining highly accurate energies and oscillator strengths comparable to existing best theoretical estimates. We computed the potential energy curves of the lowest excited states of the carbon dimer and identified the states across bond lengths by analyzing their symmetries and spins. The NES-VMC vertical excitation energies matched those obtained using the highly accurate semi-stochastic heat-bath configuration interaction (SHCI) method to within chemical accuracy for all bond lengths, whereas the adiabatic excitations were within 4 meV of experimental values on average—a fourfold improvement over SHCI. In the case of ethylene, NES-VMC correctly described the conical intersection of the twisted molecule and was in excellent agreement with highly accurate multireference configuration interaction (MR-CI) results. We also considered five challenging systems with low-lying double excitations, including multiple benzene-scale molecules. On all systems where there is good agreement between methods on the vertical excitation energies, the Psiformer was within chemical accuracy across states, including butadiene, where even the ordering of certain states has been disputed for many decades. On tetrazine and cyclopentadienone, where state-of-the-art calculations from just a few years ago were known to be inaccurate, NES-VMC results closely matched recent sophisticated diffusion Monte Carlo (DMC) and complete-active-space third-order perturbation theory (CASPT3) calculations. Finally, we considered the benzene molecule, where NES-VMC combined with the Psiformer ansatz is in substantially better agreement with theoretical best estimates compared with other methods, including neural network ansätze using penalty methods. This both validates the mathematical correctness of our approach and shows that neural networks can accurately represent excited states of molecules right at the current limit of computational approaches.

CONCLUSION: NES-VMC is a parameter-free and mathematically sound variational principle for excited states. Combining it with neural network ansätze enables marked accuracy across a wide range of benchmark problems. The development of an accurate VMC approach to excited states of quantum systems opens many possibilities and substantially expands the scope of applications of neural network wave functions. Although we considered only electronic excitations of molecular systems and neural network ansätze, NES-VMC is applicable to any quantum Hamiltonian and any ansatz, enabling accurate computational studies that could improve our understanding of vibronic couplings, optical bandgaps, nuclear physics, and other challenging problems. ■



Natural excited states. Combining neural networks with a mathematical insight enables accurate calculations of challenging excited states of molecules.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: pfau@google.com
Cite this article as D. Pfau *et al.*, *Science* **385**, eadn0137 (2024). DOI: 10.1126/science.adn0137

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adn0137

RESEARCH ARTICLE

QUANTUM CHEMISTRY

Accurate computation of quantum excited states with neural networks

David Pfau[1,2]*, Simon Axelrod[1,3,4], Halvard Sutterud[2], Ingrid von Glehn[1], James S. Spencer[1]

We present an algorithm to estimate the excited states of a quantum system by variational Monte Carlo, which has no free parameters and requires no orthogonalization of the states, instead transforming the problem into that of finding the ground state of an expanded system. Arbitrary observables can be calculated, including off-diagonal expectations, such as the transition dipole moment. The method works particularly well with neural network ansätze, and by combining this method with the FermiNet and Psiformer ansätze, we can accurately recover excitation energies and oscillator strengths on a range of molecules. We achieve accurate vertical excitation energies on benzene-scale molecules, including challenging double excitations. Beyond the examples presented in this work, we expect that this technique will be of interest for atomic, nuclear, and condensed matter physics.

The computation of excited state properties of quantum systems is a fundamental challenge in chemistry and physics. Understanding electronic excitations is critical for predicting fluorescence in quantum dots (*1*), molecular conformational changes in the presence of light (*2*), and photocatalytic activity (*3*). In condensed matter physics, excitations determine the optical bandgap of semiconductors, which is critical for predicting the behavior of solar cells, photosensors, light-emitting diodes (LEDs), and lasers. Excited states are also relevant to understanding nuclear phenomena, such as metastable isomers (*4*). Despite the importance of excited states for quantum phenomena, a full computational account of excited states remains challenging. Excited states are much more challenging to compute compared with ground states—inexpensive methods such as time-dependent density functional theory (TD-DFT) often give qualitatively incorrect results, and even gold standard methods such as multireference configuration interaction (MR-CI) can have large error bars on higher excited states. Techniques from machine learning can be applied to this problem, but most applications must be trained on expensive ab initio calculations (*5*). Far less work has focused on how insights from machine learning can improve the ab initio calculations themselves.

Recent work using neural networks as a wave function ansatz has demonstrated the ability to reach impressive levels of accuracy on ground state calculations with variational Monte Carlo (VMC) (*6*, *7*), even exceeding coupled cluster [CCSD(T)] accuracy on some bond-stretching systems. VMC (*8*, *9*) is conceptually simple—it works by finding an explicit functional form for a wave function that minimizes a variational bound—and scales as $\mathcal{O}(N^3)$–$\mathcal{O}(N^4)$ with system size, which is favorable for a wave function method. Although neural network ansätze are state of the art for variational optimization of ground states, they have not yet matched the accuracy of other methods for excited states calculations as they scale to larger systems, despite initial forays (*10*, *11*). This issue could be because of the ansatz not being sufficiently expressive, but it could also be a result of the choice of variational principle. When used to optimize ground states, there are only two variational principles for quantum Monte Carlo (QMC)—energy minimization and variance minimization. Innovations in ground state VMC primarily focus on the choice of trial wave function (*12*, *13*) or optimization method used to achieve the variational bound (*14*, *15*), but the choice of objective to optimize is well established. The same cannot be said for variational optimization of excited states.

Approaches for computing excited states by VMC either aim to find a single excited state (so-called state-targeting methods) or aim to find all of the lowest-lying excited states (called either state-averaging or state-specific depending on whether the states overlap). Methods for finding multiple states either minimize the energy of each state sequentially or minimize the (possibly weighted) total energy of all states and simultaneously maintain orthogonality. Among state-targeting methods, there are methods that target specific energy ranges (*16*, *17*), specific symmetries of the system (*10*), or a specific ordering of the roots (*18*). For state-averaging and state-specific approaches, the different states must be kept from collapsing onto one another, which can be achieved by including a penalty term in the variational bound (*10*, *11*, *19*, *20*) or by explicitly constructing orthogonal ansätze by solving a generalized eigenvalue problem, sometimes repeatedly re-orthogonalizing during optimization (*21*–*25*).

All of these approaches have drawbacks and limitations that make it difficult or impossible to use them with recently developed ansätze based on deep neural networks (*6*, *26*–*28*) on general systems. Targeting specific symmetries or energy ranges requires prior knowledge about the states of interest, which may not be available, and state-targeting by variance minimization can lose track of the desired state (*24*). Root-targeting methods are prone to root-flipping, whether they are used for QMC or other computational paradigms (*29*, *30*). Some methods require solving a generalized eigenvalue problem from stochastic estimates of the Hamiltonian and overlap matrices, which introduces statistical bias into the gradient estimates (*18*, *23*). Although this bias is usually reduced by accumulating a large number of samples of the matrix elements, it is not possible to do so in the deep learning paradigm, where parameters are optimized by a large number of small, noisy steps (*31*). Explicitly orthogonalizing ansätze by solving a generalized eigenvalue equation for linear coefficients is usually only possible when the ansatz is a linear combination of basis set functions (*32*, *33*), possibly including a shared Jastrow factor (*21*), which rules out neural networks. Other methods do not maintain true orthogonality but only keep linearized approximations to the wave functions orthogonal (*18*). Some penalty methods that have been used with neural network ansätze have problems with biased gradients (*11*), and even with unbiased gradients, convergence is only guaranteed if the strength of the penalty term is set above a certain critical value that is not known a priori. When optimizing multiple states simultaneously with unbiased penalty methods, a weighting factor must be chosen for each state, otherwise the critical penalty threshold diverges (*20*). Heuristics such as variance matching may be required to achieve good numerical results for all approaches. Despite almost four decades of work on QMC methods for excited states (*32*, *34*), no single variational principle has emerged that has no free parameters, has convergence guarantees when optimizing with noisy Monte Carlo estimates, and is applicable to all possible ansätze and all excited states regardless of symmetry.

We present a variational principle for computing the lowest excited states of a quantum system by Monte Carlo that does not suffer from any of the limitations described above. Like many state-specific and state-averaged approaches, our method minimizes the total

[1]Google DeepMind, London N1C 4DJ, UK. [2]Department of Physics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. [3]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 01238, USA. [4]Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 01239, USA.
*Corresponding author. Email: pfau@google.com

energy over states, but we make a particular choice of sampling distribution that does not require the states to be orthogonal. This choice of sampling distribution is equivalent to reformulating the problem of finding K excited states of an N-particle system into the problem of finding the ground state of a K-fermion system where each fermion is equivalent to N particles in the original system. Instead of orthogonalizing the states, the local energy is promoted from a scalar to a matrix, which gives unbiased estimates of a matrix whose eigenvalues are the energies of orthogonal states. Because wave function optimization can be done by stochastic gradient descent from unbiased noisy estimates of the total energy, the procedure is guaranteed to converge to a local minimum of the total energy over states (*35*). Because of the many desirable mathematical properties summarized above, we refer to our proposed approach as natural excited states for VMC (NES-VMC).

Natural excited states

We aim to find the K lowest eigenfunctions $\psi_1, \ldots, \psi_K$ of a Hamiltonian $\hat{H}$. To find the ground state, we can take samples $\mathbf{x} \sim \psi^2(\mathbf{x})$ of particle positions and compute unbiased estimates of the energy $\mathbb{E}_{\mathbf{x}\sim\psi^2}\left[\psi^{-1}(\mathbf{x})\hat{H}\psi(\mathbf{x})\right]$ as well as gradients of the energy and then optimize the functional form of ψ. The scalar $E_L(\mathbf{x}) \triangleq \psi^{-1}(\mathbf{x})\hat{H}\psi(\mathbf{x})$ inside the expectation is the local energy. This is the conventional energy minimization principle for VMC.

To generalize this principle to excited states, consider the function

$$\Psi\left(\mathbf{x}^1, \ldots, \mathbf{x}^K\right) \triangleq \det\begin{bmatrix} \psi_1(\mathbf{x}^1) & \cdots & \psi_K(\mathbf{x}^1) \\ \vdots & & \vdots \\ \psi_1(\mathbf{x}^K) & \cdots & \psi_K(\mathbf{x}^K) \end{bmatrix} \tag{1}$$

where $\mathbf{x}^1, \ldots, \mathbf{x}^K$ are K different sets of particle states. We call $\psi_1, \ldots, \psi_K$ the single-state ansätze and Ψ the total ansatz. The total ansatz resembles a Slater determinant, except that single-particle orbitals are substituted with many-particle single-state ansätze.

We can compute Monte Carlo estimates of the total energy over all states by taking samples $\mathbf{x}^1, \ldots, \mathbf{x}^K \sim \Psi^2(\mathbf{x}^1, \ldots, \mathbf{x}^K)$ and generalizing the local energy from a scalar to a matrix

$$\mathbf{E}_\Psi \triangleq \mathbb{E}_{\mathbf{x}^1,\ldots,\mathbf{x}^K\sim\Psi^2}\left[\mathbf{\Psi}^{-1}\hat{H}\mathbf{\Psi}\right] \tag{2}$$

$$\mathbf{\Psi}\left(\mathbf{x}^1, \ldots, \mathbf{x}^K\right) \triangleq \begin{bmatrix} \psi_1(\mathbf{x}^1) & \cdots & \psi_K(\mathbf{x}^1) \\ \vdots & & \vdots \\ \psi_1(\mathbf{x}^K) & \cdots & \psi_K(\mathbf{x}^K) \end{bmatrix} \tag{3}$$

$$\hat{H}\mathbf{\Psi}\left(\mathbf{x}^1, \ldots, \mathbf{x}^K\right) \triangleq \begin{bmatrix} \hat{H}\psi_1(\mathbf{x}^1) & \cdots & \hat{H}\psi_K(\mathbf{x}^1) \\ \vdots & & \vdots \\ \hat{H}\psi_1(\mathbf{x}^K) & \cdots & \hat{H}\psi_K(\mathbf{x}^K) \end{bmatrix} \tag{4}$$

The trace of this matrix is an unbiased estimate for the total energy over states, and unbiased gradients can be estimated as well. The determinant in the definition of Ψ guarantees that the states will not collapse onto one another during minimization, even though nothing constrains the states to be orthogonal. Minimizing $\mathbf{Tr}[\mathbf{E}_\Psi]$ is mathematically equivalent to ground state VMC for an extended system that is K times as large as the ground state of $\hat{H}$.

If the true minimum of the total energy could be found, it would give a set of single-state ansätze that are a linear combination of the lowest states of $\hat{H}$. To recover the energies of individual states, we simply diagonalize the matrix $\mathbf{E}_\Psi$ at the minimum. For other observables $\hat{O}$, a matrix similar to $\mathbf{E}_\Psi$ can be accumulated with $\hat{O}$ substituted for $\hat{H}$ and then be transformed into the same basis as the eigenvectors of $\mathbf{E}_\Psi$. We call this method NES-VMC. A derivation of NES-VMC, its equivalence to ground state VMC in an extended system, and more details on the computation of observables are given in the materials and methods section.

Although NES-VMC is fully general and can be applied to any quantum Hamiltonian, our experimental validation is focused on electronic structure in atoms and molecules because of the abundant experimental and computational literature to compare against. For all experiments, we solved the electronic Schrödinger equation in the Born-Oppenheimer approximation and in atomic units

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i>j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{iI}\frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{I>J}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{5}$$

where the indices i and j are over electrons and I and J are over atomic nuclei with fixed locations.

To try to disentangle the effect that the choice of ansatz has on performance, we investigated two different neural network architectures: the FermiNet (*27*) and the Psiformer (*36*). Although the Psiformer is more accurate on large systems, it is also slower, and for ground state calculations up to ~15 electrons, no appreciable difference in accuracy between the two has been found.

Atomic spectra

To check the correctness of our method, we investigated the excited states of first-row atoms. Although we do not aim to reach the accuracy of spectroscopic measurements, we can have high confidence in the accuracy of experimental data and do not need to worry about effects such as adiabatic relaxation and zero-point vibrational energy, which affect molecular measurements. All experimental data were taken from the energy level tables in the National Institute of Standards & Technology (NIST) Handbook of Basic Atomic Spectroscopic Data (*37*). Because we are working with the nonrelativistic Schrödinger equation, we were not able to compute fine or hyperfine structure. To remove the fine structure, experimental energy levels with different total angular momenta were averaged together weighted by the degeneracy $m_J = 2J + 1$ and treated as a single level. The hyperfine structure is too small to be of concern here. To investigate the effect of the choice of ansatz as well as the choice of number of states k to compute, we ran calculations with the FermiNet with both 5 and 10 states as well as with the Psiformer with 10 states. Results are given in Fig. 1, with numerical results in table S2.

For all atoms, NES-VMC gives results closely matching experiment on most states. From lithium up to oxygen, the error relative to experiment is far less than 1 mHa (27.2 meV) for all but the highest excited state and is often less than 0.1 mHa. On lithium, all ansätze correctly converged to the ^2S and $^2P^o$ states, which are missed by the use of the PauliNet ansatz in combination with a penalty method (*11*). The method struggled in some cases to get the highest energy state correct, but this issue seemed to be improved by computing more states—for instance, the error in the ^4P states of fluorine was cut in half by increasing the number of states from 5 to 10. In rare cases, the highest state converged to the incorrect state, such as boron with the Psiformer, which seemed to converge to the $^2P^o$ state rather than the last ^2D state. Fluorine and neon both have relatively large errors on the order of 1 to 2 mHa for low-lying states, but going from the FermiNet to the Psiformer ansatz reduced this error in all cases. The largest errors are in the highest states of fluorine and neon, on the order of 10 mHa. In this case, we suspect that the difficulty is due to the large number of different states with similar electron configurations and energies, and we hope that by computing even more states or by using even more expressive ansätze, the effects of individual states can be disentangled. The excellent performance on low-lying states gives us confidence that NES-VMC is mathematically sound.

Oscillator strengths

We are interested in the performance of NES-VMC on more complicated molecular systems as well as observable quantities other than the energy. The QUEST database (*38*–*46*) is an excellent source of benchmark vertical excited states calculations on molecules of various sizes, with consistent geometries and basis set extrapolations. Of particular interest is the subset of QUEST for which oscillator strengths have been computed (*38*) because oscillator strengths are both experimentally observable and the calculations are known to be highly sensitive to the choice of basis set (*47*).

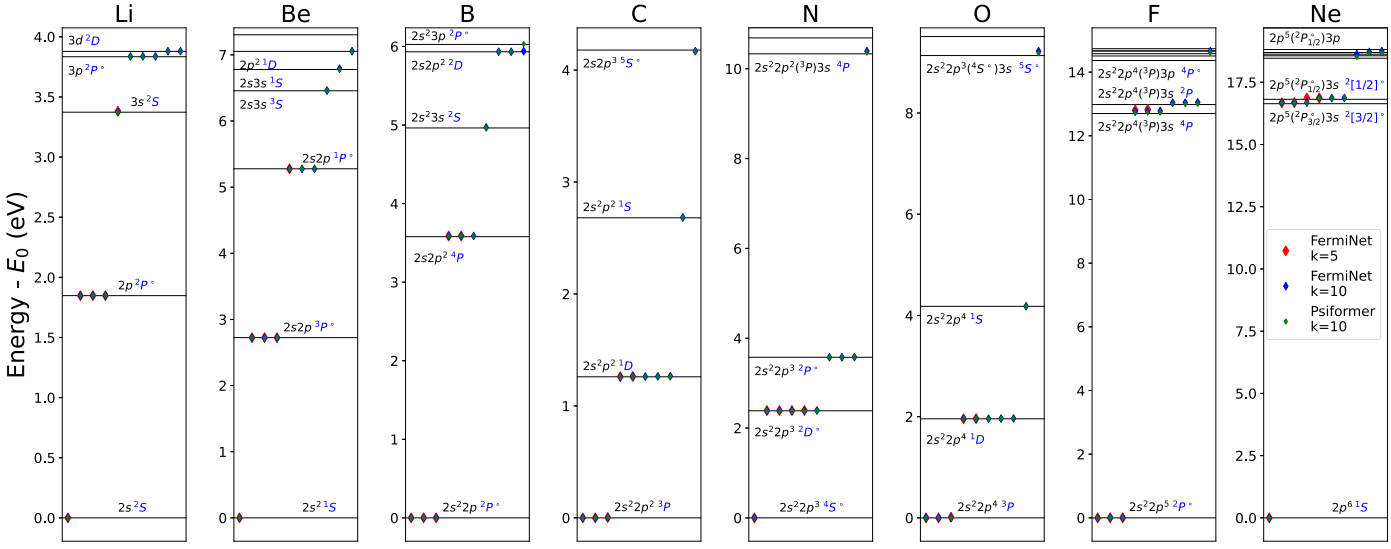


Fig. 1. Excited state energies for first-row atoms from lithium to neon. Results from NES-VMC applied to the FermiNet (10 states, blue; 5 states, red) are shown on top of experimental results (*37*). Spectral lines that match computed states are labeled with electron configurations and atomic term symbols (except for the highest levels of F and Ne, where term symbols are omitted for clarity). For all but the largest systems and highest excited states, there is excellent agreement with experiment. The discrepancy between 5 and 10 excited states is minimal except for the highest excited states of F and Ne, where computing more states increases the accuracy of a given state. Here and in other figures, error bars are too small to be seen, but are included in the tables in the supplementary materials. Complete numerical results are given in table S2.

Oscillator strengths are a measure of the probability of transition between different states occurring as a result of photon emission or absorption. The transition dipole moment between two states gives a measure of how that transition will interact with light

$$\mathbf{d}_{ij} = \left\langle \psi_i^\dagger \sum_k q_k \mathbf{r}_k \psi_j \right\rangle \tag{6}$$

where the sum over k is taken over all particles in the system with charge q_k and position $\mathbf{r}_k$. For electrons, $q_k = -e$. The transition dipole moments are vector-valued quantities, which include a complex phase, and are not directly observable. The oscillator strength of a particular transition is a dimensionless positive scalar that can be computed from the transition dipole moment and measured experimentally

$$f_{ij} = \frac{2}{3}\frac{m}{\hbar^2}(E_i - E_j)|\mathbf{d}_{ij}|^2 \tag{7}$$

Computational details are discussed in more detail in section S11.3 in the supplementary materials.

We applied NES-VMC to all of the small molecules investigated by Chrayteh *et al.* (*38*), computing the five lowest-energy states and the oscillator strengths of all transitions with both the FermiNet and Psiformer. Results are presented in Fig. 2 and table S3. Wherever possible, we took results from QUEST (*38*, *39*) to be theoretical best estimates (TBEs) for comparison, although for many of the states that we converged to, especially triplets, no results exist in QUEST. For molecules with heavier atoms (e.g., HCl, H_2S, and H_2CSi), we found that using pseudopotentials for the heaviest atoms substantially improved the accuracy of the results, likely because the total energy scale was reduced by ignoring core electrons. Where applicable, we also included a comparison against the VMC penalty method with the PauliNet ansatz of Entwistle *et al.* (*11*). We omitted N_2 because the lowest-lying excited states are all triplets. For all diatomic systems, the $^1\Pi$ state is doubly degenerate, and so the baseline oscillator strengths were divided by two to match the computed results.

In almost all cases, both the vertical excitation energies and the oscillator strengths were in excellent agreement with the TBE. The vertical excitation energies were almost all within chemical accuracy (1.6 mHa or 43 meV) of the TBE, and the oscillators strengths usually diverge from the TBE by at most an amount on the order of 0.001, comparable to the uncertainty in the calculations. We note that we did not use variance matching for any of the NES-VMC calculations.

There are a few cases where NES-VMC behaved oddly. Although the FermiNet and Psiformer found nearly identical vertical excitation energies for the $^1\Pi$ state of HCl, and the FermiNet accurately predicted the oscillator strength, the Psiformer mistakenly found this state to be a dark state. On formaldehyde (CH_2O), both the FermiNet and Psiformer failed to find the $^3A_1$ state at all, and the oscillator strength for the $^1B_2$ state diverged from the TBE by a substantial margin, although the Psiformer halved that margin relative to the FermiNet. Vertical excitation energies for systems with heavier atoms, such as H_2S, and the highest state of thioformaldehyde (CH_2S) were not quite as accurate as other results. For nitroxyl (HNO) and fluoromethylene (HCF), the network initialization must be carefully chosen to ensure convergence to the correct states (see "Double excitations" section). What is clear is that NES-VMC worked well in almost all cases and led to state-of-the-art results for neural network ansätze.

Carbon dimer

In addition to computing observable quantities, it is also desirable to be able to say something about the nature of different states. As a benchmark system for characterizing different states, we studied the carbon dimer (C_2). Despite its small size, C_2 has a complicated electronic structure with many low-lying excited states (*48*, *49*). Because of the existence of strong visible bands, C_2 is frequently detected in astrophysical measurements and can be observed in comets rich in organic materials (*50*). The exact bond order of C_2 is still a subject of some controversy—although molecular orbital theory would classify it as a double bond, valence bond calculations suggest that it may be better described as a quadruple bond (*51*). And C_2 is one of the smallest molecules to have low-lying double excitations, a class of excited states that other methods often struggle with (*41*). Correctly reconstructing the potential energy curves for different low-lying states requires correctly characterizing these different states at different geometries.

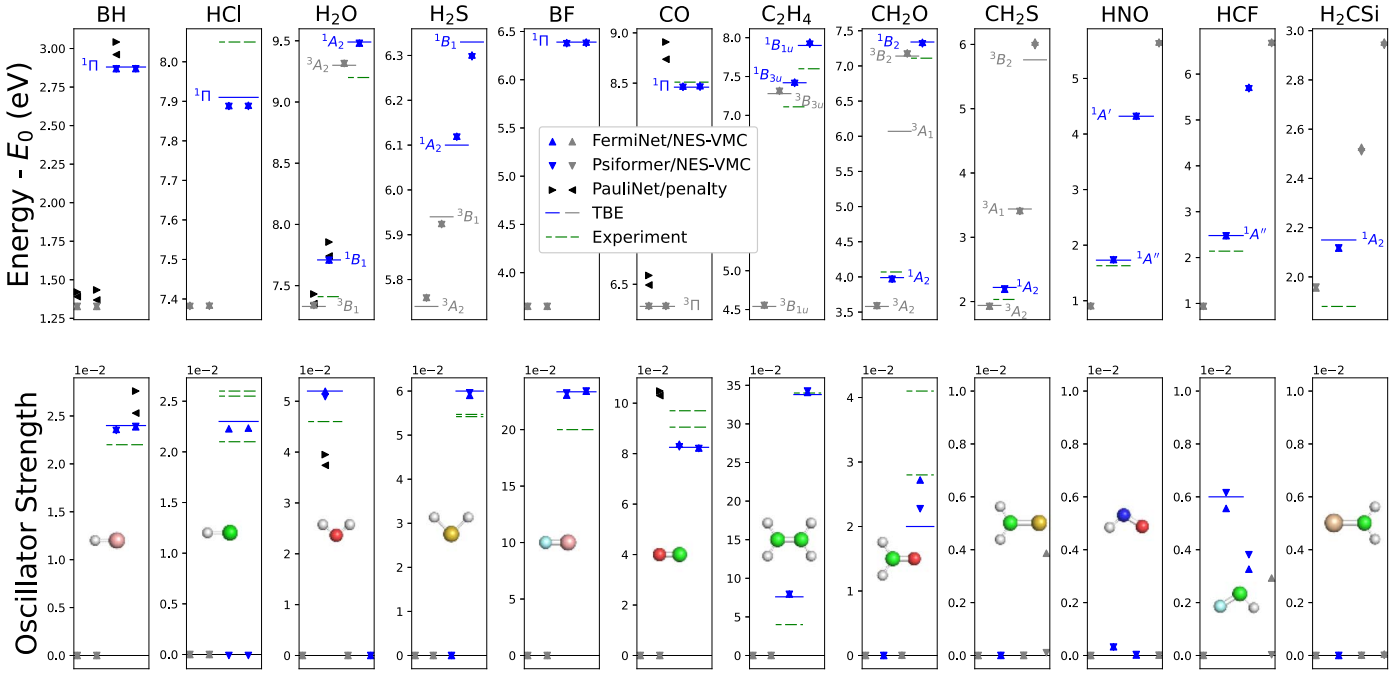


Fig. 2. Vertical excitation energies and oscillator strengths for small molecules. Singlet states are in blue, and triplet states are in gray. NES-VMC results are indicated by markers, and TBEs from Chrayteh *et al.* (*38*) or directly from QUEST (*39*) are given by the lines. When no data from QUEST are available, no TBE is given. Experimental results from Chrayteh *et al.* (*38*) and references therein are given by the dashed green lines. Where available, energies and oscillator strengths from Entwistle *et al.* (*11*) are provided (black triangles) for comparison, with (pointing left) and without (pointing right) variance matching. In almost all cases, our results on both energies and oscillator strengths agree closely with TBEs. Complete numerical results are given in table S3.

We computed the eight lowest-lying states of C_2 at several different bond lengths using the Psiformer ansatz and present the results in Fig. 3. We classified the different states by computing their spin magnitude and their parity and reflection symmetry. We did not compute the orbital angular momentum but confirmed that we see the expected degeneracy—for instance, Π states were doubly degenerate. In combination with reference energies from semistochastic heat-bath configuration interaction (SHCI) (*52*), we were able to match all computed energies to their respective states. The oscillator strengths at equilibrium showed several bright transitions, which we show in Fig. 3B. Because of the degeneracy of the Π states, we added the oscillator strengths together to give the total strength. We correctly identified the Phillips and Ballik-Ramsay systems (*53*, *54*) as well as the unnamed $B^1\Delta_g \rightarrow A^1\Pi_u$ transition. We also found that the energy of the $B^1\Delta_g$ energy closely matched the TBE in QUEST (*41*). The $A^1\Pi_u$, $c^3\Sigma_u^+$, and $b^3\Sigma_g^-$ states all have nearly the same energy at equilibrium, so correctly identifying the oscillator strengths for these transitions is challenging.

We found that our potential energy curves closely matched those computed by SHCI. The vertical energies matched SHCI to within chemical accuracy at all bond lengths, and the absolute energy of a state drifted by ~5 mHa/Å relative to SHCI as the bond length changed. We also computed the minimum energy along each interpolated potential energy curve to estimate the adiabatic energy of the excitations and found excellent agreement with experiment (*48*)—SHCI had a maximum error of 0.03 eV and mean absolute error (MAE) of 0.02 eV on the first five excitations, and our maximum error was 7 meV and MAE was 4 meV, a roughly fourfold improvement. The SHCI residual errors are correlated with the equilibrium bond length, so this 5-mHa/Å drift explains most of the difference in the results, which is strong evidence that the NES-VMC curve is closer to the ground truth than that of SHCI.

To better understand the nature of each state, we computed the occupancy of the different natural orbitals. We first computed the one-electron reduced density matrix (1-RDM) for each single-state ansatz in a large basis set and then diagonalized these matrices to find the natural orbitals, as described in more detail in section S11.2. In this case, the natural orbitals closely matched the Hartree-Fock molecular orbitals, so the 1-RDMs are nearly diagonal. We see in Fig. 3E that all states above the ground state involve excitation of electrons into the $2p_z\sigma_g$ orbital. The Π states are well described by single excitations from one of the $2p\pi_u$ orbitals, and the $c^3\Sigma_u^+$ state promotes an electron from the $2s\sigma_u^*$ orbital. Finally, both the $b^3\Sigma_g^-$ and $B^1\Delta_g$ states are double excitations of the $2p\pi_u$ electrons into the $2s\sigma_u^*$ orbital. Not only is NES-VMC able to predict double excitation energies correctly, but by having an explicit functional form for the wave function ansatz, we can compute quantities that allow us to derive insight about the nature of excitations.

Twisted ethylene

The excited states of ethylene (C_2H_4) across its potential energy surface present a challenging benchmark problem for many methods. As the carbon double bond is twisted, an avoided crossing occurs when the torsion angle is 90°. Even for ground state calculations, DFT and single-reference coupled cluster calculations predict an unphysical cusp at this location (*55*). Starting from the 90° torsion and bending the hydrogen atoms on one side inward (so-called "pyramidalization"), ethylene undergoes a conical intersection where the ground state transitions from a π to π^* highest occupied orbital (the *N* and *V* states, with term symbols ^1A_g and $^1B_{1u}$). Modeling this intersection requires multireference methods. TD-DFT struggles with this system (*56*), but MR-CI methods describe it well (*57*).

We computed the excited states of ethylene as the torsion angle is varied from 0° to 90°, followed by variation of the pyramidalization angle from 0° to 120°, enough to include the conical intersection of the *N* and *V* states. We used the geometry from previous studies (*57*). Results are shown in Fig. 4. There are also several low-lying triplet states of ethene, the $^3B_{1u}$ and $^3B_{3u}$ states, and so we calculated $K = 3$

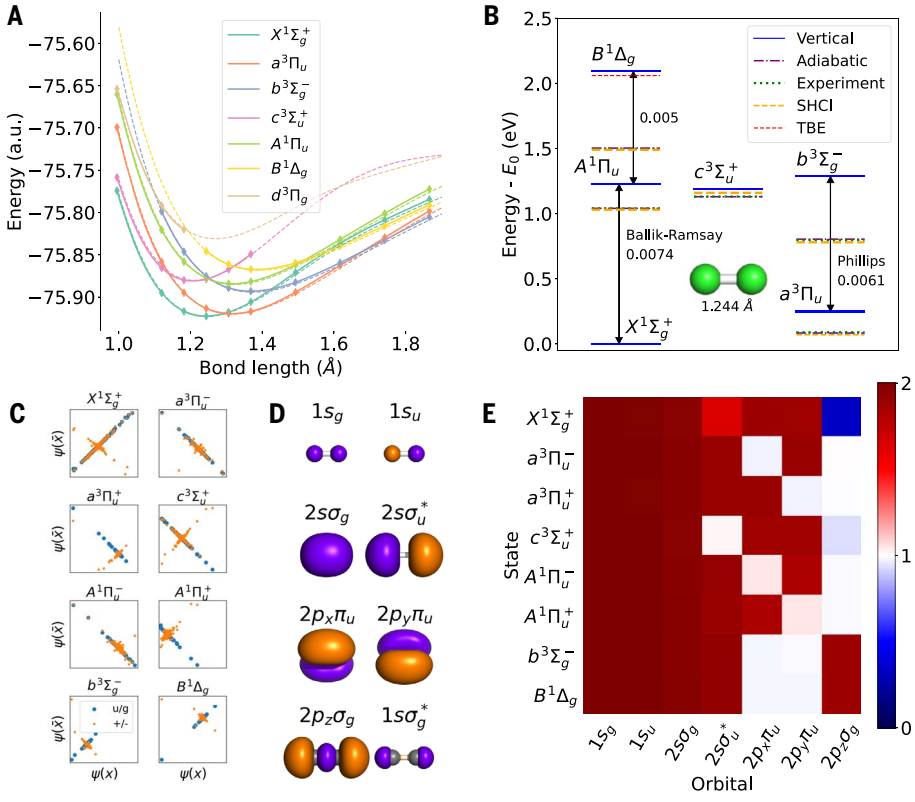


Fig. 3. Excited states of the carbon dimer (C_2). (**A**) Potential energy curves of the low-lying excited states of C_2, smoothed by cubic interpolation. The dotted lines were calculated by SHCI (*52*), shifted by 0.115 Ha to account for basis set effects. a.u., arbitrary units. (**B**) The vertical and adiabatic energies of excited states of C_2. The green line indicates experimental adiabatic energies (*48*), the orange line indicates adiabatic energies from SHCI (*52*), and the red line indicates the vertical energy of the $B^1\Delta_g$ state from QUEST (*41*). Bright transitions are labeled with their oscillator strength and, when available, their names. (**C**) The symmetries of the different states can be identified by evaluating each single-state ansatz at location **r** and −**r** for parity symmetry (u/g; blue) or by flipping **r** across the *x* axis for reflection symmetry (+/−; orange). (**D**) Visualization of the eight lowest natural orbitals of C_2. (**E**) The occupancy of the different natural orbitals for the different excited states of C_2, identified from the density matrix of each state. The $a^3\Pi_u$ through $A^1\Pi_u$ states are single excitations, and the $b^3\Sigma_g^-$ and $B^1\Delta_g$ states are double excitations. Complete numerical results are given in tables S5, S6, and S7.

excited states for all geometries, which we found was enough to find two singlet states for all geometries except at equilibrium, where we used $K = 5$ and took the highest state because the $^1B_{3u}$ state has lower energy exclusively at equilibrium. We did not find a substantial difference between the FermiNet and Psiformer, and we show the Psiformer results in this work. For comparison, in addition to TD-DFT (*58*) and MR-CI, we also compared against the PauliNet penalty method (*11*).

Qualitatively, the results from NES-VMC closely matched those of MR-CI. The spurious cusp when the torsion angle is 90° was avoided, and the error in the ground state relative to MR-CI was smaller than for the PauliNet penalty method across torsion angles. The nonparallelity error in the *V* state relative to MR-CI was lower for our method compared with the PauliNet penalty method, and our predicted location for the conical intersection (~97.5°) was closer to the MR-CI value (~96°) than the predicted PauliNet penalty method value (~100°). There was a nearly constant shift in the energy of the *V* state on the order of several tenths of an electron volt relative to MR-CI and a shift in the energy of the *N* state, which grew as the pyramidalization angle grew. Increasing the number of excited states and using a different ansatz did not seem to make a difference. We note that when using the equilibrium geometry for ethylene from QUEST in "Oscillator strengths" as opposed to the geometry from MR-CI, our results agreed with the TBEs to within chemical accuracy. The overall agreement with experimentally relevant quantities such as the location of the conical intersection was in excellent agreement with other highly accurate theoretical studies, and NES-VMC was able to capture the important behavior of this system across the potential energy surface.

Double excitations

Accurate calculation of double excitations is known to be far more challenging than single excitations (*41*). We already demonstrated that NES-VMC is effective at computing the double excitations of nitroxyl and the carbon dimer. To see how well NES-VMC scales to larger systems, we investigated five systems with 24 to 42 electrons known to have low-lying full or partial double excitations: nitrosomethane, butadiene, glyoxal, tetrazine, and cyclopentadienone. For the last two systems, both of which have the same number of electrons as benzene, the original TBEs from QUEST (*39*, *41*) were known to be unsafe for double excitations, and only recently did more accurate calculations from QMC (*59*) and complete-active-space third-order perturbation theory (CASPT3) (*60*) resolve discrepancies as large as almost 1 eV for some states. For these two especially challenging systems, we added an extra term to the Hamiltonian to push up the energy of triplet states so that the excitations of interest could be resolved with a reasonable computational budget (see section S5). Results on all systems are shown in Fig. 5 and table S9.

Butadiene is of particular interest because it is the smallest conjugated organic molecule—a class of molecules whose photochemistry is relevant to vision, photosynthesis, dyes, and photovoltaics. The exact ordering of the two lowest-lying singlet transitions, the bright $1^1A_g \rightarrow 1^1B_u$ and dark $1^1A_g \rightarrow 2^1A_g$ transitions, has been the subject of controversy for many years (*61*), only being resolved in the past decade or so (*62*) after extensive study. Whereas the 1^1B_u state is a single excitation, the 2^1A_g is known to have roughly 30% double excitation character, which makes it especially challenging to compute. For the FermiNet and Psiformer, NES-VMC was able to not only correctly predict the ordering of the states but match the TBE from QUEST to within 92 and 60 meV, respectively, for the 1^1B_u state and within 62 and 9 meV, respectively, for the 2^1A_g, a markedly high degree of agreement for such a notorious system.

On all double excitation systems with the Psiformer, and for four out of five systems with the FermiNet, NES-VMC was in excellent agreement with the best computational results. The one exception for the FermiNet, glyoxal, is at the scale where the FermiNet is known to perform worse than the Psiformer at ground state calculations (*36*), so it is not surprising that it struggles on some systems of this size. The FermiNet and Psiformer achieved a MAE relative to the TBE of 15 and 21 meV on nitrosomethane, 84 and 38 meV on butadiene, 167 and 28 meV on glyoxal, 45 and 54 meV on tetrazine, and 92 and 66 meV on cyclopentadienone, respectively. For the Psiformer, this error was within chemical accuracy (43 meV) for all systems except tetrazine and cyclopentadienone. On tetrazine, the Psiformer is within 0.1 eV of

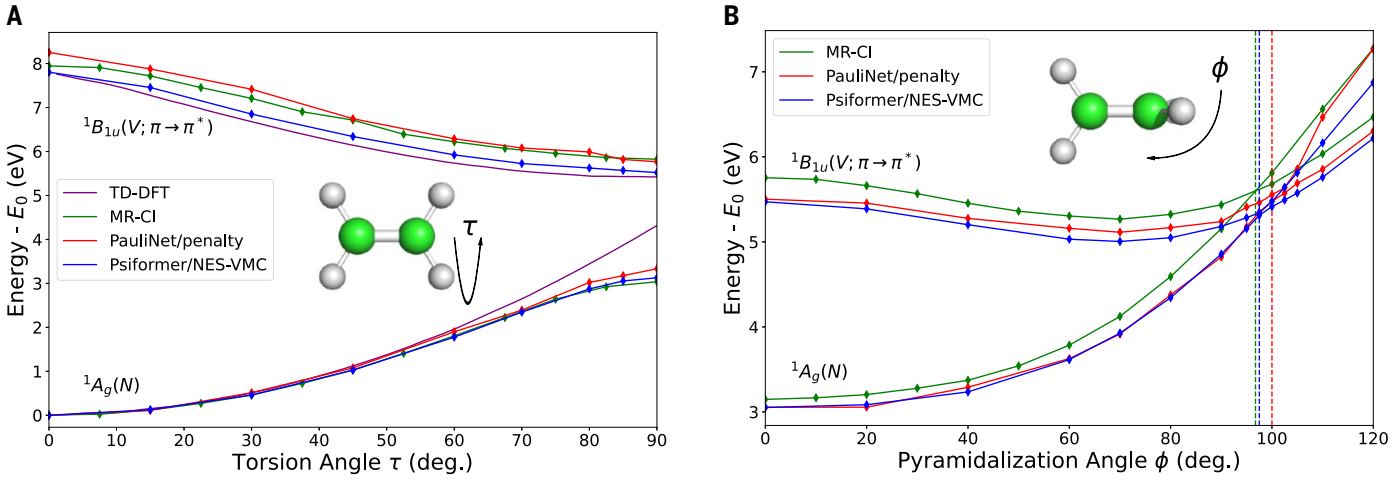


Fig. 4. Excited states and conical intersection of ethylene (C_2H_4). (**A**) Potential energy curve of the first two singlet states of ethylene under torsion around the C-C bond. Our results (blue) are compared against TD-DFT (*58*) (purple), MR-CI (*57*) (green), and a penalty method used with the PauliNet, without the variance matching correction (*11*) (red). (**B**) Potential energy curve of the first two singlet states of ethylene under pyramidalization of the C-H bonds. The best estimate of the location of the conical intersection of the *V* and *N* states for each method is given by the vertical line. Our method is in close agreement with MR-CI up to a constant shift and agrees with the location of the conical intersection better than the PauliNet penalty method. Note that the $\varphi = 0$ geometry in (B) differs slightly from the τ = 90 geometry in (A), as in Barbatti *et al.* (*57*). All results were normalized so that the ground state energy at the equilibrium geometry is 0. Complete numerical results are given in table S8.

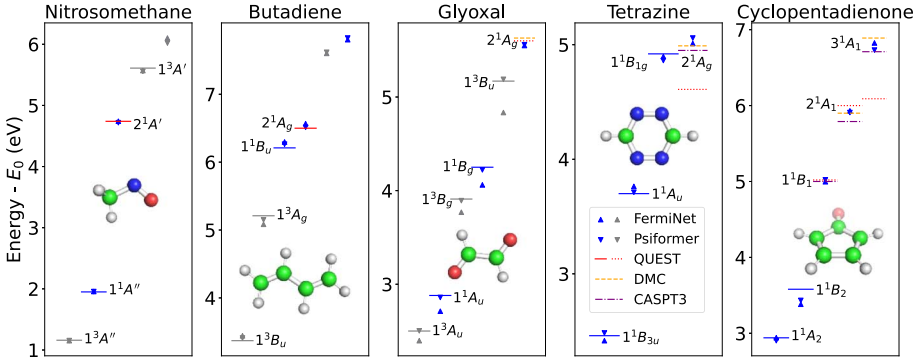


Fig. 5. Excited states of larger double excitation systems. NES-VMC with FermiNet and Psiformer for singlet (blue) and triplet (gray) systems are compared against results from the QUEST database (*39*) (singlet single excitations in blue, triplets in gray, and singlet double excitations in red), including "unsafe" results (dashed lines). For systems where the QUEST results are unsafe, more accurate results from diffusion Monte Carlo (DMC) (orange) (*59*) or with a CASPT3 correction (purple) (*60*) are given. NES-VMC closely matched the more accurate double excitation calculations on the largest and most challenging systems. Complete numerical results are given in table S9.

the best estimates of the 2^1A_g vertical excitation energy. The previous TBE in QUEST was off by nearly 1 eV. On cyclopentadienone, even the best current estimates of the 2^1A_1 and 3^1A_1 excitation energies disagree by 0.1 to 0.15 eV, a range that the Psiformer was within. These results demonstrate that NES-VMC is among the state of the art in challenging excited state calculations, where even other top methods disagree by more than chemical accuracy.

Benzene

Finally, we applied NES-VMC with both the FermiNet and Psiformer to benzene. Although benzene is the same size as tetrazine and cyclopentadienone, it is a common benchmark for medium-sized molecules, so there are more abundant data for us to compare against. For VMC, in addition to the penalty method of Entwistle *et al.* (*11*), there is also the penalty method of Pathak *et al.* (*19*), which is used with a traditional Slater-Jastrow ansatz and uses a different penalty function that allows for unbiased gradients. On top of VMC results and coupled cluster–based TBEs from QUEST, we also compared against CASPT2 (*63*) and TD-DFT with the PBE0 functional (*64*). Results are shown in Fig. 6, with complex numerical results given in table S10. For our calculations, we used the same geometry as in QUEST (*42*).

To better understand the nature of the excitations computed, we inspected the density matrices of the respective states, similarly to the analysis of C_2 in Fig. 3, D and E. The density matrices in the Hartree-Fock basis were nearly diagonal. All five excited states for benzene that we computed were single excitations from a π to π^* orbital, but they were best described by exciting half an electron from two distinct π_g orbitals into two distinct π_u^* orbitals. These orbitals are visualized in Fig. 6B.

NES-VMC with the Psiformer came close to reaching the TBE for all computed states. The FermiNet was not quite as accurate and struggled with the highest-energy $^3B_{2u}$ state. The highest excited state of the FermiNet converged to a mixture of a triplet and singlet state, which suggests that contamination from the $^1B_{1u}$ state was affecting the performance. Like with glyoxal, it is not surprising that the Psiformer is better suited for computing vertical excitation energies in this case. CASPT2 and TD-DFT methods were less accurate across the board, and CASPT2 was generally intermediate in accuracy between TD-DFT and coupled cluster. The penalty method of Pathak *et al.* (*19*) generally reached comparable levels of accuracy to those of NES-VMC with the Psiformer. Additionally, the results reported in Pathak *et al.* include a diffusion Monte Carlo correction that reduced the error by ~0.1 eV. NES-VMC did not include any postprocessing of VMC results.

The only other neural network approach that has been tried on molecules this large is the penalty method of Entwistle *et al.* in combination with the PauliNet ansatz (*11*), albeit only

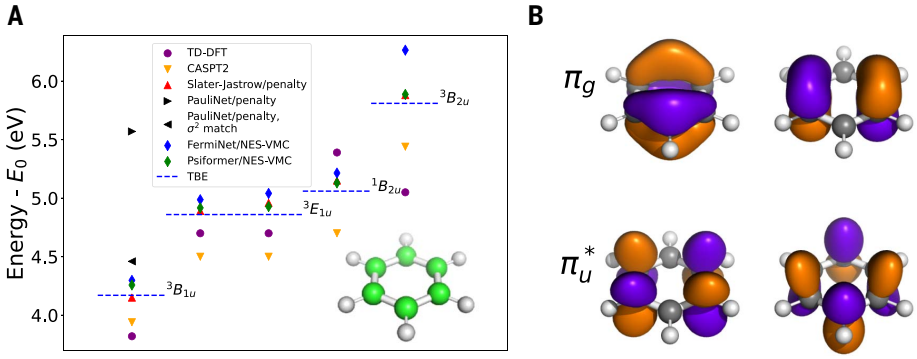


Fig. 6. Excited states of benzene. (**A**) Energy levels of benzene. The NES-VMC results (green and blue) are compared against TBEs from QUEST (*42*, *46*) alongside TD-DFT–PBE0 (*64*), CASPT2 (*63*), DMC with a Slater-Jastrow ansatz and penalty method (*19*), and the PauliNet with a penalty method, with and without variance matching (*11*). NES-VMC with the Psiformer ansatz was competitive with state-of-the-art methods and was far more accurate compared with the only other neural network result, the PauliNet penalty method. Complete numerical results are given in table S10. (**B**) The orbitals involved in the excitation of benzene. All excitations computed here are $\pi \rightarrow \pi^*$ excitations. (Top row) π orbitals occupied in the ground state. (Bottom row) π^* orbitals occupied in the excited states.

for the $^1A_{1g}$ and $^3B_{1u}$ states. Relative to the TBE, the FermiNet and Psiformer errors on the $^1A_{1g} \rightarrow {}^3B_{1u}$ transition were 0.126 and 0.088 eV, respectively. Without variance matching, the PauliNet penalty method error was 1.4 eV, and even with variance matching, the error was only reduced to 0.29 eV, substantially worse than the Psiformer. This shows that NES-VMC used in conjunction with a neural network ansatz can compute excited state properties to higher accuracy compared with other methods on large molecules.

Conclusions and outlook

We have presented a method for calculating excited state properties of quantum systems by VMC, the natural excited states method (NES-VMC). NES-VMC has no free parameters to tune and allows for unbiased estimation of energies and gradients by reformulating a multistate approach as the problem of finding the ground state of an extended system. In much the same way that sampling from ψ^2 considerably simplifies the computation of ground state properties by VMC, NES-VMC provides a simple variational principle for computing excited state properties. Additionally, it dovetails well with recent work on neural network ansätze for many-body systems.

We have demonstrated the effectiveness of NES-VMC on a number of benchmark problems ranging from small atoms and molecules up to benzene-scale molecules. In all cases, NES-VMC was competitive with TBEs for energies and oscillator strengths and could capture the behavior of challenging double excitations and conical intersections. It enabled unprecedented accuracy for neural network ansätze as system size scales. The optimized ansatz could be used in downstream analyses to characterize the nature of the electronic structure of different excited states. NES-VMC was as effective as any other method for computing excited states with QMC that we were able to compare against and is so far the only effective method for neural network ansätze on larger systems, with the added benefit of simplicity and generality.

Although we focused on applications using neural network ansätze, which can be computationally expensive, classic ansätze such as the Slater-Jastrow ansatz can be scaled to much larger systems (*23*). As the system size grows, applying different weighting to the different states is often needed to get good performance with existing QMC methods (*65*). It remains to be seen whether the same holds true for NES-VMC. Although our results suggest that more accurate ansätze are quite important for achieving good performance, we look forward to finding out how well NES-VMC works in conjunction with these classic ansätze on large problems.

Finally, although our experiments in this paper focused on electronic excitations of molecular systems, NES-VMC is fully general and can be applied to any quantum Hamiltonian. Vibronic couplings, where electronic and vibrational excitations interact, are critical for understanding the behavior of molecules near avoided crossings and conical intersections (*66*). Vibronic couplings could be investigated with neural networks by combining NES-VMC with recent work on fitting multiple geometries of molecular systems with a single neural network ansatz (*67*, *68*). Going even further, vibronic couplings could be studied by applying NES-VMC to Hamiltonians beyond the Born-Oppenheimer approximation, which deal with nuclear positions in a fully quantum manner. Beyond molecular systems, excited state calculations with QMC are an important tool for studying nuclear physics (*9*), optical bandgaps in condensed matter physics (*16*, *69*), many properties of spin systems, as well as time dynamics and finite temperature phenomena. Neural network ansätze have already been applied to ground state calculations in some of these domains (*70*, *71*). We are excited to see how NES-VMC and deep neural networks can be applied to many of the most challenging open problems in many-body quantum mechanics in the future.

Materials and methods summary

All algorithms described in the paper were implemented in JAX, a Python framework for automatic differentiation, and run on a cluster of A100 graphics processing units (GPUs). In ground state VMC, a set of walkers representing electron positions are updated by Markov chain Monte Carlo (MCMC) to sample from the wave function ansatz squared, and the wave function ansatz is updated by gradient descent to minimize an energy functional. In our method, we extended the standard VMC algorithm to track a matrix of energies and minimize the trace of this matrix, as described in the "Natural excited states" section. We then run inference with a fixed ansatz, accumulate the matrix of energies and other observables, and diagonalize the energy matrix to recover the energies and observables for individual states. We apply this algorithm to the FermiNet and Psiformer ansätze, which are deep neural network architectures that have previously been used for ground state calculations. Hyperparameters are similar to those used in ground state calculations in prior publications. Further details on the materials and methods used can be found in the supplementary materials.

REFERENCES AND NOTES

1. L. E. Brus, Electron–Electron and Electron-Hole Interactions in Small Semiconductor Crystallites: The Size Dependence of the Lowest Excited Electronic State. *J. Chem. Phys.* **80**, 4403–4409 (1984). doi: 10.1063/1.447218
2. D. Polli *et al*., Conical intersection dynamics of the primary photoisomerization event in vision. *Nature* **467**, 440–443 (2010). doi: 10.1038/nature09346; pmid: 20864998
3. C. K. Prier, D. A. Rankic, D. W. MacMillan, Visible light photoredox catalysis with transition metal complexes: Applications in organic synthesis. *Chem. Rev.* **113**, 5322–5363 (2013). doi: 10.1021/cr300503r; pmid: 23509883
4. C. J. Chiara *et al*., Isomer depletion as experimental evidence of nuclear excitation by electron capture. *Nature* **554**, 216–218 (2018). doi: 10.1038/nature25483; pmid: 29420479
5. J. Westermayr, P. Marquetand, Machine Learning for Electronically Excited States of Molecules. *Chem. Rev.* **121**, 9873–9926 (2021). doi: 10.1021/acs.chemrev.0c00749; pmid: 33211478
6. G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017). doi: 10.1126/science.aag2302; pmid: 28183973
7. J. Hermann *et al*., Ab initio quantum chemistry with neural-network wavefunctions. *Nat. Rev. Chem.* **7**, 692–709 (2023). doi: 10.1038/s41570-023-00516-8; pmid: 37558761
8. W. M. C. Foulkes, L. Mitas, R. J. Needs, G. Rajagopal, Quantum Monte Carlo Simulations of Solids. *Rev. Mod. Phys.* **73**, 33–83 (2001). doi: 10.1103/RevModPhys.73.33
9. J. Carlson *et al*., Quantum Monte Carlo Methods for Nuclear Physics. *Rev. Mod. Phys.* **87**, 1067–1118 (2015). doi: 10.1103/RevModPhys.87.1067
10. K. Choo, G. Carleo, N. Regnault, T. Neupert, Symmetries and Many-Body Excitations with Neural-Network Quantum States.

Phys. Rev. Lett. **121**, 167204 (2018). doi: 10.1103/PhysRevLett.121.167204; pmid: 30387658
11. M. T. Entwistle, Z. Schätzle, P. A. Erdman, J. Hermann, F. Noé, Electronic excited states in deep variational Monte Carlo. *Nat. Commun.* **14**, 274 (2023). doi: 10.1038/s41467-022-35534-5; pmid: 36650151
12. Y. Kwon, D. M. Ceperley, R. M. Martin, Effects of three-body and backflow correlations in the two-dimensional electron gas. *Phys. Rev. B* **48**, 12037–12046 (1993). doi: 10.1103/PhysRevB.48.12037; pmid: 10007551
13. M. Bajdich, L. Mitas, L. Wagner, K. Schmidt, Pfaffian Pairing and Backflow Wavefunctions for Electronic Structure Quantum Monte Carlo Methods. *Phys. Rev. B* **77**, 115112 (2008). doi: 10.1103/PhysRevB.77.115112
14. S. Sorella, Green Function Monte Carlo with Stochastic Reconfiguration. *Phys. Rev. Lett.* **80**, 4558–4561 (1998). doi: 10.1103/PhysRevLett.80.4558
15. J. Toulouse, C. J. Umrigar, Optimization of quantum Monte Carlo wave functions by energy minimization. *J. Chem. Phys.* **126**, 084102 (2007). doi: 10.1063/1.2437215; pmid: 17343435
16. L. Zhao, E. Neuscamman, Variational Excitations in Real Solids: Optical Gaps and Insights into Many-Body Perturbation Theory. *Phys. Rev. Lett.* **123**, 036402 (2019). doi: 10.1103/PhysRevLett.123.036402; pmid: 31386452
17. L. Otis, E. Neuscamman, A Promising Intersection of Excited-State-Specific Methods from Quantum Chemistry and Quantum Monte Carlo. *WIREs Comput. Mol. Sci.* **13**, e1659 (2023). doi: 10.1002/wcms.1659
18. P. M. Zimmerman, J. Toulouse, Z. Zhang, C. B. Musgrave, C. J. Umrigar, Excited states of methylene from quantum Monte Carlo. *J. Chem. Phys.* **131**, 124103 (2009). doi: 10.1063/1.3220671; pmid: 19791848
19. S. Pathak, B. Busemeyer, J. N. B. Rodrigues, L. K. Wagner, Excited states in variational Monte Carlo using a penalty method. *J. Chem. Phys.* **154**, 034101 (2021). doi: 10.1063/5.0030949; pmid: 33499613
20. W. A. Wheeler, K. G. Kleiner, L. K. Wagner, Ensemble Variational Monte Carlo for Optimization of Correlated Excited State Wave Functions. *Electron. Struct.* **6**, 025001 (2024). doi: 10.1088/2516-1075/ad38f8
21. F. Schautz, C. Filippi, Optimized Jastrow–Slater wave functions for ground and excited states: Application to the lowest states of ethene. *J. Chem. Phys.* **120**, 10931–10941 (2004). doi: 10.1063/1.1752881; pmid: 15268123
22. F. Cordova *et al.*, Troubleshooting time-dependent density-functional theory for photochemical applications: Oxirane. *J. Chem. Phys.* **127**, 164111 (2007). doi: 10.1063/1.2786997; pmid: 17979323
23. C. Filippi, M. Zaccheddu, F. Buda, Absorption Spectrum of the Green Fluorescent Protein Chromophore: A Difficult Case for ab Initio Methods? *J. Chem. Theory Comput.* **5**, 2074–2087 (2009). doi: 10.1021/ct900227j; pmid: 26613149
24. A. Cuzzocrea, A. Scemama, W. J. Briels, S. Moroni, C. Filippi, Variational Principles in Quantum Monte Carlo: The Troubled Story of Variance Minimization. *J. Chem. Theory Comput.* **16**, 4203–4212 (2020). doi: 10.1021/acs.jctc.0c00147; pmid: 32419451
25. M. Dash, S. Moroni, C. Filippi, A. Scemama, Tailoring CIPSI Expansions for QMC Calculations of Electronic Excitations: The Case Study of Thiophene. *J. Chem. Theory Comput.* **17**, 3426–3434 (2021). doi: 10.1021/acs.jctc.1c00212; pmid: 34029098
26. D. Luo, B. K. Clark, Backflow Transformations via Neural Networks for Quantum Many-Body Wave Functions. *Phys. Rev. Lett.* **122**, 226401 (2019). doi: 10.1103/PhysRevLett.122.226401; pmid: 31283262
27. D. Pfau, J. S. Spencer, A. G. Matthews, W. M. C. Foulkes, *Ab initio* Solution of the Many-Electron Schrödinger Equation with Deep Neural Networks. *Phys. Rev. Res.* **2**, 033429 (2020). doi: 10.1103/PhysRevResearch.2.033429
28. J. Hermann, Z. Schätzle, F. Noé, Deep-neural-network solution of the electronic Schrödinger equation. *Nat. Chem.* **12**, 891–897 (2020). doi: 10.1038/s41557-020-0544-y; pmid: 32968231
29. J. J. Dorando, J. Hachmann, G. K. Chan, Targeted excited state algorithms. *J. Chem. Phys.* **127**, 084109 (2007). doi: 10.1063/1.2768360; pmid: 17764231
30. M. Lewin, On the Computation of Excited States with MCSCF Methods. *J. Math. Chem.* **44**, 967–980 (2008). doi: 10.1007/s10910-008-9355-x
31. L. Bottou, O. Bousquet, in *Advances in Neural Information Processing Systems*, J. Platt, D. Koller, Y. Singer, S. Roweis, Eds. (Curran Associates, Inc., 2007).
32. D. M. Ceperley, B. Bernu, The Calculation of Excited State Properties with Quantum Monte Carlo. *J. Chem. Phys.* **89**, 6316–6328 (1988). doi: 10.1063/1.455398
33. M. P. Nightingale, V. Melik-Alaverdian, Optimization of ground- and excited-state wave functions and van der Waals clusters. *Phys. Rev. Lett.* **87**, 043401 (2001). doi: 10.1103/PhysRevLett.87.043401; pmid: 11461615
34. J. Carlson, V. R. Pandharipande, R. B. Wiringa, Variational Calculations of Resonant States in ^{4}He. *Nucl. Phys. A.* **424**, 47–59 (1984). doi: 10.1016/0375-9474(84)90127-1
35. H. Robbins, S. Monro, A Stochastic Approximation Method. *Ann. Math. Stat.* **22**, 400–407 (1951). doi: 10.1214/aoms/1177729586
36. I. von Glehn, J. S. Spencer, D. Pfau, A Self-Attention Ansatz for Ab-initio Quantum Chemistry, *Eleventh International Conference on Learning Representations* (ICLR, 2023).
37. J. E. Sansonetti, W. C. Martin, Handbook of Basic Atomic Spectroscopic Data. *J. Phys. Chem. Ref. Data* **34**, 1559–2259 (2005). doi: 10.1063/1.1800011
38. A. Chrayteh, A. Blondel, P.-F. Loos, D. Jacquemin, Mountaineering Strategy to Excited States: Highly Accurate Oscillator Strengths and Dipole Moments of Small Molecules. *J. Chem. Theory Comput.* **17**, 416–438 (2021). doi: 10.1021/acs.jctc.0c01111; pmid: 33256412
39. M. Véril *et al.*, QUESTDB: A Database of Highly Accurate Excitation Energies for the Electronic Structure Community. *WIREs Comput. Mol. Sci.* **11**, e1517 (2021). doi: 10.1002/wcms.1517
40. P.-F. Loos *et al.*, A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks. *J. Chem. Theory Comput.* **14**, 4360–4379 (2018). doi: 10.1021/acs.jctc.8b00406; pmid: 29966098
41. P.-F. Loos, M. Boggio-Pasqua, A. Scemama, M. Caffarel, D. Jacquemin, Reference Energies for Double Excitations. *J. Chem. Theory Comput.* **15**, 1939–1956 (2019). doi: 10.1021/acs.jctc.8b01205; pmid: 30689951
42. P.-F. Loos, F. Lipparini, M. Boggio-Pasqua, A. Scemama, D. Jacquemin, A Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Medium Sized Molecules. *J. Chem. Theory Comput.* **16**, 1711–1741 (2020). doi: 10.1021/acs.jctc.9b01216; pmid: 31986042
43. P.-F. Loos, A. Scemama, M. Boggio-Pasqua, D. Jacquemin, Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Exotic Molecules and Radicals. *J. Chem. Theory Comput.* **16**, 3720–3736 (2020). doi: 10.1021/acs.jctc.0c00227; pmid: 32379442
44. P.-F. Loos, M. Comin, X. Blase, D. Jacquemin, Reference Energies for Intramolecular Charge-Transfer Excitations. *J. Chem. Theory Comput.* **17**, 3666–3686 (2021). doi: 10.1021/acs.jctc.1c00226; pmid: 33955742
45. P.-F. Loos, D. Jacquemin, A Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Bicyclic Systems. *J. Phys. Chem. A* **125**, 10174–10188 (2021). doi: 10.1021/acs.jpca.1c08524; pmid: 34792354
46. P.-F. Loos, F. Lipparini, D. A. Matthews, A. Blondel, D. Jacquemin, A Mountaineering Strategy to Excited States: Revising Reference Values with EOM-CC4. *J. Chem. Theory Comput.* **18**, 4418–4427 (2022). doi: 10.1021/acs.jctc.2c00416; pmid: 35737466
47. R. Crossley, Fifteen Years On - the Calculation of Atomic Transition Probabilities Revisited. *Phys. Scr.* **1984**, 117–128 (1984). doi: 10.1088/0031-8949/1984/T8/020
48. M. Martin, C_2 spectroscopy and kinetics. *J. Photochem. Photobiol. Chem.* **66**, 263–289 (1992). doi: 10.1016/1010-6030(92)80001-C
49. J. G. Phillips, S. P. Davis, *The Swan System of the C2 Molecule: The Spectrum of the HgH Molecule* (Univ. California Press, 1968).
50. K. Venkataramani *et al.*, Optical Spectroscopy of Comet C/2014 Q2 (Lovejoy) from the Mount Abu Infrared Observatory. *Mon. Not. R. Astron. Soc.* **463**, 2137–2144 (2016). doi: 10.1093/mnras/stw1820
51. S. Shaik *et al.*, Quadruple bonding in C_2 and analogous eight-valence electron species. *Nat. Chem.* **4**, 195–200 (2012). doi: 10.1038/nchem.1263; pmid: 22354433
52. A. A. Holmes, C. J. Umrigar, S. Sharma, Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.* **147**, 164111 (2017). doi: 10.1063/1.4998614; pmid: 29096501
53. E. Ballik, D. Ramsay, The $A^3\Sigma_g^- X^3\Pi_u$ Band System of the C_2 Molecule. *Astrophys. J.* **137**, 61 (1963). doi: 10.1086/147485
54. E. Ballik, D. Ramsay, An Extension of the Phillips System of C_2 and a Survey of C_2 States. *Astrophys. J.* **137**, 84 (1963). doi: 10.1086/147486
55. A. I. Krylov, C. D. Sherrill, E. F. Byrd, M. Head-Gordon, Size-Consistent Wave Functions for Nondynamical Correlation Energy: The Valence Active Space Optimized Orbital Coupled-Cluster Doubles Model. *J. Chem. Phys.* **109**, 10669–10678 (1998). doi: 10.1063/1.477764
56. M. Barbatti, R. Crespo-Otero, in *Density-Functional Methods for Excited States*, N. Ferré, M. Filatov, M. Huix-Rotllant, Eds., vol. 368 of *Topics in Current Chemistry* (Springer, 2014), pp. 415–444.
57. M. Barbatti, J. Paier, H. Lischka, Photochemistry of ethylene: A multireference configuration interaction investigation of the excited-state energy surfaces. *J. Chem. Phys.* **121**, 11614–11624 (2004). doi: 10.1063/1.1807378; pmid: 15634126
58. M. Mališ, S. Luber, Trajectory Surface Hopping Nonadiabatic Molecular Dynamics with Kohn-Sham ΔSCF for Condensed-Phase Systems. *J. Chem. Theory Comput.* **16**, 4071–4086 (2020). doi: 10.1021/acs.jctc.0c00372; pmid: 32602729
59. S. Shepard, R. L. Panadés-Barrueta, S. Moroni, A. Scemama, C. Filippi, Double Excitation Energies from Quantum Monte Carlo Using State-Specific Energy Optimization. *J. Chem. Theory Comput.* **18**, 6722–6731 (2022). doi: 10.1021/acs.jctc.2c00769; pmid: 36314602
60. F. Kossoski, M. Boggio-Pasqua, P.-F. Loos, D. Jacquemin, Reference Energies for Double Excitations: Improvement and Extension. *J. Chem. Theory Comput.* **20**, 5655–5678 (2024). doi: 10.1021/acs.jctc.4c00410
61. R. J. Buenker, J. L. Whitten, *Ab initio* SCF MO and CI studies of the electronic states of butadiene. *J. Chem. Phys.* **49**, 5381–5387 (1968). doi: 10.1063/1.1670062
62. M. A. Watson, G. K.-L. Chan, Excited States of Butadiene to Chemical Accuracy: Reconciling Theory and Experiment. *J. Chem. Theory Comput.* **8**, 4013–4018 (2012). doi: 10.1021/ct300591z; pmid: 26605568
63. B. O. Roos, K. Andersson, M. P. Fülscher, Towards an Accurate Molecular Orbital Theory for Excited States: The Benzene Molecule. *Chem. Phys. Lett.* **192**, 5–13 (1992). doi: 10.1016/0009-2614(92)85419-B
64. C. Adamo, G. E. Scuseria, V. Barone, Accurate Excitation Energies from Time-Dependent Density Functional Theory: Assessing the PBE0 Model. *J. Chem. Phys.* **111**, 2889–2899 (1999). doi: 10.1063/1.479571
65. A. Cuzzocrea, S. Moroni, A. Scemama, C. Filippi, Reference excitation energies of increasingly large molecules: A QMC study of cyanine dyes. *J. Chem. Theory Comput.* **18**, 1089–1095 (2022). doi: 10.1021/acs.jctc.1c01162; pmid: 35080893
66. T. Azumi, K. Matsuzaki, What Does the Term "Vibronic Coupling" Mean? *Photochem. Photobiol.* **25**, 315–326 (1977). doi: 10.1111/j.1751-1097.1977.tb06918.x
67. N. Gao, S. Günnemann, Ab-initio Potential Energy Surfaces by Pairing GNNs with Neural Wave Functions, *Tenth International Conference on Learning Representations* (ICLR, 2022).
68. M. Scherbela, L. Gerard, P. Grohs, Towards a transferable fermionic neural wavefunction for molecules. *Nat. Commun.* **15**, 120 (2024). doi: 10.1038/s41467-023-44216-9; pmid: 38168035
69. R. J. Hunt, M. Szyniszewski, G. I. Prayogo, R. Maezono, N. D. Drummond, Quantum Monte Carlo Calculations of Energy Gaps from First Principles. *Phys. Rev. B* **98**, 075122 (2018). doi: 10.1103/PhysRevB.98.075122
70. Y. Yang, P. Zhao, Deep-Neural-Network Approach to Solving the *ab initio* Nuclear Structure Problem. *Phys. Rev. C* **107**, 034320 (2023). doi: 10.1103/PhysRevC.107.034320
71. X. Li, Z. Li, J. Chen, Ab initio calculation of real solids via neural network ansatz. *Nat. Commun.* **13**, 7895 (2022). doi: 10.1038/s41467-022-35627-1; pmid: 36550157
72. D. Pfau, S. Axelrod, H. Sutterud, I. von Glehn, J. S. Spencer, Code for "Accurate Computation of Quantum Excited States with Neural Networks," Zenodo (2024); https://doi.org/10.5281/zenodo.11937084.

ACKNOWLEDGMENTS

The authors thank M. Foulkes, G. Cassella, W. T. Lou, D. Jacquemin, M. Bearpark, A. Cohen, and A. Gaunt for helpful discussions; N. Gao for help with Folx; and J. Kirkpatrick, A. Obika, K. Lauterbach, A. Eslami, S. Mohamed, D. Rezende, E. D. Cubuk, P. Kohli, and D. Hassabis for support. **Funding:** D.P., I.v.G., and J.S.S. are all supported by Google DeepMind. S.A. was supported by a DeepMind internship during this project. H.S. is supported by the Aker Scholarship. H.S. gratefully acknowledges the Gauss Centre for Supercomputing e.V. (https://www.gauss-centre.eu/) for providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre

(JSC); the HPC RIVR consortium and EuroHPC JU for resources on the Vega high-performance computing system at IZUM, the Institute of Information Science in Maribor; and the UK Engineering and Physical Sciences Research Council for resources on the Baskerville Tier 2 HPC service. Baskerville was funded by the EPSRC and UKRI through the World Class Labs scheme (EP/T022221/1) and the Digital Research Infrastructure program (EP/W032244/1) and is operated by Advanced Research Computing at the University of Birmingham. **Author contributions:** D.P. conceived the project, wrote the code, ran the experiments, and wrote the manuscript. S.A. wrote and tested the code for computing density matrices and natural orbitals. H.S. wrote and tested the code for pseudopotentials. J.S.S. conceived and implemented the singlet targeting. I.v.G. and J.S.S. contributed to the code, and D.P., I.v.G., and J.S.S. wrote and maintained the library on which the code was based. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Data on molecular geometries and baseline values for all calculations in the paper are available in referenced publications. All other data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. Code for natural excited states, pseudopotentials, and observable operator evaluation has been added to the FermiNet GitHub repository (https://github.com/google-deepmind/ferminet) under an open source license. A snapshot of this repository is available on Zenodo (*72*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adn0137
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S10
References (*73*–*93*)

Submitted 17 November 2023; resubmitted 17 May 2024
Accepted 21 June 2024
10.1126/science.adn0137